AI-ImpactSK

Theoretical and Methodological Foundations of Artificial Intelligence Adoption Research

Structured Narrative Literature Reviews Supporting the AI-ImpactSK Project

Ekonóm
2026

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY

Publication Title: Theoretical and Methodological Foundations of Artificial Intelligence Adoption Research: Systematic Literature Reviews Supporting the AI-ImpactSK Project

Publication Category: Scientific Proceedings

Authors:

Ing. Federico Banda, Ing. Andriana Baziuk, doc. Ing. Vladimír Bolek, PhD. MBA., Ing. Adam Čardáš, Ing. Matej Černý, PhD., doc. Ing. Nora Grisáková, PhD., doc. Ing. Zuzana Hajduová, PhD., doc. Ing. Anna Hamranová, PhD., doc. Ing. Róbert Hanák, PhD., Ing. Vladimír Hojdík, PhD., Ing. Dana Hrušovská, PhD., Ing. Simona Kosztanko, PhD., Ing. Iveta Kufelová, PhD., PaedDr. Stanislava Matterová, Ing. Roman Tranžík, Ing. Miroslava Jozeková. Ing. Diana Pallérová, doc. Ing. Anita Romanová, PhD., Ing. Slávka Šagátová, PhD., Ing. Bc. Peter Štetka, PhD., Ing. Filip Stovíček, PhD., Ing. Mohammadreza Tatlari, Ing. Veronika Volfová

Reviewers:

doc. Ing. Mojmír Kokles, PhD.

Ing. Silvia Zelinová, PhD.

Edited by:

doc. Ing. Nora Grisáková, PhD., Ing. Peter Štetka, PhD., doc. RNDr. Zuzana Hajduová, PhD., Ing. Frederico Banda

Publisher: Ekonóm

Date of Publication: 31 March 2026

ISBN: 978-80-225-5312-4

License: CC BY 4.0 International

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). You are free to share and adapt the material for any purpose, provided appropriate credit is given. Full licence text: https://creativecommons.org/licenses/by/4.0/

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

© 2026 Authors. Published by Ekonóm Publishing House.

This publication was created as part of the AI-ImpactSK project: Empirical Examination of Artificial Intelligence Adoption and its Consequential Impact on Enterprise Performance, Market Dynamics, and Policy Formulation in the Slovak Economy. Funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under project No. 09I05-03-V02-00003/2025/VA.

FOREWORD

The dynamic development of artificial intelligence in recent years has fundamentally transformed the functioning of enterprises, public institutions, and society as a whole. Artificial intelligence can no longer be perceived solely as a technological phenomenon; it has evolved into a complex socio-economic construct influencing decision-making processes, organizational structures, market environments, and public policy formation. In this context, the systematic examination of AI adoption processes, together with the identification of their key determinants, becomes increasingly important.

The present scientific proceedings, *Theoretical and Methodological Foundations of Artificial Intelligence Adoption Research: Structured Narrative Literature Reviews Supporting the AI-ImpactSK Project*, represent a comprehensive collection of contributions focused on the theoretical and methodological foundations of artificial intelligence adoption research. The volume is conceived as a set of structured literature reviews grounded in secondary data analysis, which systematically synthesizes existing scientific knowledge and establishes a solid basis for subsequent empirical inquiry.

The content of the proceedings reflects the inherently multidisciplinary nature of the topic. The contributions address technology acceptance models, barriers and determinants of AI adoption, digital readiness, organizational and managerial aspects of implementation, as well as regulatory and policy-related issues. Particular emphasis is placed on methodological approaches which, through rigorous engagement with secondary data sources, enable the identification of research gaps, the formulation of hypotheses, and the conceptual advancement of the field.

This volume was developed within the framework of the AI-ImpactSK project, where secondary data analysis constitutes a key step in building its theoretical and methodological foundations. The insights presented in this publication therefore create essential preconditions for subsequent empirical research on AI adoption and its impacts on firm performance, market dynamics, and public policy in the context of the Slovak economy.

The publication is intended for the academic community, researchers, doctoral students, as well as practitioners and policymakers engaged in the field of digital transformation and artificial intelligence. We believe that this volume will contribute to advancing knowledge, fostering interdisciplinary dialogue, and supporting further research grounded in methodological rigor and societal relevance.

Finally, we would like to express our sincere gratitude to all authors and reviewers for their valuable contributions and cooperation in the preparation of this publication.

Editors

Bratislava, 2026

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

Contens

Technology Acceptance Models in AI Adoption: A Structured Literature Review and Research Agenda 5

Barriers and Drivers of Artificial Intelligence Adoption in Small and Medium-Sized Enterprises: A Systematic Literature Review 13

Structured Survey Design in Innovation Research: A Systematic Literature Approach of Methodological Foundations, Thematic Patterns, and Future Directions 26

Qualitative Methods in Technology Management Research: A Systematic Research Review 36

Best Practices in AI Implementation: A Synthesis of Case Studies Across Domains 46

Organizational Learning and AI Knowledge Transfer: A Systematic Literature Review of Emerging Paradigms and Future Directions 56

AI Policy Frameworks for National Competitiveness: A Literature Review 66

Government Interventions in Digital Transformation: A Structured Literature Review of Policies, Challenges, and Future Directions 77

The EU AI Act and Its Regulatory Implications for Businesses: A Systematic Literature Review .. 88

Cross-Country Comparison of Digital Economy Indicators: A Systematic Literature Review 97

The DESI Index as a Measure of Digital Readiness: A Systematic Literature Review of Concepts, Applications, and Future Directions 106

Bass Diffusion Model Theory and Applications: Literature Review of Foundational Frameworks, Extensions, and Emerging Research Frontiers 116

The ASSESSOR Model in New Product Forecasting: A Critical Literature Review of Methodological Foundations, Applications, and Emerging Challenges 128

Technology Diffusion in Emerging Economies: Patterns, Drivers, and Future Trajectories 139

AI Adoption and Firm Financial Performance: A Thematic Literature Review 149

Corporate performance metrics in the digital age: a systematic literature review of measurement frameworks, technological enablers, and strategic implications 163

Econometric Models for Technology Investment Analysis: A Narrative Literature Review 180

K-Means and Hierarchical Clustering in Business Research: A Systematic Synthesis of Research 191

Machine Learning Applications in Strategic Management: A Systematic Literature Review 201

Logistic Regression as a Predictive Framework in Technology Adoption Research: A Literature Review 211

Predictive Analytics for Innovation Management: A Structured Review of Machine Learning Approaches and Applications 224

A Review of the Firm-Level Determinants of Artificial Intelligence Adoption 242

Principal Component Analysis in Marketing Research: A Systematic Literature Review 253

Consumer Preferences for AI-Based Products and Services: A Systematic Research Synthesis .. 263

Sentiment Analysis Methods in Business Research: A Systematic Literature Review 273

Managerial Attitudes Toward Artificial Intelligence: Literature Review of Emerging Challenges, Opportunities, and Organizational Implications 284

Emotional Dimensions of Technology Resistance: A Thematic Literature Review 294

Open Innovation and Knowledge Dissemination in the Digital Era: A Thematic Literature Review 305

Capacity Building for Artificial Intelligence Adoption in Organizations: A Systematic Literature Review 316

Measuring Effectiveness of Knowledge Transfer Interventions: Literature Review 326

Evaluation of AI Applications in Logistics from the Perspective of ESG Criteria: A Scoping Review 337

Artificial Intelligence in Rural Businesses: A Structured Literature Review 351

Artificial Intelligence in Human Resource Management: A Structured Literature Review of Recruitment, Fairness, Talent Management, and Employee Experience 359

Artificial Intelligence in Customer Relationship Management: A Literature Review of Chatbots, Predictive Analytics, and Customer Loyalty Outcomes 372

Organizational Dynamics and Temporal Patterns in AI Adoption: An Analytical Synthesis of Leadership, Departmental Influence, and Implementation Sequencing 381

Technology Acceptance Models in AI Adoption: A Structured Literature Review and Research Agenda

Grisáková Nora [1] and Šagátová Slavka [2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: nora.grisakova@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: slavka.sagatova@euba.sk

Abstract

The rapid diffusion of artificial intelligence (AI) across organizational and consumer contexts has intensified scholarly interest in understanding how individuals and institutions decide to adopt such technologies. Technology acceptance models offer a theoretically grounded framework for examining these decisions, yet their application to AI-specific tools remains fragmented across disciplines. This structured literature review synthesizes 15 peer-reviewed studies to examine how technology acceptance frameworks have been applied in AI adoption research, with particular attention to the Technology Acceptance Model (TAM), the Unified Theory of Acceptance and Use of Technology (UTAUT), and the Technology–Organisation–Environment (TOE) framework. The analysis is organized around three thematic dimensions: individual-level determinants of AI acceptance, organizational and contextual drivers of adoption, and domain-specific adaptations of acceptance frameworks. The findings indicate that trust, perceived usefulness, and social influence consistently emerge as salient predictors, while sector-specific factors—including digital competence, digital leadership, educational context, and healthcare risk perception—shape adoption outcomes in important ways. The review also highlights key gaps related to longitudinal acceptance dynamics, ethical considerations, and cross-cultural generalizability, and outlines a corresponding research agenda. These insights carry direct implications for technology designers, organizational managers, and policymakers seeking to facilitate responsible AI integration.

Key Words: artificial intelligence, technology acceptance, AI adoption, TAM, UTAUT, structured literature review

JEL Classification: O33, M15

1. Introduction

Artificial intelligence has evolved from a speculative technology into a pervasive infrastructure that increasingly reshapes decision-making, communication, and service delivery across virtually every sector of the global economy. From conversational agents embedded in customer service operations to AI-powered diagnostic tools in healthcare and generative language models transforming educational practice, the scope of AI applications continues to expand at an unprecedented pace. Yet deployment does not automatically translate into adoption; the history of information technology offers many examples of sophisticated systems that were rejected, underused, or used in unintended ways by their target users. Understanding the factors that drive or inhibit the acceptance of AI technologies has therefore become a matter of both practical urgency and theoretical interest.

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

Technology acceptance research has a well-established intellectual lineage, originating in Davis's foundational Technology Acceptance Model and subsequently extended through frameworks such as UTAUT and TOE. These models have generated substantial cumulative knowledge about the cognitive, social, and structural determinants of technology use intentions and behaviors. However, AI introduces a set of distinctive attributes—including perceived anthropomorphism, machine autonomy, opacity of decision logic, and variable reliability—that challenge assumptions embedded in models developed for more conventional software tools. This mismatch between existing theoretical apparatus and the unique characteristics of AI systems remains an important issue in current research.

Despite a growing body of empirical work examining AI acceptance across specific domains—such as marketing and consumer psychology (Mariani et al., 2022), education (Chocarro et al., 2023; Galindo-Domínguez et al., 2024), healthcare (Choudhury & Shamszare, 2023; Shahsavar & Choudhury, 2023), construction (Na et al., 2022), and public administration (Neumann et al., 2023)—there has been only limited effort to synthesize these contributions into a coherent intellectual map. Individual studies tend to remain within disciplinary silos, adapting acceptance frameworks to local concerns without fully engaging with parallel developments in adjacent fields. This fragmentation limits the accumulation of generalizable knowledge and obscures the theoretical debates that merit more focused scholarly attention.

The objective of this structured literature review is threefold. First, it documents the theoretical models most frequently employed in AI acceptance research and assesses their conceptual adequacy. Second, it synthesizes empirical findings thematically in order to identify areas of consensus, contradiction, and productive debate. Third, it proposes a research agenda addressing the most consequential gaps in current understanding. In doing so, this review targets both researchers seeking to position new empirical contributions and practitioners looking to apply evidence-based insights to technology deployment decisions.

The remainder of the paper is organized as follows. Section 2 establishes the theoretical background, defining key constructs and introducing the foundational frameworks that structure AI acceptance research. Section 3 presents a thematic synthesis of the empirical literature. Section 4 identifies research gaps and outlines future directions. Section 5 offers concluding remarks.

2. Theoretical Background

This paper adopts a structured literature review approach to synthesize how technology acceptance frameworks have been applied in research on AI adoption. The aim is not to provide an exhaustive mapping of all available studies, but rather to identify and analytically organize conceptually relevant peer-reviewed research addressing the acceptance, adoption, or intended use of AI-based technologies. Particular attention is given to studies engaging with the Technology Acceptance Model (TAM), the Unified Theory of Acceptance and Use of Technology (UTAUT/UTAUT2), and the Technology–Organisation–Environment (TOE) framework.

The final sample consists of 15 peer-reviewed studies from different application domains, including education, healthcare, construction, public administration, marketing, and digital consumer contexts. The selected studies were examined through thematic analysis and grouped into three analytical clusters: (1) individual-level determinants of AI acceptance, (2) organizational and contextual drivers of adoption, and (3) domain-specific adaptations of acceptance frameworks. Although the review is not exhaustive and the thematic grouping

involves a degree of interpretive judgment, this approach provides a coherent basis for identifying recurring constructs, emerging AI-specific extensions, and key research gaps.

Foundations of Technology Acceptance Research

Technology acceptance research is anchored in a small set of theoretical models that have demonstrated enduring utility across multiple technological generations and disciplinary contexts. The Technology Acceptance Model (TAM), introduced by Davis in the late 1980s, posits that an individual's behavioral intention to use a technology is primarily determined by two cognitive perceptions: perceived usefulness (PU), defined as the degree to which a person believes that using the technology will enhance their performance, and perceived ease of use (PEOU), defined as the degree to which the technology is expected to be free of effort. These two constructs have been validated across thousands of empirical studies and continue to serve as a central theoretical foundation of technology acceptance research. In the context of AI adoption, Chocarro et al. (2023) explicitly apply TAM to the study of teachers' acceptance of educational chatbots, demonstrating that both PU and PEOU remain robust predictors even when the technology in question exhibits social and conversational characteristics not anticipated by the original model.

A more comprehensive theoretical architecture was later introduced through the Unified Theory of Acceptance and Use of Technology (UTAUT), which consolidated constructs from eight earlier models into four core determinants: performance expectancy, effort expectancy, social influence, and facilitating conditions. UTAUT was subsequently extended into UTAUT-2 to accommodate consumer contexts, adding hedonic motivation, price-saving orientation, and habit as additional predictors. Palau-Saumell et al. (2019) employ an expanded UTAUT-2 framework to examine mobile application adoption, providing empirical evidence that hedonic and habitual motivations hold explanatory power beyond the original model's instrumental constructs. More recently, Menon and Shilpa (2023) adapt UTAUT to the analysis of ChatGPT adoption, introducing perceived interactivity and privacy concerns as domain-relevant extensions that significantly enrich the model's explanatory scope.

Organizational and Environmental Frameworks

While TAM and UTAUT focus principally on individual-level cognitive processes, AI adoption in organizational settings requires frameworks capable of accommodating structural and environmental influences. The Technology–Organisation–Environment (TOE) framework addresses this need by conceptualizing adoption as a function of technological characteristics, organizational readiness, and external environmental pressures. Na et al. (2022) integrate TAM and TOE in their study of AI-based technologies in construction firms, illustrating the value of combining perceptual variables with organizational context. Similarly, Neumann et al. (2023) demonstrate that AI adoption in public organizations unfolds in stages influenced by internal capacities, regulatory uncertainty, and broader policy environments.

This organizational perspective is especially important in the context of AI because adoption decisions frequently involve collective actors, strategic investments, and institutional accountability. Unlike many earlier information technologies, AI systems often require not only user-level acceptance but also managerial endorsement, infrastructure readiness, data governance capacity, and tolerance for operational risk. As a result, purely individual-level models may capture only part of the broader causal architecture underlying AI uptake.

AI-Specific Extensions to Acceptance Theory

The distinctive characteristics of AI systems have prompted scholars to revisit and adapt conventional acceptance constructs. One recurring theme is trust, particularly in contexts where AI outputs are probabilistic, opaque, or consequential. Choudhury and Shamszare (2023) show that trust significantly shapes healthcare professionals' willingness to engage with ChatGPT-based systems, while Pitardi and Marriott (2021) demonstrate that anthropomorphic cues and perceived warmth influence consumer trust in voice-based AI assistants. These findings suggest that trust in AI is not reducible to traditional notions of system reliability, but may also involve social-cognitive evaluations more commonly associated with human interaction.

Another emerging line of research concerns interactivity and social presence. In studies of chatbots and conversational AI, users appear to respond not only to functional utility but also to whether the system feels responsive, socially competent, and contextually aware. Chocarro et al. (2023) find that the social language used by educational chatbots affects user attitudes, while Menon and Shilpa (2023) identify perceived interactivity as a key extension to UTAUT in explaining ChatGPT usage intention.

A third area of theoretical adaptation involves perceived risk and ethical concern. AI systems often raise questions about privacy, autonomy, fairness, and accountability, especially when deployed in domains involving sensitive data or high-stakes decisions. Although these dimensions remain underdeveloped in many empirical studies, they constitute a growing frontier in the attempt to explain why users may simultaneously perceive AI as useful and yet remain reluctant to adopt it.

3. Thematic Synthesis of the Literature

At the individual level, the reviewed literature consistently confirms the continued relevance of perceived usefulness and effort-related perceptions in explaining AI acceptance. Across educational settings, consumer applications, and healthcare-related uses, individuals are more likely to adopt AI systems when they believe those systems will improve task performance or provide meaningful support. Chocarro et al. (2023), Menon and Shilpa (2023), and Shahsavar and Choudhury (2023) each provide evidence that instrumental beliefs remain central, even where AI introduces novel relational or ethical concerns.

Trust emerges as an especially salient determinant in AI contexts. Unlike conventional software, AI systems often generate outputs that users cannot fully verify or explain, increasing reliance on heuristic judgments regarding credibility and intent. In healthcare, trust becomes intertwined with perceptions of diagnostic safety and professional responsibility; in consumer settings, it may be linked to perceived humanity, transparency, or brand legitimacy. The reviewed studies therefore suggest that trust should be treated not merely as an ancillary variable, but as a core component of AI acceptance theory.

User characteristics also matter. Recent educational AI research also suggests that digital competence is positively associated with teachers' attitudes toward AI, while familiarity with digital technologies remains relevant for acceptance-related outcomes (Galindo-Domínguez et al., 2024; Chocarro et al., 2023). These findings indicate that AI acceptance is filtered through broader user capabilities and dispositions, reinforcing the need to move beyond one-size-fits-all models.

Organizational and Contextual Drivers

The literature also makes clear that AI adoption is deeply embedded in organizational and contextual settings. Within firms, adoption depends not only on employee perceptions but also

on leadership support, strategic vision, technical infrastructure, and the perceived compatibility of AI systems with existing routines. Chatterjee et al. (2023) highlight the moderating role of digital leadership in shaping organizational performance in digitally transformed workplaces. Although the study does not focus directly on AI adoption, it offers a useful broader context for understanding the organizational conditions that may facilitate AI uptake.

Sector-specific context significantly shapes the salience of acceptance determinants. In healthcare, perceived risk and professional accountability temper enthusiasm for AI tools, while in education, pedagogical compatibility and concerns about student dependence become more central. Construction firms emphasize organizational readiness and technological fit, whereas public organizations face constraints related to legal legitimacy, procurement complexity, and policy uncertainty. These contextual differences caution against overly universalistic applications of acceptance models and support the case for domain-sensitive adaptation.

The TOE-based studies reviewed further underscore that environmental pressures—such as regulatory expectations, competitive dynamics, and ecosystem maturity—may exert substantial influence on adoption decisions. AI adoption is thus not only an outcome of beliefs and attitudes but also a strategic response to institutional conditions.

Domain-Specific Adaptations and Cross-Disciplinary Applications

One of the most striking patterns in the literature is the proliferation of domain-specific model extensions. In educational research, acceptance models are adapted to include pedagogical fit, student learning outcomes, and social language characteristics. In healthcare, trust, perceived diagnostic risk, and self-efficacy become central. In conversational commerce and voice assistant research, anthropomorphism, warmth, and perceived intelligence enter the explanatory framework. These adaptations are not merely additive; they reveal that the meaning of "acceptance" itself may differ across contexts.

At the same time, certain cross-domain regularities remain visible. Perceived usefulness, social influence, and trust repeatedly appear as influential variables across diverse settings. This suggests that although AI adoption is highly context-sensitive, some constructs retain broader explanatory value while others require contextual specification.

The reviewed literature also reveals an unresolved tension between continuity and novelty. On the one hand, classical acceptance models continue to offer a stable conceptual foundation. On the other hand, the distinctive features of AI systems—autonomy, opacity, interactional richness, and ethical sensitivity—consistently generate pressure for theoretical extension. This tension should not be seen as a weakness of the field, but rather as a sign of ongoing conceptual transition.

4. Research Gaps and Future Directions

A first major gap in the literature concerns the temporal dynamics of AI acceptance. Most reviewed studies rely on cross-sectional survey designs that capture user attitudes at a single point in time. Yet acceptance of AI may evolve as users gain experience, encounter failures, or recalibrate expectations. Longitudinal research would be especially valuable for distinguishing initial curiosity from sustained use and for tracing how trust, habit, and perceived usefulness evolve over time.

A second gap lies in the relative underdevelopment of ethical and normative dimensions. Ethical dimensions of AI acceptance represent a critically undertheorized domain. While

Choudhury and Shamszare (2023) and Shahsavar and Choudhury (2023) raise important concerns about the risks of over-reliance on ChatGPT in healthcare contexts, the broader question of how ethical perceptions—including concerns about algorithmic fairness, data privacy, and accountability—shape adoption decisions has received only peripheral attention. Menon and Shilpa (2023) introduce privacy concerns as an extended UTAUT construct, but this treatment remains theoretically underdeveloped. Future research should systematically examine how users' ethical perceptions of AI systems mediate or moderate established acceptance relationships, drawing on theoretical resources from technology ethics and responsible AI scholarship. This is particularly urgent given that AI systems increasingly operate in domains such as healthcare, education, and public administration, where ethical violations may carry serious societal consequences.

Cross-cultural generalizability constitutes a third major gap. The majority of empirical studies reviewed draw on samples from a limited range of national and cultural contexts—including Spain (Palau-Saumell et al., 2019), India (Menon & Shilpa, 2023), Switzerland (Neumann et al., 2023), and Greece (Kitsios et al., 2021)—and few make systematic comparisons across cultural boundaries. Given that AI adoption is a global phenomenon shaped by culturally variable attitudes toward technology, authority, privacy, and uncertainty, the cross-cultural portability of findings derived from TAM, UTAUT, and TOE cannot be assumed. Future research should employ multi-country comparative designs that explicitly theorize cultural moderators, drawing on frameworks such as Hofstede's cultural dimensions or GLOBE to generate culturally sensitive predictions about acceptance dynamics.

A fourth direction concerns the integration of acceptance and post-adoption behavior. Most acceptance frameworks treat behavioral intention as the terminal dependent variable, with few studies examining the relationship between acceptance and the quality, depth, or sustainability of actual AI use. Future research should move beyond intention to examine how acceptance translates into patterns of active engagement, habitual use, or strategic underutilization, while also considering the practical implications of these behavioral outcomes for the organizations and sectors in which AI is deployed. This extension would also facilitate closer engagement with the organizational performance literature, connecting technology acceptance research to questions of productivity, service quality, and innovation outcomes.

5. Conclusion

This structured literature review has synthesized 15 peer-reviewed studies to provide an analytically organized account of how technology acceptance frameworks have been applied, adapted, and extended in the context of AI adoption research. Three principal conclusions emerge from this synthesis.

First, foundational acceptance frameworks—TAM, UTAUT, and TOE—retain substantial explanatory value in AI adoption contexts, but they consistently require domain-specific extension in order to capture features of AI that conventional software lacks. Trust, anthropomorphism, perceived interactivity, privacy sensitivity, and digital competence all emerge as theoretically significant constructs that enrich the explanatory power of established models. Chocarro et al. (2023), Menon and Shilpa (2023), Balakrishnan and Dwivedi (2024), and Galindo-Domínguez et al. (2024) each illustrate how established models can be extended or complemented in adjacent AI-related contexts, though no unified framework has yet emerged.

Second, organizational and contextual factors exert a meaningful influence on AI adoption that individual-level models tend to underestimate. The staged adoption dynamics

documented by Neumann et al. (2023), the leadership moderation effects identified by Chatterjee et al. (2023), and the sectoral heterogeneity evident across construction (Na et al., 2022), banking (Kitsios et al., 2021), and healthcare (Choudhury & Shamszare, 2023; Shahsavar & Choudhury, 2023) collectively argue for multi-level theoretical frameworks that articulate the interplay between individual cognition and structural context.

Third, the field continues to face significant gaps in longitudinal research, ethical theorization, and cross-cultural generalization. Addressing these gaps will require methodological diversification, theoretical integration across disciplines, and stronger engagement with the ethical and societal dimensions of AI deployment.

This review is subject to certain limitations. The corpus of 15 papers, while selected to represent intellectual diversity across the field, does not constitute an exhaustive review of all available literature. Additionally, the review's thematic structure inevitably involves interpretive judgments that alternative reviewers might make differently. Despite these limitations, the review makes a meaningful contribution by identifying the theoretical frameworks, empirical patterns, and research priorities that may guide the next generation of scholarship on AI acceptance.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Balakrishnan, J., & Dwivedi, Y. K. (2024). Conversational commerce: Entering the next stage of AI-powered digital assistants. Annals of Operations Research, 333, 653–687. https://doi.org/10.1007/s10479-021-04049-5

Chatterjee, S., Chaudhuri, R., Vrontis, D., & Giovando, G. (2023). Digital workplace and organization performance: Moderating role of digital leadership capability. Journal of Innovation & Knowledge, 8(1), 100334. https://doi.org/10.1016/j.jik.2023.100334

Chocarro, R., Cortiñas, M., & Marcos-Matás, G. (2023). Teachers' attitudes towards chatbots in education: A technology acceptance model approach considering the effect of social language, bot proactiveness, and users' characteristics. Educational Studies, 49(2), 295–313. https://doi.org/10.1080/03055698.2020.1850426

Choudhury, A., & Shamszare, H. (2023). Investigating the impact of user trust on the adoption and use of ChatGPT: Survey analysis. Journal of Medical Internet Research, 25, e47184. https://doi.org/10.2196/47184

Galindo-Domínguez, H., Delgado, N., Campo, L., & Losada, D. (2024). Relationship between teachers' digital competence and attitudes towards artificial intelligence in education. International Journal of Educational Research, 128, 102381. https://doi.org/10.1016/j.ijer.2024.102381

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. https://doi.org/10.3390/joitmc7030204

Mariani, M. M., Perez-Vega, R., & Wirtz, J. (2022). AI in marketing, consumer research and psychology: A systematic literature review and research agenda. Psychology & Marketing, 39(4), 755–776. https://doi.org/10.1002/mar.21619

Menon, D., & Shilpa, K. (2023). "Chatting with ChatGPT": Analyzing the factors influencing users' intention to use the OpenAI's ChatGPT using the UTAUT model. Heliyon, 9(11), e20962. https://doi.org/10.1016/j.heliyon.2023.e20962

Na, S., Heo, S., Han, S., Shin, Y.-S., & Roh, Y.-S. (2022). Acceptance model of artificial intelligence (AI)-based technologies in construction firms: Applying the Technology Acceptance Model (TAM) in combination with the Technology–Organisation–Environment (TOE) framework. Buildings, 12(2), 90. https://doi.org/10.3390/buildings12020090

Neumann, O., Guirguis, K., & Steiner, R. (2023). Exploring artificial intelligence adoption in public organizations: A comparative case study. Public Management Review, 26(1), 114–141. https://doi.org/10.1080/14719037.2022.2048685

Nicolescu, L., & Tudorache, M. T. (2022). Human-computer interaction in customer service: The experience with AI chatbots—A systematic literature review. Electronics, 11(10), 1579. https://doi.org/10.3390/electronics11101579

Palau-Saumell, R., Forgas-Coll, S., Sánchez, J., & Robres, E. (2019). User acceptance of mobile apps for restaurants: An expanded and extended UTAUT-2. Sustainability, 11(4), 1210. https://doi.org/10.3390/su11041210

Patel, V., Chesmore, A., Legner, C. M., & Pandey, S. (2022). Trends in workplace wearable technologies and connected-worker solutions for next-generation occupational safety, health, and productivity. Advanced Intelligent Systems, 4(1), 2100099. https://doi.org/10.1002/aisy.202100099

Pitardi, V., & Marriott, H. R. (2021). Alexa, she's not human but… Unveiling the drivers of consumers' trust in voice-based artificial intelligence. Psychology & Marketing, 38(4), 626–642. https://doi.org/10.1002/mar.21457

Shahsavar, Y., & Choudhury, A. (2023). User intentions to use ChatGPT for self-diagnosis and health-related purposes: Cross-sectional survey study. JMIR Human Factors, 10, e47564. https://doi.org/10.2196/47564

Barriers and Drivers of Artificial Intelligence Adoption in Small and Medium-Sized Enterprises: A Systematic Literature Review

Hamranová Anna[1] and Štetka Peter[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: anna.hamranova@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: peter.stetka@euba.sk

Abstract

This paper presents a systematic literature review focusing on the barriers and drivers of artificial intelligence (AI) adoption in small and medium-sized enterprises (SMEs). Based on the PRISMA 2020 methodology, the study synthesizes empirical and conceptual insights from peer-reviewed publications published in Europe between 2021 and 2025, addressing technological, organizational, and strategic determinants of AI implementation. The findings indicate that SMEs face persistent constraints, including high investment costs, a lack of digital skills, insufficient data quality, infrastructural limitations, and cultural resistance to change. Despite these barriers, several enabling factors have been identified, including efficiency gains, competitive pressure, customer expectations, leadership support, digital maturity, and access to external innovation ecosystems. The review also reveals significant divergence among authors regarding the relative importance of specific barriers, the role of big data, the significance of workforce capabilities, and the relevance of ethical considerations. Several research gaps are identified, particularly the lack of longitudinal studies, limited attention to sustainability, and insufficient examination of human-related factors such as trust and explainability. Overall, the study provides an integrated framework that deepens the understanding of AI adoption dynamics in SMEs and outlines implications for future research, policymaking, and managerial practice.

Key Words: artificial intelligence, SME barriers, adoption drivers, technology adoption, digital transformation

JEL Classification: O33, L25, M15

1. Introduction

The rapid proliferation of artificial intelligence technologies has fundamentally altered the competitive landscape across virtually every sector of the global economy. From automated decision-making systems to machine learning algorithms capable of processing vast datasets in real time, AI has emerged as a general-purpose technology with the potential to augment and, in some cases, replace human cognitive and physical labor at scale (Dwivedi et al., 2019). While large enterprises have been at the forefront of AI integration, leveraging significant capital resources and specialized talent pools to operationalize these technologies, small and medium-sized enterprises occupy a far more precarious position in this evolving landscape. SMEs constitute the backbone of most national economies, accounting for the majority of employment and a substantial share of economic output across developed and developing nations alike, yet their engagement with transformative digital technologies such as AI remains inconsistent, fragmented, and poorly understood.

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

The significance of this gap extends beyond academic interest. SMEs that fail to adopt AI risk falling behind in productivity, customer responsiveness, and innovation capacity relative to larger, better-resourced competitors. Conversely, those that successfully integrate AI into their operations stand to gain disproportionate competitive advantages through enhanced efficiency, new venture opportunities, and improved market positioning (Chalmers et al., 2020). Understanding what enables or impedes this adoption process is therefore a matter of considerable practical and theoretical urgency.

Despite a growing body of literature on AI adoption at the enterprise level and an extensive discourse on digital transformation more broadly (Kraus et al., 2021), relatively few studies have systematically examined the specific challenges and enablers that characterize SME contexts. Much of the existing research conflates SMEs with large organizations or treats firm size as merely a control variable rather than a substantively important moderating factor. This represents a meaningful gap in the literature, given that the organizational, financial, and human resource constraints faced by SMEs are qualitatively distinct from those encountered by large corporations.

The remainder of this article is organized as follows. Section 2 establishes the theoretical background by defining core concepts and situating the review within relevant theoretical frameworks. Section 3 presents methodology of systematic literature review , Section 4 presents results of systematic literature review including research gaps and Section 5 proposes future directions. Section 6 concludes with a summary of contributions and an acknowledgment of the review's limitations.

2. Theoretical Background

Theoretical background consists of the characteristics of artificial intelligence in the context of the enterprise, digital transformation as a theoretical framework, the characteristics of human-centered AI and organizational readiness, digital literacy as an enabling construct and sustainability and competitive advantage.

2.1 Defining Artificial Intelligence in an Organizational Context

Artificial intelligence, as a field and as a set of technologies, encompasses a remarkably broad range of tools, systems, and methodologies. Dwivedi et al. (2019) offer a foundational multidisciplinary perspective, describing AI as a set of technologies capable of performing tasks that, when performed by humans, would require intelligence — encompassing perception, reasoning, learning, and autonomous action. In organizational contexts, AI typically manifests through machine learning algorithms, natural language processing systems, robotic process automation, and predictive analytics platforms, among other configurations. Critically, Dwivedi et al. (2019) argue that AI is not merely a productivity-enhancing tool but a genuinely transformative technology with implications for organizational structure, competitive strategy, and the nature of work itself.

This transformative framing is echoed in the literature on entrepreneurship, where Chalmers et al. (2020) characterize AI as a foundational element of the Fourth Industrial Revolution — a technological paradigm shift capable of reshaping how new ventures are created, how markets are structured, and how firms generate and capture value. For SMEs specifically, this framing is simultaneously promising and daunting: the same technologies that enable radical innovation may also impose new demands that smaller organizations are structurally ill-equipped to meet.

2.2 Digital Transformation as a Theoretical Frame

A dominant theoretical lens through which AI adoption in SMEs can be understood is that of digital transformation. Kraus et al. (2021) define digital transformation as a process by which digital technologies create disruptions, triggering strategic responses from organizations seeking to alter their value creation paths. This definition is relevant to AI adoption because it positions AI not as an isolated tool but as an integral component of a broader organizational transformation process. Importantly, Kraus et al. (2021) note that digital transformation entails not only technological change but also shifts in culture, capabilities, and organizational design — dimensions that are particularly challenging for resource-constrained SMEs.

Correani et al. (2020) extend this perspective by emphasizing the critical role of strategy implementation in digital transformation. Examining three large-scale digital transformation projects, these authors identify a persistent disconnection between strategy formulation and execution, a problem that is arguably more acute in SMEs where strategic planning capacity is often limited. Their framework highlights the need for firms to develop dynamic capabilities — the ability to sense opportunities, seize them, and reconfigure internal resources accordingly — a requirement that maps directly onto the AI adoption challenge in SMEs.

Brunetti et al. (2020) further contribute to this theoretical foundation by adopting a multi-stakeholder perspective on digital transformation challenges. Their analysis of a regional innovation system reveals that effective digital transformation requires coordinated effort across companies, educational institutions, and public administrators — a systemic view that underscores the importance of external ecosystems in enabling SME-level technology adoption.

2.3 Human-Centered AI and Organizational Readiness

Beyond the organizational strategy literature, the concept of human-centered AI provides a second important theoretical strand. Garibay et al. (2023) articulate six grand challenges for AI development, emphasizing the need for AI systems to be ethical, fair, explainable, and beneficial to human users. For SMEs, the implications are twofold: on the demand side, employees and managers must develop the capacity to work alongside AI systems that they can trust and interpret; on the supply side, AI developers must create tools that are accessible and comprehensible to non-specialist users. Both dimensions are directly relevant to the adoption problem.

2.4 Digital Literacy as an Enabling Construct

Digital literacy represents a third theoretical pillar of this review. Reddy et al. (2020) define digital literacy as the set of abilities, competencies, and skills required to function effectively in a technology-mediated environment. In the context of AI adoption, digital literacy encompasses not only basic ICT proficiency but also higher-order capacities for evaluating, selecting, and applying AI tools to specific organizational challenges. The absence of adequate digital literacy among SME owners, managers, and employees constitutes one of the most frequently cited barriers to AI adoption, and Reddy et al. (2020) provide a theoretical grounding for understanding both its dimensions and its developmental pathways.

2.5 Sustainability and Competitive Advantage

Finally, the relationship between AI adoption and sustainability-oriented competitive advantage is theorized by Martínez-Peláez et al. (2023), who examine how micro, small, and medium-sized enterprises can leverage digital transformation to achieve sustainability goals.

Their framework positions AI-enabled digital transformation as a strategic resource that allows SMEs to attract socially conscious customers and investors, reduce operational costs, and enhance organizational reputation — all of which constitute drivers of adoption when framed from a strategic management perspective.

3. Systematic Literature Review – Methodology

The aim of this study is to synthesize the existing scientific literature, identify the main barriers and drivers of AI adoption in enterprises—with a particular focus on SMEs—and provide an overview of the current state of knowledge in this field.

The systematic literature review was conducted in accordance with the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (PRISMA, 2020). The objective was to identify, evaluate, and synthesize both empirical and conceptual studies addressing the barriers and drivers (enablers) of AI adoption in SMEs.

In line with the PRISMA methodology, the main research question guiding this study was formulated as follows:

"What barriers and drivers influence the adoption of artificial intelligence in small and medium-sized enterprises?"

For the systematic analysis of scientific sources, the Web of Science database was used. To ensure a focused and structured selection of relevant literature, the PCC framework (Population–Concept–Context) was applied, alongside clearly defined inclusion and exclusion criteria.

The PCC framework was defined as follows:

Population: Small and medium-sized enterprises (SMEs), as defined by the respective authors in individual studies

Concept: Adoption, implementation, use, or readiness for artificial intelligence technologies (including machine learning, natural language processing, computer vision, predictive analytics, etc.)

Context: Organizational, technological, economic, strategic, and regulatory environments of business practice

The inclusion criteria for selecting studies were: focus on SMEs, AI adoption, barriers and drivers/enablers, publication period 2021–2025, peer-reviewed journal articles, English language, and European context.

The exclusion criteria included: studies not primarily focused on SMEs (e.g., large firms without SME-specific analysis), studies addressing digital transformation without explicit reference to AI, non-peer-reviewed sources (e.g., reports, blogs, media articles), inaccessible full texts, and duplicate records.

4. Systematic Literature Review - Results

The adoption of artificial intelligence (AI) in small and medium-sized enterprises (SMEs) represents a complex process shaped by a wide range of internal and external factors. Based on the analyzed sources, these factors can be categorized into key barriers and drivers, while also allowing for the identification of areas of scholarly consensus, divergence, and existing research gaps.

The process of literature selection is illustrated in Figure 1.

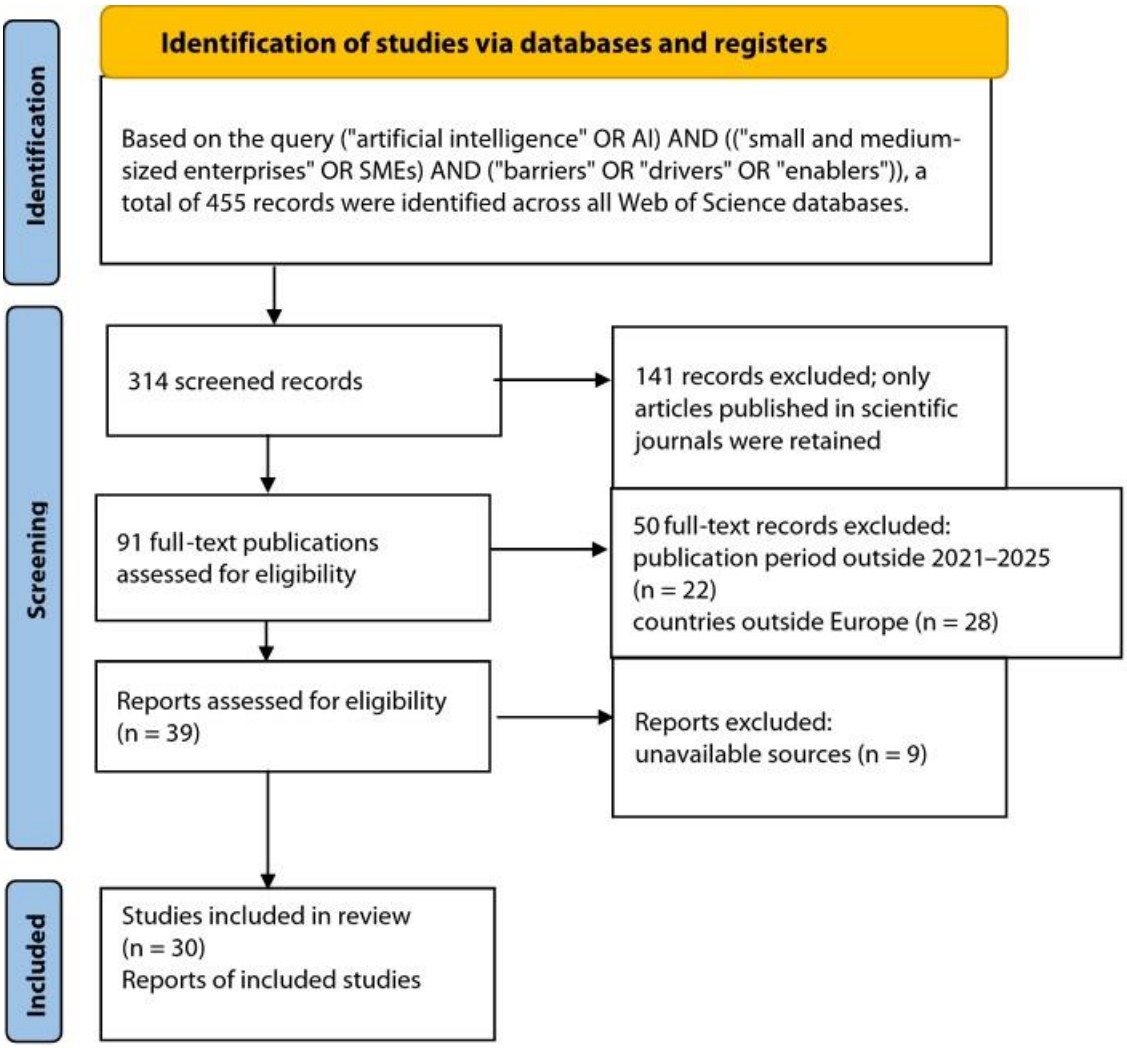


Figure 1: PRISMA flow diagram

4.1 Barriers to AI Implementation in SMEs

The main barriers to AI adoption in SMEs are closely linked to their limited resources and specific risk profiles. One of the most significant obstacles is the presence of high costs and financial constraints. Several authors indicate that SMEs perceive the substantial initial investments required for technologies, hardware, and maintenance as a critical challenge (Oldemeyer et al., 2025; Molina-Abril et al., 2025; Sanchez et al., 2025; Mousa & Bouraoui, 2025). Furthermore, uncertainty regarding return on investment leads managers to view AI implementation as excessively risky (Oldemeyer et al., 2025; Bąk et al., 2024).

Another major barrier is the lack of expertise and digital skills. The literature highlights that SMEs often lack access to technical specialists and that their employees typically exhibit low levels of digital literacy (Silva, 2022; Apostoaie et al., 2025; Bąk et al., 2024; Oldemeyer et al., 2025; Sanchez et al., 2025). In addition, SMEs face difficulties competing with large enterprises in attracting and retaining highly skilled professionals (Huseyn et al., 2024; Molina-Abril et al., 2025).

Data-related challenges also represent a critical barrier. AI systems require large volumes of high-quality, structured data, which SMEs often struggle to collect, manage, and analyse effectively (Oldemeyer et al., 2025; Bąk et al., 2024; Tawil et al., 2024; Jafarzadeh et al., 2025). Implementation is further hindered by resistance rooted in organizational culture and concerns related to change. Employee fears of job displacement and reluctance to adopt new practices are frequently identified as significant barriers (Oldemeyer et al., 2025; Bąk et al., 2024; Mousa & Bouraoui, 2025).

Technical complexity and outdated infrastructure also play an important role. The integration of AI into existing systems requires technically demanding adjustments and modernization, which poses a considerable challenge for SMEs (Oldemeyer et al., 2025; Bąk et al., 2024). Finally, legal and ethical concerns, including limited awareness of data protection, regulatory requirements, and ethical risks associated with AI use, may significantly discourage adoption (Bąk et al., 2024; Oldemeyer et al., 2025).

4.2 Drivers/enablers of AI adoption in SMEs

Despite these barriers, several factors motivate and facilitate AI adoption in SMEs. One of the most prominent drivers is the pursuit of increased efficiency and productivity, particularly through the automation of routine tasks, acceleration of decision-making processes, and optimization of operational costs (Oldemeyer et al., 2025; Bąk et al., 2024; Gabelaia & Hendieh, 2025).

Competitive pressure is another significant driver, as firms adopt AI technologies to remain competitive and keep pace with market innovations (Apostoaie et al., 2025; Cimino et al., 2025; Bąk et al., 2024). Customer expectations also play a crucial role, especially in terms of service personalization, improved customer experience, and faster communication, for example through chatbots, which further increases demand for AI solutions (Apostoaie et al., 2025; Bąk et al., 2024).

Leadership support and strategic vision are critical success factors. The commitment of top management, openness to innovation, and the ability to recognize the long-term strategic benefits of AI implementation significantly facilitate adoption (Apostoaie et al., 2025; Bąk et al., 2024; Cimino et al., 2025).

The level of digital maturity within the organization is also crucial. Firms already utilizing advanced digital technologies, such as cloud computing or the Internet of Things, are generally better prepared for AI adoption and can integrate these technologies more rapidly (Arroyabe et al., 2024; Bąk et al., 2024).

Another important enabler is access to external support ecosystems. Collaboration with universities, technology partners, and the use of public funding or digital innovation hubs significantly facilitates AI implementation in SMEs (Apostoaie et al., 2025; Bąk et al., 2024). Additionally, startups and scale-ups benefit from higher levels of agility, flexibility, and innovation orientation, which contributes to higher adoption rates compared to traditional SMEs (Segarra-Blasco et al., 2025).

Overall, despite numerous barriers, SMEs possess several enabling factors that support AI adoption. The combination of technological readiness, competitive pressures, customer expectations, and both internal and external support creates an environment in which AI implementation becomes not only feasible but increasingly strategically necessary.

4.3 Areas of Significant Scholarly Consensus in Research on AI Adoption in SMEs

Across the literature, authors demonstrate substantial agreement in several key areas concerning the adoption of artificial intelligence in small and medium-sized enterprises. They consistently point not only to similar barriers and motivating factors identified in the preceding sections, but also to shared assumptions regarding the conditions that shape AI implementation and to recurring shortcomings in the existing body of research on this topic.

Most authors identify a similar set of obstacles that hinder SMEs from using AI technologies effectively. The most frequently cited barrier is high cost and financial constraint, as AI implementation requires substantial initial investment in technologies, infrastructure, and specialized capabilities, which many firms find difficult to bear (Oldemeyer et al., 2025; Sanchez et al., 2025; Mousa & Bouraoui, 2025). Authors also emphasize that SMEs often suffer from a lack of specialized expertise. The absence of internal technical experts and the generally low level of digital literacy among employees represent some of the most common barriers to adoption (Bąk et al., 2024; Oldemeyer et al., 2025; Sanchez et al., 2025). Other repeatedly identified problems include low data quality and limited data availability, as SMEs typically do not possess a sufficient volume of high-quality, structured data required to train artificial intelligence models (Oldemeyer et al., 2025; Bąk et al., 2024; Jafarzadeh et al., 2025). The adoption process is also significantly slowed by cultural factors, including resistance to change, employee fears of being replaced by technology, and a broader lack of innovation-oriented culture within the organization (Bąk et al., 2024; Mousa & Bouraoui, 2025).

Authors also agree that the motivation of SMEs to adopt AI is often closely tied to their efforts to ensure long-term survival and support firm growth. These enterprises seek to enhance operational efficiency through the automation of routine tasks, process optimization, and cost reduction (Oldemeyer et al., 2025; Trunina et al., 2025; Gabelaia & Hendieh, 2025). Competitive and market pressure is also identified as a major stimulus, pushing firms to adopt AI technologies in order to respond to changing customer needs and market dynamics (Apostoaie et al., 2025; Bąk et al., 2024; Cimino et al., 2025). Another commonly cited motivating factor is the ambition to improve customer experience, with AI tools such as chatbots and personalized marketing solutions enabling firms to better understand and serve customer needs (Apostoaie et al., 2025; Trunina et al., 2025).

There is also broad agreement among authors that not all SMEs possess the same preconditions for successful AI implementation. Firms with higher levels of digital maturity, particularly those already experienced in using other advanced digital technologies, generally demonstrate greater readiness for AI adoption (Arroyabe et al., 2024; Segarra-Blasco et al., 2025). Similarly, the importance of leadership support is consistently emphasized, with the engagement and strategic vision of top management identified as a critical success factor (Apostoaie et al., 2025; Bąk et al., 2024). The likelihood of adoption is also significantly influenced by firm size and age, as larger medium-sized enterprises and younger, more agile firms such as startups and scale-ups tend to adopt AI technologies more rapidly than micro-enterprises or older organizations (Huseyn et al., 2024; Segarra-Blasco et al., 2025).

Finally, authors share a critical view of the current state of research. They note that the majority of existing studies on artificial intelligence have focused primarily on large corporations, while the specific needs and constraints of SMEs have long been overlooked (Oldemeyer et al., 2025; Apostoaie et al., 2025). At the same time, they emphasize that SMEs constitute the backbone of the economy, and that their successful transformation in line with Industry 4.0 principles is crucial for maintaining competitiveness at both the national and international levels (Dvořáková et al., 2021; Oldemeyer et al., 2025).

In most cases, authors agree on the principal barriers and motivating factors that shape AI adoption in SMEs. The common denominator across their findings is the need to overcome financial, technical, data-related, and cultural barriers, while successful implementation depends on the level of digital readiness, the availability of relevant skills, and managerial support. Authors also stress that, despite the growing importance of SMEs, these firms remain less extensively studied in the academic literature than large enterprises, thereby creating space for further knowledge development and a more precise understanding of their specific needs.

4.4 Areas of Significant Divergence in the Literature on AI Adoption in SMEs

Although the literature on artificial intelligence adoption in small and medium-sized enterprises demonstrates considerable agreement in many areas, authors also diverge on several key issues, highlighting the complexity of the phenomenon under investigation. One of the most prominent areas of disagreement concerns the identification of the most critical barrier to adoption. While several authors consider the lack of knowledge and specialized skills to be the primary obstacle (Kumar & Kalse, 2021; Ulrich et al., 2022), other studies emphasized the high costs associated with AI implementation (Barton et al., 2022). Additional research points to challenges related to user acceptance of technology (Joerg & Carlos, 2022) or insufficient data quality and availability (Bauer et al., 2020). This variability in perspectives suggests that the relative importance of different barriers is highly context-dependent and may vary according to firm size, industry, or organizational environment.

Divergence is also evident in discussions concerning data availability and the role of so-called big data. Some authors argue that SMEs suffer from a significant lack of data, which must be addressed prior to the implementation of AI systems (Putnik et al., 2021). In contrast, other researchers suggest that the existing level of data within SMEs may be sufficient, particularly when firms utilize external training datasets or adopt modelling approaches that function effectively with smaller volumes of proprietary data (Ingalagi et al., 2021). These differing interpretations reflect variations in modelling approaches and expectations regarding system performance and accuracy.

Disagreement among authors also emerges in assessing the role of employee capabilities in successful AI adoption. While the majority of the literature assumes that digital and technical skills are essential prerequisites for implementation (Almashawreh et al., 2024), some empirical studies do not confirm this assumption (Apostoaie et al., 2025). One possible explanation is that certain firms opt to outsource AI solutions rather than develop internal capabilities, thereby bypassing the need for internal training and skills development. Further differences concern the importance of costs and technical complexity.

Although these factors are frequently identified as major barriers, their statistically significant impact on adoption decisions has not been confirmed in some contexts, such as among Romanian SMEs (Apostoaie et al., 2025). This suggests that firms may identify alternative ways to mitigate these challenges or may prioritize other factors in their decision-making processes.

Significant divergence is also observed in the evaluation of AI adoption's impact on sustainable performance. While its contribution to technological, economic, and collaborative performance is generally assessed positively (Cimino et al., 2025), a direct relationship between AI adoption and sustainable performance has not been consistently confirmed in empirical models. This indicates that the effects of AI on environmental and social outcomes may be more complex and less predictable.

Finally, authors differ in their assessment of the importance of ethics, trust, and transparency in the SME context. While these issues represent central themes in discussions concerning large corporations, they receive comparatively limited attention in SME-focused research (Joerg & Carlos, 2022). It remains unclear whether this is due to the early stage of AI adoption among SMEs—where ethical concerns have not yet become salient—or whether such issues are objectively less relevant for smaller firms (Oldemeyer et al., 2025).

The differences outlined above highlight areas that remain underexplored or require deeper theoretical and empirical investigation. Accordingly, the following Section 4.5 focuses on identifying research gaps and outlining directions for future research on AI adoption in small and medium-sized enterprises.

4.5 Research Gaps

The existing literature on artificial intelligence adoption in small and medium-sized enterprises consistently highlights several significant research gaps that hinder a comprehensive understanding of the conditions, processes, and impacts of AI implementation within this segment. One of the most frequently identified gaps is the general lack of research explicitly focused on SMEs. As noted by Oldemeyer et al. (2025), the majority of existing studies concentrate on large corporations, while the specific characteristics and needs of smaller firms remain largely overlooked. Similarly, Bąk et al. (2024) emphasize the absence of systematic research comparing the conditions of AI adoption between SMEs and large enterprises, while Fülöp & Topor (2025) points out that the challenges and opportunities faced by SMEs in the context of digital transformation remain insufficiently explored.

Research gaps are also evident in the area of AI adoption drivers. Fülöp & Topor (2025) highlights the need for a deeper understanding of how a culture of technological innovation acts as a moderating factor in the relationship between implemented technologies and firms' ESG (Environmental, Social, and Governance) performance. Apostoaie et al. (2025) further draw attention to the geographical imbalance in the literature, noting that most studies are conducted in developed economies, while the context of developing regions, including Eastern Europe, remains significantly underrepresented. Mousa and Bouraoui (2025) add that there is still an insufficient understanding of how barriers influence the "optimal sequencing" of capability development within firms, particularly during periods of crisis.

Another area characterized by substantial gaps concerns barriers to AI adoption. Oldemeyer et al. (2025) note the absence of studies systematically examining the specific requirements of SMEs in the application of machine learning. The same authors also identify a lack of research focused on measuring return on investment (ROI) specifically for AI solutions in SME environments, representing a critical practical and theoretical gap. Furthermore, inconsistency persists regarding the most appropriate indicators for defining and categorizing SMEs in the context of AI, whether based on number of employees, turnover, or level of digital maturity.

Significant gaps are also present at the methodological level. Oldemeyer et al. (2025) emphasize the need to extend the analysis of AI adoption barriers to additional business functions, such as accounting, quality management, and human resource management, which may significantly influence AI utilization. Arroyabe et al. (2024) highlights the lack of longitudinal studies, as the majority of existing research is cross-sectional and therefore unable to capture the dynamic evolution of AI adoption over time. Apostoaie et al. (2025) further note that research rarely compares actual firm behavior—namely, the real extent of AI adoption—with declared intentions, representing a significant methodological limitation with implications for the interpretation of findings.

5. Implications for Future Research

The most important implication for future research is the need to explicitly focus on SMEs, as the majority of existing studies analyze large enterprises and therefore fail to adequately capture the specific financial, organizational, and technological constraints faced by smaller firms. Equally important is the systematic comparison of SMEs with large enterprises in order to identify differences in AI adoption approaches, barriers, and expected benefits. A third key direction involves a deeper examination of the role of innovation culture, which may significantly influence the relationship between AI implementation and firms' ESG (Environmental, Social, and Governance) performance. A fourth critical aspect is the inclusion of the context of developing economies, particularly Eastern Europe, where firms operate under different conditions and where insights derived from developed countries cannot be directly generalized. Finally, greater attention should be devoted to the long-term analysis of AI adoption through longitudinal studies, which would enable a better understanding of dynamic changes, the evolution of barriers, and the real impacts of AI over time.

Additional important implications include the need to examine the influence of crisis periods on the sequencing and pace of capability development within firms; to analyze specific barriers related to machine learning adoption in SMEs; to develop robust methodologies for measuring the return on investment (ROI) of AI initiatives; to standardize the criteria used to define SMEs in the context of digital transformation; to extend the study of barriers to specific business functions such as accounting, human resource management, and quality management; and to systematically compare the gap between firms' declared intentions and their actual behavior in adopting AI technologies.

6. Conclusion

AI adoption in small and medium-sized enterprises represents a multidimensional and dynamic process shaped by a combination of technological, organizational, economic, and cultural factors. The systematic literature review indicates that SMEs face persistent barriers, most notably high investment costs, a lack of specialized digital skills, low data quality and limited data availability, outdated infrastructure, and resistance rooted in organizational culture. These obstacles significantly hinder their ability to effectively integrate AI technologies into everyday business processes.

Despite these constraints, the analysis also identifies several key drivers that support AI adoption. These include the pursuit of increased efficiency, competitive pressure, and rising customer expectations, as well as strong leadership support and higher levels of digital maturity within firms. External innovation ecosystems, such as technology partners, universities, and support programs, also play a positive role. Together, these factors create a business environment in which AI implementation becomes not only feasible but often strategically necessary.

The findings further highlight areas of strong scholarly consensus, particularly regarding the dominant role of financial, technical, data-related, and cultural barriers in slowing AI adoption. At the same time, significant differences exist in how authors assess the relative importance of specific barriers, the role of data availability and skills, and the extent to which AI implementation contributes to sustainable performance. These divergences confirm that contextual factors—such as industry, firm size, level of digitalisation, and economic environment—play a crucial role in shaping SME decision-making processes.

The review also identifies several research gaps that must be addressed to achieve a more comprehensive understanding of AI adoption in SMEs. These include the lack of studies specifically focused on SMEs, the absence of systematic comparisons with large enterprises, geographical imbalances in the literature, and insufficient attention to the long-term impacts of AI. Additional gaps persist in the measurement of return on investment, the definition and classification of SMEs in the context of AI, and the analysis of discrepancies between firms' declared intentions and their actual behavior.

Overall, the study demonstrates that while SMEs face significant challenges, they also possess increasing opportunities for effective AI adoption. Advancing knowledge in the identified areas is essential not only for academia but also for policymakers and managers seeking to support the digital transformation of SMEs and to leverage AI as a source of competitive advantage.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Almashawreh, R. E., Talukder, M., Charath, S. K., & Khan, M. I. (2024). AI adoption in Jordanian SMEs: The influence of technological and organizational orientations. Global Business Review, 09721509241250273.

Apostoaie, C. M., Roman, T., Maxim, A., & Jijie, D. T. (2025). Determinants of AI adoption intention in SMEs: Romanian case study. Journal of Business Economics and Management, 26(2), 277–296.

Arroyabe, M. F., Arranz, C. F., de Arroyabe, I. F., & de Arroyabe, J. C. F. (2024). Analyzing AI adoption in European SMEs: A study of digital capabilities, innovation, and external environment. Technology in Society, 79, 102733.

Bąk, S., Jedynak, P., & Kaczmaryk, P. (2024). Adaptation determinants of artificial intelligence in small and medium enterprises. European Management Studies, 22(1), 76–97.

Barton, M., Budjac, R., Tanuska, P., Gaspar, G., & Schreiber, P. (2022). Identification overview of industry 4.0 essential attributes and resource-limited embedded artificial-intelligence-of-things devices for small and medium-sized enterprises. Applied Sciences, 12(11), 5672.

Bauer, M., van Dinther, C., & Kiefer, D. (2020). Machine learning in SME: an empirical study on enablers and success factors.

Brunetti, F., Matt, D. T., Bonfanti, A., De Longhi, A., & Pedrini, 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

Chalmers, D., MacKenzie, N., & Carter, S. (2020). Artificial intelligence and entrepreneurship: Implications for venture creation in the Fourth Industrial Revolution. Entrepreneurship Theory and Practice, 45(5), 1028–1053. https://doi.org/10.1177/1042258720934581

Cimino, A., Corvello, V., Troise, C., Thomas, A., & Tani, M. (2025). Artificial intelligence adoption for sustainable growth in SMEs: An extended dynamic capability framework. Corporate Social Responsibility and Environmental Management, 32(5), 6120-6138.

Correani, A., De Massis, A., Frattini, F., & Messeni Petruzzelli, A. (2020). Implementing a digital strategy: Learning from the experience of three digital transformation projects. California Management Review, 62(4), 37–56. https://doi.org/10.1177/0008125620934864

Dvořáková, L., Horák, J., Caha, Z., Machová, V., Hašková, S., Rowland, Z., & Krulicky, T. (2021). Adaptation of small and medium-sized enterprises in the service sector to the conditions of Industry 4.0 and Society 4.0: Evidence from the Czech Republic. Economic Annals-XXI, 191.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., & Coombs, C. (2019). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Fülöp, M. T., & Topor, D. I. (2025). The digital transformation of SMEs: opportunities and challenges for sustainability reporting in Romania. Journal of Global Responsibility.

Gabelaia, I., & Hendieh, J. (2025). The transformative power of AI and its impact on business strategy, financial operations, and marketing decision-making: A case study method. International Journal of Innovation Science, 1–21.

Garibay, O. O., Winslow, B., Andolina, S., Antona, M., & Bodenschatz, A. (2023). Six human-centered artificial intelligence grand challenges. International Journal of Human-Computer Interaction, 39(3), 391–436. https://doi.org/10.1080/10447318.2022.2153320.

Huseyn, M., Ruiz-Gandara, A., González-Abril, L., & Romero, I. (2024). Adoption of artificial intelligence in small and medium-sized enterprises in Spain: The role of competences and skills. Amfiteatru Economic, 26(67), 848–866.

Ingalagi, S. S., Mutkekar, R. R., & Kulkarni, P. M. (2021). Artificial Intelligence (AI) adaptation: Analysis of determinants among Small to Medium-sized Enterprises (SME's). In IOP Conference Series: Materials Science and Engineering (Vol. 1049, No. 1, p. 012017). IOP Publishing.

Jafarzadeh, P., Vähämäki, T., Nevalainen, P., Tuomisto, A., & Heikkonen, J. (2025). Supporting SME companies in mapping out AI potential: A Finnish AI development case. The Journal of Technology Transfer, 50(3), 1016–1035.

Joerg, G., & Carlos, J. (2022). Design Framework for the Implementation of AI-based (Service) Business Models for Small and Medium-sized Manufacturing Enterprises (Mar, 10.1007/s13132-022-01003-z, 2022). JOURNAL OF THE KNOWLEDGE ECONOMY, 14(3), 3570-3571.

Kraus, S., Kanbach, D. K., Krysta, P. M., Steinhoff, M. M., & Tomini, N. (2022). Facebook and the creation of the metaverse: Radical business model innovation or incremental transformation? International Journal of Entrepreneurial Behaviour & Research, 28(9), 52–77. https://doi.org/10.1108/ijebr-12-2021-0984.

Kumar, A., & Kalse, A. (2021). WITHDRAWN: Usage and adoption of artificial intelligence in SMEs.

Martínez-Peláez, R., Ochoa-Brust, A., Rivera Manrique, S. I., & Velarde-Alvarado, P. (2023). Role of digital transformation for achieving sustainability. Sustainability, 15(14), 11221. https://doi.org/10.3390/su151411221.

Molina-Abril, G., Calvet, L., Juan, A. A., & Riera, D. (2025). Strategic decision-making in SMEs: A review of heuristics and machine learning for multi-objective optimization. Computation, 13(7), 173.

Mousa, S., & Bouraoui, T. (2025). Staging digital capability activation for innovation resilience: A dynamic capability perspective on sequencing AI, big data, IoT and blockchain. International Journal of Organizational Analysis, 1–25.

Oldemeyer, L., Jede, A., & Teuteberg, F. (2025). Investigation of artificial intelligence in SMEs: A systematic review of the state of the art and the main implementation challenges. Management Review Quarterly, 75(2), 1185–1227.

PRISMA. (2020). PRISMA 2020 statement. https://www.prisma-statement.org/prisma-2020

Putnik, G. D., Shah, V., Putnik, Z., & Ferreira, L. (2021). Machine learning in cyber-physical systems and manufacturing singularity–it does not mean total automation, human is still in the centre: part II–in-CPS and a view from community on industry 4.0 impact on society. Journal of Machine Engineering, 21(1), 133-153.

Reddy, P., Sharma, B., & Chaudhary, K. (2020). Digital literacy: A review of literature. International Journal of Technoethics, 11(2), 1–21. https://doi.org/10.4018/ijt.20200701.oa1

Sánchez, E., Calderón, R., & Herrera, F. (2025). Artificial intelligence adoption in SMEs: Survey based on TOE–DOI framework, primary methodology and challenges. Applied Sciences, 15(12), 6465.

Segarra-Blasco, A., Tomàs-Porres, J., & Teruel, M. (2025). AI, robots and innovation in European SMEs: A. Segarra-Blasco. Small Business Economics, 65(1), 719-745.

Silva, A. G. C. (2022). Práticas de letramentos com o uso de tecnologias digitais: vozes a partir de professores/pesquisadores (Master's thesis, Universidade Regional de Blumenau (Brazil)).

Tawil, A. R. H., Mohamed, M., Schmoor, X., Vlachos, K., & Haidar, D. (2024). Trends and challenges towards effective data-driven decision making in UK Small and Medium-sized Enterprises: Case studies and lessons learnt from the analysis of 85 Small and Medium-sized Enterprises. Big Data and Cognitive Computing, 8(7), 79.

Trunina, I., Pryakhina, K., Bilyk, M., & Moroz, O. (2025). AI as a digital transformation tool for competitive business development. Marketing and Management of Innovations, 16(3), 15–27. https://doi.org/10.21272/mmi.2025.3-02.

Ulrich, P. S., Timmermann, A., & Frank, V. (2022). Organizational aspects of cybersecurity in German family firms–Do opportunities or risks predominate? Organizational Cybersecurity Journal: Practice, Process and People, 2(1), 21-40.

Structured Survey Design in Innovation Research: A Systematic Literature Approach of Methodological Foundations, Thematic Patterns, and Future Directions

Hanák Róbert[1] and Grisáková Nora[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: robert.hanak@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: nora.grisakova@euba.sk

Abstract

Structured survey research has become an increasingly prominent methodology across innovation studies, yet the field lacks a consolidated synthesis of its methodological foundations and application patterns. This literature review examines the theoretical underpinnings and practical dimensions of structured survey design as deployed across diverse innovation-related domains, drawing on a corpus of fifteen peer-reviewed sources. The review identifies three principal thematic clusters: the role of perception-based survey instruments in capturing stakeholder orientations toward emerging challenges, the methodological challenges of ensuring sampling adequacy and instrument validity, and the integration of consensus-building and multi-round approaches to refine research findings. Key findings indicate that survey design in innovation research is frequently driven by technology acceptance frameworks, sustainability imperatives, and behavioral intention models, each of which imposes distinct demands on questionnaire architecture and sampling strategy. However, significant gaps remain in the cross-contextual standardization of instruments, longitudinal design adoption, and the reconciliation of expert-elicitation approaches with large-scale quantitative surveys. Implications for researchers and practitioners are discussed, with attention to how methodological rigor can be strengthened to improve the generalizability and actionability of innovation survey findings.

Key Words: structured survey design, innovation research methodology, questionnaire development, quantitative methods, empirical research

JEL Classification: C83, O31, M20

1. Introduction

Survey-based research occupies a foundational position in the study of innovation, offering a systematic means by which researchers may capture among others the attitudes, intentions, behaviors, and perceptions of individuals and organizations navigating complex, rapidly evolving environments. From investigations into corporate responses to sustainability pressures, to studies of technology adoption in educational and industrial settings, the structured survey has proved to be an extraordinarily versatile instrument. Yet precisely because of this versatility, the field risks methodological fragmentation: different research communities have developed distinct conventions for instrument design, sampling procedures, and analytical frameworks, often without adequate dialogue and level of agreement across disciplinary boundaries.

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

The significance of this fragmentation should not be underestimated. As innovation research increasingly intersects with domains such as digital transformation, and organizational behavior, the comparability and replicability of survey-based findings become critical concerns. For example, institutional investors assessing climate-related financial risks, firms evaluating the implementation of Industry 4.0 and industry 5.0 technologies, and educators appraising the uptake of digital learning platforms all represent populations whose complex orientations must be measured with precision and rigor. When survey instruments are constructed without reference to validated theoretical frameworks, or when sampling strategies are selected without systematic justification, the resulting evidence base is compromised in ways that may not be immediately apparent to consumers of that research.

A further concern relates to the underexplored tension between large-scale, standardized survey instruments and expert-elicitation methodologies such as the Delphi technique. Both approaches have legitimate claims to generate actionable knowledge in innovation contexts, yet they rest on quite different epistemological assumptions and impose different demands on respondents and researchers.. A review that brings these approaches into dialogue can illuminate where they complement one another and where their divergence reflects genuine theoretical disagreement rather than mere methodological preference.

The objective of this literature review is therefore threefold. First, it aims to present the theoretical frameworks that have been most influential in shaping structured survey design within innovation research. Second, it synthesizes methodological findings of selected body of recent literature across three thematic clusters: perception and risk assessment in innovation contexts; technology acceptance and behavioral intention modeling; and methodological rigor, sampling, and consensus processes. Third, it identifies persistent gaps in the literature and proposes specific future research directions intended to advance both methodological and substantive understanding.

The remainder of this article is organized as follows. Section 2 presents the theoretical background, defining key constructs and situating the review within established conceptual traditions. Section 3 offers a thematically organized synthesis of the extant literature. Section 4 identifies research gaps and future directions. Section 5 concludes with a summary of key contributions and a candid acknowledgment of the review's limitations.

2. Theoretical Background

2.1 Defining Structured Survey Research in Innovation Contexts

Structured survey research, in the context of innovation studies, refers to the systematic collection of standardized data from a defined population through the use of fixed-format questionnaires, scaled response items, and pre-specified analytical protocols. Unlike qualitative interview studies or ethnographic approaches, structured surveys prioritize comparability, reproducibility, and statistical generalizability. Their application within innovation research is motivated by the need to measure constructs—such as perceived usefulness, risk tolerance, behavioral intention, or organizational readiness—that are inherently latent and therefore cannot be directly observed.

A central theoretical challenge in this domain concerns the operationalization of complex, multidimensional constructs. Theoreticcal constructs such as "technology acceptance" or "climate risk perception" are not directly observable; they must be inferred from patterns of response across carefully designed indicator items. This requirement situates structured survey design squarely within the tradition of psychometric measurement theory, which demands

attention to the validity such as construct validity, convergent and discriminant validity, also to the reliability, and finally to the measurement invariance across subgroups.

2.2 Technology Acceptance and Behavioral Intention Frameworks

The Technology Acceptance Model (TAM) and its successors occupy an important position in the innovation survey literature (Davis, 1989). Scherer, Siddiq, and Tondeur (2018) conducted a comprehensive meta-analytic structural equation modeling study demonstrating that TAM's core constructs—perceived usefulness and perceived ease of use—remain robust predictors of technology adoption across diverse educational contexts, even after accounting for moderating variables such as teacher experience and technology type. This finding has been widely influential because it suggests that TAM-derived survey instruments can achieve a degree of cross-contextual transferability that many other theoretical frameworks cannot.

Building upon TAM, the Unified Theory of Acceptance and Use of Technology (UTAUT) incorporates additional constructs including performance expectancy, effort expectancy, social influence, and facilitating conditions (Venkatesh et al., 2003). Chao (2019) extended the UTAUT framework to the domain of mobile learning, adding moderators such as perceived enjoyment, mobile self-efficacy, and trust, and demonstrating that this enriched model explains a substantially greater proportion of variance in behavioral intention than the original formulation. The implication for survey design is significant: richer theoretical models demand more complex instruments, which in turn raise important questions about respondent burden, instrument parsimony, and the risk of construct proliferation.

2.3 Perception-Based Surveys and Risk Assessment

A second foundational tradition within innovation survey research concerns the use of structured instruments to capture stakeholder perceptions of risk and opportunity. Krueger, Sautner, and Starks (2019) provide an illustration of this approach in their survey of institutional investors regarding climate risk perceptions. By combining closed-ended Likert-scale items with structured open-ended follow-up probes, their instrument was able to capture not only the prevalence of risk awareness but also the qualitative texture of investor concerns and the mechanisms through which perceived risks translate into portfolio decisions. Their finding that regulatory risks are perceived as more immediate than physical climate risks has important implications for how future surveys structure their risk typologies and response scales.

Müller, Kiel, and Voigt (2018) similarly employed a structured survey to investigate the drivers of Industry 4.0 implementation, situating their instrument within a theoretical framework that distinguishes between opportunity-driven and challenge-driven motivations. Their differentiated treatment of company size, industry sector, and value chain position as moderating variables demonstrates how theoretical sophistication can be embedded directly into survey architecture through stratified sampling and moderated regression analysis.

2.4 Consensus Methods and Expert Elicitation

A third theoretical tradition relevant to structured survey design in innovation research involves iterative, consensus-building methodologies, most notably the Delphi technique. Niederberger and Spranger (2020) provide a systematic mapping of Delphi applications in the health sciences that is broadly applicable to innovation research, demonstrating that the technique is particularly valuable when the knowledge base is incomplete, when stakeholder perspectives are heterogeneous, and when the goal is not merely description but the

identification of actionable consensus. The Delphi approach differs from conventional cross-sectional surveys in that it involves multiple rounds of structured questioning, with aggregated feedback from earlier rounds informing the refinement of items and response options in subsequent rounds.

This iterative logic has important implications for the design of structured surveys in emerging innovation domains, where the very constructs of interest may not yet be sufficiently well understood to permit confident a priori operationalization. Sovacool, Axsen, and Sorrell (2018) argue persuasively that innovation social science suffers from a tendency toward premature methodological standardization, which forecloses the possibility of genuinely novel inquiry. Their prescription—that researchers invest more carefully in the conceptual foundations of their research designs before committing to particular instruments or analytic strategies—applies with particular force to survey-based work in rapidly evolving innovation contexts.

2.5 Sample Size and Statistical Power

Finally, the question of sample size determination constitutes a critical yet frequently undertheorized dimension of structured survey design. Memon, Ting, Cheah, Ramayah, and Chuah (2020) provide a comprehensive review of sample size guidelines in survey research, demonstrating that the field is characterized by a proliferation of competing rules of thumb—including minimum case-to-variable ratios, minimum absolute sample sizes, and power-analytic approaches—which are often applied without adequate justification. Their recommendation that sample size decisions be grounded in formal a priori power analysis, with explicit reference to the complexity of the structural model, the expected effect size, and the acceptable rates of Type I and Type II error, represents an important standard against which the methodological quality of innovation surveys can be assessed.

3. Literature Review

3.1 Perception, Risk, and Stakeholder Orientation in Innovation Surveys

One of the most extensively documented applications of structured survey methodology in innovation research concerns the measurement of how different stakeholder groups perceive and respond to novel risks and opportunities (citácia). This strand of literature is unified by its concern with the gap between objective risk exposure and subjective risk perception, and by its use of structured instruments to map the contours of this gap across heterogeneous populations.

Krueger et al. (2019) make an important contribution to this theme through their large-scale survey of institutional investors regarding climate-related financial risks. Their findings reveal a striking pattern: the majority of surveyed investors believe that climate risks carry material financial implications for their portfolio companies, and that regulatory risks in particular have already begun to affect valuations. Importantly, however, this widespread acknowledgment of risk does not translate uniformly into portfolio action; many investors, particularly those with shorter time horizons or smaller asset bases, report that the costs and uncertainties associated with climate-responsive investment strategies outweigh the perceived benefits. This finding has significant implications for survey design, as it highlights the importance of distinguishing between attitudinal and behavioral outcomes—a distinction that is frequently collapsed in less carefully constructed instruments.

Müller et al. (2018) extend this theme into the domain of digital industrial transformation, examining how manufacturing firms perceive and respond to the opportunities and challenges associated with Industry 4.0 implementation. Their survey reveals that opportunity-driven motivations—including efficiency gains, new business model development, and supply chain optimization—are consistently rated as more important than challenge-related concerns such as cybersecurity risks and workforce displacement, though this pattern is moderated significantly by firm size and sector. Larger firms and those operating in more technology-intensive sectors are more likely to adopt a comprehensive, opportunity-oriented framing, while smaller firms and those in traditional sectors exhibit a more risk-averse, challenge-driven orientation. This finding underscores the importance of theoretically motivated sampling stratification in innovation surveys: without deliberate attention to firm-level heterogeneity, aggregate findings may obscure substantively important subgroup differences. It can also explain small firms lagging larger ones in the AI adoption.

The measurement of perception in social media research contexts further illustrates this theme. Kapoor, Tamilmani, Rana, Patil, and colleagues (2017) synthesize a substantial body of research demonstrating that users' perceptions of social media platforms—particularly their assessments of platform utility, information credibility, and social connectivity—are powerful determinants of engagement behavior and innovation diffusion. Crucially, however, they note that the constructs used to operationalize these perceptions vary substantially across studies, making cumulative synthesis difficult. This finding resonates with the broader methodological argument advanced by Sovacool et al. (2018) that innovation social science would benefit from greater terminological and conceptual standardization, without sacrificing the flexibility needed to capture contextually specific phenomena.

The cross-sectional survey conducted by Dost, Hossain, Shehab, Abdelwahed, and Al-Nusair (2020) among UK medical students during the COVID-19 pandemic illustrates how large-scale perception surveys can generate actionable findings in contexts of rapid innovation-driven disruption. Their survey of 2,721 students regarding perceptions of online teaching documented significant variation in satisfaction, perceived learning effectiveness, and concerns about professional development, with important implications for the design and delivery of digitally mediated education. While the healthcare education context may appear distant from core innovation research concerns, the methodological lessons are directly transferable: the study's use of stratified national sampling, validated attitudinal scales, and multivariate analysis of subgroup differences represents a methodological template that could be productively applied in corporate innovation surveys.

3.2 Technology Acceptance Modeling and Behavioral Intention in Survey-Based Innovation Research

The second major thematic cluster in the literature concerns the application of technology acceptance and behavioral intention frameworks to the study of innovation adoption. This body of work is distinguished by its commitment to theory-driven instrument design, its reliance on structural equation modeling as the primary analytic strategy, and its engagement with debates about model specification, construct validity, and the appropriate inclusion of moderating and mediating variables.

Scherer et al. (2018) provide the most methodologically comprehensive treatment of this theme, synthesizing findings from a large corpus of TAM-based studies through meta-analytic structural equation modeling. Their analysis confirms that perceived usefulness and perceived ease of use are robust predictors of behavioral intention across diverse technology adoption

contexts, but also identifies significant heterogeneity in effect sizes that is attributable to moderating variables including the type of technology, the level of mandatory versus voluntary use, and the demographic characteristics of the user population. For survey designers, this finding carries a clear message: the selection and weighting of predictor constructs should be informed by the specific characteristics of the population and technology under investigation, rather than by mechanical application of a standard model.

Chao (2019) complements this finding by demonstrating how the UTAUT framework can be extended to capture the motivational complexity of mobile learning adoption. His survey-based study reveals that constructs not included in the original UTAUT model—including perceived enjoyment, mobile self-efficacy, and perceived risk—contribute unique explanatory variance to behavioral intention, particularly among younger and more digitally experienced user groups. This finding has important implications for the design of innovation surveys in rapidly evolving technology environments, where the psychological and social dimensions of adoption may shift faster than established theoretical frameworks can accommodate.

The behavioral intention literature also raises important questions about the temporal structure of survey-based research. Both Scherer et al. (2018) and Chao (2019) rely on cross-sectional data, which limits their ability to establish the causal direction of relationships between constructs and to track how adoption intentions translate into actual behavior over time. This limitation is widely acknowledged in the technology acceptance literature but has yet to prompt a systematic shift toward longitudinal or experience-sampling designs, a gap that represents a significant opportunity for methodological innovation.

Kapoor et al. (2017), in their review of social media research, add a further dimension to this theme by documenting how the motivational antecedents of technology adoption vary across cultural and institutional contexts. Their synthesis of 132 papers reveals that constructs such as social norms, privacy concerns, and institutional trust—which play a secondary role in the original TAM and UTAUT formulations—become considerably more important in non-Western and organizationally heterogeneous contexts. This finding reinforces the argument that survey instruments designed for one cultural or institutional context should not be deployed in another without systematic validation, a point that has particular relevance for the globalized context of contemporary innovation research.

3.3 Methodological Rigor, Sampling Strategy, and Consensus Building

The third identified thematic cluster addresses the procedural and methodological dimensions of structured survey design, with particular attention to the challenges of achieving adequate sampling, ensuring measurement quality, and selecting appropriate analytic strategies. This literature is marked by a sustained tension between the pragmatic demands of field research—which often preclude ideal sampling conditions and impose constraints on instrument length and complexity—and the methodological standards required for findings to be considered scientifically credible.

Memon et al. (2020) offer the most systematic treatment of sample size determination, cataloguing the various rules of thumb that have been proposed in the survey research literature and demonstrating that their application is frequently inconsistent and insufficiently theorized. They argue that the widespread reliance on minimum-case-to-variable ratios and absolute minimum thresholds—such as the oft-cited recommendation that surveys achieve a minimum of 100 to 200 responses—is problematic because these heuristics do not account for the complexity of the structural model, the reliability of the constructs, or the expected effect size. Their prescription for formal a priori power analysis represents a methodological standard

that would substantially improve the quality of innovation surveys, but its adoption requires a level of statistical sophistication that is not universally available among innovation researchers.

Sovacool et al. (2018) approach methodological rigor from a broader perspective, arguing that innovation social science suffers from a series of interrelated weaknesses that extend beyond sample size to encompass research question formulation, conceptual engagement, and the quality of writing and argumentation. Their code of practice for appropriate research methods includes specific recommendations for survey-based work, including the need to justify instrument choices with reference to the research question and theoretical framework, to report non-response rates and potential non-response biases, and to triangulate survey findings with complementary data sources wherever feasible. These recommendations are broadly consistent with established standards in survey methodology but take on particular force in the innovation research context, where methodological shortcuts are often rationalized by the novelty and complexity of the phenomena under investigation.

Niederberger and Spranger (2020) contribute an important complementary perspective through their mapping of Delphi applications, demonstrating that iterative consensus-building approaches can address some of the limitations of conventional cross-sectional surveys, particularly in domains where the relevant constructs are not yet well established. Their review identifies significant heterogeneity in Delphi protocols across studies, including variation in the number of rounds, the composition of expert panels, the criteria for consensus, and the methods used to aggregate and feed back responses. This heterogeneity undermines the comparability of Delphi-based findings and suggests that the field would benefit from more standardized reporting guidelines, analogous to those developed for observational epidemiology by Skrivankova and colleagues (2021) through the STROBE-MR initiative.

The connection to reporting standards is further reinforced by Patricios and colleagues (2023), whose consensus statement on concussion in sport illustrates how rigorous, multi-round expert consultation processes can be organized and reported in a way that is both transparent and reproducible. Although the substantive domain is distinct from innovation research, the methodological architecture—including the use of systematic reviews to inform expert deliberations, the pre-registration of consensus criteria, and the transparent reporting of dissenting views—represents a template that could be productively adapted for consensus-building surveys in innovation contexts.

It is also worth noting that the methodological challenges of survey design are not confined to quantitative instruments. Dost et al. (2020) demonstrate that large-scale national surveys can achieve high methodological quality even within constrained timeframes, provided that instrument development, sampling, and analysis are conducted with appropriate rigor. Their study's use of validated scales, systematic non-response analysis, and transparent reporting of limitations provides a useful model for innovation researchers who must balance methodological ideals against the practical demands of field research in dynamic contexts.

4. Research Gaps and Future Directions

Despite the substantial selected body of literature reviewed in the preceding sections, several important gaps remain that warrant attention from researchers and practitioners alike. In our research, we identified several gaps that should be addressed.

The first and most pressing gap concerns the lack of longitudinal survey designs in innovation research. The overwhelming majority of the studies reviewed in this article are cross-sectional in nature, capturing attitudes, perceptions, and intentions at a single point in time.

While cross-sectional surveys are efficient and well-suited to the descriptive mapping of population characteristics, they are fundamentally limited in their ability to track how innovation-related orientations evolve over time, how intentions translate into behaviors, and how the feedback loops between adoption, experience, and attitude revision operate in practice. Future research should invest in panel-based survey designs that follow the same respondents across multiple time points, enabling longitudinal analysis of innovation adoption trajectories and the identification of critical transition points. Scherer et al. (2018) and Chao (2019) both acknowledge the limitations of cross-sectional designs within their respective frameworks, and the field as a whole would benefit from a systematic shift toward temporal depth.

A second gap concerns the cross-cultural and cross-institutional validity of dominant survey instruments. As Kapoor et al. (2017) demonstrate, the motivational antecedents of technology adoption and innovation engagement vary substantially across cultural and institutional contexts, yet the instruments most widely used in the field—including TAM-derived scales and risk perception batteries—were developed and validated primarily in Western, organizationally homogeneous settings. This raises serious questions about the validity of findings generated by these instruments when applied in different cultural or institutional contexts. Future research should systematically investigate the measurement invariance of established innovation survey instruments across diverse cultural, sectoral, and organizational contexts, drawing on multigroup structural equation modeling and cross-cultural validation protocols to assess where adaptation is necessary.

A third gap concerns the integration of structured survey methods with complementary data sources and methodological traditions. Sovacool et al. (2018) argue forcefully for methodological pluralism in innovation social science, noting that the field's predominant reliance on single-method designs—whether quantitative surveys or qualitative case studies—limits the complexity and nuance of the knowledge it generates. Future research should develop mixed-methods designs that combine the statistical power and generalizability of large-scale structured surveys with the contextual richness of qualitative approaches, including interviews, ethnographic observation, and document analysis. The Delphi technique, as systematically reviewed by Niederberger and Spranger (2020), represents one important bridge between expert elicitation and quantitative survey methods, and its integration with conventional survey designs in innovation research deserves greater methodological attention.

A related practical implication concerns the need for standardized reporting guidelines for structured survey studies in innovation research, analogous to the STROBE-MR checklist developed by Skrivankova et al. (2021) for epidemiological studies. Such guidelines would enhance the transparency, comparability, and reproducibility of survey-based findings, and would provide a shared methodological language that could facilitate cumulative knowledge synthesis across innovation research subfields.

5. Conclusion

This literature review has synthesized a selected and methodologically diverse body of research bearing on structured survey design in innovation contexts, organized around three principal thematic clusters: perception and risk assessment, technology acceptance and behavioral intention modeling, and methodological rigor and consensus-building processes. The review makes several contributions to the existing literature.

First, it demonstrates that the theoretical frameworks most influential in shaping innovation survey design—including the TAM, UTAUT, and various risk perception models—are well validated within their original domains of application but face important challenges of cross-contextual transferability. The finding that constructs such as social norms, perceived risk, and institutional trust become more important as theoretical sophistication increases has important implications for instrument design across diverse innovation contexts and across diverse countries (Chao, 2019; Kapoor et al., 2017; Scherer et al., 2018).

Second, the review highlights that methodological rigor in innovation survey research remains uneven, with particular weaknesses in sample size justification, non-response bias analysis, and the longitudinal tracking of innovation-related attitudes and behaviors. The recommendations advanced by Memon et al. (2020) and Sovacool et al. (2018) provide a coherent framework for addressing these weaknesses, but their adoption requires deliberate investment in methodological capacity building across the field.

Third, the review identifies the Delphi technique and related expert-elicitation approaches as underutilized complements to conventional cross-sectional surveys in innovation research (Niederberger & Spranger, 2020). Their integration into mixed-methods designs represents a promising avenue for future work.

Several limitations of this review should be acknowledged. The corpus of sources is small and spans multiple disciplinary domains, which limits the depth of engagement with any single subfield of innovation research. Additionally, the review is necessarily selective in its thematic organization and may not capture all relevant methodological debates within the field. These limitations notwithstanding, the review offers a valuable synthesis of cross-disciplinary insights into the design, conduct, and reporting of structured surveys in innovation research, and provides a foundation for future methodological development.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Chao, C.-M. (2019). Factors determining the behavioral intention to use mobile learning: An application and extension of the UTAUT model. Frontiers in Psychology, 10, 1652. https://doi.org/10.3389/fpsyg.2019.01652

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly, 13(3), 319-340.

Dost, S., Hossain, A., Shehab, M., Abdelwahed, A., & Al-Nusair, L. (2020). Perceptions of medical students towards online teaching during the COVID-19 pandemic: A national cross-sectional survey of 2721 UK medical students. BMJ Open, 10(11), e042378. https://doi.org/10.1136/bmjopen-2020-042378

Kapoor, K. K., Tamilmani, K., Rana, N. P., Patil, P. P., Dwivedi, Y. K., & Nerur, S. (2017). Advances in social media research: Past, present and future. Information Systems Frontiers, 20(3), 531–558. https://doi.org/10.1007/s10796-017-9810-y

Krueger, P., Sautner, Z., & Starks, L. T. (2019). The importance of climate risks for institutional investors. Review of Financial Studies, 33(3), 1067–1111. https://doi.org/10.1093/rfs/hhz137

Memon, M. A., Ting, H., Cheah, J.-H., Ramayah, T., & Chuah, F. (2020). Sample size for survey research: Review and recommendations. Journal of Applied Structural Equation Modeling, 4(2), i–xx. https://doi.org/10.47263/jasem.4(2)01

Müller, J. M., Kiel, D., & Voigt, K.-I. (2018). What drives the implementation of Industry 4.0? The role of opportunities and challenges in the context of sustainability. Sustainability, 10(1), 247. https://doi.org/10.3390/su10010247

Niederberger, M., & Spranger, J. (2020). Delphi technique in health sciences: A map. Frontiers in Public Health, 8, 457. https://doi.org/10.3389/fpubh.2020.00457

Patricios, J., Schneider, K., Dvořák, J., Ahmed, O. H., Blauwet, C., & colleagues. (2023). Consensus statement on concussion in sport: The 6th International Conference on Concussion in Sport–Amsterdam, October 2022. British Journal of Sports Medicine, 57(11), 695–711. https://doi.org/10.1136/bjsports-2023-106898

Scherer, R., Siddiq, F., & Tondeur, J. (2018). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. Computers & Education, 128, 13–35. https://doi.org/10.1016/j.compedu.2018.09.009

Skrivankova, V., Richmond, R. C., Woolf, B., Davies, N. M., Higgins, J. P. T., & colleagues. (2021). Strengthening the reporting of observational studies in epidemiology using Mendelian randomisation (STROBE-MR): Explanation and elaboration. BMJ, 375, n2233. https://doi.org/10.1136/bmj.n2233

Sovacool, B. K., Axsen, J., & Sorrell, S. (2018). Promoting novelty, rigor, and style in energy social science: Towards codes of practice for appropriate methods and research design. Energy Research & Social Science, 45, 12–42. https://doi.org/10.1016/j.erss.2018.07.007

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view1. MIS quarterly, 27(3), 425-478.

Qualitative Methods in Technology Management Research: A Systematic Research Review

Peter Štetka[1] and Dana Hrušovská[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: peter.stetka@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: dana.hrusovska@euba.sk

Abstract

Qualitative methods have become indispensable tools in technology management research, offering nuanced insights into complex organizational, social, and technological phenomena that quantitative approaches alone cannot adequately capture. This review examines how qualitative methodologiesincluding semi-structured interviews, thematic analysis, mixed-methods designs, and emerging computational techniqueshave been applied across technology management contexts, from digital transformation and human-AI collaboration to remote work and educational technology adoption. Drawing on twenty peer-reviewed sources, the study identifies three principal thematic clusters: (1) the design and rigor of qualitative data collection instruments, (2) analytical frameworks for interpreting technology-mediated experiences, and (3) the integration of qualitative methods within mixed-methods and computational research paradigms. Key findings indicate that while semi-structured interviews and thematic analysis remain dominant, researchers increasingly integrating inductive and deductive approaches and incorporate computational text analysis to manage large qualitative datasets. Significant methodological gaps persist, particularly in relation to longitudinal qualitative designs, cross-cultural validity, and the systematic integration of stakeholder perspectives in rapidly evolving technological environments. These findings carry important implications for the methodological advancement of technology management as a research discipline.

Key Words: case study method, grounded theory, qualitative research, technology management, , thematic analysis,

JEL Classification: O32, M15, C18

1. Introduction

Technology management has expanded sygnificantly over the past two decades, driven by the accelerating pace of digital transformation, the proliferation of artificial intelligence, and the growing complexity of sociotechnical systems within organizations and society. While quantitative approaches have historically dominated the discipline—offering reproducible metrics, statistical generalizability, and comparative benchmarking—there has been a progressive and increasingly well-documented turn toward qualitative inquiry. This turn reflects a recognition that the human dimensions of technology adoption, organizational change, and innovation cannot be fully understood through numerical data alone. Processes such as sense-making, identity negotiation, trust formation, and stakeholder engagement are inherently interpretive phenomena that demand methodologies capable of capturing depth, context, and meaning.

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

Despite growing interest in qualitative approaches within technology management research, the field continues to grapple with questions of methodological rigor, transparency, and appropriate design. Researchers working at the intersection of technology, management, and social science often face competing pressures: the demand for the statistical rigor associated with natural science traditions , and the need for contextually rich, interpretive accounts on the other. This tension is further complicated by the emergence of hybrid methodological approaches—such as inductive-deductive thematic analysis and computational topic modeling—that blur traditional disciplinary boundaries. While methodological contributions exist across related domains, a consolidated review of qualitative methods as applied specifically to technology management research remains absent from the literature.

This review addresses that gap. The primary objective is to synthesize existing scholarship on qualitative and mixed-methods approaches as employed in technology management research contexts, examining both the methodological principles underpinning these approaches and their practical applications across a range of substantive domains. In doing so, the review draws on twenty published studies spanning health informatics, educational technology, digital transformation, human-AI interaction, organizational behavior, and energy social science. While these domains are diverse, they converge around a shared concern: how qualitative methods can be used to generate valid, rigorous, and theoretically meaningful knowledge about the relationship between technology and human actors.

The article is structured as follows. The theoretical background section defines key methodological constructs and situates them within broader epistemological traditions. The literature review section synthesizes findings across three thematic clusters. A subsequent section identifies research gaps and proposes future directions, before a concluding section summarizes the review's contributions and acknowledges its limitations.

2. Theoretical Background

2.1 Epistemological Foundations of Qualitative Inquiry

Qualitative research is grounded in interpretivist and constructivist epistemologies, which hold that knowledge is socially constructed and that understanding requires engagement with the subjective meanings that individuals and communities assign to their experiences. In technology management research, this epistemological orientation is particularly relevant given that the uptake, resistance, and transformation of technology is mediated by human interpretation, institutional norms, and relational dynamics. The goal of qualitative inquiry, therefore, is not to measure outcomes across large populations but to generate in-depth of understanding within specific social and organizational contexts.

Semistructured interviewing represents one of the most widely deployed qualitative data collection instruments in technology-focused research. DeJonckheere and Vaughn (2019) provide a comprehensive account of this method, characterizing it as a dialogue between researcher and participant guided by a flexible interview protocol and enriched through follow-up questions, probes, and reflective comments. This approach enables researchers to collect open-ended data while retaining sufficient structure to ensure that key topics are addressed consistently across participants. The balance between structure and flexibility is central to its value: too rigid a protocol forecloses emergent insight; too loose a framework risks losing analytical coherence. In technology management contexts, where the phenomena under

investigation are often novel, fast-moving, and poorly theorized, this balance is especially important.

2.2 Thematic Analysis and Its Variations

Thematic analysis—the systematic identification, organization, and interpretation of patterns of meaning within qualitative data—has become one of the most widely applied analytical frameworks in social and organizational research. Within technology management, thematic analysis has been used to examine everything from teachers' perceptions of online learning platforms (Rasmitadila et al., 2020) to the challenges of digital transformation for regional stakeholders (Brunetti et al., 2020). The method offers considerable flexibility and is compatible with a range of epistemological positions, from essentialist to constructivist.

A significant methodological development in recent years has been the emergence of inductive/deductive hybrid thematic analysis, which Proudfoot (2022) argues that this approach offers particular advantages within mixed-methods and critical realist research paradigms. Rather than committing exclusively to data-driven induction or theory-driven deduction, the hybrid approach allows researchers to move iteratively between raw data and theoretical frameworks, using the critical realist concepts of abduction and retroduction to generate explanatory depth. This is especially relevant in technology management research, where empirical data frequently challenges or extends existing theoretical models.

2.3 Mixed-Methods Designs

Mixed-methods research, which deliberately integrates qualitative and quantitative data within a single study, has gained considerable traction in technology management and related fields. Zhou (2018) demonstrates this approach in a study of creative climates within project-organized groups, combining quantitative survey data from 126 participants with qualitative interviews from 15 respondents to triangulate findings and achieve a more comprehensive understanding of the phenomena under investigation. Similarly, Lindahl et al. (2017) employ a mixed-methods framework to investigate meditation-related challenges among practitioners, illustrating how qualitative data can illuminate the experiential dimensions of phenomena that quantitative instruments measure only partially.

Sovacool, Axsen, and Sorrell (2018) offer an influential critique of methodological practices in energy social science that has considerable resonance across technology management research more broadly. They argue that many studies suffer from poorly constructed research questions, absence of explicit research design, and a failure to match methods to epistemological commitments. Their proposed codes of practice—emphasizing novelty, rigor, and stylistic clarity—serve as a normative benchmark for qualitative and mixed-methods work in technology-adjacent social sciences.

2.4 Computational and Novel Analytical Approaches

Beyond conventional qualitative methods, technology management researchers have increasingly drawn on computational techniques to analyze large bodies of textual data. Hannigan et al. (2019) provide a foundational account of topic modeling as a method for rendering phenomenon-based constructs and conceptual relationships from textual data in management research. By conceptualizing topic modeling as a process of construct rendering rather than a purely algorithmic procedure, the authors bridge the gap between computational and interpretive traditions, opening productive methodological territory for researchers seeking to analyze large qualitative datasets without sacrificing theoretical grounding.

The emergence of qualitative approaches in human-AI research represents another frontier. Cai et al. (2019) conduct a qualitative laboratory study with 21 pathologists to uncover the informational needs of medical practitioners engaging with AI-assisted diagnostic systems—a study design that combines interview methods with in situ observation to capture the nuanced dynamics of human-AI collaboration. Such approaches are increasingly relevant as technology management researchers grapple with AI's integration into organizational decision-making.

3. Structured Literature Review

3.1 Thematic Cluster One: Qualitative Data Collection in Technology Contexts

The design of data collection instruments is a foundational concern in qualitative technology management research. Across the reviewed literature, semistructured interviewing emerges as the preferred method for capturing stakeholder perspectives on technology adoption, use, and resistance. DeJonckheere and Vaughn (2019) articulate the epistemological rationale for this choice with particular clarity, emphasizing that the semistructured format enables researchers to explore participant thoughts, feelings, and beliefs about a particular topic while retaining the flexibility to follow unanticipated lines of inquiry. This characteristic is especially valuable in technology management research, where the phenomena under investigation—whether digital transformation, AI integration, or online learning—are often experienced differently across organizational roles, demographic groups, and institutional settings.

Licorish et al. (2018) exemplify this approach in their study of students' perceptions of the game-based response system Kahoot!, in which semistructured interviews were combined with observational data to produce a rich account of how technology-enhanced learning tools reshape classroom dynamics and student engagement. The study illustrates how interview-based qualitative methods can surface fine-grained experiential data that structured surveys systematically exclude—a point reinforced by Kim, Lee, and Cho (2022), who employ in-depth interviews with ten Korean educators to explore the curriculum and pedagogical design requirements for student-AI collaboration in K-12 settings. In both studies, the researchers' ability to probe and redirect during the interview process yielded insights into unanticipated dimensions of technology use that would have been invisible to quantitative instruments.

The appropriateness of particular data collection strategies is not, however, universally agreed upon in the literature. Nadarzynski et al. (2019), in their mixed-methods study of AI-led health chatbot acceptability, argue that interview data alone are insufficient for capturing the behavioral dimensions of technology acceptance, necessitating complementary quantitative measures. This position resonates with Ferri, Grifoni, and Guzzo (2020), who adopt a two-stage qualitative design in their study of emergency remote teaching—first analyzing online discussion forums with international experts, then drawing on secondary sources and opinion leader statements—to triangulate findings and strengthen their interpretive claims. The diversity of data collection strategies employed across these studies points to an emerging consensus that no single method is universally appropriate, and that the choice of qualitative instrument must be calibrated to the specific research question and social context under investigation.

3.2 Thematic Cluster Two: Analytical Frameworks and the Construction of Meaning

Beyond data collection, the literature reveals considerable variation and innovation in the analytical frameworks applied to qualitative data in technology management research. Thematic analysis, in its various forms, is ubiquitous—but how it is conducted and theorized

differs significantly across studies. Rasmitadila et al. (2020) employ a relatively conventional inductive thematic analysis to identify four main themes—instructional strategies, challenges, support, and motivation—from interview and survey data collected from 67 Indonesian primary school teachers during the COVID-19 pandemic. This study exemplifies the utility of thematic analysis for generating descriptive, policy-relevant findings about technology adoption in educational contexts, where the primary audience may be practitioners as much as academic researchers.

Proudfoot's (2022) contribution, by contrast, is explicitly methodological rather than substantive: it proposes an inductive/deductive hybrid approach to thematic analysis that is designed to integrate qualitative and quantitative data within a mixed-methods framework grounded in critical realism. The significance of this contribution for technology management research lies in its demonstration that thematic analysis need not be confined to purely inductive or exploratory work; it can be deployed in dialogue with theoretical constructs to test, refine, or challenge existing models of technology adoption, diffusion, or governance. The articulation of abduction and retroduction as analytical moves within this framework offers technology management researchers a more philosophically sophisticated set of tools for generating explanatory, rather than merely descriptive, qualitative accounts.

Hannigan et al. (2019) extend this methodological territory into computational analysis, proposing topic modeling as a method for revealing phenomenon-based constructs from large bodies of textual data in management research. Their framework explicitly addresses the risk that computational methods become epistemologically opaque—what they describe as the "black box" problem—by insisting that topic modeling be understood as a theoretically engaged process of construct rendering rather than a purely algorithmic procedure. This is a critical methodological contribution for technology management, where the availability of large digitally generated datasets (user reviews, social media posts, policy documents) makes computational text analysis increasingly feasible and attractive.

The affordance perspective deployed by Waizenegger et al. (2020) in their qualitative study of knowledge workers' experiences of enforced remote working during COVID-19 represents yet another analytical framework gaining traction in technology management research. By conceptualizing technology as a bundle of affordances—relational properties that enable or constrain particular human actions—this approach enables researchers to move beyond simplistic adoption or resistance narratives and engage with the situated, contingent ways in which individuals and organizations appropriate digital tools. The study, based on interviews with 29 knowledge workers, demonstrates how the affordance lens can generate theoretically grounded qualitative accounts of the human-technology relationship in organizational settings.

3.3 Thematic Cluster Three: Mixed-Methods Designs and Multi-Stakeholder Approaches

A third and increasingly prominent theme in the qualitative technology management literature concerns the strategic integration of qualitative and quantitative methods, as well as the importance of incorporating multiple stakeholder perspectives into research designs. Brunetti et al. (2020) offer a compelling example of the multi-stakeholder approach in their exploratory qualitative study of digital transformation challenges in the Tyrol-Veneto macroregion. By engaging with companies, public administrators, and educational organizations simultaneously, the study captures the systemic and relational dimensions of digital transformation that single-stakeholder studies systematically miss. The authors argue

that a place-based, multi-stakeholder qualitative design is particularly well suited to understanding regional innovation systems, where transformation is shaped by interactions among heterogeneous actors operating across institutional levels.

Zhou (2018) demonstrates the productive complementarity of qualitative and quantitative methods in a study combining survey data from 126 participants with qualitative interviews from 15 respondents to investigate creative climates in project-organized groups. The integration of methods at both the data collection and analytical stages enables the researcher to identify patterns at the population level while simultaneously illuminating the processes and meanings that underlie those patterns—a combination that neither method alone could achieve. This methodological logic is echoed by Lindahl et al. (2017), whose mixed-methods study of contemplative experience challenges uses qualitative interview data to explain and contextualize quantitative patterns in self-reported experiences of meditation-related difficulties.

Sovacool, Axsen, and Sorrell (2018) occupy a distinctive position in this thematic cluster as methodological critics rather than empirical contributors. Their proposed codes of practice for appropriate methods and research design in energy social science constitute a normative framework that applies equally to technology management research. They argue that the selection of research methods must be explicitly justified in relation to research questions, theoretical commitments, and the nature of the phenomena under investigation—a principle that many empirical studies in the reviewed literature honor more in theory than in practice. Their critique of methodological boilerplate—the tendency to rehearse standard justifications for qualitative approaches without engaging substantively with their epistemological implications—serves as an important corrective for the field.

The role of qualitative methods in capturing stakeholder perspectives on technology in healthcare contexts is addressed by both Donaghy et al. (2019) and Nadarzynski et al. (2019). Donaghy et al. (2019) conduct a qualitative study of the acceptability, benefits, and challenges of video consulting in primary care, finding that the visual component of video consultations offers distinct advantages over telephone-based alternatives while also generating new technical and relational challenges. The study illustrates how qualitative inquiry can surface the nuanced experiential and contextual factors that determine whether technological innovations are accepted and sustained in practice—factors that are largely invisible to trial-based or survey-based research designs. Nadarzynski et al. (2019), in their mixed-methods study of AI-led chatbot acceptability in healthcare, arrive at a similar conclusion: intervention designers need to employ user-centered and theory-based approaches that take patients' perspectives, motivations, and capabilities into account. Together, these studies demonstrate that qualitative methods are not merely supplementary to quantitative technology evaluations but are often epistemologically essential for understanding the human dimensions of technology acceptance and implementation.

Cai et al.'s (2019) qualitative investigation of the onboarding needs of medical practitioners for human-AI collaborative decision-making adds another dimension to this cluster. Through interviews conducted before, during, and after exposure to a deep neural network diagnostic system, the study maps the informational needs and cognitive processes of expert users navigating novel AI tools—a research design that is both methodologically sophisticated and practically significant. The finding that accurate algorithmic predictions alone are insufficient for effective human-AI collaboration underscores the importance of qualitative inquiry for understanding the relational and communicative dimensions of technology-mediated work.

4. Research Gaps and Future Directions

Despite the breadth and sophistication of qualitative work reviewed here, several significant methodological and substantive gaps remain. Identifying these gaps is essential not only for advancing the theoretical development of technology management as a discipline but also for ensuring that qualitative research generates practically actionable knowledge.

The first and perhaps most significant gap is the lack of the longitudinal qualitative research in technology management . The overwhelming majority of studies in the reviewed literature adopt cross-sectional designs, capturing participants' experiences and perspectives at a single point in time. While this approach is often appropriate for exploratory research, it is poorly suited to understanding how technology adoption processes, digital transformations, and human-AI relationships evolve over time. As Brunetti et al. (2020) acknowledge, digital transformation is a dynamic and nonlinear process shaped by ongoing negotiation among stakeholders across institutional levels—a process that cross-sectional qualitative designs can only partially illuminate. Future research should develop and apply longitudinal qualitative frameworks, including repeated-interview designs and ethnographic approaches, to capture the temporal dynamics of technology management processes.

A second gap concerns the cross-cultural generalizability of qualitative findings in technology management. Several studies in this review draw on single-country or single-institutional samples—Kim, Lee, and Cho (2022) focus exclusively on Korean educators, Rasmitadila et al. (2020) on Indonesian primary school teachers, and Brunetti et al. (2020) on a specific European macroregion—without systematically examining how cultural context shapes the phenomena under investigation. While qualitative research does not seek statistical generalizability, the systematic comparison of qualitative findings across cultural and institutional settings can generate important theoretical insights into the contextual conditions that enable or constrain particular patterns of technology adoption and management. Future research should prioritize comparative qualitative designs that explicitly theorize the role of cultural context.

A third gap concerns the methodological treatment of emerging technologies, particularly generative AI, within qualitative research frameworks. Budhwar et al. (2023) and Michel-Villarreal et al. (2023) both identify the transformative potential of generative AI for organizations and educational institutions, yet neither study employs qualitative methods to examine how organizational actors make sense of, adapt to, or resist these technologies in practice. The rapid organizational uptake of generative AI tools creates an urgent need for qualitative research that can capture the experiential, relational, and institutional dimensions of this transformation. Future research should develop qualitative and mixed-methods designs specifically suited to investigating human-generative AI collaboration in organizational settings, drawing on established frameworks such as the affordance perspective deployed by Waizenegger et al. (2020) while adapting them to the distinctive characteristics of generative AI systems.

A fourth, cross-cutting gap concerns the systematic application of methodological quality standards to qualitative technology management research. Sovacool, Axsen, and Sorrell's (2018) critique of methodological shortcomings in energy social science—including poorly constructed research questions, absent or implicit research designs, and insufficient engagement with theory—resonates broadly across the technology management literature. Future methodological work should develop domain-specific quality standards for qualitative and mixed-methods research in technology management, drawing on existing frameworks

from health services research (DeJonckheere & Vaughn, 2019) and management science (Hannigan et al., 2019) while attending to the distinctive epistemological challenges posed by technology-focused inquiry.

5. Conclusion

This literature review has synthesized qualitative and mixed-methods research across twenty peer-reviewed studies to map the current state of qualitative methodology in technology management research. Three principal thematic clusters were identified and examined: the design and implementation of qualitative data collection instruments, particularly semistructured interviewing; the diversity of analytical frameworks applied to qualitative data, from thematic analysis to computational topic modeling and affordance-based approaches; and the strategic integration of qualitative methods within mixed-methods and multi-stakeholder research designs.

The review's central finding is that qualitative methods are not merely supplementary tools for explaining what quantitative designs leave unexplained. Rather, they are foundational to understanding the human, relational, and institutional dimensions of technology adoption, governance, and transformation. Across domains as diverse as emergency remote teaching, AI-assisted medical decision-making, digital transformation strategy, and educational technology, qualitative approaches have generated insights that are both theoretically significant and practically consequential.

The review also identifies meaningful methodological tensions in the literature. The demand for rigor, transparency, and replicability—articulated most forcefully by Sovacool, Axsen, and Sorrell (2018)—sits in productive tension with the interpretive openness and contextual sensitivity that qualitative methods require. Navigating this tension productively, rather than dissolving it in favor of either pole, is perhaps the defining methodological challenge for the field.

Several limitations of this review should be acknowledged. The sample of twenty sources, while carefully selected for methodological and thematic relevance, is not exhaustive, and important contributions may have been excluded. The review does not constitute a systematic review in the formal sense, as no explicit inclusion criteria or database search protocols were applied. Additionally, the thematic clusters identified are not mutually exclusive, and alternative organizing frameworks would have yielded different analytical emphases. These limitations notwithstanding, the review contributes a structured and theoretically grounded account of qualitative methodology in technology management research that should be of value to researchers, methodologists, and practitioners alike.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Brunetti, F., Matt, D. T., Bonfanti, A., De Longhi, A., & Pedrini, 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

Budhwar, P., Chowdhury, S., Wood, G., Aguinis, H., & Bamber, G. J. (2023). Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT. Human Resource Management Journal, 33(3), 606–659. https://doi.org/10.1111/1748-8583.12524

Cai, C. J., Winter, S., Steiner, D. F., Wilcox, L., & Terry, M. (2019). "Hello AI": Uncovering the onboarding needs of medical practitioners for human-AI collaborative decision-making. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW), 1–24. https://doi.org/10.1145/3359206

Dawson, N. M., Coolsaet, B., Sterling, E. J., Loveridge, R., & Gross-Camp, N. D. (2021). The role of Indigenous peoples and local communities in effective and equitable conservation. Ecology and Society, 26(3), 19. https://doi.org/10.5751/es-12625-260319

DeJonckheere, M., & Vaughn, L. M. (2019). Semistructured interviewing in primary care research: A balance of relationship and rigour. Family Medicine and Community Health, 7(2), e000057. https://doi.org/10.1136/fmch-2018-000057

Donaghy, E., Atherton, H., Hammersley, V., McNeilly, H., & Bikker, A. (2019). Acceptability, benefits, and challenges of video consulting: A qualitative study in primary care. British Journal of General Practice, 69(686), e586–e594. https://doi.org/10.3399/bjgp19x704141

Ferri, F., Grifoni, P., & Guzzo, T. (2020). Online learning and emergency remote teaching: Opportunities and challenges in emergency situations. Societies, 10(4), 86. https://doi.org/10.3390/soc10040086

Hannigan, T. R., Haans, R. F. J., Vakili, K., Tchalian, H., & Glaser, V. L. (2019). Topic modeling in management research: Rendering new theory from textual data. Academy of Management Annals, 13(2), 586–632. https://doi.org/10.5465/annals.2017.0099

Keiler, L. S. (2018). Teachers' roles and identities in student-centered classrooms. International Journal of STEM Education, 5(1), 34. https://doi.org/10.1186/s40594-018-0131-6

Kim, J., Lee, H.-K., & Cho, Y. H. (2022). Learning design to support student-AI collaboration: Perspectives of leading teachers for AI in education. Education and Information Technologies, 27(5), 6803–6825. https://doi.org/10.1007/s10639-021-10831-6

Licorish, S. A., Owen, H., Daniel, B. K., & George, J. L. (2018). Students' perception of Kahoot!'s influence on teaching and learning. Research and Practice in Technology Enhanced Learning, 13(1), 9. https://doi.org/10.1186/s41039-018-0078-8

Lindahl, J. R., Fisher, N. E., Cooper, D. J., Rosen, R. K., & Britton, W. B. (2017). The varieties of contemplative experience: A mixed-methods study of meditation-related challenges in Western Buddhists. PLoS ONE, 12(5), e0176239. https://doi.org/10.1371/journal.pone.0176239

Lukashin, Yu. P., & Rakhlina, L. I. (2021). On world development indicators. Vestnik MIRBIS, 2, 1–12. https://doi.org/10.25634/mirbis.2021.2.1

Michel-Villarreal, R., Vilalta-Perdomo, E. L., & Salinas-Navarro, D. E. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. Education Sciences, 13(9), 856. https://doi.org/10.3390/educsci13090856

Nadarzynski, T., Miles, O., Cowie, A., & Ridge, D. (2019). Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study. Digital Health, 5, 1–12. https://doi.org/10.1177/2055207619871808

Proudfoot, K. (2022). Inductive/deductive hybrid thematic analysis in mixed methods research. Journal of Mixed Methods Research, 17(3), 308–326. https://doi.org/10.1177/15586898221126816

Rasmitadila, R., Aliyyah, R. R., Rachmadtullah, R., & Samsu, A. (2020). The perceptions of primary school teachers of online learning during the COVID-19 pandemic period: A case study in Indonesia. Journal of Ethnic and Cultural Studies, 7(2), 90–109. https://doi.org/10.29333/ejecs/388

Sovacool, B. K., Axsen, J., & Sorrell, S. (2018). Promoting novelty, rigor, and style in energy social science: Towards codes of practice for appropriate methods and research design. Energy Research & Social Science, 45, 12–42. https://doi.org/10.1016/j.erss.2018.07.007

Waizenegger, L., McKenna, B., Cai, W., & Bendz, T. (2020). An affordance perspective of team collaboration and enforced working from home during COVID-19. European Journal of Information Systems, 29(4), 429–442. https://doi.org/10.1080/0960085x.2020.1800417

Zhou, C. (2018). A study on creative climate in project-organized groups (POGs) in China and implications for sustainable pedagogy. Sustainability, 10(1), 114. https://doi.org/10.3390/su10010114

Best Practices in AI Implementation: A Synthesis of Case Studies Across Domains

Kosztanko Simona[1] and Kufelová Iveta[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: simona.kosztanko@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: iveta.kufelova@euba.sk

Abstract

Artificial intelligence (AI) has rapidly permeated diverse sectors including healthcare, marketing, education, and engineering, yet systematic evidence regarding effective implementation practices remains fragmented across disciplinary silos. This review synthesizes findings from twenty scholarly works to identify convergent best practices, persistent challenges, and emerging research priorities in AI implementation. The analysis is organized thematically around three major clusters: technical and architectural considerations, ethical and governance imperatives, and domain-specific deployment strategies. Key findings indicate that successful AI implementation consistently requires alignment between algorithmic transparency, robust data governance, stakeholder engagement, and context-sensitive adaptation. Ethical concerns particularly around bias, explainability, and privacy emerge as cross-cutting barriers that transcend individual application domains. Furthermore, the integration of federated learning, edge intelligence, and explainable AI frameworks is increasingly recognized as foundational to responsible deployment. These findings carry significant implications for practitioners, policymakers, and researchers seeking to operationalize AI systems that are simultaneously effective, trustworthy, and scalable. The review concludes by identifying critical gaps in the literature and proposing targeted future research directions.

Keywords Words: AI implementation best practices, case studies, organizational learning, digital adoption, knowledge management

JEL Classification: O33, M15, L20

1. Introduction

Artificial intelligence is no longer a peripheral or speculative technology. Across industries ranging from clinical medicine to financial services, AI systems are being embedded into operational workflows, decision-support tools, and customer-facing interfaces at an unprecedented pace. This proliferation has generated a substantial body of literature documenting the technical achievements of AI, yet comparatively less attention has been directed toward the organizational, ethical, and procedural conditions that determine whether AI deployments succeed or fail in practice. The gap between algorithmic capability and real-world implementation effectiveness represents one of the most pressing challenges in contemporary AI research.

The urgency of this gap is underscored by several converging trends. Deep learning architectures have demonstrated performance exceeding human benchmarks on a range of

complex cognitive tasks (Alzubaidi et al., 2021), and large language models have begun encoding clinical-grade knowledge that was previously the exclusive province of trained professionals (Singhal et al., 2023). Yet the translation of these technical achievements into reliable, equitable, and institutionally accepted AI systems is far from automatic. Problems of model interpretability (Tjoa & Guan, 2020; Carvalho et al., 2019), data privacy (Kairouz et al., 2020), and ethical accountability (Maiti et al., 2025) continue to impede deployment across high-stakes domains. Meanwhile, sectors such as education and marketing are navigating not only technical complexity but also profound questions about the social and normative implications of AI adoption (Alali et al., 2024; Davenport et al., 2019).

Despite a growing volume of domain-specific AI implementation studies, the literature lacks a consolidated synthesis that identifies cross-domain best practices grounded in empirical case evidence. Existing reviews tend to focus on technical advancements within narrow fields or on theoretical frameworks for responsible AI, rarely integrating both dimensions in a way that is actionable for practitioners. This review addresses that gap by conducting a structured thematic synthesis of twenty scholarly works drawn from diverse disciplines, with the explicit aim of identifying principles, practices, and conditions that characterize effective and responsible AI implementation.

The objective of this literature review is threefold: first, to establish a theoretical grounding for understanding AI implementation as a multidimensional challenge; second, to synthesize empirical and conceptual findings from the available literature into coherent thematic clusters; and third, to identify research gaps and propose future directions that can advance both the science and practice of AI deployment. The remainder of the article is organized as follows. Section 2 presents the theoretical background, defining key constructs and foundational frameworks. Section 3 provides a thematic literature review across three major clusters. Section 4 identifies research gaps and proposes future directions. Section 5 offers a concluding synthesis.

2. Theoretical Background

Understanding AI implementation as a scholarly and practical domain requires clarity regarding the foundational constructs that govern both the design and deployment of intelligent systems. At its most fundamental level, AI implementation refers to the process through which machine learning models, algorithmic decision-making systems, or intelligent automation tools are integrated into real-world operational environments. This process is substantially more complex than model training or technical benchmarking; it encompasses organizational change, stakeholder alignment, regulatory compliance, and ongoing performance monitoring.

Deep learning constitutes one of the most consequential technological paradigms underlying contemporary AI implementation. Alzubaidi et al. (2021) provide a comprehensive taxonomy of deep learning architectures, including convolutional neural networks (CNNs), recurrent neural networks, and generative adversarial networks, each of which presents distinct implementation considerations regarding computational requirements, data volume, and interpretability. The ability of deep learning systems to extract hierarchical representations from massive datasets has driven adoption across vision, language, and control tasks, but has simultaneously rendered these systems opaque to end-users and domain experts alike.

Explainability and interpretability represent perhaps the most theoretically significant constructs in the AI implementation literature. Tjoa and Guan (2020) distinguish between

interpretable models whose internal logic is inherently comprehensible and post-hoc explanation methods that generate human-readable rationales for black-box model outputs. This distinction has direct implications for implementation: in regulated domains such as healthcare and finance, stakeholders may require explanations not merely for epistemic satisfaction but as a legal or ethical requirement. Carvalho et al. (2019) further elaborate a taxonomy of interpretability methods, differentiating model-agnostic approaches such as LIME and SHAP from model-specific techniques, and introducing evaluation metrics that span fidelity, stability, and comprehensibility.

Federated learning offers a foundational architectural paradigm for privacy-preserving AI implementation. Kairouz et al. (2020) define federated learning as a setting in which multiple clients collaboratively train a shared model without centralizing their respective datasets, thereby addressing systemic privacy risks inherent in conventional data aggregation. This approach is particularly relevant in healthcare and financial services, where data sensitivity and regulatory constraints prohibit the pooling of raw data. The federated paradigm also connects directly to the concept of edge intelligence, wherein AI inference and even training are performed at the network edge rather than in centralized cloud infrastructures. Zhou et al. (2019) frame edge intelligence as a necessary evolution to accommodate the computational demands of AI deployment on resource-constrained IoT and mobile devices, arguing that it represents the final frontier of AI operationalization.

Data governance constitutes another foundational construct, particularly as AI systems grow more dependent on large, heterogeneous datasets. Chukwurah et al. (2024) defines data governance as an ensemble of policies, processes, roles, and metrics designed to ensure data quality, privacy, and security across organizational contexts. Effective data governance is not merely a technical undertaking but an institutional one, requiring cross-functional stakeholder engagement and alignment with sector-specific regulatory frameworks. The importance of governance is amplified in AI contexts because model performance is tightly coupled to data quality, and failures in governance upstream can propagate into discriminatory, unreliable, or non-compliant model outputs downstream.

Ethical AI constitutes a theoretical domain, addressing the normative dimensions of AI deployment. Maiti et al. (2025) identify a cluster of interrelated ethical concerns that function as barriers to AI adoption in business contexts: privacy and data protection, algorithmic bias and fairness, transparency and explainability, job displacement, accountability, and ethical decision-making. These concerns are not merely academic; they carry practical weight in shaping public trust, regulatory scrutiny, and organizational risk management. Together, these constructs—deep learning architecture, explainability, federated learning, edge intelligence, data governance, and ethical AI form the theoretical scaffolding upon which this review's thematic synthesis is built.

3. Literature Review

3.1 Technical Architectures and Infrastructure for AI Deployment

The technical dimension of AI implementation is shaped by a complex interplay of model architecture choices, infrastructure design, and the computational constraints imposed by deployment environments. Alzubaidi et al. (2021) document the trajectory of deep learning architectures from foundational CNNs to sophisticated variants designed for specific perceptual and generative tasks, emphasizing that architectural choice must be calibrated to the nature of the input data, the computational budget, and the performance requirements of

the target application. This is not a purely engineering decision; it has downstream consequences for interpretability, maintenance, and scalability all of which are central to long-term implementation success.

A particularly salient technical challenge in AI implementation is the management of computational resources in resource-constrained environments. Zhou et al. (2019) advances the concept of edge intelligence as a response to the latency, bandwidth, and privacy limitations of cloud-centric AI deployment. By pushing inference and in some architectures, training to the network edge, organizations can reduce dependency on centralized infrastructure, improve real-time responsiveness, and mitigate data transmission risks. This has direct implications for implementation best practices: rather than treating cloud deployment as a default, practitioners are advised to conduct a systematic assessment of the latency tolerance, connectivity reliability, and data sensitivity of each deployment context before selecting an infrastructure model.

The domain of fluid mechanics provides an instructive case study in the implementation of machine learning for high-dimensional scientific problems. Brunton et al. (2019) demonstrate that ML techniques including neural networks, Gaussian process regression, and reinforcement learning can be effectively applied to turbulence modeling, flow control, and reduced-order modeling when domain knowledge is actively incorporated into model design. This work illustrates a broader principle: AI implementation in technical domains achieves superior performance and interpretability when domain expertise is treated as a first-class input rather than as an afterthought. The hybridization of physics-based models with data-driven ML approaches commonly referred to as physics-informed machine learning exemplifies this integration and has been cited as a best practice for scientific AI deployment.

Digital twin technology further extends the architecture of AI implementation to encompass continuous, bidirectional coupling between physical assets and their virtual representations. Rasheed et al. (2020) describe digital twins as virtual models enabled by data and simulation that support real-time prediction, optimization, and decision-making. From an implementation standpoint, digital twins require not only advanced ML components but also robust data pipelines, Multiphysics solvers, and cybersecurity infrastructure. The authors identify significant challenges including model validation, computational cost, and interoperability that must be systematically addressed in any implementation of roadmap. The digital twin paradigm represents a high-complexity implementation scenario in which best practices for data governance, model monitoring, and stakeholder communication are especially critical.

Federated learning introduces an additional layer of technical complexity that must be accounted for in implementation planning. Kairouz et al. (2020) enumerate a range of open problems in federated learning that directly affect its deplorability, including statistical heterogeneity across clients, communication efficiency, and the difficulty of verifying model integrity in adversarial settings. Despite these challenges, the federated paradigm is increasingly recognized as a best-practice architectural choice in privacy-sensitive domains, with implementations reported in healthcare, mobile computing, and financial analytics. The practical implication is that AI implementation frameworks must be sufficiently flexible to accommodate distributed training workflows, necessitating new tooling, governance structures, and performance evaluation protocols.

3.2 Ethical, Governance, and Regulatory Dimensions

The ethical and governance dimensions of AI implementation have attracted growing scholarly attention, driven by high-profile failures of algorithmic fairness, accountability gaps,

and regulatory developments such as the European Union's AI Act. Maiti et al. (2025) present a structured analysis of ethical barriers to AI adoption in business, identifying privacy and data protection as the most frequently cited concern, followed closely by bias and fairness. Their analysis further reveals that organizations with explicit AI ethics frameworks encompassing documented principles, accountability mechanisms, and regular audits report higher success rates in AI deployment than those relying on ad hoc governance arrangements. This finding aligns with the broader literature on responsible innovation, which emphasizes that ethical considerations must be embedded in the design process rather than applied retrospectively.

Explainability constitutes a particularly contested terrain in the AI ethics literature. Tjoa and Guan (2020) argue that explainability is not a single property but a multidimensional requirement whose specifications differ by stakeholder: a data scientist may require feature-level attribution, a clinician may require a case-based rationale, and a regulator may require a formal audit trail. This heterogeneity of explainability needs presents a significant implementation challenge, as no single method satisfies all stakeholder requirements simultaneously. Carvalho et al. (2019) complement this analysis by showing that existing interpretability methods involve inherent trade-offs between fidelity how accurately the explanation represents the model's actual reasoning and comprehensibility how easily the explanation is understood by non-expert users. Navigating these trade-offs in a principled way is a hallmark of mature AI implementation practice.

Data governance provides institutional infrastructure through which ethical AI principles are operationalized. Chukwurah et al. (2024) identify a set of cross-industry best practices for data governance that are directly applicable to AI contexts: establishing clearly defined data ownership and stewardship roles, implementing tiered data access controls, maintaining audit trails for data lineage, and instituting regular data quality assessments. The authors note, however, that implementation challenges are substantial, particularly in organizations characterized by siloed data architectures, legacy systems, and limited cross-functional coordination. Their comparative analysis across financial services, healthcare, manufacturing, and the public sector reveals that governance frameworks must be adapted to sector-specific regulatory environments rather than applied as generic templates.

The integration of AI into healthcare presents some of the most ethically complex implementation scenarios documented in the literature. Singhal et al. (2023) demonstrate that large language models can encode clinical knowledge at a level sufficient to outperform human clinicians on standardized medical question-answering benchmarks, yet they simultaneously emphasize that benchmark performance is an insufficient proxy for clinical trustworthiness. Their proposed MultiMedQA framework for human evaluation introduces additional assessment of axes of factuality, reasoning quality, potential for harm, and evidence of grounding that reflect the multidimensional nature of clinical AI accountability. This work implicitly advocates for a best practice in which technical performance evaluation is systematically supplemented by domain-expert human evaluation before deployment in high-stakes settings.

In the educational domain, Alali et al. (2024) examine the implementation of generative AI tools in classroom settings and identify a set of ethical and pedagogical best practices: maintaining transparency with students about AI assistance, establishing clear boundaries for permissible AI use, prioritizing learning objectives over efficiency gains, and incorporating AI literacy into the curriculum. The authors document case studies in which the absence of these practices led to outcomes ranging from academic integrity violations to student disengagement. The security dimension of AI implementation is addressed by Desai and Nisha

(2021), who document DevSecOps best practices for integrating security into AI-driven software development pipelines, arguing that security must be treated as a continuous process rather than a deployment-phase checkpoint. These findings collectively reinforce the principle that effective AI implementation requires governance structures that are proactive, institutionalized, and capable of evolving alongside the technology.

3.3 Domain-Specific Implementation Strategies and Lessons

Beyond generalized best practices, the literature reveals that AI implementation is strongly conditioned by domain-specific factors—including regulatory environments, data characteristics, user expectations, and the nature of the decisions being supported. Healthcare represents the most extensively documented implementation domain, and its lessons carry broad transferable value. The deployment of AI in Alzheimer's disease management, for instance, intersects with the epidemiological realities documented by Villemagne et al. (2023), who report that an estimated 6.7 million Americans currently live with Alzheimer's dementia and that this number could reach 13.8 million by 2060. This scale of need creates powerful incentives for AI-driven solutions in early detection, care coordination, and caregiver support, but also demands rigorous implementation standards given the vulnerability of the patient population.

The management of acute coronary syndromes and atrial fibrillation—complex, time-sensitive clinical conditions governed by evidence-based guidelines (Collet et al., 2020; Joglar et al., 2023) illustrates the challenge of aligning AI decision-support tools with dynamic, evolving clinical protocols. Implementation in these contexts requires not only accurate models but also seamless integration with clinical workflows, real-time updating mechanisms when guidelines change, and mechanisms for communicating uncertainty to clinicians. The broader literature on big data in healthcare (Dash et al., 2019) underscores that the sheer volume and heterogeneity of clinical data—ranging from electronic health records and imaging to wearables and genomics—necessitates robust preprocessing pipelines and data harmonization strategies as foundational implementation prerequisites.

In marketing, Davenport et al. (2019) proposes a multidimensional framework for AI implementation that accounts for intelligence levels, task types, and the embodiment of AI in physical agents such as robots. Their empirical analysis of industry practices identifies a progression from task automation through analytics-driven personalization to autonomous decision-making, with each stage requiring distinct organizational capabilities and change management strategies. A key insight is that successful marketing AI implementation depends critically on the quality and integration of customer data, reinforcing the centrality of data governance as a cross-domain best practice. The authors also observe that AI implementations tend to succeed when they augment human judgment rather than replace it wholesale, a finding consistent with broader literature on human-AI collaboration.

The metaverse represents an emerging and distinctive implementation environment in which AI capabilities including deep learning-based recognition models and natural language generation are embedded within immersive, socially networked virtual worlds (Park & Kim, 2022). The implementation challenges identified in this context are notably different from those in traditional enterprise AI deployments: they include the management of user-generated content at scale, the design of AI-mediated social interactions, and the ethical governance of AI behavior in environments where the boundary between real and virtual identity is deliberately blurred. These challenges anticipate a new generation of AI implementation problems that existing best-practice frameworks are not fully equipped to address.

The educational domain offers rich case study evidence for the importance of pedagogical alignment in AI implementation. Rudolph et al. (2023) examines the disruptive potential of ChatGPT in higher education assessment, documenting both the capabilities and limitations of generative AI in producing academically credible text. Their analysis highlights a critical implementation of tension: the same capabilities that make AI tools valuable for personalized learning and feedback generation also create conditions for misuse that can undermine assessment integrity. Alali et al. (2024) respond to this tension by proposing a set of implementation best practices grounded in constructive alignment ensuring that AI tools are deployed in ways that reinforce rather than circumvent the intended learning outcomes. This domain-specific insight has broader applicability: effective AI implementation requires a clear articulation of the intended outcomes it is designed to serve, against which deployment decisions, governance structures, and success metrics can be consistently evaluated.

4. Research Gaps and Future Directions

Despite the substantial and growing body of literature on AI implementation, several significant gaps remain that limit both theoretical understanding and practical guidance. Three priority areas for future research are identified here.

The first and perhaps most consequential gap concerns the absence of longitudinal implementation studies. The overwhelming majority of existing work captures AI deployment at a single point in time, documenting initial performance, adoption rates, or user perceptions. What remains poorly understood is how AI implementations evolve over time: how model performance degrades as data distributions shift, how organizational capabilities for AI governance mature, and how stakeholder trust is built or eroded through sustained interaction with AI systems. Longitudinal studies are particularly needed in high-stakes domains such as healthcare where the epidemiological burden documented by Villemagne et al. (2023) creates ongoing pressure for AI-assisted care and in clinical decision support for conditions governed by periodically updated guidelines (Collet et al., 2020; Joglar et al., 2023). Without longitudinal evidence, best practice recommendations remain essentially static for prescriptions applied to an inherently dynamic reality.

The second gap concerns the underrepresentation of organizational and sociotechnical factors in AI implementation research. The literature on technical architectures (Alzubaidi et al., 2021; Zhou et al., 2019; Kairouz et al., 2020) and ethical frameworks (Tjoa & Guan, 2020; Maiti et al., 2025) is comparatively well-developed, but research that systematically examines how organizational culture, change management practices, and leadership commitment interact with technical design choices to produce implementation outcomes remains sparse. Chukwurah et al. (2024) gesture toward this dimension in their discussion of cultural obstacles to data governance, but a dedicated research agenda drawing on organizational theory, sociotechnical systems theory, and implementation science is needed to fill this gap. Practically, this means that organizations cannot rely solely on technical best practices but must invest in organizational capacity building as a coequal implementation priority.

The third gap concerns the need for cross-domain comparative studies that rigorously test the generalizability of AI implementation best practices. Each domain reviewed in this article healthcare, marketing, education, fluid mechanics, DevOps has developed its own implementation guidance, often in relative isolation from adjacent fields. Brunton et al. (2019) demonstrate how domain knowledge can be productively hybridized with ML methods in fluid mechanics; Davenport et al. (2019) develop marketing-specific frameworks; Desai and Nisha (2021) address DevSecOps security practices. A synthetic research program that systematically

compares implementation outcomes across these domains controlling for contextual factors and isolating the contribution of specific practices would generate higher-order principles capable of informing a genuinely general theory of AI implementation. Such a research program would also be well-positioned to identify boundary conditions: the contextual factors that determine which best practices are universally applicable, and which are domain contingent.

5. Conclusion

This review has synthesized findings from twenty scholarly works to develop a structured, thematically organized account of best practices in AI implementation. The analysis has identified three major thematic clusters technical and infrastructural considerations, ethical and governance imperatives, and domain-specific deployment strategies each of which yields distinct and actionable insights for practitioners and researchers.

With respect to technical implementation, the literature converges on the importance of context-sensitive architecture selection, the integration of domain knowledge into model design, and the adoption of privacy-preserving paradigms such as federated learning and edge intelligence (Alzubaidi et al., 2021; Kairouz et al., 2020; Zhou et al., 2019; Brunton et al., 2019). With respect to ethics and governance, the evidence consistently shows that effective AI implementation requires proactive and institutionalized governance frameworks, multidimensional explainability mechanisms, and robust data stewardship practices (Tjoa & Guan, 2020; Carvalho et al., 2019; Maiti et al., 2025; Chukwurah et al., 2024). With respect to domain-specific implementation, the review reveals that context including regulatory environment, data characteristics, and stakeholder expectations is a critical determinant of implementation strategy, as illustrated by cases spanning healthcare (Singhal et al., 2023; Dash et al., 2019), marketing (Davenport et al., 2019), education (Alali et al., 2024; Rudolph et al., 2023), and engineering (Rasheed et al., 2020).

Several limitations of this review merit acknowledgment. The scope was bounded by a predefined set of twenty sources, which, while diverse, do not exhaustively represent the full breadth of AI implementation literature. Certain important application domains including financial services, autonomous systems, and public administration are less thoroughly represented than healthcare and education. Additionally, the review is primarily analytical rather than meta-analytic; the absence of a formal quality appraisal protocol means that findings vary in their level of empirical grounding. Future reviews would benefit from systematic search protocols and formal inclusion criteria. Notwithstanding these limitations, this review makes a meaningful contribution by providing a coherent, cross-domain synthesis of AI implementation best practices at a moment when such guidance is urgently needed by practitioners and policymakers alike.

Funding

This paper is funded by the EU Next Generation EU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Alali, R., Wardat, Y., Al-Saud, K., & Alhayek, K. A. (2024). Generative AI in education: Best practices for successful implementation. International Journal of Religion, 5(11), 1–12. https://doi.org/10.61707/pkwb8402

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A. Q., & Duan, Y. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data, 8(1), 53. https://doi.org/10.1186/s40537-021-00444-8

Brunton, S. L., Noack, B. R., & Koumoutsakos, P. (2020). Machine learning for fluid mechanics. Annual Review of Fluid Mechanics, 52, 477–508. https://doi.org/10.1146/annurev-fluid-010719-060214

Carvalho, D. V., Pereira, E. M., & Cardoso, J. S. (2019). Machine learning interpretability: A survey on methods and metrics. Electronics, 8(8), 832. https://doi.org/10.3390/electronics8080832

Chukwurah, N., Ige, A. B., Adebayo, V. I., & Eyebonye, O. G. (2024). Frameworks for effective data governance: Best practices, challenges, and implementation strategies across industries. Computer Science & IT Research Journal, 5(7), 1601–1620. https://doi.org/10.51594/csitrj.v5i7.1351

Collet, J.-P., Thiele, H., Barbato, E., & Barthélémy, O. (2020). 2020 ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation. European Heart Journal, 42(14), 1289–1367. https://doi.org/10.1093/eurheartj/ehaa575

Dash, S., Shakyawar, S. K., Sharma, L., & Kaushik, S. (2019). Big data in healthcare: Management, analysis, and prospects. Journal of Big Data, 6(1), 54. https://doi.org/10.1186/s40537-019-0217-0

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. https://doi.org/10.1007/s11747-019-00696-0

Desai, R., & Nisha, T. N. (2021). Best practices for ensuring security in DevOps: A case study approach. Journal of Physics: Conference Series, 1964(4), 042045. https://doi.org/10.1088/1742-6596/1964/4/042045

Joglar, J. A., Chung, M. K., Armbruster, A. L., & Benjamin, E. J. (2024). 2023 ACC/AHA/ACCP/HRS Guideline for the diagnosis and management of atrial fibrillation. Circulation, 149(1), e1–e156. https://doi.org/10.1161/cir.0000000000001193

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., & Bennis, M. (2021). Advances and open problems in federated learning. Foundations and Trends® in Machine Learning, 14(1–2), 1–210. https://doi.org/10.1561/2200000083

Maiti, M., Kayal, P., & Vujko, A. (2025). A study on ethical implications of artificial intelligence adoption in business: Challenges and best practices. Future Business Journal, 11(1), 62. https://doi.org/10.1186/s43093-025-00462-5

Park, S., & Kim, Y.-G. (2022). A metaverse: Taxonomy, components, applications, and open challenges. IEEE Access, 10, 4209–4251. https://doi.org/10.1109/access.2021.3140175

Rasheed, A., San, O., & Kvamsdal, T. (2020). Digital twin: Values, challenges and enablers from a modeling perspective. IEEE Access, 8, 21980–22012. https://doi.org/10.1109/access.2020.2970143

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? Journal of Applied Learning & Teaching, 6(1), 342–363. https://doi.org/10.37074/jalt.2023.6.1.9

Singhal, K., Azizi, S., Tu, T., Mahdavi, S. S., & Lee, J. (2023). Large language models encode clinical knowledge. Nature, 620, 172–180. https://doi.org/10.1038/s41586-023-06291-2

Tjoa, E., & Guan, C. (2021). A survey on explainable artificial intelligence (XAI): Toward medical XAI. IEEE Transactions on Neural Networks and Learning Systems, 32(11), 4793–4813. https://doi.org/10.1109/tnnls.2020.3027314

Villemagne, V., Burnham, S., Bourgeat, P., & Reiman, E. (2023). 2023 Alzheimer's disease facts and figures. Arzheimers & Dementia, 19(4), 1598–1695. https://doi.org/10.1002/alz.13016

Zhou, Z., Chen, X., Li, E., Zeng, L., & Luo, K. (2019). Edge intelligence: Paving the last mile of artificial intelligence with edge computing. Proceedings of the IEEE, 107(8), 1738–1762. https://doi.org/10.1109/jproc.2019.2918951

Organizational Learning and AI Knowledge Transfer: A Systematic Literature Review of Emerging Paradigms and Future Directions

Romanová Anita[1] and Kosztanko Simona[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: anita.romanova@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: simona.kosztanko@euba.sk

Abstract

The intersection of organizational learning and artificial intelligence (AI) knowledge transfer has emerged as a critical domain of inquiry as organizations increasingly integrate intelligent systems into their operational and strategic frameworks. This review synthesizes 21 peer-reviewed sources to examine how foundational concepts of organizational learning interact with contemporary AI methodologies, including deep learning, federated learning, and continual learning, to reshape how knowledge is acquired, retained, and transferred within and across organizational boundaries. The review identifies three principal thematic clusters: the theoretical foundations linking absorptive capacity and organizational learning capabilities to AI-enabled innovation; the mechanisms through which AI systems facilitate or constrain knowledge transfer; and the human-centric challenges arising from automation-augmentation tensions in learning environments. Key findings indicate that while AI substantially amplifies the velocity and scale of knowledge dissemination, critical gaps persist regarding privacy-preserving knowledge transfer, catastrophic forgetting in adaptive AI systems, and the organizational antecedents that enable effective AI integration. The review concludes by proposing directions for future research that bridge technical AI advances with organizational behavior theory, with significant implications for practitioners designing learning-enabled enterprises.

Key Words: organisational learning, knowledge transfer, artificial intelligence, knowledge management, innovation capability

JEL Classification: O33, M54, D83

1. Introduction

The rapid diffusion of artificial intelligence across industries has fundamentally altered the conditions under which organizations acquire, process, and deploy knowledge. As intelligent systems become embedded in organizational routines, the classical frameworks of organizational learning—rooted in concepts such as single-loop and double-loop learning, absorptive capacity, and dynamic capabilities—must be reconsidered in light of AI's distinctive properties as both a knowledge-generating and knowledge-transferring agent. The question of how organizations learn with and through AI, rather than simply about it, represents one of the most consequential and least resolved problems in contemporary management scholarship.

The significance of this inquiry is underscored by the scale and pace of AI adoption. Deep learning systems now achieve performance levels comparable to or exceeding human

specialists across numerous cognitive tasks, from medical diagnosis to natural language understanding (Alzubaidi et al., 2021; Singhal et al., 2023). Yet the organizational infrastructures required to absorb, institutionalize, and leverage the knowledge generated by such systems remain poorly understood. Organizations that deploy AI without corresponding investment in learning capabilities risk what Raisch and Krakowski (2021) describe as the automation-augmentation paradox—a condition in which the efficiency gains of automation undermine the cognitive engagement necessary for deeper organizational learning.

A research gap emerges at precisely the junction of technical AI scholarship and organizational learning theory. While the machine learning literature has advanced considerably in addressing how AI models transfer knowledge between tasks and domains—including through federated learning (Kairouz et al., 2020) and selective knowledge distillation (Yu et al., 2024)—this work largely proceeds in isolation from the organizational processes that govern how such capabilities are institutionalized and exploited. Conversely, organizational learning scholars have been slow to incorporate the structural properties of AI systems—including their tendencies toward catastrophic forgetting, distributional shift, and model opacity—into their theoretical frameworks.

This literature review addresses that gap by synthesizing findings across both technical and management-oriented research streams to answer the following primary question: How do organizational learning processes interact with AI knowledge transfer mechanisms, and what theoretical and practical implications follow from this interaction? Secondary questions include: What theoretical constructs from organizational learning are most pertinent to AI-enabled knowledge transfer? What tensions and contradictions emerge across the literature? And what research directions remain underexplored?

The remainder of this article is organized as follows. Section 2 establishes the theoretical background by reviewing foundational constructs from organizational learning and AI knowledge transfer research. Section 3 presents a thematic synthesis of the literature, organized around three substantive clusters. Section 4 identifies research gaps and proposes future directions. Section 5 concludes with a summary of contributions and limitations.

2. Theoretical Background

2.1 Organizational Learning: Core Constructs

Organizational learning has been a central concern of management theory for several decades, with roots in behavioral theories of the firm and the resource-based view. At its core, organizational learning refers to the process through which organizations acquire, interpret, distribute, and institutionalize knowledge in ways that modify organizational behavior and enhance adaptive capacity. The construct of absorptive capacity—defined as an organization's ability to recognize the value of new external knowledge, assimilate it, and apply it to commercial ends—has proven particularly durable and has been extended to encompass both potential and realized dimensions.

Makhloufi et al. (2021) provide contemporary empirical grounding for these constructs by demonstrating that organizational learning capabilities (OLC) and absorptive capacity (AC) jointly mediate the relationship between entrepreneurial orientation and innovation capability. Their findings confirm that learning capability is not a passive outcome of organizational experience but an active mediating mechanism that channels strategic orientations into innovation outputs. This framing is consequential for understanding AI integration:

organizations that possess strong learning capabilities are better positioned to absorb the knowledge encoded in AI systems and translate it into competitive advantage.

AlSaied and Alkhoraif (2024) extend this reasoning to the domain of ambidextrous innovation, arguing that innovative organizational culture and organizational learning jointly enable firms to pursue both exploitative and exploratory innovation simultaneously. Their research, conducted in the context of Saudi Arabian public-sector firms, reveals that the organizational learning–innovation relationship is moderated by cultural factors, suggesting that the generalizability of AI-enabled knowledge transfer depends substantially on organizational context.

2.2 AI as a Knowledge-Generating and Knowledge-Transferring Agent

From a technical standpoint, AI systems—particularly deep learning models—function as knowledge compressors and transfer mechanisms. A deep learning model trained on large corpora encodes statistical regularities from its training data into millions of learned parameters, which can subsequently be applied to new domains through a process called transfer learning (Alzubaidi et al., 2021; Shrestha & Mahmood, 2019). This technical capability maps onto the organizational learning concept of knowledge transfer, albeit with important structural differences: whereas human knowledge transfer is mediated by social processes such as mentoring, communities of practice, and organizational routines, AI knowledge transfer is mediated by architectural choices, training regimes, and fine-tuning protocols.

Federated learning represents a particularly instructive case at the intersection of AI knowledge transfer and organizational theory. Kairouz et al. (2020) define federated learning as a collaborative training paradigm in which many clients—which may be entire organizations—jointly train a model without sharing raw data. This architecture institutionalizes a form of distributed organizational learning in which knowledge is pooled across organizational boundaries while preserving data sovereignty. The organizational implications are substantial: federated learning enables inter-organizational knowledge transfer at scale while managing the privacy and competitive sensitivity concerns that typically constrain such transfer.

2.3 The Automation-Augmentation Tension

A foundational tension running through the literature concerns the relationship between automation and augmentation in organizational learning. Raisch and Krakowski (2021) theorize this as a paradox: automation and augmentation are not discrete alternatives but interdependent poles of a continuum, and organizations that treat them as mutually exclusive risk sub-optimal outcomes. The paradox is directly relevant to AI knowledge transfer because automated AI systems that generate knowledge may simultaneously displace the human cognitive processes through which that knowledge would otherwise be developed, critically attenuating the organization's long-term learning capacity.

This tension is also manifest in the economics of labor and technology. Acemoğlu and Restrepo (2019) demonstrate that automation systematically shifts the task content of production away from labor, with potential negative consequences for human skill development and knowledge accumulation. In organizational learning terms, automation may reduce opportunities for learning-by-doing, mentoring, and the tacit knowledge formation that underpins adaptive capacity. This displacement effect must be weighed against the augmentation benefits of AI—the enhancement of human cognitive capacities through intelligent decision support, knowledge retrieval, and pattern recognition.

2.4 E-Learning Platforms as Organizational Learning Infrastructure

The infrastructure through which organizational learning is operationalized has itself been transformed by digital technologies. Giannakos et al. (2021) conduct a systematic review of e-learning capabilities and their role in enhancing organizational learning, finding that effective e-learning systems must exhibit adaptivity, interactivity, and integration with organizational knowledge management processes. Their analysis identifies a critical alignment problem: many organizations deploy e-learning technology without redesigning the organizational processes that govern knowledge flow, resulting in a decoupling of technological capability from organizational learning outcomes. This finding has direct implications for AI integration, where a similar alignment problem frequently manifests.

3. Literature Review

3.1 Theme One: Absorptive Capacity, Learning Capabilities, and AI-Enabled Innovation

The first thematic cluster concerns the organizational antecedents of effective AI knowledge absorption and the mechanisms through which AI augments innovation capability. The literature in this cluster draws on absorptive capacity theory to analyze how organizational readiness shapes the outcomes of AI deployment.

Makhloufi et al. (2021) provide empirical evidence that absorptive capacity mediates the pathway from entrepreneurial orientation to innovation capability, confirming that external knowledge—including that generated by AI systems—requires organizational learning infrastructure to translate into innovation outcomes. Their findings suggest that organizations with high absorptive capacity are better equipped to recognize the strategic value of AI-generated insights, integrate them into existing knowledge structures, and apply them to the development of new products, processes, or services. This finding is theoretically consistent with the view that AI should be understood not as an autonomous innovation engine but as a knowledge amplifier whose organizational impact is conditioned by learning capability.

AlSaied and Alkhoraif (2024) reinforce this argument by demonstrating that innovative organizational culture moderates the learning-innovation relationship, implying that the organizational context in which AI is deployed shapes its knowledge-transfer effectiveness. In organizations characterized by bureaucratic inertia, knowledge hoarding, or risk aversion, AI systems may generate knowledge that is never fully absorbed or applied, resulting in what might be termed an AI knowledge gap—a disparity between the knowledge encoded in AI systems and the knowledge actually available for organizational use.

The clinical AI literature offers a striking illustration of this gap. Singhal et al. (2023) demonstrate that large language models can encode substantial clinical knowledge, achieving performance on medical question-answering benchmarks that approaches or matches human clinician scores. However, the translation of this encoded knowledge into institutional learning and improved patient outcomes depends critically on organizational processes of validation, integration, and professional acceptance—processes that vary widely across healthcare institutions and that the AI systems themselves cannot determine. This example illustrates a broader principle: the knowledge-generating capacity of AI systems is a necessary but insufficient condition for organizational learning; the organizational learning capacity to absorb and institutionalize AI-generated knowledge is equally essential.

Giannakos et al. (2021) extend this analysis to the e-learning domain, finding that organizational learning outcomes depend not on the sophistication of e-learning technology

per se but on the degree to which such technology is integrated into broader organizational learning processes. Their identification of key e-learning affordances—including personalization, asynchronous access, and analytics-driven feedback—maps closely onto the capabilities of contemporary AI learning systems, suggesting that the organizational learning literature on e-learning offers a valuable but underutilized resource for theorizing AI integration.

3.2 Theme Two: AI Knowledge Transfer Mechanisms—Technical Architectures and Organizational Implications

The second thematic cluster examines the specific mechanisms through which AI systems transfer knowledge, focusing on the organizational implications of architectural choices in machine learning.

Deep learning architectures represent the foundational mechanism of AI knowledge encoding and transfer. Alzubaidi et al. (2021) provide a comprehensive review of deep learning concepts and architectures, documenting how convolutional neural networks (CNNs) and related models extract hierarchical feature representations from data that can subsequently be transferred to new domains through pre-training and fine-tuning procedures. Shrestha and Mahmood (2019) complement this account by reviewing the full landscape of deep learning algorithms, emphasizing the role of architectural innovations—including residual connections, attention mechanisms, and generative adversarial networks—in expanding the scope and effectiveness of knowledge transfer across domains.

The organizational implications of these technical capabilities are substantial. When a pre-trained model is fine-tuned on organization-specific data, the resulting system embodies a form of hybrid knowledge—a combination of general statistical regularities encoded during pre-training and organization-specific patterns learned during fine-tuning. This hybrid character raises important questions about knowledge ownership, provenance, and institutional memory that have not been systematically addressed in the organizational learning literature. Yu et al. (2024) address one dimension of this problem through their Selective Dual-Teacher Knowledge Transfer framework, which uses both a pre-trained and a fine-tuned model as dual teachers to preserve previously learned knowledge while adapting to new tasks—a technical solution to the problem of catastrophic forgetting that has direct analogues in organizational learning theory, where the challenge of preserving institutional memory while adapting to new strategic demands is a perennial concern.

Federated learning, as analyzed by Kairouz et al. (2020), represents a qualitatively distinct knowledge transfer mechanism with profound organizational implications. By enabling model training across distributed data sources without centralized data pooling, federated learning instantiates a form of inter-organizational learning that preserves competitive and regulatory boundaries while enabling collective knowledge accumulation. The authors identify numerous open problems in federated learning, including communication efficiency, robustness to non-independent and identically distributed (non-IID) data, and differential privacy—challenges that translate, in organizational terms, into questions about the equity of knowledge sharing, the management of heterogeneous organizational knowledge bases, and the governance of collaborative learning arrangements.

The application of AI knowledge transfer in domain-specific organizational contexts further illustrates both the potential and the limitations of current approaches. Brunton et al. (2019) demonstrate how machine learning transfers knowledge from large-scale simulation data to real-time fluid mechanics applications, an example that generalizes to industrial settings where

AI bridges the gap between experimental knowledge and operational decision-making. Similarly, Váradi et al. (2023) illustrate how the AlphaFold protein structure prediction system—an AI model trained on evolutionary sequence data—has transferred structural biology knowledge at unprecedented scale, with over 214 million predicted protein structures made available to the global research community, fundamentally reorganizing the collective learning infrastructure of structural biology.

Digital twin technology, which Rasheed et al. (2020) characterize as a virtual representation of a physical asset enabled through data and simulators, represents another organizational knowledge transfer mechanism that integrates AI. Digital twins enable real-time monitoring, predictive maintenance, and optimization by continuously learning from sensor data and updating their internal models—a form of organizational learning that operates at machine speed and scale. The organizational challenge is to ensure that the knowledge encoded in digital twins is accessible, interpretable, and actionable by human decision-makers, raising the augmentation questions theorized by Raisch and Krakowski (2021).

3.3 Theme Three: Human-Centric Challenges—Automation, Augmentation, and Organizational Learning Capacity

The third thematic cluster addresses the human dimensions of AI knowledge transfer, focusing on the tensions between automation and augmentation and their implications for organizational learning capacity.

Raisch and Krakowski (2021) provide the most theoretically developed account of the automation-augmentation paradox, arguing that organizations face an irreducible tension between efficiency-oriented automation and learning-oriented augmentation. Drawing on paradox theory, they contend that this tension cannot be resolved through a simple strategic choice but must be managed dynamically—a claim that has significant implications for organizational learning. If automation consistently reduces opportunities for human learning-by-doing, organizations may achieve short-term efficiency gains at the cost of long-term adaptive capacity, creating a learning debt that accumulates invisibly until a discontinuous environmental change renders existing knowledge stocks obsolete.

The economic analysis of Acemoğlu and Restrepo (2019) provides complementary evidence for this concern. Their labor displacement-reinstatement framework demonstrates that automation technologies reduce the labor share in value added and may reduce labor demand even as they raise productivity, with the net employment effect depending on the pace and nature of new task creation. In organizational learning terms, the displacement of workers from tasks previously performed by humans eliminates the experiential learning that those tasks would have generated, while the new tasks that automation reinstates may require qualitatively different learning processes that organizations are not immediately equipped to support.

Nahavandi (2019) extends this analysis to the Industry 5.0 framework, arguing for a human-centric approach to AI integration that preserves cognitive engagement and collaborative human-machine problem-solving. This perspective aligns with the augmentation pole of Raisch and Krakowski's (2021) paradox and suggests that Industry 5.0 represents not merely a technological evolution but an organizational learning philosophy that privileges human agency alongside machine intelligence. The brain-machine interface technologies discussed by Nahavandi (2019) represent an extreme case of augmentation in which AI knowledge transfer occurs through direct cognitive enhancement—a prospect that highlights the extent to which the boundaries between human and machine knowledge are becoming increasingly permeable.

Service organizations face distinctive instantiations of the automation-augmentation tension. Wirtz et al. (2018) examine the deployment of service robots in frontline service contexts, documenting both the efficiency gains that automation enables and the risks to service quality and customer relationship management that arise when human judgment is displaced. Their analysis emphasizes that service robots must be understood as components of sociotechnical systems in which organizational learning about customer needs, service recovery, and relationship dynamics cannot be fully encoded in machine behavior—a finding that resonates with the broader argument that AI knowledge transfer is most effective when augmenting rather than replacing human organizational learning.

Marketing AI presents a further case in which the automation-augmentation tension manifests with strategic consequences. Huang and Rust (2020) develop a three-stage strategic framework distinguishing mechanical AI (automating repetitive functions), thinking AI (processing data for decision support), and feeling AI (analyzing human emotions and interactions). Each stage implies a different relationship between AI knowledge transfer and organizational learning: mechanical AI requires minimal organizational learning investment, thinking AI demands capability development in data interpretation and algorithmic management, and feeling AI necessitates deep integration of AI insights with human relational knowledge—the most challenging and organizationally demanding form of augmentation.

Across these diverse settings—industrial manufacturing, service delivery, and marketing—a consistent pattern emerges: the effectiveness of AI knowledge transfer in enhancing organizational learning depends on the deliberate design of human-AI interfaces that preserve cognitive engagement, support sense-making, and maintain organizational memory in forms that remain accessible and actionable to human decision-makers.

4. Research Gaps and Future Directions

Despite the substantial progress documented in this review, significant gaps persist in the understanding of how organizational learning and AI knowledge transfer interact.

First, the literature lacks integrative frameworks that bridge the technical mechanics of AI knowledge transfer with organizational learning theory. Existing work tends to proceed within disciplinary silos: machine learning researchers address catastrophic forgetting and transfer learning efficiency without reference to organizational learning constructs, while management scholars analyze absorptive capacity and learning culture without engaging with the structural properties of the AI systems organizations are deploying. Future research should develop cross-disciplinary models that map technical AI properties—such as the distinction between pre-training and fine-tuning, or the trade-off between model generalization and specialization—onto organizational learning constructs such as exploitation and exploration, knowledge codification and personalization, and organizational memory. The Selective Dual-Teacher Knowledge Transfer framework of Yu et al. (2024) offers one promising point of connection, as its preservation of pre-trained generalization capability alongside domain adaptation closely parallels the organizational ambidexterity concept developed by AlSaied and Alkhoraif (2024).

Second, the governance and equity dimensions of federated and distributed organizational learning remain undertheorized. Kairouz et al. (2020) identify numerous technical challenges in federated learning, but the organizational challenges—including the governance of inter-organizational learning consortia, the equitable distribution of learning benefits across participating organizations, and the management of asymmetric knowledge contributions—

have received minimal scholarly attention. Future research should examine how organizations structure federated learning partnerships, how they negotiate knowledge sharing agreements, and how power asymmetries between large and small organizational participants shape the distribution of learning outcomes. The practical implications of such research are considerable, particularly for small and medium enterprises that may wish to benefit from collective AI learning without possessing the data scale of larger competitors.

Third, the longitudinal dynamics of AI-enabled organizational learning remain poorly understood. Most existing studies adopt cross-sectional designs that capture the state of organizational learning at a single point in time, obscuring the dynamic processes through which AI systems and organizational learning capabilities co-evolve. Future research should employ longitudinal methodologies—including process studies, computational modeling, and natural experiments—to examine how organizational learning capabilities develop in response to AI integration over time, how the knowledge encoded in AI systems depreciates or becomes obsolete as environmental conditions change, and how organizations manage the transition from AI-assisted to AI-dependent knowledge processes. The digital twin literature, exemplified by Rasheed et al. (2020), offers a useful methodological analogy: just as digital twins continuously update their internal models in response to real-time sensor data, organizational learning theories of AI integration should capture the continuous, feedback-driven character of human-machine knowledge co-evolution.

The practical implications of these research directions are directly relevant to organizations designing AI learning architectures. Specifically, practitioners would benefit from evidence-based guidance on the organizational learning investments required to realize AI knowledge transfer benefits, the governance structures that enable equitable inter-organizational AI learning, and the change management processes that support longitudinal adaptation to AI-driven knowledge environments.

5. Conclusion

This literature review has synthesized research across organizational learning theory, machine learning, and management science to examine how organizational learning processes interact with AI knowledge transfer mechanisms. Three principal thematic clusters were identified and analyzed: the role of absorptive capacity and organizational learning capabilities in enabling AI-driven innovation; the organizational implications of specific AI knowledge transfer architectures, including deep learning, federated learning, and continual learning; and the human-centric challenges arising from the automation-augmentation paradox.

The central finding of this review is that the effectiveness of AI as a knowledge transfer mechanism is critically conditioned by organizational learning capacity. Technical advances in deep learning (Alzubaidi et al., 2021; Shrestha & Mahmood, 2019), clinical AI (Singhal et al., 2023), and knowledge distillation (Yu et al., 2024) provide powerful tools for encoding and transferring knowledge across domains and organizations. However, the organizational translation of AI-generated knowledge into competitive advantage, improved service quality, or enhanced innovation capability depends on the absorptive capacity (Makhloufi et al., 2021), learning culture (AlSaied & Alkhoraif, 2024), and human-AI interface design (Raisch & Krakowski, 2021) that organizations bring to bear.

Several limitations of this review should be acknowledged. The selection of 21 sources, while encompassing substantial disciplinary breadth, does not claim to be exhaustive, and relevant work in adjacent fields—including knowledge management, information systems, and human-

computer interaction—may have been underrepresented. Additionally, the review is necessarily constrained by the disciplinary diversity of the sources, which span technical AI research, management theory, and domain-specific applications, creating synthesis challenges that a more focused corpus might avoid. Future reviews that apply systematic bibliometric methods to a larger, more targeted literature would complement and extend the thematic analysis offered here.

Notwithstanding these limitations, this review makes a substantive contribution by identifying the theoretical bridges, empirical gaps, and practical implications that arise at the intersection of organizational learning and AI knowledge transfer—a domain whose importance will only grow as intelligent systems become more deeply integrated into organizational life.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Acemoğlu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. Journal of Economic Perspectives, 33(2), 3–30. https://doi.org/10.1257/jep.33.2.3

AlSaied, M., & Alkhoraif, A. (2024). The role of organizational learning and innovative organizational culture for ambidextrous innovation. The Learning Organization, 31(3), 289–308. https://doi.org/10.1108/tlo-06-2023-0101

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A. Q., & Duan, Y. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data, 8(1), 53. https://doi.org/10.1186/s40537-021-00444-8

Brunton, S. L., Noack, B. R., & Koumoutsakos, P. (2019). Machine learning for fluid mechanics. Annual Review of Fluid Mechanics, 52, 477–508. https://doi.org/10.1146/annurev-fluid-010719-060214

Giannakos, M., Mikalef, P., & Pappas, I. (2021). Systematic literature review of e-learning capabilities to enhance organizational learning. Information Systems Frontiers, 24(2), 619–635. https://doi.org/10.1007/s10796-020-10097-2

Huang, M.-H., & Rust, R. T. (2020). A strategic framework for artificial intelligence in marketing. Journal of the Academy of Marketing Science, 49(1), 30–50. https://doi.org/10.1007/s11747-020-00749-9

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., & Bennis, M. (2020). Advances and open problems in federated learning. Foundations and Trends® in Machine Learning, 14(1–2), 1–210. https://doi.org/10.1561/2200000083

Makhloufi, L., Laghouag, A., Sahli, A. A., & Belaîd, F. (2021). Impact of entrepreneurial orientation on innovation capability: The mediating role of absorptive capability and organizational learning capabilities. Sustainability, 13(10), 5399. https://doi.org/10.3390/SU13105399

Nahavandi, S. (2019). Industry 5.0—A human-centric solution. Sustainability, 11(16), 4371. https://doi.org/10.3390/su11164371

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. Academy of Management Review, 46(1), 192–210. https://doi.org/10.5465/amr.2018.0072

Rasheed, A., San, O., & Kvamsdal, T. (2020). Digital twin: Values, challenges and enablers from a modeling perspective. IEEE Access, 8, 21980–22012. https://doi.org/10.1109/access.2020.2970143

Shrestha, A., & Mahmood, A. (2019). Review of deep learning algorithms and architectures. IEEE Access, 7, 53040–53065. https://doi.org/10.1109/access.2019.2912200

Singhal, K., Azizi, S., Tu, T., Mahdavi, S. S., & Lee, J. (2023). Large language models encode clinical knowledge. Nature, 620(7972), 172–180. https://doi.org/10.1038/s41586-023-06291-2

Váradi, M., Bertoni, D., Magaña, P., Paramval, U., & Pidruchna, I. (2023). AlphaFold Protein Structure Database in 2024: Providing structure coverage for over 214 million protein sequences. Nucleic Acids Research, 52(D1), D368–D375. https://doi.org/10.1093/nar/gkad1011

Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., & Lu, V. N. (2018). Brave new world: Service robots in the frontline. Journal of Service Management, 29(5), 907–931. https://doi.org/10.1108/josm-04-2018-0119

Yu, Y.-C., Huang, C.-P., Chen, J.-J., Chang, K.-P., & Lai, Y.-H. (2024). Select and distill: Selective dual-teacher knowledge transfer for continual learning on vision-language models. In Proceedings of the European Conference on Computer Vision (ECCV). https://doi.org/10.48550/arXiv.2403.09296

AI Policy Frameworks for National Competitiveness: A Literature Review

Černý Matej[1] and Grisáková Nora[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: matej.cerny@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: nora.grisakova@euba.sk

Abstract

The rapid proliferation of artificial intelligence (AI) technologies has compelled governments worldwide to develop comprehensive policy frameworks aimed at sustaining and enhancing national competitiveness. This literature review examines the theoretical underpinnings, empirical evidence, and policy debates surrounding AI governance strategies at the national level. Drawing upon 21 peer-reviewed sources, the review synthesizes findings across three thematic clusters: the governance and ethical dimensions of AI policy, the economic and labor market implications of AI-driven automation, and the role of digital transformation ecosystems in enabling competitive AI strategies. Key findings indicate that while leading nations have made substantial progress in articulating AI strategies, significant gaps remain in regulatory coherence, cross-border governance harmonization, and the operationalization of ethical principles. Emerging economies, as illustrated by case studies from India and Uzbekistan, demonstrate both the aspirational potential and the structural barriers inherent in late-mover AI policy contexts. The review concludes that effective AI policy frameworks must integrate ethical governance, workforce transition mechanisms, and robust digital infrastructure investment to achieve durable national competitiveness. Future research should prioritize comparative institutional analysis, longitudinal policy evaluation, and interdisciplinary approaches to AI strategy design.

Key Words: AI policy framework, national competitiveness, innovation policy, regulatory governance, digital strategy

JEL Classification: O38, H11, O31

1. Introduction

Artificial intelligence has emerged as one of the most consequential technological forces shaping economic productivity, national security, and social development in the twenty-first century. Governments across the globe have recognized that their capacity to harness AI technologies constitutes a foundational dimension of national competitiveness, prompting a wave of national AI strategies, regulatory initiatives, and public investment programs. Yet the translation of technological potential into sustained competitive advantage remains far from straightforward. The complexity of AI ecosystems, the pace of algorithmic innovation, and the deeply interdisciplinary nature of AI governance present formidable challenges for policymakers, researchers, and practitioners alike.

Despite considerable scholarly attention to AI technologies in individual domains—ranging from healthcare and manufacturing to education and communication—the literature on AI

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

policy frameworks as instruments of national competitiveness remains comparatively fragmented. Much of the existing work addresses either the technical dimensions of AI development or the ethical and philosophical challenges of AI governance in isolation, without adequately connecting these concerns to the broader strategic imperatives that motivate national-level policy action. This fragmentation constitutes a meaningful gap in our understanding of how governments can design, implement, and evaluate AI policies in ways that generate durable competitive advantages.

This review addresses a gap at the intersection of AI governance, national competitiveness, and implementation capacity, where these dimensions are often studied separately rather than as an integrated policy problem. The present literature review is motivated by this gap. Its primary objective is to synthesize the current state of knowledge regarding AI policy frameworks for national competitiveness, drawing on a diverse body of literature that spans AI governance, digital transformation, labor economics, and strategic management. In doing so, the review aims to identify thematic convergences and tensions across the literature, highlight areas of empirical and theoretical underexploration, and propose directions for future research that can inform both scholarship and practice.

The review is structured as follows. Section 2 establishes the theoretical background and Section 3 outlines the methodological approach. Section 4 presents the substantive literature review, organized around three thematic clusters. Section 5 identifies research gaps and proposes directions for future inquiry. Section 6 concludes with a summary of the review's contributions and an acknowledgment of its limitations.

2. Theoretical Background

2.1 Defining Artificial Intelligence in the Policy Context

Any rigorous examination of AI policy frameworks must begin with a clear conceptualization of artificial intelligence itself. Dwivedi et al. (2019) offer a multidisciplinary perspective, characterizing AI as a transformative technological force capable of augmenting or replacing human tasks across industrial, intellectual, and social domains. This framing is particularly instructive for policy analysis because it foregrounds the breadth of AI's societal implications, extending far beyond any single sector or application domain. For the purposes of national policy, AI encompasses a constellation of technologies—including machine learning, natural language processing, computer vision, and autonomous decision systems—each of which generates distinct governance challenges and strategic opportunities.

The concept of national competitiveness, in turn, draws on a rich tradition in strategic management and economic geography. At its core, competitiveness in the AI era refers to a nation's capacity to develop, deploy, and derive economic and social value from AI technologies at a scale and quality that sustains or improves its relative position in the global economy. This capacity is not merely a function of technological endowment; it is equally dependent on the institutional environment, regulatory infrastructure, human capital base, and international partnerships that collectively constitute a nation's AI ecosystem.

2.2 The Governance Imperative: Ethics and Accountability

A central theoretical concern in the AI policy literature is the relationship between governance, ethics, and legitimacy. Hagendorff (2020) provides a systematic evaluation of 22 AI ethics guidelines, revealing substantial convergence around principles such as transparency, fairness, accountability, and privacy, while simultaneously identifying critical omissions—

particularly regarding enforcement mechanisms and the accountability of corporate actors. This finding has important implications for national AI strategies: the widespread articulation of ethical principles in policy documents does not, in itself, constitute effective governance unless accompanied by institutional structures capable of operationalizing those principles.

The tension between innovation and regulation is a recurring theoretical theme. Policymakers face a fundamental dilemma in that overly restrictive governance frameworks may impede the development and deployment of AI technologies, thereby undermining competitive positioning, while insufficiently regulated AI ecosystems generate social harms that erode public trust and long-term sustainability. This tension has been described as the "governance paradox" in AI policy, and its resolution requires sophisticated institutional design that can adapt to rapidly evolving technological conditions.

2.3 Automation, Labor, and the Political Economy of AI

Acemoğlu and Restrepo (2019) provide a foundational economic framework for understanding the labor market implications of AI-driven automation. Their analysis centers on the allocation of tasks between capital and labor, demonstrating that automation displaces workers from tasks they previously performed while potentially reinstating labor demand through the creation of new tasks. This framework is essential for AI policy analysis because it clarifies the conditions under which AI-driven growth is inclusive or exclusionary—a distinction with profound implications for the political sustainability of national AI strategies. Policies that fail to address labor displacement risk generating distributional conflicts that undermine the broader legitimacy of AI-oriented governance.

2.4 Digital Transformation as Strategic Infrastructure

The theoretical literature on digital transformation provides a further essential backdrop for understanding AI policy frameworks. Kraus et al. (2021) characterize digital transformation as a pervasive process of organizational and institutional change driven by digitalization, encompassing not only business processes but also environmental, societal, and institutional dimensions. For national AI strategies, this perspective underscores the importance of systemic approaches that address the full range of enabling conditions—including digital infrastructure, data governance, organizational capacity, and regulatory agility—rather than focusing narrowly on AI research and development investment.

The Industrie 4.0 paradigm, as reviewed by Kagermann and Wahlster (2022), offers a particularly instructive example of how national policy frameworks can be designed to integrate AI with broader digital transformation agendas. The German experience demonstrates that competitive AI strategies must be embedded within ecosystems that encompass advanced manufacturing, connectivity infrastructure, and human-centered design principles, suggesting that AI policy cannot be meaningfully separated from the wider digital economy strategy of which it forms a part.

2.5 Explainability, Interpretability, and Regulatory Legitimacy

A final theoretical thread of considerable importance for AI policy concerns the interpretability and explainability of AI systems. Murdoch et al. (2019) articulate a comprehensive framework for interpretable machine learning, emphasizing that the ability to understand and explain AI decisions is a prerequisite for responsible deployment across high-stakes domains. This concern is directly relevant to AI policy because regulatory frameworks increasingly require that AI systems deployed in sensitive contexts—such as criminal justice, healthcare, and public administration—be capable of providing meaningful explanations for

their outputs. Gunning and Aha (2019) reinforce this point in their account of DARPA's Explainable AI program, which explicitly frames explainability as a national security imperative. These frameworks collectively suggest that explainability is not merely a technical challenge but a governance requirement with direct implications for AI policy design.

3. Methodology

This study adopts a narrative, semi-structured literature review design aimed at synthesizing the emerging scholarly debate on artificial intelligence (AI) policy frameworks in relation to national competitiveness. Given the conceptual focus and interdisciplinary character of the topic, the review was transparent and analytically structured as the synthesis of peer-reviewed literature.

The literature search was conducted across the databases Scopus, Web of Science, and Google Scholar, which together provide broad coverage of public policy, management, information systems, and interdisciplinary technology research. The search strategy combined keywords related to the core constructs examined in this article, including "artificial intelligence policy," "national AI strategy," "AI governance," "national competitiveness," "digital transformation," "AI ethics," "explainable AI," "automation and labor," and "digital infrastructure." These terms were used both individually and in Boolean combinations in order to identify studies addressing AI as a policy, governance, and competitiveness issue rather than as a purely technical phenomenon.

The review focused primarily on literature published between 2019 and 2025, a period corresponding to the rapid expansion of national AI strategies and the consolidation of scholarly debate on AI governance and policy implementation. Earlier foundational works were considered only where they were directly relevant to the conceptual framing of governance or competitiveness. To ensure analytical consistency, the inclusion criteria prioritized peer-reviewed journal articles and scholarly book chapters published in English that explicitly addressed at least one of the following dimensions: national AI strategy design, AI governance and ethics, labor-market implications of AI, or the infrastructural and ecosystemic conditions of AI competitiveness. Excluded were purely technical engineering studies without policy implications, non-scholarly commentary, duplicated records, and sources whose relevance could not be substantiated through title, abstract, and full-text screening.

The source selection proceeded in three stages. First, potentially relevant records were identified through database searches and backward checking of references in key papers. Second, titles and abstracts were screened for topical relevance. Third, full texts were reviewed to assess conceptual fit, analytical rigor, and contribution to the article's research objective. The final corpus comprised 20 peer-reviewed sources, which were then subjected to qualitative thematic synthesis. Following an iterative reading process, the literature was categorized into three principal thematic clusters: national AI strategies and institutional capacity, ethics and governance, and digital ecosystems as enabling conditions of competitiveness. This thematic structuring made it possible to compare findings across heterogeneous studies while preserving the conceptual coherence of the review and identifying recurrent tensions, convergences, and research gaps in the current literature (Chatterjee, 2020; Radu, 2021; Guenduez & Mettler, 2023).

4. Literature Review

4.1 National AI Strategies: Architectures, Ambitions, and Institutional Realities

The first thematic cluster encompasses the comparative analysis of national AI strategies, with particular attention to the institutional architectures through which governments seek to translate policy ambitions into competitive outcomes. Chatterjee (2020) provides a detailed examination of India's AI strategy, employing focus group discussions and documentary analysis to identify a charter of policy recommendations for the Indian government. The study reveals that India's AI policy framework is characterized by ambitious aspirational goals—including the aspiration to leverage AI for economic growth and social inclusion—but faces significant adoption challenges rooted in inadequate digital infrastructure, skills shortages, and fragmented institutional coordination. Chatterjee (2020) concludes that effective AI governance in developing economies requires a phased approach that prioritizes foundational infrastructure investment, human capital development, and regulatory capacity building before attempting to compete directly with technologically advanced nations on AI innovation.

Rakha (2023) extends this comparative lens to Uzbekistan, analyzing the country's national AI strategy within the context of broader modernization objectives. The study finds that Uzbekistan conceptualizes AI primarily as an enabler of economic growth, social development, and institutional modernization, and has articulated formal strategies aimed at positioning the country as a regional AI hub. However, Rakha (2023) identifies substantial challenges related to data infrastructure, the scarcity of AI expertise, and the absence of robust legal frameworks governing AI deployment. These findings resonate with Chatterjee's (2020) observations regarding India, suggesting that the constraints facing emerging economy AI strategies are structural rather than idiosyncratic—a point with important implications for international AI governance and development assistance.

This broader strategic logic is also evident in Guenduez and Mettler's (2023) analysis of governmental AI policies, which shows that national AI strategies are structured through recurring policy narratives that frame AI simultaneously as an economic opportunity, a governance challenge, and a domain of active state involvement. Their findings suggest that such strategies do not merely define policy priorities, but also construct specific governmental roles, including those of enabler, regulator, leader, and user.

A critical tension emerges in the comparative literature between the ambition of national AI strategies and the institutional capacity required to implement them effectively. While both India and Uzbekistan have articulated formal policy frameworks, the gap between stated objectives and operational capacity is substantial. This observation raises important questions about the role of institutional design in determining whether AI strategies function as genuine instruments of competitive development or as symbolic expressions of modernity with limited practical effect.

Evidence from other developing-country contexts points in a similar direction. Demaidi (2025), in an analysis of national AI strategy formation in Palestine, shows that the effectiveness of AI policy depends not only on strategic ambition but also on baseline readiness across education, entrepreneurship, government, and research systems. The study further indicates that weak legal frameworks, limited institutional preparedness, and low levels of AI awareness can substantially constrain implementation even where policy intent is clearly articulated.

Dwivedi et al. (2019) situate these country-specific observations within a broader global context, arguing that the AI revolution is generating a multi-speed world in which nations with

advanced technological and institutional infrastructure are able to capitalize on AI-driven productivity gains while others risk being left further behind. This framing implies that national AI strategies must be evaluated not only in terms of their internal coherence but also in relation to the competitive dynamics of the global AI landscape.

4.2 Ethics, Governance, and the Regulatory Frontier

The second thematic cluster addresses the ethical and regulatory dimensions of AI policy, which have received sustained scholarly and policy attention in recent years, particularly in response to concerns over accountability, transparency, and institutional oversight. Hagendorff's (2020) evaluation of AI ethics guidelines reveals a paradox at the heart of contemporary AI governance: while the field has produced an impressive volume of ethical frameworks and principles, these documents frequently lack the specificity, enforceability, and institutional grounding necessary to translate normative commitments into practice. Hagendorff (2020) identifies significant omissions in existing guidelines, including inadequate treatment of power asymmetries between AI developers and affected communities, insufficient attention to systemic risks, and the conspicuous absence of mechanisms for holding corporate actors accountable for ethical violations.

This critique resonates with broader debates in the AI governance literature about the limitations of soft law approaches to AI regulation. The predominance of voluntary guidelines and self-regulatory initiatives in the AI governance landscape has been interpreted by some scholars as indicative of industry influence over the governance agenda, particularly where ethical commitments are not accompanied by enforceable institutional mechanisms. This reading is consistent with Radu's (2021) comparative analysis, which finds that national AI strategies tend to prioritize ethics-oriented rather than rule-based approaches and often leave public and private roles only loosely specified. Whether this interpretation is persuasive remains contested, but the observation underscores the importance of institutional design in determining the practical effectiveness of AI ethics frameworks.

This implementation gap is further emphasized by Papagiannidis et al. (2023), who argue that effective AI governance depends on the development of procedural, structural, and relational practices capable of aligning AI deployment with organizational objectives while mitigating both technical and non-technical risks. Their analysis indicates that governance effectiveness depends not only on the articulation of ethical principles, but on their institutional embedding in organizational routines and oversight mechanisms.

Murdoch et al. (2019) and Gunning and Aha (2019) converge on the view that interpretability and explainability are central requirements for trustworthy AI governance, particularly in high-stakes domains. Gunning and Aha (2019) argue that AI systems deployed in high-stakes national security applications must be capable of providing intelligible explanations for their decisions, and that achieving this goal requires fundamental advances in machine learning methodology as well as new approaches to human-AI interaction design. This perspective suggests that the technical and governance dimensions of AI policy are deeply intertwined: regulatory requirements for explainability create demand for new technical capabilities, which in turn shape the direction of AI research and development investment.

These limitations become even more pronounced at the international level. Schmitt (2022) argues that global AI governance remains a nascent and fragmented regime characterized by overlapping initiatives, dispersed authority, and relatively weak cross-border coordination. This suggests that national AI governance frameworks operate within an international environment in which institutional coherence remains limited.

The digital transformation literature adds a further dimension to the governance debate by highlighting the importance of organizational and institutional change processes that accompany AI adoption. Kraus et al. (2021) emphasize that digital transformation is not merely a technological phenomenon but a socio-institutional process that requires the active management of change at multiple levels of analysis. For AI governance, this perspective implies that effective regulatory frameworks must address not only the properties of AI systems themselves but also the organizational contexts in which those systems are embedded and the social processes through which they are adopted and legitimated.

4.3 Digital Ecosystems, Infrastructure, and the Enabling Conditions for AI Competitiveness

The third thematic cluster examines the enabling conditions—technological, organizational, and infrastructural—that determine whether national AI strategies can generate durable competitive advantages. The literature in this cluster converges on the view that AI competitiveness is fundamentally an ecosystemic phenomenon, requiring the coordinated development of multiple interdependent capabilities rather than isolated investment in any single dimension of the AI stack.

Kagermann and Wahlster (2022) offer a particularly rich account of this ecosystemic logic in their retrospective analysis of the Industrie 4.0 initiative. Over a decade of implementation, Germany's approach to AI-integrated manufacturing has demonstrated that competitive advantage in AI-driven production depends critically on the integration of physical and digital infrastructure, the cultivation of human-centered design principles, and the development of flexible governance frameworks that can adapt to rapidly evolving technological conditions. Kagermann and Wahlster (2022) observe that the most significant advances have occurred at the intersection of AI, data infrastructure, and organizational innovation, suggesting that policy interventions must target the interfaces between these domains rather than treating them as distinct silos.

The role of advanced communication infrastructure as an enabling condition for AI competitiveness is addressed in several papers reviewed here. Wang et al. (2023) analyze the requirements and enabling technologies for sixth-generation (6G) wireless networks, arguing that the ultra-low latency, massive connectivity, and high reliability characteristics of 6G will be essential for enabling AI applications that require real-time data processing and actuation. This perspective aligns with Wu et al. (2021), whose analysis of intelligent reflecting surfaces highlights the potential for AI-driven optimization of wireless network performance, and with Jiang et al. (2021), who situate these technical advances within the broader context of global competition to define and deploy next-generation communication standards. These contributions suggest that national AI competitiveness strategies must encompass investment in communication infrastructure as a foundational enabling condition.

Acemoğlu and Restrepo (2019) introduce a critical counterpoint to the predominantly optimistic narrative of AI-driven competitiveness by foregrounding the labor market disruptions that accompany automation. Their framework demonstrates that the competitive benefits of AI adoption at the firm or national level can be accompanied by significant distributional costs, particularly for workers in routine cognitive and manual tasks. This observation has important policy implications: national AI strategies that fail to address workforce transition will generate political resistance and social instability that ultimately undermine competitive objectives. The most effective AI policy frameworks, from this

perspective, are those that integrate investment in AI development with robust mechanisms for workforce upskilling, social protection, and labor market adjustment.

This perspective can be further refined by Agrawal, Gans, and Goldfarb (2019), who conceptualize recent advances in artificial intelligence primarily as the automation of prediction rather than the wholesale replacement of human judgment. From this viewpoint, the labor market effects of AI are inherently ambiguous, since automated prediction may substitute for some tasks while simultaneously augmenting decision-making and enabling the emergence of new forms of work.

Kraus et al. (2021) reinforce this systemic perspective by documenting the institutional and organizational dimensions of digital transformation that are frequently underweighted in AI policy discussions. Their review emphasizes that the competitive benefits of digital technologies, including AI, are mediated by organizational capability, managerial practices, and institutional environment—factors that are often more difficult to develop than the technologies themselves. This finding suggests that national AI strategies must devote substantial attention to organizational and human capital development, not only at the level of individual firms but across the institutional ecosystem as a whole.

The emergence of generative AI tools, as analyzed by Baidoo-Anu and Ansah (2023) in the context of education, provides a further illustration of the ecosystemic dynamics of AI competitiveness. The rapid proliferation of tools such as ChatGPT has created new challenges and opportunities for educational systems, with implications for the human capital pipeline that ultimately determines the long-term sustainability of national AI strategies. Baidoo-Anu and Ansah (2023) observe that generative AI has the potential to significantly transform educational practice, raising fundamental questions about curriculum design, assessment validity, and the skills that educational systems should prioritize in preparing students for an AI-intensive economy. These questions are directly relevant to AI policy, since the quality and orientation of educational systems constitute a critical determinant of national AI competitiveness over the medium to long term.

Singh et al. (2021) contribute a further dimension to the ecosystem analysis through their examination of digital twin technology, which provides governments and industry actors with powerful tools for modeling and optimizing complex sociotechnical systems. The deployment of digital twins in industrial and urban planning contexts creates new data infrastructure assets that can, in turn, support more sophisticated AI applications and more evidence-based AI governance. Taken together, these contributions paint a picture of AI competitiveness as fundamentally dependent on the depth and coherence of national digital ecosystems rather than on AI capability alone.

5. Research Gaps and Future Directions

Notwithstanding the substantial body of literature reviewed above, several significant gaps in the current state of knowledge merit explicit attention.

The first and perhaps most consequential gap concerns the absence of rigorous longitudinal evaluations of national AI strategies. The comparative studies of India and Uzbekistan by Chatterjee (2020) and Rakha (2023), while informative, are largely cross-sectional in design and therefore unable to track the evolution of AI policy frameworks over time or to establish causal connections between policy interventions and competitive outcomes. Future research should prioritize longitudinal comparative case studies that follow national AI strategies from formulation through implementation, documenting the institutional learning processes,

political dynamics, and feedback mechanisms that shape policy trajectories. Such research would generate the empirical foundations necessary for more prescriptive guidance on AI policy design.

A second gap concerns the limited theoretical integration between the AI governance literature and the broader field of comparative political economy. While scholars such as Acemoğlu and Restrepo (2019) have made important contributions to understanding the macroeconomic dynamics of AI-driven automation, the political processes through which distributional conflicts over AI are mediated within national institutional frameworks remain poorly understood. Future research should draw more systematically on comparative institutionalism and varieties of capitalism theory to analyze how different national institutional configurations—including labor market institutions, innovation systems, and welfare state architectures—shape the design and effectiveness of AI policy frameworks. This theoretical integration would help explain observed cross-national variation in AI strategy design and implementation effectiveness.

A third gap relates to the governance of AI in the context of international technology competition. The existing literature on national AI strategies tends to treat each country's policy framework as a largely autonomous endeavor, without adequately theorizing the competitive dynamics between nations that shape both the incentives for AI investment and the constraints on regulatory action. As Wang et al. (2023) and Jiang et al. (2021) suggest in the context of 6G communications, the competition to define and deploy next-generation technologies has significant geopolitical dimensions that interact with national AI strategies in complex ways. Future research should develop analytical frameworks for understanding how international competitive dynamics—including technology export controls, standards competition, and AI research alliances—interact with national AI policy frameworks to shape competitive outcomes.

Beyond these three directions, the practical implications of the identified gaps are substantial. Policymakers who lack longitudinal evidence on the effectiveness of AI strategies are likely to make suboptimal investments and to miss opportunities for institutional learning. The absence of theoretical frameworks that integrate political economy with AI governance may lead to policy designs that underestimate distributional conflicts and their political consequences. And the failure to adequately theorize international competitive dynamics may result in national AI strategies that are internally coherent but strategically naive in a geopolitically complex world.

6. Conclusion

This literature review has synthesized current scholarly knowledge regarding AI policy frameworks for national competitiveness across three thematic clusters: national AI strategy architectures and institutional realities, the ethics and governance of AI, and the enabling conditions of digital ecosystems for AI competitiveness. Three main findings emerge from this synthesis.

First, the literature reveals a persistent gap between the ambition of national AI strategies and the institutional capacity required to implement them effectively. This gap is particularly visible in emerging economies, but it is not absent from advanced industrial nations. Effective AI policy frameworks therefore depend not only on strategic vision, but also on the sustained development of institutional capacity, human capital, and enabling infrastructure.

Second, the governance of AI is shaped by a persistent tension between the imperative to promote innovation and the imperative to prevent harm. While ethical principles and

guidelines have proliferated, the enforcement mechanisms and institutional structures required to make AI governance effective remain underdeveloped in most national contexts.

Third, AI competitiveness should be understood as an ecosystemic phenomenon rather than a narrowly technological one. The competitive benefits of AI cannot be realized through investment in AI alone; they require the coordinated development of communication infrastructure, data governance frameworks, organizational capabilities, workforce skills, and educational systems.

Several limitations of the present review should be acknowledged. The analysis is based on a purposively selected corpus of 20 sources that, although sufficiently diverse for the purposes of thematic synthesis, does not capture the full scope of the expanding literature on AI policy. In addition, the review was not conducted in accordance with a formal systematic review protocol, which may limit the transparency and replicability of the search and selection procedure. Moreover, the broad comparative orientation of the study inevitably restricts the extent to which individual national settings and particular policy dimensions can be explored in greater analytical depth. Future research addressing the gaps outlined in Section 4 will therefore be necessary to refine the evidence base and to strengthen both academic debate and policy practice in this dynamically developing area.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Acemoğlu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. The Journal of Economic Perspectives, 33(2), 3–30. https://doi.org/10.1257/jep.33.2.3

Agrawal, A., Gans, J. S., & Goldfarb, A. (2019). Artificial intelligence: The ambiguous labor market impact of automating prediction. The Journal of Economic Perspectives, 33(2), 31–50. https://doi.org/10.1257/jep.33.2.31

Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. Journal of AI, 7(1). https://doi.org/10.61969/jai.1337500

Chatterjee, S. (2020). AI strategy of India: Policy framework, adoption challenges and actions for government. Transforming Government: People, Process and Policy, 14(5), 757–775. https://doi.org/10.1108/tg-05-2019-0031

Demaidi, M. N. (2025). Artificial intelligence national strategy in a developing country. AI & Society, 40, 423–435. https://doi.org/10.1007/s00146-023-01779-x

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., & Coombs, C. (2019). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Guenduez, A. A., & Mettler, T. (2023). Strategically constructed narratives on artificial intelligence: What stories are told in governmental artificial intelligence policies? Government Information Quarterly, 40(1), 101719. https://doi.org/10.1016/j.giq.2022.101719

Gunning, D., & Aha, D. W. (2019). DARPA's explainable artificial intelligence program. AI Magazine, 40(2), 44–58. https://doi.org/10.1609/aimag.v40i2.2850

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. Minds and Machines, 30(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Jiang, W., Han, B., Habibi, M. A., & Schotten, H. D. (2021). The road towards 6G: A comprehensive survey. IEEE Open Journal of the Communications Society, 2, 334–366. https://doi.org/10.1109/ojcoms.2021.3057679

Kagermann, H., & Wahlster, W. (2022). Ten years of Industrie 4.0. Sci, 4(3), 26. https://doi.org/10.3390/sci4030026

Kraus, S., Jones, P., Kailer, N., Weinmann, A., & Chaparro-Banegas, N. (2021). Digital transformation: An overview of the current state of the art of research. SAGE Open, 11(3). https://doi.org/10.1177/21582440211047576

Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Definitions, methods, and applications in interpretable machine learning. Proceedings of the National Academy of Sciences, 116(44), 22071–22080. https://doi.org/10.1073/pnas.1900654116

Papagiannidis, E., Enholm, I. M., Dremel, C., Mikalef, P., & Krogstie, J. (2023). Toward AI governance: Identifying best practices and potential barriers and outcomes. Information Systems Frontiers, 25, 123–141. https://doi.org/10.1007/s10796-022-10251-y

Radu, R. (2021). Steering the governance of artificial intelligence: National strategies in perspective. Policy & Society, 40(2), 178–193. https://doi.org/10.1080/14494035.2021.1929728

Rakha, N. A. (2023). Artificial intelligence strategy of Uzbekistan: Policy framework, preferences, and challenges. International Journal of Law and Policy, 1(2). https://doi.org/10.59022/ijlp.27

Schmitt, L. (2022). Mapping global AI governance: A nascent regime in a fragmented landscape. AI and Ethics, 2(4), 883–899. https://doi.org/10.1007/s43681-021-00083-y

Singh, M., Fuenmayor, E., Hinchy, E. P., Qiao, Y., & Murray, N. (2021). Digital twin: Origin to future. Applied System Innovation, 4(2), 36. https://doi.org/10.3390/asi4020036

Wang, C.-X., You, X., Gao, X., Zhu, X., & Li, Z. (2023). On the road to 6G: Visions, requirements, key technologies, and testbeds. IEEE Communications Surveys & Tutorials, 25(3), 1–56. https://doi.org/10.1109/comst.2023.3249835

Wu, Q., Zhang, S., Zheng, B., You, C., & Zhang, R. (2021). Intelligent reflecting surface-aided wireless communications: A tutorial. IEEE Transactions on Communications, 69(5), 3313–3351. https://doi.org/10.1109/tcomm.2021.3051897

Government Interventions in Digital Transformation: A Structured Literature Review of Policies, Challenges, and Future Directions

Štetka Peter[1] and Bolek Vladimír[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: peter.stetka@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: vladimir.bolek@euba.sk

Abstract

Digital transformation has emerged as one of the most consequential forces reshaping economies, industries, and social systems in the twenty-first century. Despite growing scholarly attention to organizational and technological dimensions of this phenomenon, the role of government interventions in steering, enabling, and regulating digital transformation remains underexplored as a cohesive area of inquiry. This literature review synthesizes evidence from seventeen peer-reviewed publications to examine how public policies, regulatory frameworks, and institutional mechanisms influence the trajectory of digital transformation across sectors. The review identifies three principal thematic clusters: the governance of emerging technologies such as artificial intelligence and financial technology; the role of public investment and crisis-driven policy responses in accelerating digitalization; and the intersection of digital transformation with sustainability and inclusive development agendas. The review applies a qualitative thematic synthesis based on predefined inclusion and exclusion criteria. Key findings indicate that governments play a decisive role in shaping the pace, equity, and direction of digital transitions, yet significant gaps persist in our understanding of coordinated multi-level governance, policy evaluation, and long-term societal outcomes. The review concludes with recommendations for future research and highlights implications for policymakers seeking to design effective interventions in an era of accelerating technological change.

Key Words: digital transformation, government intervention, public policy, ICT adoption, institutional support

JEL Classification: H11, O38, O33

1. Introduction

The rapid proliferation of digital technologies is fundamentally altering the structure of economies, the organization of firms, and the lived experiences of citizens worldwide. From the deployment of artificial intelligence across industrial and service sectors to the digitalization of financial systems and the emergence of smart manufacturing paradigms, the scope and pace of digital transformation present profound opportunities and equally significant governance challenges. While considerable scholarly effort has been devoted to understanding digital transformation from organizational, managerial, and technological perspectives, the role that governments and public institutions play in shaping these transitions has received comparatively less systematic attention in the academic literature.

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

This gap is consequential. Digital transformation does not unfold in a policy vacuum. Regulatory frameworks define permissible uses of data and algorithmic systems; public investment programs fund infrastructure and research; industrial policies such as Germany's Industrie 4.0 initiative actively promote the adoption of advanced manufacturing technologies; and crisis-response packages, as demonstrated during the COVID-19 pandemic, can either entrench or disrupt existing digital trajectories. The interplay between state action and technological change is therefore central to understanding not merely how digital transformation happens, but for whom and with what consequences.

Despite the relevance of these questions, the existing literature tends to treat government intervention as a background condition rather than as a primary object of analysis. Reviews of digital transformation research typically foreground firm-level strategies, leadership behaviors, and technology adoption decisions, while policy dimensions remain peripheral or undertheorized. Similarly, literature on industrial policy and innovation governance rarely engages substantively with the specific dynamics of digital transformation. This fragmentation limits the capacity of scholars and policymakers to draw on an integrated body of knowledge when designing and evaluating public interventions.

The present review addresses this gap by synthesizing findings from a carefully selected body of literature spanning artificial intelligence governance, financial technology regulation, Industry 4.0 policy, pandemic fiscal responses, sustainability transitions, and related domains. The objective is to map the current state of knowledge regarding government interventions in digital transformation, to identify areas of scholarly consensus and debate, and to delineate the most pressing gaps that future research should address.

The article contributes to the literature by integrating fragmented evidence across industrial policy, crisis governance, fintech regulation, and sustainability-oriented digital transformation into a single public-policy-oriented analytical framework. The remainder of the article is structured as follows. Section 2 establishes the theoretical background by defining key constructs and outlining foundational frameworks relevant to understanding state intervention in technological transitions. Section 3 presents a thematically organized synthesis of the literature across three clusters. Section 4 identifies research gaps and proposes future directions. Section 5 concludes with a summary of contributions, limitations, and broader implications.

2. Theoretical Background

2.1 Defining Digital Transformation

Digital transformation is a multidimensional concept that has attracted a wide range of definitional approaches in the scholarly literature. At its broadest, digital transformation refers to the process by which digital technologies are integrated into virtually all areas of economic and social life, fundamentally altering how value is created, delivered, and captured. Kraus et al. (2021) characterize digital transformation as extending well beyond the organizational level, encompassing environmental, societal, and institutional implications that arise from the deep digitalization of economies. This framing is significant because it elevates digital transformation from a matter of firm-level strategy to a systemic, economy-wide phenomenon that inherently invites public policy responses.

The concept of digital transformation is closely related to, but distinct from, the narrower notions of digitization and digitalization. Digitization refers to the conversion of analog information into digital formats, while digitalization denotes the use of digital technologies to

improve or modify existing processes. Digital transformation, by contrast, implies a more thoroughgoing reconstitution of organizational structures, business models, and societal arrangements, driven by but not reducible to technological change (Kraus et al., 2021). Almeida et al. (2020) further emphasize that this transformative process accelerated dramatically in the context of the COVID-19 pandemic, as firms and public institutions were compelled to adopt digital solutions across labor relations, marketing, and operational processes with unprecedented speed.

2.2 Theories of Government Intervention in Technological Change

Understanding why and how governments intervene in processes of technological transformation requires engagement with several bodies of theoretical literature. The most foundational of these is market failure theory, which holds that public intervention is justified when markets fail to produce socially optimal outcomes. In the context of digital transformation, market failures may arise from information asymmetries, network externalities, coordination problems, and the public-good character of knowledge and infrastructure. These conditions create a rationale for government action in the form of regulation, subsidies, standards-setting, and public investment.

A complementary theoretical lens is provided by the innovation systems perspective, which emphasizes the role of institutions, organizations, and their interactions in shaping the direction and pace of technological change. From this perspective, governments are not merely correctors of market failures but active architects of the environments in which technological innovation occurs. The German Industrie 4.0 initiative offers a paradigmatic example: Kagermann and Wahlster (2022) describe how a deliberate, government-supported national strategy succeeded in establishing a globally dominant paradigm for the digital transformation of manufacturing, integrating research institutions, industry associations, and federal policy instruments into a coherent national innovation system.

The concept of digital leadership at the organizational level provides a micro-level complement to these macro-level theories. AlNuaimi et al. (2022) demonstrate that digital transformational leadership and organizational agility are significant drivers of digital transformation, with digital strategy functioning as a critical moderating variable. While this work focuses on firms, its insights are transferable to the public sector: government agencies that aspire to lead digital transformation must cultivate adaptive capacities and strategic clarity comparable to those required of private organizations.

2.3 Artificial Intelligence as a Governance Challenge

Artificial intelligence occupies a central place in contemporary discussions of digital transformation, both as a driver of change and as an object of governance. Dwivedi et al. (2019) provide a comprehensive multidisciplinary account of AI's transformative potential across industrial, intellectual, and social domains, and articulate an extensive research and policy agenda addressing issues of accountability, transparency, labor displacement, and ethical responsibility. These concerns directly implicate governments, which must develop regulatory frameworks capable of managing the risks and distributing the benefits of AI-driven transformation without stifling innovation.

The financial sector represents a particularly instructive domain for understanding the governance of AI-enabled digital transformation. Mhlanga (2020) examines how AI applications are reshaping digital financial inclusion, enabling fintech companies to extend formal financial services to populations previously excluded from mainstream markets. Demir

et al. (2022) corroborate this finding using quantitative analysis, demonstrating that fintech adoption is associated with reductions in income inequality, particularly among lower-income populations. However, both studies implicitly acknowledge that the realization of these benefits is contingent on supportive regulatory environments and deliberate policy design, underscoring the indispensable role of government in mediating the social outcomes of digital transformation.

2.4 Sustainability and the Social Dimensions of Digital Transitions

A third theoretical strand relevant to understanding government intervention concerns the relationship between digital transformation and sustainable development. Hariram et al. (2023) argue that existing economic paradigms have proven inadequate for addressing the interlocking challenges of social equity, environmental sustainability, and economic development, and propose an integrated model—sustainalism—that places quality of life and social justice at the center of development policy. This framework has direct implications for the governance of digital transformation, as it demands that policymakers consider not merely the economic efficiency of digital interventions but their distributional and ecological consequences. Similarly, Suchek et al. (2021) demonstrate that the transition to a circular economy requires firms to rethink their business models and innovation strategies, a process that is significantly shaped by public policy incentives and regulatory requirements.

3. Methodology

The aim of this study is to systematically examine the academic literature concerning the role of government interventions in the process of digital transformation and to identify the principal thematic areas, dominant approaches, and gaps in current research. Methodologically, the article is conceived as a systematic literature review, with the objective of collecting, selecting, and synthesising relevant scholarly publications in a transparent and reproducible manner.

The review was conducted in accordance with the principles of a systematic review, with particular emphasis on an explicitly defined search strategy, clearly established inclusion and exclusion criteria, and the thematic synthesis of findings. Given the interdisciplinary nature of the subject under investigation, which brings together the fields of public policy, the digital economy, innovation systems, regulation, and management, a broader conceptual search framework was adopted in order to capture literature from several interrelated domains.

3. 1 Source selection

The literature search focused on scholarly, peer-reviewed articles published in international academic journals. Particular attention was paid to studies addressing the relationship between digital transformation and public policy, state regulation, public investment, institutional support, or crisis-driven interventions affecting processes of digitalisation. Relevant sources were identified through combinations of keywords such as digital transformation, government intervention, public policy, digitalisation, artificial intelligence governance, fintech regulation, Industry 4.0 policy, public investment and digital transformation, digital inclusion, and institutional support. The search strategy was designed to capture not only studies examining explicit government policies related to digital transformation, but also broader forms of public intervention shaping the conditions of technological change, including regulatory frameworks, innovation support, industrial policy, crisis-response measures, and policies promoting inclusive and sustainable development.

Source selection was guided by clearly defined inclusion and exclusion criteria. Studies were included in the review if they were peer-reviewed academic publications, published in the English language, thematically related to digital transformation or closely associated processes of digital change, examined the role of the state, public institutions, regulatory frameworks, or public policy, and provided theoretical, empirical, or systematically structured insights relevant to the topic under investigation. By contrast, studies were excluded if they focused exclusively on internal organisational processes without any connection to public policy or the regulatory environment, were purely technical in nature without a socio-economic or institutional dimension, had not undergone peer review, or did not demonstrate a sufficiently clear link to the issue of government interventions. These criteria were intended to ensure that the final body of literature consisted of studies relevant to understanding the state as an actor in digital transformation, while avoiding the inclusion of works that, although concerned with digitalisation, did not address the public policy dimension of the research problem.

3. 2 Analytical Procedure and Thematic Synthesis

Following the selection of relevant studies, a qualitative thematic synthesis was employed. The individual publications were analysed in terms of the specific domain of digital transformation under consideration, the type of government intervention or public policy instrument involved, the benefits and risks identified, the broader social, economic, and institutional implications, and the recommendations offered for future research and public policy. On the basis of this analysis, three principal thematic clusters were identified, which also structure the core of the literature review. The first cluster encompasses national strategies, industrial policies, and institutional coordination in digital transformation. The second focuses on crisis-driven state interventions and their accelerating effect on digitalisation. The third examines regulatory challenges associated with artificial intelligence, fintech, and inclusive digital development. This approach made it possible not only to summarise the existing body of knowledge, but also to compare systematically different approaches to government intervention and to identify areas in which current research remains fragmented or methodologically underdeveloped.

4. Literature Review

4.1 Industrial Policy and National Digital Transformation Strategies

One of the most prominent themes in the literature on government interventions in digital transformation concerns the design and implementation of national industrial policies aimed at accelerating the adoption of digital technologies in manufacturing and other sectors. The Industrie 4.0 initiative, launched in Germany in 2011 with substantial government support, represents the most extensively studied example of this type of intervention. Kagermann and Wahlster (2022) document a decade of progress under this initiative, noting that it has established Industrie 4.0 as the dominant global paradigm for the digital transformation of manufacturing. The initiative's success is attributed to its systematic integration of research-based knowledge with practical industrial experience, its human-centered philosophy, and its emphasis on data-driven value creation and innovative business models. Importantly, the authors identify continued government coordination as essential to addressing emerging challenges in areas such as data sovereignty, sustainability, and the reorganization of work.

The broader scholarly debate on Industry 4.0 illuminates both the opportunities and the governance challenges associated with state-led digital industrial policy. Büchi et al. (2019) provide empirical evidence that a greater openness to Industry 4.0 enabling technologies is

associated with improved firm performance, measured in terms of business opportunities realized. However, their findings also suggest that the relationship between technology adoption and performance is heterogeneous across firm types, implying that blanket industrial policies may not be equally effective for all categories of enterprises. This heterogeneity has important implications for the design of government support programs, which may need to be differentiated by firm size, sector, and technological readiness.

Machado et al. (2019) extend this analysis by examining the intersection of Industry 4.0 and sustainable manufacturing, arguing that the full potential of digital industrial policy can only be realized when sustainability considerations are integrated into its design from the outset. Their systematic review identifies a nascent but growing research agenda linking digital transformation in manufacturing to environmental and social sustainability objectives, an integration that mirrors broader policy trends toward green industrial strategies. Longo et al. (2020) further enrich this perspective by introducing the concept of Industry 5.0, which emphasizes the reconciliation of human and machine capabilities and raises critical ethical questions about the design of automated production systems. Their work implicitly calls for regulatory and governance frameworks that extend beyond economic efficiency to encompass worker well-being, social equity, and ethical technology design.

The university sector occupies a distinctive position within national digital transformation strategies, serving simultaneously as a producer of knowledge, a trainer of talent, and an innovator in its own right. Compagnucci and Spigarelli (2020) analyze what they term the "Third Mission" of universities—their contribution to society beyond teaching and research—and identify both the potential of and the constraints on university engagement in digital innovation ecosystems. Their findings suggest that government policy plays a critical role in creating the conditions under which universities can effectively fulfil this third mission, through funding mechanisms, intellectual property frameworks, and collaborative infrastructure. This insight highlights the systemic character of effective digital transformation governance, which must coordinate actors across public, private, and academic spheres.

4.2 Crisis-Driven Policy Responses and the Acceleration of Digitalization

A second major thematic cluster in the literature concerns the role of exceptional, crisis-driven government interventions in accelerating digital transformation. The COVID-19 pandemic provided an unprecedented natural experiment in this regard, compelling governments worldwide to mobilize massive fiscal resources in response to an acute economic shock while simultaneously creating conditions for rapid digital adoption across sectors.

Hepburn et al. (2020) examine how COVID-19 fiscal recovery packages could be designed to serve dual objectives of economic recovery and climate progress. Surveying 231 central bank officials, finance ministry officials, and economic experts from G20 countries, they identify a set of fiscal archetypes with high potential on both economic multiplier and climate impact dimensions. Although their primary focus is on climate change, their analysis has direct implications for digital transformation governance, as several of the high-performing policy archetypes—including investments in digital infrastructure, clean energy systems, and research and development—are simultaneously drivers of digitalization. The authors' findings suggest that crisis recovery packages represent a critical juncture at which governments can either entrench incumbent technological systems or catalyze transformative shifts toward more sustainable and digitally advanced economies.

Almeida et al. (2020) provide a complementary analysis of the digitalization pressures generated by the COVID-19 crisis at the firm level. Their exploratory study documents the

accelerated adoption of digital technologies across labor relations, marketing and sales, and operational processes, and identifies both the opportunities this acceleration creates and the challenges it poses for organizations that lack the absorptive capacity or infrastructure to manage rapid digital transitions. The policy implication is clear: governments that invest proactively in digital infrastructure, workforce skills, and regulatory readiness are better positioned to help their economies navigate crises with digital tools and to ensure that the benefits of accelerated digitalization are broadly distributed rather than concentrated among already-advantaged actors.

Juergensen et al. (2020) focus specifically on the vulnerability of small and medium-sized enterprises (SMEs) in European manufacturing to the COVID-19 shock and draw out the policy implications for digital transformation support. Their analysis distinguishes between the immediate effects of the pandemic—including supply chain disruptions, demand collapses, and liquidity crises—and its longer-term implications for SMEs' positioning within global value chains. They argue that effective policy responses must address both the acute financial vulnerabilities of SMEs and the structural barriers to digital adoption that the crisis has brought into sharp relief. This dual-track argument underscores the importance of government interventions that combine emergency support with long-term investment in digital capabilities, particularly for the smaller enterprises that comprise the backbone of many European economies.

The intersection of crisis response, digital transformation, and public health governance is illustrated by the National COVID Cohort Collaborative (N3C) initiative analyzed by Haendel et al. (2020). This federally supported infrastructure project demonstrated that multiorganizational clinical data networks—built on principles of data sharing, collaborative governance, and scalable digital infrastructure—can generate rapid, actionable insights for clinical and research communities facing a health emergency. While the N3C is primarily a public health initiative, it exemplifies a broader model of government-supported digital transformation in which public investment in data infrastructure creates spillover benefits for innovation, evidence-based policymaking, and service delivery.

4.3 Regulatory Governance of Fintech, AI, and Inclusive Digital Finance

A third thematic cluster addresses the regulatory governance of digitally enabled financial services and artificial intelligence, with particular attention to the equity implications of these transformations. This cluster is notable for the convergence of technological, economic, and social policy concerns, and for the prominence of distributional questions that are often absent from more technology-centered accounts of digital transformation.

Dwivedi et al. (2019) situate the governance of AI within a broader account of its transformative potential, arguing that the speed and scope of AI-driven change require an equally ambitious policy and regulatory response. They identify accountability, transparency, data privacy, and labor market impacts as among the most pressing policy challenges, and call for multidisciplinary research and policy communities to develop governance frameworks adequate to the complexity of AI systems. This call resonates strongly in the context of government-led digital transformation initiatives, where the deployment of AI in public services—from welfare administration to healthcare—raises acute questions about accountability, fairness, and citizen rights.

Mhlanga (2020) examines the specific case of AI-enabled digital financial inclusion in the context of Industry 4.0, arguing that fintech companies leveraging AI can extend formal financial services to low-income earners, women, youth, and small businesses that have

historically been excluded from mainstream financial systems. This argument frames AI-driven digitalization as a potential instrument of social policy, capable of advancing financial inclusion objectives that have traditionally been pursued through regulatory mandates and public programs. However, Mhlanga (2020) also acknowledges that the realization of these inclusion benefits is not automatic and depends on the presence of enabling regulatory environments, digital literacy infrastructure, and consumer protection frameworks—all of which are fundamentally matters of public policy.

Demir et al. (2022) provide rigorous quantitative corroboration for the inclusion-enhancing potential of fintech, demonstrating through quantile regression analysis that fintech adoption is associated with significant reductions in income inequality, with the largest effects concentrated among the poorest income groups. This finding has direct implications for government policy: regulatory frameworks that enable responsible fintech innovation, and public investments that extend digital access to underserved populations, may be among the most effective instruments available for addressing income inequality in digitally transforming economies. However, the authors also caution that the distributional benefits of fintech are conditional on broader institutional quality and financial market development, suggesting that fintech regulation cannot be treated in isolation from wider economic governance reforms.

The tension between innovation promotion and risk regulation in the fintech domain illustrates a broader challenge facing governments across all domains of digital transformation governance. Regulatory frameworks that are too restrictive may impede the development and diffusion of beneficial technologies, while frameworks that are too permissive may allow harms to proliferate before corrective action can be taken. This dynamic is evident not only in financial services but also in areas such as AI governance, data protection, and platform regulation, where governments worldwide are actively experimenting with different regulatory models. The literature reviewed here does not offer a single optimal regulatory model but consistently underscores the importance of adaptive, evidence-based governance that can evolve in response to the rapidly changing technological landscape.

AlNuaimi et al. (2022) contribute to this discussion by demonstrating that digital strategy plays a critical moderating role in the relationship between leadership, agility, and digital transformation outcomes. Although their study is set in the private sector, its findings have implications for public sector digital transformation, suggesting that government agencies require not only technical capabilities and political will but also coherent, strategically articulated digital transformation roadmaps if they are to manage the transition effectively. The parallel with Kagermann and Wahlster's (2022) account of the Industrie 4.0 strategy is instructive: in both cases, the availability of a clear, well-resourced strategic framework appears to be a significant determinant of successful digital transformation outcomes.

5. Research Gaps and Future Directions

Despite the growing volume of scholarship on digital transformation and the increasing policy salience of government interventions in this domain, the literature exhibits several important gaps that merit attention from future researchers.

A first and fundamental gap concerns the evaluation of government digital transformation policies. The reviewed literature documents a wide range of policy instruments—from national industrial strategies to crisis recovery packages to fintech regulatory frameworks—but provides limited systematic evidence on the effectiveness of these interventions relative to counterfactual scenarios or alternative policy designs. Kagermann and Wahlster (2022) offer a

rich descriptive account of the Industrie 4.0 initiative's progress, but the causal identification of policy effects remains methodologically challenging. Future research should develop more rigorous evaluation frameworks, drawing on natural experiments, quasi-experimental designs, and comparative case studies, to generate credible causal evidence on the conditions under which government interventions in digital transformation succeed or fail. This evidence base is urgently needed to inform the design of the next generation of digital industrial policies, particularly as governments worldwide are committing substantial public resources to digital transformation initiatives.

A second significant gap concerns the governance of digital transformation across multiple levels of government and across national borders. The literature tends to focus either on national-level industrial policies or on firm-level responses to technological change, leaving the intermediate levels of regional and municipal governance, as well as the transnational dimensions of digital regulation, undertheorized. Hepburn et al. (2020) touch on the G20 dimension of climate-digital policy coordination, and Juergensen et al. (2020) analyze European-level SME policy responses, but a systematic account of how digital transformation governance is coordinated—or fails to be coordinated—across governmental levels and international jurisdictions remains to be developed. Future research should examine the conditions for effective multi-level governance of digital transformation, drawing on comparative political economy and European integration scholarship to understand how governance architectures shape policy outcomes.

A third gap relates to the equity and inclusion dimensions of government-led digital transformation. While Demir et al. (2022) and Mhlanga (2020) provide important evidence on the distributional effects of fintech and AI-enabled financial inclusion, the literature lacks comprehensive assessments of whether and how government digital transformation policies address the needs of disadvantaged populations, including the elderly, the digitally illiterate, rural communities, and low-income households. The sustainalism framework proposed by Hariram et al. (2023) offers a normative foundation for equity-centered digital transformation governance, but its empirical operationalization in policy evaluation contexts remains underdeveloped. Future research should develop indicators and methodologies for assessing the equity impacts of government digital transformation interventions, and should examine how policy design choices—including the targeting of support programs, the sequencing of regulatory reforms, and the allocation of public investment—affect the distributional outcomes of digital transitions.

These gaps collectively suggest a practical implication for policymakers: the design of effective government interventions in digital transformation requires not only ambition and resources but also investment in the evidence infrastructure—data collection systems, evaluation capacities, and inter-agency coordination mechanisms—that makes learning and adaptation possible over time.

6. Conclusion

This literature review has examined the role of government interventions in digital transformation, synthesizing evidence from seventeen peer-reviewed publications spanning industrial policy, crisis response, fintech regulation, AI governance, and sustainability transitions. The review has identified three principal thematic clusters that together constitute the current state of knowledge in this domain: national digital industrial strategies and their governance, crisis-driven policy accelerators, and the regulatory governance of digitally enabled services and technologies.

Across these themes, the reviewed literature consistently indicates that government interventions matter significantly for the pace, direction, and equity of digital transformation. National strategies such as Industrie 4.0 can establish globally influential paradigms for digital industrial change; crisis recovery packages represent critical junctures at which path dependencies can be disrupted or reinforced; and regulatory frameworks for fintech and AI determine whether the distributional benefits of digital innovation are broadly shared or captured by already-privileged actors. At the same time, the literature makes clear that effective government intervention is neither straightforward nor guaranteed: the heterogeneity of firm capabilities, the complexity of multi-level governance, and the speed of technological change all pose substantial challenges to policy design and implementation.

This review is subject to limitations that should be acknowledged. The selection of seventeen sources, while carefully chosen for relevance and quality, does not constitute an exhaustive survey of the literature on this topic. The review prioritizes synthesis and theoretical coherence over systematic coverage of all available evidence. Furthermore, the thematic clusters identified here inevitably involve some simplification of a complex and multidimensional field of inquiry. Despite these limitations, the review makes a meaningful contribution by articulating a coherent research agenda at the intersection of digital transformation and public policy—an agenda that holds significant promise for both academic advancement and practical governance improvement.

Future scholarship that addresses the identified gaps—particularly regarding rigorous policy evaluation, multi-level governance coordination, and equity impact assessment—will be essential for equipping governments with the knowledge they need to navigate one of the most consequential transitions of the twenty-first century.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

AlNuaimi, B. K., Singh, S. K., Ren, S., Budhwar, P., & Vorobyev, D. (2022). Mastering digital transformation: The nexus between leadership, agility, and digital strategy. Journal of Business Research, 145, 636–648. https://doi.org/10.1016/j.jbusres.2022.03.038

Almeida, F., Santos, J. D., & Monteiro, J. (2020). The challenges and opportunities in the digitalization of companies in a post-COVID-19 world. IEEE Engineering Management Review, 48(3), 97–103. https://doi.org/10.1109/emr.2020.3013206

Büchi, G., Cugno, M., & Castagnoli, R. (2019). Smart factory performance and Industry 4.0. Technological Forecasting and Social Change, 150, 119790. https://doi.org/10.1016/j.techfore.2019.119790

Compagnucci, L., & Spigarelli, F. (2020). The Third Mission of the university: A systematic literature review on potentials and constraints. Technological Forecasting and Social Change, 161, 120284. https://doi.org/10.1016/j.techfore.2020.120284

Demir, A., Pesqué-Cela, V., Altunbaş, Y., & Murinde, V. (2022). Fintech, financial inclusion and income inequality: A quantile regression approach. European Journal of Finance, 28(1), 86–107. https://doi.org/10.1080/1351847x.2020.1772335

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., & Coombs, C. (2019). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Haendel, M., Chute, C. G., Bennett, T. D., Eichmann, D., & Saltz, J. (2020). The National COVID Cohort Collaborative (N3C): Rationale, design, infrastructure, and deployment. Journal of the American Medical Informatics Association, 28(3), 427–443. https://doi.org/10.1093/jamia/ocaa196

Hariram, N. P., Mekha, K. B., Suganthan, V., & Sudhakar, K. (2023). Sustainalism: An integrated socio-economic-environmental model to address sustainable development and sustainability. Sustainability, 15(13), 10682. https://doi.org/10.3390/su151310682

Hepburn, C., O'Callaghan, B., Stern, N., Stiglitz, J. E., & Zenghelis, D. (2020). Will COVID-19 fiscal recovery packages accelerate or retard progress on climate change? Oxford Review of Economic Policy, 36(Supplement 1), S359–S381. https://doi.org/10.1093/oxrep/graa015

Juergensen, J., Guimón, J., & Narula, R. (2020). European SMEs amidst the COVID-19 crisis: Assessing impact and policy responses. Journal of Industrial and Business Economics, 47(3), 499–510. https://doi.org/10.1007/s40812-020-00169-4

Kagermann, H., & Wahlster, W. (2022). Ten years of Industrie 4.0. Sci, 4(3), 26. https://doi.org/10.3390/sci4030026

Kraus, S., Jones, P., Kailer, N., Weinmann, A., & Chaparro-Banegas, N. (2021). Digital transformation: An overview of the current state of the art of research. SAGE Open, 11(3). https://doi.org/10.1177/21582440211047576

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. https://doi.org/10.3390/app10124182

Machado, C. G., Winroth, M., & Ribeiro da Silva, E. (2019). Sustainable manufacturing in Industry 4.0: An emerging research agenda. International Journal of Production Research, 58(5), 1462–1484. https://doi.org/10.1080/00207543.2019.1652777

Mhlanga, D. (2020). Industry 4.0 in finance: The impact of Artificial Intelligence (AI) on digital financial inclusion. International Journal of Financial Studies, 8(3), 45. https://doi.org/10.3390/ijfs8030045

Suchek, N., Fernandes, C., Kraus, S., Filser, M., & Sjögrén, H. (2021). Innovation and the circular economy: A systematic literature review. Business Strategy and the Environment, 30(8), 3686–3702. https://doi.org/10.1002/bse.2834

The EU AI Act and Its Regulatory Implications for Businesses: A Systematic Literature Review

Grisáková Nora[1] and Hojdik Vladimír[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: nora.grisakova@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: vladimir.hojdik@euba.sk

Abstract

The rapid proliferation of artificial intelligence across industries has prompted the European Union to establish the AI Act, a landmark regulatory framework designed to govern the development, deployment, and use of AI systems. This literature review examines the regulatory implications of the EU AI Act for businesses, drawing upon a corpus of fifteen peer-reviewed publications to synthesize current knowledge and identify critical research gaps. The review addresses three principal thematic clusters: the conceptual foundations of AI trustworthiness and risk governance, sector-specific compliance challenges, and the broader socioeconomic consequences of AI regulation for business innovation. Key findings reveal that the Act's risk-based classification system imposes significant compliance burdens on high-risk AI operators, particularly in sectors such as healthcare and digital finance, while its conceptualization of "trustworthiness" has been critiqued as theoretically underdeveloped. Furthermore, the regulation's territorial scope raises questions about global competitiveness and the innovation capacity of small and medium-sized enterprises. The review concludes by identifying underexplored areas—including regulatory impacts on generative AI business models and SME compliance capacity—and proposes specific directions for future empirical research.

Keywords: EU AI Act, AI regulation, business compliance, legal framework, responsible AI

JEL Classification: K23, O38, M15

1. Introduction

Artificial intelligence has emerged as one of the most transformative forces in contemporary economic and social life, reshaping industries from healthcare and finance to agriculture and public administration. As AI systems grow in both complexity and deployment scale, concerns surrounding their safety, fairness, and accountability have intensified, prompting regulatory authorities worldwide to develop governance frameworks commensurate with these challenges. In this context, the European Union's Artificial Intelligence Act (EU AI Act), formally adopted in 2024, represents a watershed moment in AI governance—establishing what is widely considered the world's most comprehensive binding legal framework for AI (Laux et al., 2023; Aboy et al., 2024).

The EU AI Act introduces a tiered, risk-based classification system that categorizes AI applications according to their potential for harm, imposing obligations ranging from minimal transparency requirements to outright prohibitions. For businesses operating within or interacting with the EU single market, the Act carries profound implications: it mandates

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

conformity assessments, documentation obligations, human oversight mechanisms, and ongoing post-market surveillance for systems classified as high-risk. These requirements introduce substantial compliance costs and operational uncertainties, particularly for organizations that have built their value propositions around AI-driven automation or algorithmic decision-making (Laux et al., 2023; Aboy et al., 2024).

Despite the growing policy significance of the EU AI Act, academic scholarship has not yet produced a consolidated synthesis of its implications for business practice. The existing literature addresses AI governance in fragmented domains—medical AI ethics (Zhang & Zhang, 2023), generative AI risks (Wach et al., 2023), cybersecurity implications (Gupta et al., 2023), and labor productivity effects (Damioli et al., 2021)—without sufficiently examining how the EU's regulatory architecture interacts with business strategy and innovation management across sectors. This fragmentation constitutes a meaningful research gap that the present review seeks to address.

The aim of this literature review is to synthesize current knowledge on how the EU AI Act reshapes the regulatory environment in which businesses develop, deploy, and govern AI systems. The paper focuses especially on the conceptual foundations of the Act, the practical compliance implications emerging across business domains, and the research gaps that remain insufficiently addressed in the intersection of AI regulation and business strategy.

The paper contains five main chapters. The second chapter establishes key constructs and reviews foundational frameworks relevant to AI regulation. The third chapter presents the thematic review organized around key analytical clusters. The fourth chapter identifies gaps in the current literature and proposes future research directions. The final chapter makes conclusion of key contributions and an acknowledgment of the review's limitations.

2. Regulatory and Conceptual Foundations of the EU AI Act

2.1 Defining AI Governance and the Risk-Based Regulatory Paradigm

AI governance refers to the set of norms, institutions, and mechanisms through which societies manage the development and deployment of AI technologies to align these with societal values and legal standards. Within academic discourse, AI governance frameworks are generally distinguished by their philosophical orientation—whether they prioritize precautionary prohibition, ex ante certification, or ex post liability—and by the organizational level at which they operate, from individual firm policies to supranational legislation (Yiğitcanlar et al., 2021). The EU AI Act adopts a predominantly ex ante, risk-based approach, in which the regulatory burden on a given AI system is proportionate to the severity and probability of harm it may cause (Laux et al., 2023).

The risk-based paradigm as an organizational concept draws from established traditions in product safety law and financial regulation, yet its application to AI presents novel challenges. Unlike physical products, AI systems are dynamic, context-sensitive, and capable of producing emergent behaviors that may not be foreseeable at the time of deployment. This introduces what Laux et al. (2023) describe as a fundamental conceptual tension: the Act attempts to operationalize "trustworthiness" through the acceptability of measurable risks, yet trust in institutional contexts is a far more complex socio-relational construct that cannot be reduced to probabilistic harm assessment alone. The conflation of these two concepts, they argue, results in a regulatory framework that may be technically compliant yet substantively insufficient to build genuine public confidence in AI systems.

2.2 Trustworthiness, Accountability, and Human-Centered AI

A foundational construct in the EU AI Act—and in the broader literature on AI ethics—is trustworthiness. Zhang and Zhang (2023) propose that trustworthy medical AI requires a multi-layered framework encompassing ethical value alignment, legal accountability structures, and technical mechanisms for algorithmic transparency and bias mitigation. Their framework situates human health promotion as the overarching normative objective and emphasizes that regulatory accountability must remain with human actors rather than AI systems, given the current absence of legal personhood for machines.

This position resonates with the human-centered AI framework articulated by Garibay et al. (2023), who identify six grand challenges for AI development: ensuring equitable access, supporting human autonomy, promoting algorithmic fairness, enabling explainability, fostering social well-being, and guaranteeing safety. These challenges map closely onto the EU AI Act's requirements for transparency, human oversight, and non-discrimination, suggesting that the Act is at least partially responsive to scholarly calls for more human-centered governance frameworks.

The behavioral and systemic dimensions of AI governance are further illuminated by Chater and Loewenstein (2022), whose analysis of i-frame versus s-frame policy interventions provides a valuable theoretical lens. They argue that policy frameworks focusing exclusively on individual-level behavioral change—what they term the "i-frame"—systematically underperform relative to system-level ("s-frame") interventions that restructure the institutional environment itself. Applied to AI regulation, this distinction suggests that compliance-focused, firm-level requirements (the dominant mode of the AI Act) may prove insufficient unless accompanied by structural interventions in AI market design, data infrastructure, and computational resource distribution.

2.3 The Business Context: Innovation, Productivity, and Regulatory Compliance

Understanding the regulatory implications of the EU AI Act for businesses requires a grounding in the economic literature on AI and firm-level performance. Damioli et al. (2021) provide empirical evidence that AI patent activity correlates with enhanced labor productivity, particularly among small and medium-sized enterprises (SMEs) and service-sector firms. This finding underscores the economic stakes of AI regulation: insofar as the Act's compliance requirements disproportionately burden smaller firms, it risks attenuating the productivity dividends that AI adoption can deliver across the business landscape.

In the domain of innovation management, Mariani and Dwivedi (2024) identify generative AI as a transformative force reshaping product development, knowledge creation, and competitive strategy. They argue that generative AI introduces novel forms of co-innovation between humans and machines, fundamentally altering the innovation value chain. The EU AI Act's requirements for transparency and documentation in high-risk AI systems may impose constraints on this co-innovation dynamic, particularly when proprietary model architectures must be disclosed to regulatory authorities.

The sociotechnical perspective advanced by Sartori and Theodorou (2022) further enriches the theoretical background by emphasizing that AI systems are not merely technical artifacts but socially embedded entities whose design choices reflect and reproduce existing power structures, inequalities, and narratives. They argue that effective AI governance must therefore engage with the sociological dimensions of AI deployment, including the distributional consequences of algorithmic decision-making for marginalized communities. This perspective

aligns with Capraro et al. (2024), who demonstrate that generative AI has the potential both to exacerbate and to ameliorate socioeconomic inequalities, depending critically on the governance frameworks within which it operates.

3. Business Implications Across Compliance Domains

3.1 Risk Classification, Compliance Architecture, and Legal Interpretation

A foundational theme in the literature concerns the architecture of the EU AI Act's risk-based classification system and the interpretive challenges it generates for businesses seeking compliance. Laux et al. (2023) offer the most direct critical engagement with the Act's conceptual design, arguing that its operationalization of trustworthiness through risk acceptability is both theoretically impoverished and practically problematic. Their systematic review of institutional trust research demonstrates that trust is constituted through a complex interplay of competence perception, value alignment, and procedural legitimacy—dimensions that a risk calculus based primarily on harm probability and severity cannot fully capture. For businesses, this critique carries an important implication: meeting the letter of the Act's requirements may not be sufficient to generate the public and regulatory trust necessary for sustained market acceptance of their AI products.

Aboy et al. (2024) extend this analysis to the specific case of regulated digital medical products, providing the most detailed available examination of how the AI Act interacts with existing product safety legislation. Their analysis reveals significant complexity at the intersection of the AI Act and the EU Medical Device Regulation (MDR) and In Vitro Diagnostic Regulation (IVDR). AI-enabled medical devices are simultaneously subject to two overlapping regulatory regimes, each imposing distinct conformity assessment pathways, documentation requirements, and notified body oversight obligations. They find that the AI Act's definition of a "provider"—the entity responsible for placing an AI system on the market—does not map cleanly onto the MDR's concept of a medical device "manufacturer," creating compliance ambiguities that firms will need to resolve through regulatory guidance that has not yet materialized. This dual-regulation challenge exemplifies a broader pattern in which the Act's horizontal design, intended to apply uniformly across sectors, may in practice generate regulatory friction when superimposed upon existing vertical regulatory frameworks.

The question of territorial scope is equally significant for international businesses. The AI Act applies not only to AI systems deployed within the EU but also to systems developed outside the EU whose outputs are used within the Union, making it a de facto global standard for any company seeking access to the European market (Aboy et al., 2024). This extraterritorial reach echoes the "Brussels Effect" observed in relation to EU data protection law, whereby EU standards effectively become global standards due to the economic importance of market access. For multinational enterprises, this dynamic may reduce the compliance cost of global divergence but simultaneously raises concerns about regulatory imperialism and the marginalization of non-European normative perspectives on AI governance (Laux et al., 2023).

3.2 Sector-Specific Implications—Healthcare, Cybersecurity, and Beyond

The EU AI Act's implications are particularly acute in sectors where AI systems interact directly with human welfare, safety, or fundamental rights. Healthcare is perhaps the most extensively analyzed sector in this regard. Zhang and Zhang (2023) argue that trustworthy medical AI requires not only technical robustness but also clear ethical governance structures, transparent algorithmic processes, and mechanisms for multi-stakeholder accountability. They

emphasize that data quality management is a prerequisite for algorithmic fairness in medical contexts, a concern directly addressed by the AI Act's requirements for high-quality training data in high-risk systems. However, they also note that regulatory frameworks must be adaptive to the rapid pace of medical AI innovation, cautioning against overly rigid ex ante requirements that may inadvertently stifle beneficial applications.

The intersection of AI governance and cybersecurity is addressed by Gupta et al. (2023), whose comprehensive analysis of generative AI's cybersecurity implications highlights a domain of particular regulatory complexity. They document how generative AI models can be weaponized for sophisticated cyberattacks, including the creation of highly convincing phishing content, automated malware development, and the generation of disinformation at scale. For businesses deploying generative AI, these findings underscore the dual-use character of the technology: the same capabilities that drive innovation can be exploited for malicious purposes. The EU AI Act's treatment of general-purpose AI models, including large language models, has been one of the most contested aspects of the legislative process, reflecting precisely this dual-use tension. Gupta et al.'s (2023) analysis suggests that cybersecurity risks from AI may be systematically underweighted in the Act's current risk classification, as the primary harm scenarios used to justify high-risk designation tend to focus on direct discrimination or safety violations rather than indirect harms mediated through malicious third-party use.

Wach et al. (2023) contribute a broader critical analysis of the dark side of generative AI in business contexts, identifying dimensions of risk that span ethical, economic, and societal domains. Their conceptual framework encompasses risks related to misinformation, intellectual property violations, job displacement, and the reinforcement of cognitive biases among users who over-rely on AI-generated content. These risks are relevant to the EU AI Act insofar as they inform the design of transparency and accountability obligations for general-purpose AI systems deployed in business contexts. Wach et al. (2023) argue that existing governance frameworks—including the AI Act—tend to focus on technical risks while inadequately addressing the organizational and behavioral risks that emerge when AI systems are embedded in business processes and decision-making cultures.

The agricultural sector, though less frequently discussed in the AI regulation literature, offers an instructive case of a domain undergoing rapid AI-driven transformation with distinct governance needs. Finger (2023) demonstrates that digital innovations in agriculture can yield substantial sustainability and productivity benefits, including precision resource management and enhanced supply chain resilience. Yet he also identifies significant risks related to digital divides, data ownership concentration, and the displacement of small-scale farmers excluded from digital infrastructure. These challenges are broadly relevant to the EU AI Act's regulatory agenda, as they illustrate how sector-specific contextual factors can mediate the impacts of general-purpose AI regulation on different business populations.

3.3 Socioeconomic Consequences, Inequality, and Innovation Dynamics

The third thematic cluster addresses the broader socioeconomic implications of AI regulation, including its effects on inequality, labor markets, and the trajectory of business innovation. Capraro et al. (2024) provide a comprehensive interdisciplinary overview of how generative AI may reshape socioeconomic inequalities across domains including information access, work, education, and healthcare. They argue that AI's distributional consequences are not technologically determined but are instead shaped by the policy frameworks within which AI is developed and deployed. Regulatory choices, including the EU AI Act's obligations,

therefore directly influence whether AI's productivity benefits are broadly shared or concentrated among already-advantaged actors.

The labor market implications identified by Damioli et al. (2021) are directly relevant in this context. Their analysis of AI patenting activity demonstrates that AI investment generates labor productivity gains, but these gains are concentrated in specific firm sizes—particularly SMEs—and in services industries rather than manufacturing. This finding suggests that the EU AI Act's compliance burden, if not carefully calibrated to firm size and sector, could paradoxically disadvantage the very business populations that stand to gain most from AI adoption. Small firms may lack the legal and technical resources to navigate complex conformity assessment procedures, creating a regulatory barrier that reinforces market concentration in favor of large technology incumbents.

Mariani and Dwivedi (2024) examine the implications of generative AI for innovation management, identifying ten major research themes at the intersection of generative AI and innovation. They argue that generative AI is fundamentally altering the economics of knowledge creation, enabling non-expert actors to participate in creative and analytical processes previously restricted to specialists. From a regulatory perspective, this democratizing potential must be weighed against the accountability challenges it creates: when AI systems generate outputs that influence business decisions, the question of who bears responsibility for errors or harms becomes significantly more complex. The AI Act's provider-deployer distinction attempts to address this accountability gap, but Mariani and Dwivedi's (2024) analysis suggests that the organizational realities of AI-enabled innovation management may not map cleanly onto these binary categories.

The sociotechnical perspective articulated by Sartori and Theodorou (2022) provides a critical counterpoint to economistic analyses of AI's impact. They argue that discussions of AI's transformative potential frequently reproduce techno-solutionist narratives that obscure the political choices embedded in AI design and governance. In particular, they emphasize that algorithmic biases are not random technical errors but structural reflections of the social inequalities present in training data and design decisions. This argument reinforces Garibay et al.'s (2023) call for human-centered AI governance that actively addresses fairness and inclusion, rather than treating these as secondary concerns relative to safety and transparency. The EU AI Act's prohibitions on discriminatory AI uses and its requirements for bias testing in high-risk systems represent partial responses to these concerns, but critics note that the enforcement mechanisms remain underdeveloped relative to the ambition of the norms themselves.

The privacy dimensions of AI regulation constitute another important thread in the literature. Meden et al. (2021) examine the specific case of face recognition technology, documenting how biometric AI systems can expose sensitive personal attributes including age, gender, race, and health status without users' knowledge or consent. Their comprehensive survey of privacy-enhancing approaches to biometric AI demonstrates that technical solutions exist to mitigate these risks, but that their adoption requires both regulatory mandates and market incentives. The EU AI Act's prohibition of real-time remote biometric identification in public spaces for law enforcement purposes—one of its most politically contested provisions—directly engages with these concerns, and its implications for businesses in the security, retail, and access management sectors are substantial. Chater and Loewenstein's (2022) framework further suggests that privacy-protective behaviors are unlikely to emerge spontaneously from individual business decisions and require systemic regulatory intervention to become the default.

4. Existing Gaps for Future Business-Related AI Regulation Research

Despite the growing body of scholarship on AI governance and its business implications, the existing literature exhibits several significant gaps that warrant dedicated future research attention.

The most conspicuous gap concerns the empirical measurement of compliance costs and their distributional effects across firm types. While Damioli et al. (2021) establish that AI adoption generates productivity gains particularly for SMEs, and while Aboy et al. (2024) document the complexity of compliance requirements in the medical device sector, no study to date has provided a systematic empirical assessment of the compliance burden imposed by the EU AI Act across firm sizes, sectors, and national contexts. Future research should employ survey-based and case study methodologies to quantify the financial and organizational costs of conformity assessment, technical documentation, and ongoing monitoring obligations, with particular attention to how these costs vary across the business population. Such evidence would be directly valuable to policymakers considering graduated compliance regimes or targeted support mechanisms for SMEs.

Next significant gap relates to the regulation of generative AI and its implications for business models built on large language models and other foundation model architectures. Gupta et al. (2023), Wach et al. (2023), and Mariani and Dwivedi (2024) collectively document the transformative and disruptive potential of generative AI across numerous business domains, yet none of these studies engages directly with the EU AI Act's provisions for general-purpose AI models introduced in the final legislative text. Future research should examine how businesses developing or deploying foundation models are responding to the Act's novel transparency and systemic risk assessment requirements, and whether these requirements are calibrated appropriately to the unique risk profile of models that can be applied across an essentially unlimited range of contexts and use cases.

Another gap concerns the interaction between the EU AI Act and existing sector-specific regulations, beyond the medical device context analyzed by Aboy et al. (2024). Sectors including financial services, critical infrastructure, autonomous transport, and employment management are all subject to both the AI Act's high-risk provisions and their own regulatory frameworks. The potential for regulatory redundancy, contradiction, or gap between these overlapping regimes has not been systematically mapped. Future research adopting a comparative regulatory analysis methodology could provide significant value to both policymakers engaged in regulatory coordination and practitioners seeking compliance clarity.

Finally, the literature is notably silent on the governance implications of the AI Act for non-European businesses operating in global markets, as well as on the potential for regulatory arbitrage whereby firms relocate AI development activities to jurisdictions with less stringent oversight. Capraro et al.'s (2024) analysis of how AI governance shapes socioeconomic outcomes globally suggests that the extraterritorial dimension of the AI Act deserves more sustained scholarly attention, particularly in relation to its implications for the global distribution of AI development capacity and the risk of a fragmented international regulatory landscape.

5. Conclusion

Viewed from the perspective of business strategy, the EU AI Act does not merely regulate AI systems; it redistributes the costs, risks, and capabilities required to compete with them. The reviewed literature points to three closely connected domains of impact: the conceptual and

risk-based architecture of the regulation, its differentiated compliance implications across sectors, and its broader consequences for innovation, competitiveness, and inequality among firms. Taken together, these findings suggest that the Act should be understood not only as a legal framework, but also as a structural force shaping the future conditions of AI-enabled business activity in Europe.

All in all, the EU AI Act's conceptual framework, while broadly aligned with scholarly calls for human-centered and trustworthy AI, exhibits theoretical limitations in its conflation of trustworthiness with risk acceptability, as documented by Laux et al. (2023). Moreover, the Act's compliance architecture imposes differentiated burdens across sectors, with healthcare and digital medical products facing particularly complex dual-regulation challenges (Aboy et al., 2024; Zhang & Zhang, 2023). Then, the Act's distributional implications for business innovation are significant and potentially regressive unless accompanied by targeted support mechanisms for smaller firms (Damioli et al., 2021; Capraro et al., 2024). In summary, the regulation of generative AI and biometric systems remains an area of active scholarly debate, with cybersecurity risks and privacy implications potentially underweighted in the current regulatory design (Gupta et al., 2023; Meden et al., 2021).

This review carries several limitations that should be acknowledged. The corpus of fifteen sources, while carefully selected, does not exhaust the relevant literature, and the absence of sources focusing explicitly on legal doctrine limits the depth of the regulatory analysis. Additionally, as the EU AI Act entered into force in 2024, empirical evidence of its practical business impacts remains nascent, constraining the conclusions that can be drawn about real-world compliance outcomes. Future reviews should incorporate a broader range of legal scholarship, empirical compliance studies, and comparative international perspectives to build upon the thematic foundations established here.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Aboy, M., Minssen, T., & Vayena, E. (2024). Navigating the EU AI Act: Implications for regulated digital medical products. npj Digital Medicine, 7(1). https://doi.org/10.1038/s41746-024-01232-3

Capraro, V., Lentsch, A., Acemoğlu, D., Akgün, S., & Akhmedova, A. (2024). The impact of generative artificial intelligence on socioeconomic inequalities and policy making. PNAS Nexus, 3(6). https://doi.org/10.1093/pnasnexus/pgae191

Chater, N., & Loewenstein, G. (2022). The i-frame and the s-frame: How focusing on individual-level solutions has led behavioral public policy astray. Behavioral and Brain Sciences, 46, e30. https://doi.org/10.1017/s0140525x22002023

Damioli, G., Van Roy, V., & Vértesy, D. (2021). The impact of artificial intelligence on labor productivity. Eurasian Economic Review, 11(1), 1–25. https://doi.org/10.1007/s40821-020-00172-8

Fedyk, A., Hodson, J., Khimich, N. V., & Fedyk, T. (2022). Is artificial intelligence improving the audit process? Review of Accounting Studies, 27(3), 938–985. https://doi.org/10.1007/s11142-022-09697-x

Finger, R. (2023). Digital innovations for sustainable and resilient agricultural systems. European Review of Agricultural Economics, 50(4), 1277–1309. https://doi.org/10.1093/erae/jbad021

Garibay, Ö. Ö., Winslow, B., Andolina, S., Antona, M., & et al. (2023). Six human-centered artificial intelligence grand challenges. International Journal of Human-Computer Interaction, 39(3), 391–437. https://doi.org/10.1080/10447318.2022.2153320

Gupta, M., Akiri, C., Aryal, K., Parker, E., & Praharaj, L. (2023). From ChatGPT to ThreatGPT: Impact of generative AI in cybersecurity and privacy. IEEE Access, 11, 80218–80245. https://doi.org/10.1109/access.2023.3300381

Laux, J., Wachter, S., & Mittelstadt, B. (2023). Trustworthy artificial intelligence and the European Union AI Act: On the conflation of trustworthiness and acceptability of risk. Regulation & Governance, 18(3), 771–787. https://doi.org/10.1111/rego.12512

Mariani, M. M., & Dwivedi, Y. K. (2024). Generative artificial intelligence in innovation management: A preview of future research developments. Journal of Business Research, 175, 114542. https://doi.org/10.1016/j.jbusres.2024.114542

Meden, B., Rot, P., Terhörst, P., Damer, N., & Kuijper, A. (2021). Privacy-enhancing face biometrics: A comprehensive survey. IEEE Transactions on Information Forensics and Security, 16, 4147–4183. https://doi.org/10.1109/tifs.2021.3096024

Sartori, L., & Theodorou, A. (2022). A sociotechnical perspective for the future of AI: Narratives, inequalities, and human control. Ethics and Information Technology, 24(1), 4. https://doi.org/10.1007/s10676-022-09624-3

Wach, K., Duong, C. D., Ejdys, J., Kazlauskaitė, R., & Korzyński, P. (2023). The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT. Entrepreneurial Business and Economics Review, 11(2), 7–32. https://doi.org/10.15678/eber.2023.110201

Yiğitcanlar, T., Corchado, J. M., Mehmood, R., Li, R. Y. M., & Mossberger, 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, 7(1), 71. https://doi.org/10.3390/joitmc7010071

Zhang, J., & Zhang, Z. (2023). Ethics and governance of trustworthy medical artificial intelligence. BMC Medical Informatics and Decision Making, 23(1), 7. https://doi.org/10.1186/s12911-023-02103-9

Cross-Country Comparison of Digital Economy Indicators: A Systematic Literature Review

Zuzana Hajduová[1] and Štetka Peter[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: zuzana.hajduova@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: peter.stetka@euba.sk

Abstract

The digital economy has emerged as a transformative force reshaping production, governance, and social interaction across national boundaries. Despite growing scholarly interest, the comparative assessment of digital economy indicators across countries remains fragmented, with inconsistent methodologies and contested frameworks impeding rigorous cross-country analysis. This review synthesizes the existing literature on digital economy measurement and cross-country comparison, drawing on empirical studies, conceptual frameworks, and policy-oriented research. The central objective is to identify dominant thematic clusters, methodological approaches, and knowledge gaps in how digital economy indicators are conceptualized and compared internationally. The review finds that the literature converges around three principal themes: the measurement and classification of digital development parameters, the relationship between information and communication technology (ICT) diffusion and economic growth, and the institutional and organizational dimensions of digital transformation. Significant disagreements persist regarding the selection and weighting of composite indicators, the role of governance structures in enabling digitalization, and the extent to which findings derived from advanced economies can be generalized to developing regions. The review concludes by identifying critical research gaps and proposing future directions for comparative digital economy research with direct implications for policymakers and international development institutions.

Keywords: digital economy indicators, cross-country comparison, ICT measurement, digital competitiveness, international benchmarking

JEL Classification: O33, O57, F69

1. Introduction

The rapid proliferation of digital technologies over the past two decades has fundamentally altered the architecture of national economies, prompting scholars, policymakers, and international organizations to develop frameworks capable of capturing the scope and depth of this transformation. The concept of the digital economy-broadly understood as economic activity that is facilitated by digital technologies-has become a central concern for comparative economic research, particularly as disparities between highly digitalized and less digitalized nations continue to widen. Understanding how countries differ in their digital economy performance requires robust, internationally comparable indicators that can illuminate both current states and trajectories of development.

Despite substantial progress in constructing composite indices such as the Digital Economy and Society Index (DESI) and the ICT Development Index, the literature remains fragmented. Studies differ markedly in their selection of indicators, their methods for aggregating multidimensional data, and their underlying theoretical assumptions about what the digital economy encompasses. This heterogeneity complicates cross-country comparison and limits the policy relevance of individual studies. Furthermore, the majority of rigorous empirical work has concentrated on European Union member states or high-income economies, leaving significant gaps in the comparative literature concerning developing regions and transition economies.

A further complication arises from the dynamic nature of digital transformation itself. As technologies evolve and new dimensions of digitalization - such as artificial intelligence, platform economies, and digital public services - gain prominence, indicator frameworks developed even a decade ago risk becoming obsolete. Researchers and policymakers alike must therefore address the dual challenge of methodological rigor and conceptual currency.

This literature review addresses these challenges by systematically examining and synthesizing the scholarly literature on cross-country digital economy indicators. The specific objectives are threefold: first, to map the theoretical and conceptual foundations underpinning comparative digital economy research; second, to identify and analyze the principal thematic clusters in the empirical literature; and third, to delineate research gaps and propose directions for future inquiry. By doing so, this review aims to provide a consolidated reference point for researchers and practitioners engaged in international comparisons of digital development.

The remainder of the article is structured as follows. Section 2 establishes the theoretical background, defining key constructs and presenting foundational frameworks relevant to digital economy measurement. Section 3 presents a thematic synthesis of the literature, organized around three distinct clusters. Section 4 identifies critical research gaps and future research directions. Section 5 concludes the review with a summary of key contributions and an acknowledgment of limitations.

2. Theoretical Background

This study adopts a structured literature review approach to systematically identify and synthesize research on cross-country digital economy indicators. Relevant publications were collected from major academic databases, including Web of Science, Scopus, and Google Scholar, using keyword combinations such as “digital economy indicators,” “ICT measurement,” and “cross-country comparison.” The selection process focused on peer-reviewed studies with international or comparative relevance, while excluding purely technical or single-country analyses without broader implications. The identified literature was analyzed using thematic synthesis, involving coding of key concepts, grouping studies into thematic clusters, and interpreting relationships and research gaps. Although the review does not strictly follow a formal PRISMA protocol and is limited to English-language sources, it ensures a transparent and analytically consistent synthesis of current knowledge in the field.

2.1 Defining the Digital Economy

The digital economy is a multifaceted construct that has resisted singular, universally accepted definition. At its broadest, it encompasses the totality of economic transactions and productive activities conducted through or enabled by digital networks, platforms, and technologies. More narrowly, some scholars focus exclusively on information and communication technologies (ICT) as the infrastructural foundation of digital economic activity, while others incorporate the behavioral and organizational transformations that digital adoption precipitates. Bilozubenko et al. (2020) define the digital economy in terms of the widespread use of ICT and the subsequent transformations in production, distribution, and governance structures, noting that the European Union has positioned itself as a supranational subject actively engaged in building such an economy among its member states.

This definitional complexity has direct implications for indicator selection and cross-country comparison. If the digital economy is understood primarily as an infrastructure phenomenon, then indicators related to broadband penetration, internet access, and ICT investment dominate comparative frameworks. If, however, the digital economy is conceived more broadly to include digital public services, e-commerce, and digitalized marketing practices, the indicator set necessarily expands. Špak et al. (2020) illustrate this broader conception in their analysis of enterprise-level digitalization, demonstrating that the digital economy manifests through online sales channels, social media platforms, and mobile applications in ways that aggregate national statistics may fail to capture adequately.

2.2 Theoretical Frameworks for Digital Development

Several theoretical frameworks inform the comparative study of digital economies. Modernization theory, adapted to the digital context, posits that countries converge toward higher levels of digital development as they accumulate capital, human resources, and institutional capacity. This perspective underpins much of the development economics literature on ICT diffusion, including the work of Saba et al. (2023), who analyze the ICT diffusion–economic growth nexus across 73 countries in sub-Saharan Africa, the Middle East and North Africa, and Latin America, finding heterogeneous effects that challenge simple convergence narratives.

An alternative framework draws on institutional economics and highlights the role of governance structures, regulatory environments, and public administration capacity in shaping digital outcomes. Dobrolyubova (2021) situates digital transformation within the broader discourse of public value creation, arguing that measuring outcomes of digital transformation in public administration requires indicators that go beyond infrastructure to encompass service quality, citizen satisfaction, and governance efficiency. This institutional perspective is particularly salient for cross-country comparisons that span different political systems and administrative traditions.

Capability maturity models (CMM), originally developed in the context of software engineering, have been adapted to assess organizational and sectoral readiness for digital transformation. Hortoványi et al. (2023) apply this framework to manufacturing firms in a B2B context, demonstrating that digital transformation capability develops incrementally and depends on a firm's existing resource base and dynamic capabilities. While primarily applied at the firm level, this framework offers insights relevant to national-level comparisons, particularly regarding the absorptive capacity that countries must develop to benefit from ICT diffusion.

2.3 Composite Indicators and Index Construction

The construction of composite indicators for digital economy benchmarking represents a significant methodological challenge. Such indices typically aggregate multiple dimensions-including infrastructure, human capital, business digitalization, and digital public services-into a single score or ranking. The selection of component indicators, their weighting, and the normalization procedures employed can substantially affect outcomes and country rankings, raising questions of validity and reliability that are central to the comparative literature. Bilozubenko et al. (2020) employ a feature vector of twenty indicators in their cluster analysis of EU member states, demonstrating how different indicator configurations produce different groupings and yield distinct policy implications regarding the digital divide.

The concept of the digital divide the gap between those with meaningful access to digital technologies and those without provides a normative as well as analytical framework for cross-country comparison. Originally conceptualized as a binary distinction between connected and unconnected populations, the digital divide has evolved into a multidimensional construct encompassing access, skills, and the quality and nature of use. This evolution necessitates more sophisticated composite indicators capable of capturing qualitative differences in digital engagement, not merely quantitative differences in infrastructure provision.

Behavioral dimensions of the digital economy add further complexity. Silva et al. (2019) demonstrate how digital behavioral data-specifically Google Trends search volumes - can serve as proxies for consumer activity and economic forecasting, suggesting that non-traditional data sources may complement official statistical indicators in cross-country comparisons. This methodological insight points toward a broader diversification of the digital economy indicator toolkit.

3. Literature Review

3.1 Thematic Cluster One: Measurement, Classification, and the Digital Divide

The measurement of digital economy parameters and the classification of countries according to their level of digital development constitute the most directly relevant strand of literature for cross-country comparison. Bilozubenko et al. (2020) provide a foundational contribution through their cluster analysis of EU member states, using a composite feature vector of twenty digital economy indicators drawn from Eurostat and the DESI framework. Their analysis identifies several distinct country clusters: a leading group comprising Nordic states and the Netherlands, a middle tier of Western European economies, and a lagging group predominantly constituted by Central and Eastern European member states. This clustering reveals systematic patterns of digital inequality within the EU that persist despite decades of cohesion policy, and the authors identify connectivity infrastructure and digital skills as the most discriminating indicators separating leading from lagging economies.

The identification of a persistent digital divide within Europe resonates with broader debates in the comparative digital economy literature. The widening gap between digital leaders and laggards has practical consequences for economic convergence, labor market outcomes, and the capacity of governments to deliver efficient public services. Dobrolyubova (2021) highlights that high-income, digitally advanced countries demonstrate strong correlations between governance indicators and digital transformation outcomes, while middle-income countries exhibit more ambiguous relationships, suggesting that the benefits of digital investment are mediated by institutional quality in ways that aggregate indicators often obscure.

Šhpak et al. (2020) extend the measurement discussion to the enterprise level, developing a comparative system for estimating the parameters of digitalization's influence on marketing activities. Although their case study focuses on Ukrainian service enterprises, their methodological contribution a two-part comparative system distinguishing structural and dynamic parameters of digital marketing-offers a template applicable to cross-country enterprise-level benchmarking. Their findings indicate that firms in less digitalized economies face structural barriers to digital marketing adoption that are qualitatively different from those encountered by firms in leading digital economies, a distinction that country-level aggregate indicators frequently fail to capture.

The use of non-traditional data sources in digital economy measurement represents an emerging and contested strand within this cluster. Silva et al. (2019) demonstrate the feasibility of using Google Trends data to forecast consumer behavior, validating singular spectrum analysis as a superior univariate forecasting technique compared with conventional parametric methods. While their application focuses on the fashion retail sector, the broader methodological implication for cross-country digital economy comparison is significant: Internet search data provide real-time, high-frequency proxies for economic activity that can supplement or, in data-scarce environments, substitute for official statistics. This is particularly relevant for developing economies where statistical infrastructure is limited.

3.2 Thematic Cluster Two: ICT Diffusion, Economic Growth, and Regional Heterogeneity

A substantial body of literature investigates the relationship between ICT diffusion and economic growth, typically employing panel data methods across multiple countries and regions. This literature broadly affirms a positive association between ICT adoption and economic development, yet significant debates persist regarding the magnitude, conditionality, and direction of this relationship. Saba et al. (2023) contribute a particularly comprehensive comparative analysis, examining the ICT-growth nexus across 73 countries in three developing regions over the period 2000-2018. Their findings reveal pronounced regional heterogeneity: ICT diffusion generates more consistent growth effects in Latin America than in sub-Saharan Africa or the Middle East and North Africa, a divergence the authors attribute to differences in human capital, institutional quality, and complementary infrastructure investment.

The implication of these findings for the construction of cross-country digital economy indicators is substantial. If ICT's growth effects are conditional on contextual factors, then composite indices that treat all countries within a single ranking framework may generate misleading comparisons. Regional stratification of indicator systems may therefore be both methodologically justified and practically necessary. Batrancea et al. (2023) provide supporting evidence from the European context, demonstrating that wellbeing-related infrastructure-including digital infrastructure-has heterogeneous effects on economic growth across 212 NUTS 2 regions in the EU, with Central and Eastern European regions exhibiting different response patterns than Western European regions. This regional differentiation reinforces the argument that cross-country comparisons must account for structural and institutional context rather than assuming uniform relationships between digital indicators and economic outcomes.

The COVID-19 pandemic introduced an exogenous shock that temporarily accelerated digital adoption across many economies while simultaneously exposing the vulnerabilities of digitally lagging nations. Buszko et al. (2021) analyze the pandemic's impact on stock market stability across sectors, finding that digitally-enabled service sectors demonstrated greater

resilience than traditional industries. While their focus is financial rather than explicitly digital economy comparative, their findings implicitly reinforce the economic value of digitalization and underscore the importance of tracking digital economy indicators across economic cycles and crisis periods.

At the subnational and organizational level, Hortományi et al. (2023) demonstrate that digital maturity in manufacturing firms is a function of both resource endowment and dynamic capability, with significant variation across countries in Central and Eastern Europe. This firm-level heterogeneity has national-level implications: countries with higher average firm-level digital maturity are likely to exhibit stronger aggregate digital economy performance, suggesting that enterprise-level indicators should be incorporated more systematically into national digital economy benchmarking frameworks. Their application of capability maturity modeling to cross-country B2B manufacturing contexts represents a methodological innovation with potential for broader adoption in comparative digital economy research.

3.3 Thematic Cluster Three: Institutional Dimensions, Digital Governance, and Strategic Autonomy

The third thematic cluster in the cross-country digital economy literature concerns the institutional, regulatory, and strategic dimensions of digital development. This cluster recognizes that digital economy performance is not solely a function of technology diffusion or investment levels but is shaped by governance structures, regulatory frameworks, and strategic policy choices at both national and supranational levels. Dobrolyubova (2021) argues that measuring the outcomes of digital transformation in public administration requires a fundamental expansion of the indicator framework to include public value creation, institutional capacity, and the quality of digital service delivery. Her review identifies significant limitations in current measurement approaches, particularly the predominance of input and output indicators—such as the number of available e-government services—relative to outcome indicators that capture whether digital public services actually improve citizens' lives and administrative efficiency.

The European Union's strategic response to global digital disruption is analyzed by Miró (2022), whose examination of the concept of "open strategic autonomy" provides a critical frame for understanding cross-country digital economy dynamics within the EU. Miró argues that the EU's pursuit of strategic autonomy—including in digital technologies, data governance, and critical digital infrastructure—represents a fundamental reassessment of its approach to globalization. This strategic orientation has direct implications for how digital economy indicators are constructed and interpreted: indicators that merely measure adoption of globally available technologies may understate the strategic dimension of digital development, particularly as geopolitical competition over digital standards and infrastructure intensifies.

The organizational and human dimensions of digital transformation provide further complexity for cross-country comparison. Ďurana et al. (2020) investigate the behavioral incentives of innovative business management, identifying managerial decision-making patterns and innovation-performance linkages across Central European firms. While not exclusively focused on digitalization, their findings highlight that behavioral factors-manager risk tolerance, organizational culture, and innovation orientation—constitute important determinants of digital adoption that composite national indicators typically do not capture. This behavioral dimension is particularly relevant for explaining why countries with similar

levels of digital infrastructure investment can exhibit markedly different outcomes in terms of enterprise digitalization and economic impact.

The construction industry provides a sector-specific lens through which institutional barriers to digitalization can be examined. Nagy et al. (2021) analyze the challenges of Construction 4.0 at the organizational level, finding that technological adoption is constrained by organizational fragmentation, resistance to change, and the absence of standardized digital workflows. Their cross-organizational analysis within the construction sector mirrors national-level patterns, where countries with strong innovation ecosystems and coordinated industrial policy demonstrate more advanced integration of digital technologies even in traditionally conservative sectors. The parallel between organizational and national barriers to digital transformation is suggestive, implying that institutional variables-rather than purely technological or financial ones-may be decisive in explaining cross-country differences in digital economy performance.

4. Research Gaps and Future Directions

Despite the substantial scholarly activity reviewed above, significant gaps remain in the literature on cross-country digital economy indicators. Three research directions are particularly pressing.

The first gap concerns the insufficient integration of enterprise-level and firm-level digital data into national comparative frameworks. The existing literature demonstrates convincingly that aggregate national indicators may obscure important structural heterogeneity at the firm, sector, and regional levels (Hortoványi et al., 2023; Špak et al., 2020). However, efforts to systematically link enterprise-level digital maturity assessments with national composite indices remain nascent. Future research should develop multi-level measurement frameworks that aggregate enterprise digital capability data in ways that are internationally comparable and statistically representative. Such frameworks would enable more nuanced diagnosis of the sources of cross-country digital performance differences, distinguishing between cases where national-level gaps reflect inadequate infrastructure and those where they reflect organizational and managerial barriers.

The second gap concerns the limited coverage of developing and transition economies in rigorous comparative digital economy research. While Saba et al. (2023) make a valuable contribution by extending the ICT–growth analysis to developing regions, the broader literature on digital economy indicator construction and cross-country classification is heavily dominated by European Union and high-income country contexts (Bilozubenko et al., 2020; Batrancea et al., 2023). This geographic skew limits the generalizability of existing frameworks and may embed assumptions-about infrastructure quality, statistical capacity, and institutional development—that are inappropriate for application in lower-income contexts. Future research should develop comparative digital economy indicator frameworks specifically calibrated to the institutional and developmental contexts of sub-Saharan Africa, South and Southeast Asia, and Latin America, potentially incorporating non-traditional data sources such as search engine data (Silva et al., 2019) to address statistical infrastructure limitations.

The third gap concerns the temporal and dynamic dimensions of digital economy comparison. Most existing comparative studies employ cross-sectional or relatively short panel data designs, which are poorly suited to capturing the long-run trajectories of digital transformation and the distributional dynamics of convergence and divergence. Dobrolyubova (2021) notes that the measurement of digital transformation outcomes requires longitudinal

frameworks that can track changes in public value creation over time, yet such frameworks are rare in the comparative literature. Future research should invest in developing harmonized longitudinal digital economy datasets that enable analysis of convergence and divergence dynamics, path dependencies, and the long-run effects of digital policy interventions. The incorporation of wellbeing-outcome indicators alongside purely economic performance measures-consistent with the broader framework proposed by Batrancea et al. (2023)-would further enrich such longitudinal frameworks. From a policy standpoint, closing these research gaps would provide digital economy planners, development banks, and international organizations with more reliable evidence for prioritizing interventions and allocating resources across countries at different stages of digital development.

5. Conclusion

This literature review has synthesized the scholarly literature on cross-country comparison of digital economy indicators, organizing the evidence around three principal thematic clusters: measurement, classification, and the digital divide; ICT diffusion, economic growth, and regional heterogeneity; and institutional dimensions, digital governance, and strategic autonomy. Across these clusters, the literature converges on several central findings.

First, the digital divide within and between regions is persistent and multidimensional, driven by differences in infrastructure, human capital, institutional quality, and organizational capacity that composite national indicators capture only partially (Bilozubenko et al., 2020; Saba et al., 2023). Second, the economic returns to ICT investment and digital development are heterogeneous across countries and regions, conditioned on institutional context, absorptive capacity, and complementary investments in skills and governance (Batrancea et al., 2023; Dobrolyubova, 2021). Third, the strategic and institutional dimensions of digital transformation-including regulatory frameworks, public administration quality, and geopolitical considerations-are underrepresented in existing indicator frameworks relative to their demonstrated importance for digital outcomes (Miró, 2022; Hortoványi et al., 2023).

This review is subject to several limitations. The sources available for synthesis were not exhaustively identified through a systematic database search, which may introduce selection bias in the coverage of the field. Moreover, the scope of the review necessarily omits important adjacent literatures in computer science, political science, and sociology that bear on digital economy measurement. Future reviews should employ fully systematic search protocols and broader disciplinary scope to complement the thematic synthesis presented here. Notwithstanding these limitations, this review provides a consolidated analytical foundation for researchers and policymakers engaged in the comparative study of digital economic development, identifying both the state of knowledge and the directions in which it must advance.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Batrancea, L. M., Nichita, A., Balcı, M. A., & Akgüller, Ö. (2023). Empirical investigation on how wellbeing-related infrastructure shapes economic growth: Evidence from the European Union regions. PLoS ONE, 18(3), e0283277. https://doi.org/10.1371/journal.pone.0283277

Bilozubenko, V., Yatchuk, O., Wolanin, E., Serediuk, T., & Korneyev, M. V. (2020). Comparison of the digital economy development parameters in the EU countries in the context of bridging the digital divide. Problems and Perspectives in Management, 18(2), 206–218. https://doi.org/10.21511/ppm.18(2).2020.18

Buszko, M., Orzeszko, W., & Stawarz, M. (2021). COVID-19 pandemic and stability of stock market—A sectoral approach. PLoS ONE, 16(5), e0250938. https://doi.org/10.1371/journal.pone.0250938

Dobrolyubova, E. (2021). Measuring outcomes of digital transformation in public administration: Literature review and possible steps forward. NISPAcee Journal of Public Administration and Policy, 14(1), 61–86. https://doi.org/10.2478/nispa-2021-0003

Ďurana, P., Valášková, K., Vagner, L., Zadnanova, S., & Podhorská, I. (2020). Disclosure of strategic managers' factotum: Behavioral incentives of innovative business. International Journal of Financial Studies, 8(1), 17. https://doi.org/10.3390/ijfs8010017

Hortoványi, L., Morgan, R. E., Vuksanović Herceg, I., Djuričin, D., & Roóz, J. (2023). Assessment of digital maturity: The role of resources and capabilities in digital transformation in B2B firms. International Journal of Production Research, 61(14), 4693–4711. https://doi.org/10.1080/00207543.2022.2164087

Miró, J. (2022). Responding to the global disorder: The EU's quest for open strategic autonomy. Global Society, 37(2), 185–203. https://doi.org/10.1080/13600826.2022.2110042

Nagy, O., Papp, I., & Szabó, R. Z. (2021). Construction 4.0 organisational level challenges and solutions. Sustainability, 13(21), 12321. https://doi.org/10.3390/su132112321

Saba, C. S., Ngepah, N., & Odhiambo, N. M. (2023). Information and communication technology (ICT), growth and development in developing regions: Evidence from a comparative analysis and a new approach. Journal of the Knowledge Economy, 15, 973–1011. https://doi.org/10.1007/s13132-023-01571-8

Silva, E. S., Hassani, H., Madsen, D. Ø., & Gee, L. (2019). Googling fashion: Forecasting fashion consumer behaviour using Google Trends. Social Sciences, 8(4), 111. https://doi.org/10.3390/socsci8040111

Šhpak, N., Kuzmin, O., Dvulit, Z., Onysenko, T., & Sroka, W. (2020). Digitalization of the marketing activities of enterprises: Case study. Information, 11(2), 109. https://doi.org/10.3390/info11020109

The DESI Index as a Measure of Digital Readiness: A Systematic Literature Review of Concepts, Applications, and Future Directions

Stovíček Filip[1] and Grisáková Nora[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: filip.stovicek@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: nora.grisakova@euba.sk

Abstract

The rapid advancement of digital technologies has intensified the need for robust, standardized frameworks to assess national and regional digital readiness. The Digital Economy and Society Index (DESI), developed by the European Commission, has emerged as a principal instrument for measuring digital competitiveness across member states and beyond. This article presents a systematic literature review examining the conceptual foundations, methodological dimensions, and empirical applications of the DESI index as a measure of digital readiness. Drawing on 24 peer-reviewed sources, the review synthesizes scholarly evidence across three thematic clusters: the conceptual and structural underpinnings of DESI, its application in cross-country comparative analyses, and its relationship with broader socioeconomic outcomes, including sustainable development and macroeconomic performance. Key findings indicate that while DESI offers a multidimensional and policy-relevant framework for evaluating digital transformation, significant methodological limitations persist, including weighting subjectivity and insufficient coverage of organizational-level indicators. Future research should address these gaps through enhanced composite index design, longitudinal analysis, and integration with complementary frameworks. The review contributes to the growing body of knowledge on digital readiness measurement and offers implications for researchers, policymakers, and practitioners engaged in the digital transformation agenda.

Key Words: DESI index, digital readiness, digital economy, EU measurement, e-government

JEL Classification: O33, O52, H10

1. Introduction

The digital transformation of economies and societies represents one of the most consequential structural shifts of the twenty-first century. From the reorganization of public services and industrial production to the restructuring of labor markets and educational systems, digital technologies are reshaping the foundational architecture of contemporary life. Within this context, the ability to measure, compare, and monitor the progression of digital readiness across countries and sectors has become a critical priority for policymakers, academics, and international organizations alike. The challenge, however, lies in the multidimensional nature of digital readiness, which encompasses connectivity infrastructure, human capital endowments, internet service adoption, business digitalization, and the provision of digital public services—dimensions that resist reduction to any single metric.

The Digital Economy and Society Index (DESI), introduced by the European Commission as part of the Digital Single Market strategy, was designed precisely to address this complexity.

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

By aggregating performance indicators across five principal dimensions into a composite score, DESI enables systematic tracking of digital progress and cross-country benchmarking within the European Union. Since its inception, DESI has attracted substantial scholarly attention, serving as both a dependent variable in studies of macroeconomic performance and sustainable development, and as an independent analytical tool for diagnosing digital skills gaps, competitiveness deficits, and transformation barriers. Despite its widespread adoption, DESI's methodological rigor and representational scope remain subjects of ongoing academic debate.

A clear research gap exists in the literature: while numerous studies have applied DESI as an empirical instrument, relatively few have systematically reviewed its conceptual foundations, examined the coherence of its dimensional architecture, or synthesized the breadth of its applications and limitations across disciplinary contexts. This gap is particularly significant given the evolving nature of digital transformation itself and the expanding policy expectations placed upon composite indices as tools of governance and strategic planning.

This literature review addresses that gap by synthesizing the existing body of evidence on DESI as a measure of digital readiness. Specifically, it aims to clarify the theoretical and methodological underpinnings of the index, evaluate its applications in comparative and sectoral research, and identify the outstanding challenges and research opportunities that remain. By organizing the review thematically rather than sequentially, the article seeks to offer an integrative and critical perspective that goes beyond descriptive summary.

The remainder of this paper is structured as follows. Section 2 establishes the theoretical background by defining key constructs and situating DESI within broader frameworks of digital transformation. Section 3 presents the synthesized literature review, organized into three thematic clusters. Section 4 identifies research gaps and proposes future directions. Section 5 concludes by summarizing the key contributions and noting the limitations of the review.

2. Theoretical Background

The assessment of digital readiness requires a clear theoretical grounding, since the concept itself is complex, multidimensional, and closely related to several adjacent constructs. In academic discourse, digital readiness is not understood merely as the availability of digital infrastructure, but rather as a broader condition reflecting the capacity of societies, economies, institutions, and firms to adopt, use, and benefit from digital technologies. Because of this breadth, any meaningful discussion of the DESI index must be situated within wider theoretical debates on digital transformation, composite measurement, and the role of digital capabilities in contemporary development. This section therefore establishes the conceptual foundations necessary for understanding DESI as a measure of digital readiness. It first clarifies the meaning of digital readiness and digital transformation, then outlines the structure of the DESI framework, discusses the methodological logic of composite indices, and finally highlights the central role of digital skills as a cross-cutting dimension of digital development.

2.1 Defining Digital Readiness and Digital Transformation

Digital readiness is a multidimensional construct that captures the extent to which an economy, organization, or society is prepared to engage with, adopt, and benefit from digital

technologies. It is conceptually distinct from, though closely related to, digital maturity, digital competence, and digital transformation. Digital transformation, as a broader concept, refers to the process of integrating digital technologies into all areas of economic and social activity, fundamentally altering how value is created, delivered, and captured. Jovanović et al. (2018) characterize this process as one of the primary catalysts of contemporary competitiveness, observing that tracking technological development has become crucial at both micro and macro levels. Importantly, digital transformation is not a binary state but a continuous process whose trajectory can be measured and compared across contexts.

The concept of digital readiness draws from multiple theoretical traditions. At the macroeconomic level, it is informed by endogenous growth theory, which recognizes technological capability and human capital as endogenous drivers of productivity and economic growth. Jiao and Sun (2021) provide empirical grounding for this perspective, demonstrating that digital economy development—assessed from the dimensions of internet development, digital literacy, and industrial efficiency—exerts a significant and positive impact on economic growth in China. At the organizational level, digital readiness relates to dynamic capabilities theory, which emphasizes the capacity of organizations to sense and seize digital opportunities while reconfiguring internal competencies. Neri et al. (2023) extend this theoretical lens to the intersection of digital technologies and circular economy practices, showing that digital-enabled dynamic capabilities are foundational to organizational adaptation in the context of sustainability transitions.

2.2 The DESI Framework: Structure and Dimensions

The Digital Economy and Society Index is a composite indicator developed by the European Commission to monitor the digital performance of EU member states. It is structured around five principal dimensions: Connectivity, Human Capital (including digital skills), Use of Internet Services, Integration of Digital Technology by businesses, and Digital Public Services. Each dimension is operationalized through a set of sub-indicators drawn from harmonized European datasets, and the resulting scores are aggregated using fixed weights into an overall index value. Jovanović et al. (2018) were among the first scholars to describe DESI systematically, positioning it as a representative instrument for measuring the digital competitiveness of nations and exploring its relationship to sustainable development.

The multidimensional structure of DESI reflects an acknowledgment that digital readiness cannot be reduced to infrastructure provision alone. Laitsou et al. (2020) employ all five DESI dimensions in their analysis of Greece's digital competitiveness, demonstrating that disaggregated dimensional analysis can reveal structural weaknesses that aggregate scores obscure. Similarly, Małkowska et al. (2021) use DESI data to conduct comparative assessments of digital transformation across European countries, noting that cross-country disparities are most pronounced in the dimensions of connectivity and integration of digital technology.

2.3 Composite Indices and Methodological Foundations

The broader literature on composite indices provides an important methodological context for evaluating DESI. Composite indices aggregate heterogeneous sub-indicators into a scalar value, which facilitates communication and policy benchmarking but also introduces risks of distortion through inappropriate weighting, double-counting, and loss of dimensional nuance. Sidorov and Senchenko (2020) address these concerns in the context of regional digital economy assessment, proposing a composite index model that structures indicators through a functional network and uses additive convolution, deliberately avoiding expert-determined

weights to reduce subjective bias. This methodological alternative highlights a fundamental tension in composite index design between interpretability and rigor.

Olczyk and Kuc-Czarnecka (2022) conduct perhaps the most technically sophisticated investigation of DESI's methodology, employing sensitivity-based analysis to evaluate whether the weights assigned to DESI's constituent elements are optimal and whether the index genuinely captures digital transformation trajectories. Their findings suggest that current DESI weighting structures are suboptimal and that methodological revisions could substantially improve the index's ability to explain variation in GDP per capita across EU economies. This contributes to a broader scholarly consensus that while DESI provides a valuable policy tool, its methodological architecture requires ongoing refinement to maintain scientific validity.

2.4 Digital Skills as a Cross-Cutting Dimension

Within the DESI framework, human capital and digital skills occupy a particularly significant theoretical position. The digital skills gap—the discrepancy between the digital competencies demanded by the economy and those available in the workforce—represents a structural barrier to digital readiness. Carlisle et al. (2021) provide extensive empirical evidence of this gap in the European tourism and hospitality sector, documenting through a large mixed-methods study that digital skills deficits are pervasive and sector-differentiated. Their findings underscore that national-level DESI scores may mask significant variation in digital readiness at the sectoral and occupational levels, raising questions about the granularity and representational adequacy of aggregate indices.

The theoretical literature on digital skills is also informed by human capital theory, which treats workforce competencies as productive assets accumulated through education and training. Elo et al. (2024) contribute to this theoretical discussion by examining the expectation-performance gap in digital skill development among accounting students, finding that technological skills are among the most urgently needed yet least adequately developed through formal education. This finding has direct implications for the Human Capital dimension of DESI, suggesting that educational systems may be lagging behind labor market demands in ways that composite index scores do not fully reflect.

3. Literature Review

This study employs a systematic literature review (SLR) approach to examine the Digital Economy and Society Index (DESI) as a measure of digital readiness. Given the multidimensional and interdisciplinary nature of the topic, a structured and transparent methodological design is necessary to ensure analytical rigor, reproducibility, and consistency of findings. The research methodology is grounded in established principles of systematic review research and is designed to provide a comprehensive and critical synthesis of existing academic knowledge.

The methodological procedure consists of three main stages: (1) identification and selection of relevant literature, (2) data extraction and analytical classification, and (3) thematic synthesis of findings. This structured approach enables the integration of diverse research contributions while maintaining clarity and comparability across studies. The analytical framework of the review is organized around three thematic domains: cross-country comparative applications of

DESI, methodological aspects of composite index construction, and the relationship between digital readiness and socioeconomic outcomes.

3.1 Literature Selection and Data Collection

The most extensively developed application of the DESI index in the academic literature is its deployment as an instrument for comparative analysis of digital competitiveness across European Union member states. A consistent theme in this body of research is the pronounced heterogeneity of digital readiness across EU countries, which DESI's disaggregated structure is well positioned to illuminate. Małkowska et al. (2021) use DESI data to assess the impact of digital transformation across EU members, identifying systematic clusters of high-performing and low-performing countries and linking digital advancement to macroeconomic and social indicators. Their analysis reveals that Northern and Western European countries consistently outperform Southern and Eastern counterparts, a finding that is replicated across multiple studies and time periods.

Borowiecki et al. (2021) extend this comparative analysis by examining DESI data from 2015 and 2020, employing cluster analysis to classify EU countries according to the similarity of their index values and structural profiles. Their findings indicate that while absolute DESI scores have increased across most member states, relative disparities have not diminished significantly, raising questions about the effectiveness of EU digital cohesion policies. This is a critically important observation: aggregate convergence in DESI scores does not necessarily imply structural convergence in the underlying determinants of digital readiness, a distinction that has substantial implications for both research design and policy formulation.

Stanković et al. (2021) introduce a multiple-criteria decision-making approach to supplement DESI-based analysis, arguing that the composite index's fixed weighting scheme may fail to reflect the differential priorities and contexts of individual countries. Their methodology allows for the re-ranking of EU countries based on context-specific weights, yielding results that sometimes diverge substantially from official DESI rankings. This contribution constitutes an important methodological challenge to the uncritical adoption of DESI as an authoritative measure, suggesting that its utility as a comparative tool depends significantly on the analytical assumptions embedded in its construction.

Laitsou et al. (2020) apply DESI specifically to the Greek context, using the index's five dimensions to diagnose the structural determinants of Greece's relatively poor digital competitiveness position within the EU. Their analysis finds that Greece's weakest performance occurs in the Integration of Digital Technology and Digital Public Services dimensions, and that these deficits correlate with broader indicators of economic development and institutional quality. The Greek case illustrates how DESI can function as a diagnostic instrument at the country level, translating abstract composite scores into actionable policy insights.

3.2 Methodological Dimensions and Index Quality

A second major thematic cluster in the literature concerns the methodological integrity of DESI as a composite index. While DESI's practical utility is widely acknowledged, several scholars have raised substantive concerns about its design, weighting, and capacity to accurately reflect the multidimensional phenomenon of digital readiness. Olczyk and Kuc-Czarnecka (2022) represent the most direct engagement with these concerns, demonstrating through sensitivity analysis that the official DESI weighting scheme produces suboptimal results and that alternative weighting configurations would better explain variation in

economic outcomes. Their study also introduces a revised version of the index, suggesting that the academic community has a role to play in the ongoing refinement of DESI as a measurement instrument.

Kutnjak et al. (2020) contribute to this discussion by applying decision tree methodology to identify the key DESI indicators that most powerfully differentiate countries in their digital readiness assessments. Their data-driven approach identifies a subset of indicators—particularly those related to digital skills and internet use—as having disproportionate discriminatory power, implying that the uniform weighting of all indicators within a dimension may not reflect their differential empirical importance. This finding has direct implications for the construction of future index iterations and highlights the value of machine learning methods in composite index validation.

Sidorov and Senchenko (2020) address similar concerns in the context of regional digital economy measurement, proposing a hierarchically structured composite index that uses additive convolution rather than expert weighting to aggregate sub-indicators. Their approach, which explicitly targets the elimination of subjective bias in the weighting process, offers a methodological counterpoint to DESI's current architecture and demonstrates the feasibility of data-driven index construction. Together, these methodological contributions suggest that DESI, while a valuable and widely recognized instrument, would benefit from more systematic sensitivity testing and transparent justification of its structural choices.

Dobrolyubova (2021) approaches the measurement question from the perspective of public administration, reviewing the literature on measuring digital transformation outcomes in government contexts. She finds that existing indices, including DESI, tend to focus on input and process indicators rather than outcome measures, creating a gap between the ambitions of digital transformation programs and the measurement frameworks used to evaluate them. This critique resonates with arguments made by Agostino and Costantini (2021), who develop a framework for assessing the digital transformation of cultural institutions and argue that organizational-level digital assessment requires substantially different methodological approaches than those applied at the national level.

3.3 DESI in Relation to Socioeconomic Outcomes

The third and perhaps most policy-relevant thematic cluster concerns the relationship between DESI scores and broader socioeconomic outcomes, including economic growth, macroeconomic stability, and sustainable development. Research in this cluster consistently finds positive and statistically significant associations between digital readiness as measured by DESI and a range of desirable outcomes, though the causal mechanisms and boundary conditions of these relationships remain subjects of scholarly debate.

Tiutiunyk et al. (2021) investigate the role of digital transformation in achieving macroeconomic stability in EU countries, using DESI-based measures to classify countries into clusters according to the character and direction of the relationship between digitalization and stability indicators. Their findings confirm that digital transformation is associated with improved macroeconomic performance but also identify significant heterogeneity across country clusters, with some groups exhibiting a negative relationship between digital transformation speed and short-term stability. This nuanced result challenges simplistic narratives of digital transformation as uniformly beneficial and calls for more sophisticated econometric approaches to causal identification.

Imran et al. (2022) provide a more direct test of the relationship between DESI dimensions and sustainable development indicators in EU countries, using panel regression analysis to

estimate the effects of each DESI component on a composite sustainable development index. Their results indicate that connectivity and integration of digital technology are the most powerful predictors of sustainable development outcomes, while the effects of digital public services are more modest and context-dependent. These differential effects at the dimensional level reinforce arguments made by Olczyk and Kuc-Czarnecka (2022) that the aggregate DESI score may obscure important variation in the mechanisms through which digital readiness translates into economic and social outcomes.

Jiao and Sun (2021) extend the analysis of digital economy measurement to the Chinese context, constructing a three-dimensional digital economy development index that parallels the DESI structure in several respects. Using spatial econometric models, they find significant positive spillover effects of digital economy development on economic growth, suggesting that the benefits of digital readiness extend beyond national boundaries—a finding with important implications for EU-level digital cohesion policy. Rupeika-Apoga et al. (2022) contribute a firm-level perspective, examining the public support requirements of Latvian SMEs engaged in digital transformation and finding that many firms lack the organizational capabilities necessary to benefit from digital infrastructure investments, regardless of national DESI scores. This finding highlights a critical disconnect between macro-level digital readiness as measured by DESI and micro-level organizational capacity, a gap that aggregate indices are structurally ill-equipped to capture.

The relationship between DESI and higher education readiness is explored by Grīnberga-Zālīte and Zvirbule (2020), who examine the digital preparedness of EU higher education institutions in the context of the COVID-19 pandemic. Their analysis demonstrates that national DESI scores are correlated with the speed and effectiveness of higher education institutions' transition to remote learning, lending further validity to DESI as a macro-level indicator while simultaneously revealing that within-country variation in digital readiness can be substantial. Satalkina and Steiner (2020) broaden the scope further by exploring the indicators of digital entrepreneurship as an outcome of digital transformation, arguing that composite indices like DESI must be complemented by indicators specifically designed to capture innovation and entrepreneurial dynamism within digital ecosystems.

4. Research Gaps and Future Directions

Despite the substantial and growing body of research on DESI as a measure of digital readiness, several important gaps remain that limit both the theoretical development of the field and its practical utility for policy. This section identifies three principal gaps and outlines corresponding future research directions.

The first and most significant gap concerns the methodological foundations of the DESI composite index. While scholars such as Olczyk and Kuc-Czarnecka (2022) and Kutnjak et al. (2020) have begun to interrogate the weighting and structural choices embedded in DESI, a comprehensive, independent methodological audit of the index has not been conducted. The current weighting scheme is not derived from empirical analysis of the relative importance of different dimensions to digital readiness outcomes but rather from a combination of policy judgment and stakeholder consultation. Future research should develop and validate alternative weighting frameworks using robust statistical methods, including principal component analysis, structural equation modeling, and Bayesian model averaging, and should systematically compare the results of such frameworks against the official index. Furthermore, the temporal stability of the index's dimensional structure should be examined, as the relative importance of connectivity, digital skills, and business integration may shift substantially as

digital technologies evolve. A validated, empirically grounded weighting methodology would significantly enhance DESI's scientific credibility and its practical value as a policy instrument.

The second gap relates to the scalar level at which DESI is applied. As Rupeika-Apoga et al. (2022) and Agostino and Costantini (2021) demonstrate, organizational-level digital readiness is not well captured by national-level composite indices. DESI aggregates country-level data, rendering it insensitive to the profound within-country heterogeneity in digital readiness that exists across sectors, firm sizes, urban and rural areas, and demographic groups. Carlisle et al. (2021) illustrate this limitation vividly in the tourism sector, where digital skills deficits coexist with relatively high national DESI scores in some EU countries. Future research should develop downscaled or sector-specific adaptations of the DESI framework that preserve its dimensional logic while operating at more granular levels of analysis. Such adaptations would provide a more accurate picture of digital readiness as it is experienced by businesses, institutions, and individuals, and would enable the identification of specific intervention points for policy and practice.

The third gap concerns the dynamic and longitudinal dimensions of digital readiness. Most existing applications of DESI rely on cross-sectional or short-panel data, limiting the capacity to identify long-term trends, structural breaks, and feedback dynamics in the relationship between digital readiness and socioeconomic outcomes. Borowiecki et al. (2021) provide a partial exception by comparing DESI scores from 2015 and 2020, but even a five-year panel is insufficient to capture the generational dynamics of digital transformation. Future research should leverage the growing historical depth of DESI data to conduct genuinely longitudinal analyses, employing time-series econometrics and dynamic panel methods to investigate how digital readiness develops over time, whether it exhibits path dependency or convergence, and how external shocks—such as the COVID-19 pandemic—affect its trajectory. Grīnberga-Zālīte and Zvirbule (2020) provide a suggestive initial analysis in the higher education context, but a broader, economy-wide longitudinal study remains an important priority for the field.

5. Conclusion

This literature review has synthesized the existing scholarly evidence on the DESI index as a measure of digital readiness, examining its conceptual foundations, methodological architecture, empirical applications, and relationship to socioeconomic outcomes. The review has demonstrated that DESI occupies a central and well-established position in the academic and policy discourse on digital transformation, offering a structured, multidimensional framework that enables systematic cross-country comparison and policy benchmarking. The index's five dimensions—connectivity, human capital, use of internet services, integration of digital technology, and digital public services—collectively capture the principal vectors along which digital readiness is constituted at the national level.

At the same time, the review has identified significant limitations and unresolved debates in the literature. The methodological foundations of the index, particularly its weighting structure, remain open to revision, and sensitivity analyses suggest that the current configuration may not optimally capture the variance in digital transformation outcomes that DESI is intended to explain. Furthermore, the aggregate, national-level character of the index renders it insufficiently sensitive to within-country and organizational-level variation in digital readiness—a limitation that becomes increasingly problematic as the practical demands of digital transformation policy become more granular and context-specific. Finally, the predominantly cross-sectional character of DESI-based research constrains causal inference and limits understanding of the temporal dynamics of digital readiness.

The principal limitations of this review itself should be acknowledged. The analysis is restricted to 24 sources, which, while carefully selected to represent the breadth of the literature, cannot constitute a fully comprehensive survey of all relevant scholarship. Additionally, the review focuses primarily on European contexts, reflecting DESI's own geographic scope, and therefore may not fully account for comparative insights available from non-EU measurement frameworks.

Notwithstanding these limitations, this review contributes to the field by offering a structured, thematically organized synthesis of current knowledge, by clearly articulating the methodological and substantive debates surrounding DESI, and by identifying concrete directions for future research. The continued development of rigorous, valid, and contextually sensitive digital readiness measurement remains a vital priority for researchers, policymakers, and practitioners navigating the complex landscape of digital transformation.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Agostino, D., & Costantini, C. (2021). A measurement framework for assessing the digital transformation of cultural institutions: The Italian case. Meditari Accountancy Research, 29(4), 1–24. https://doi.org/10.1108/medar-02-2021-1207

Borowiecki, R., Siuta-Tokarska, B., Maroń, J., Suder, M., Thier, A., & Żmija, K. (2021). Developing digital economy and society in the light of the issue of digital convergence of the markets in the European Union countries. Energies, 14(9), 2717. https://doi.org/10.3390/en14092717

Carlisle, S., Ivanov, S., & Dijkmans, C. (2021). The digital skills divide: Evidence from the European tourism industry. Journal of Tourism Futures, 8(2), 134–148. https://doi.org/10.1108/jtf-07-2020-0114

Dobrolyubova, E. (2021). Measuring outcomes of digital transformation in public administration: Literature review and possible steps forward. NISPAcee Journal of Public Administration and Policy, 14(1), 39–66. https://doi.org/10.2478/nispa-2021-0003

Elo, T., Pätäri, S., Sjögrén, H., & Mättö, M. (2024). Transformation of skills in the accounting field: The expectation–performance gap perceived by accounting students. Accounting Education, 33(3), 237–273. https://doi.org/10.1080/09639284.2023.2191289

Grīnberga-Zālīte, G., & Zvirbule, A. (2020). Digital readiness and competitiveness of the EU higher education institutions: The COVID-19 pandemic impact. Emerging Science Journal, 4(4), 297–304. https://doi.org/10.28991/esj-2020-01232

Imran, M., Liu, X., Wang, R., Saud, S., & Zhao, Y. (2022). The influence of Digital Economy and Society Index on sustainable development indicators: The case of European Union. Sustainability, 14(18), 11130. https://doi.org/10.3390/su141811130

Jiao, S., & Sun, Q. (2021). Digital economic development and its impact on economic growth in China: Research based on the perspective of sustainability. Sustainability, 13(18), 10245. https://doi.org/10.3390/su131810245

Jovanović, M., Dlačić, J., & Okanović, M. (2018). Digitalization and society's sustainable development – Measures and implications. Zbornik radova Ekonomskog fakulteta u Rijeci, 36(2), 905–928. https://doi.org/10.18045/zbefri.2018.2.905

Kutnjak, A., Hrustek, L., & Križanić, S. (2020). Applying the decision tree method in identifying key indicators of the Digital Economy and Society Index (DESI). In Proceedings of the International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). https://doi.org/10.23919/MIPRO48935.2020.9245142

Laitsou, E., Kárgas, A., & Varoutas, D. (2020). Digital competitiveness in the European Union era: The Greek case. Economies, 8(4), 85. https://doi.org/10.3390/economies8040085

Małkowska, A., Urbaniec, M., & Kosała, M. (2021). The impact of digital transformation on European countries: Insights from a comparative analysis. Equilibrium: Quarterly Journal of Economics and Economic Policy, 16(2), 325–358. https://doi.org/10.24136/eq.2021.012

Neri, A., Negri, M., Cagno, E., Kumar, V., & Garza-Reyes, J. A. (2023). What digital-enabled dynamic capabilities support the circular economy? A multiple case study approach. Business Strategy and the Environment, 32(8), 1–18. https://doi.org/10.1002/bse.3409

Obermayer, N., Csizmadia, T., & Hargitai, D. M. (2022). Influence of Industry 4.0 technologies on corporate operation and performance management from human aspects. Meditari Accountancy Research, 30(4), 1143–1164. https://doi.org/10.1108/medar-02-2021-1214

Olczyk, M., & Kuc-Czarnecka, M. (2022). Digital transformation and economic growth – DESI improvement and implementation. Technological and Economic Development of Economy, 28(3), 775–803. https://doi.org/10.3846/tede.2022.16766

Rupeika-Apoga, R., Bule, L., & Petrovska, K. (2022). Digital transformation of small and medium enterprises: Aspects of public support. Journal of Risk and Financial Management, 15(2), 45. https://doi.org/10.3390/jrfm15020045

Satalkina, L., & Steiner, G. (2020). Digital entrepreneurship: A theory-based systematization of core performance indicators. Sustainability, 12(10), 4018. https://doi.org/10.3390/su12104018

Sidorov, A., & Senchenko, P. V. (2020). Regional digital economy: Assessment of development levels. Mathematics, 8(12), 2143. https://doi.org/10.3390/math8122143

Stanković, J., Marjanović, I., Drezgić, S., & Popović, Ž. (2021). The digital competitiveness of European countries: A multiple-criteria approach. Journal of Competitiveness, 13(2), 117–134. https://doi.org/10.7441/joc.2021.02.07

Tiutiunyk, I., Drábek, J., Antoniuk, N. A., & Navickas, V. (2021). The impact of digital transformation on macroeconomic stability: Evidence from EU countries. Journal of International Studies, 14(3), 220–234. https://doi.org/10.14254/2071-8330.2021/14-3/14

Bass Diffusion Model Theory and Applications: Literature Review of Foundational Frameworks, Extensions, and Emerging Research Frontiers

Štetka Peter[1] and Pallérová Diana[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: peter.stetka@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: diana.pallerova@euba.sk

Abstract

The Bass Diffusion Model, first introduced in 1969, remains one of the most cited and widely applied frameworks in marketing science for forecasting new product adoption and diffusion. Despite its enduring relevance, the model has undergone substantial theoretical development and empirical extension over the past five decades, raising important questions about how contemporary technological and market dynamics challenge its foundational assumptions. This literature review examines the theoretical underpinnings of the Bass Diffusion Model, its principal extensions, and its diverse applications across domains including dynamic pricing, healthcare technology, digital transformation, and sustainability assessment. Drawing on 17 carefully selected sources, the review identifies three primary thematic clusters: model structure and mathematical generalizations, applications in technology adoption and market diffusion, and integration with artificial intelligence and decision-support systems. The review reveals that while utility-based extensions and AI-driven adaptations have substantially enhanced the model's prescriptive power, critical gaps remain in cross-cultural validation, metaverse-era market modeling, and lifecycle-oriented diffusion assessment. Future research should address these gaps to ensure the continued practical relevance and theoretical precision of diffusion modeling frameworks.

Key Words: Bass diffusion model, technology diffusion, innovation adoption forecasting, diffusion curve, market penetration

JEL Classification: O33, C53, M31

1. Introduction

The study of how new products, technologies, and innovations spread through populations represents one of the most fundamental challenges in marketing and management science. Understanding diffusion dynamics is essential not only for academic theorists but also for practitioners who must make high-stakes decisions regarding pricing strategies, resource allocation, promotional investments, and market entry timing. At the center of this intellectual tradition stands the Bass Diffusion Model (BDM), a parsimonious yet powerful mathematical framework that conceptualizes adoption as a function of two social influence processes: innovation, driven by individual exposure to external communications, and imitation, driven by word-of-mouth contact between adopters and non-adopters. The model's simplicity, empirical tractability, and broad applicability have rendered it a cornerstone of quantitative marketing research.

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

Yet the landscape in which the BDM must operate has changed dramatically since its introduction. The emergence of digital ecosystems, AI-powered decision systems, metaverse environments, and wearable health technologies has introduced market complexities that challenge several of the model's classical assumptions, particularly those related to market homogeneity, static pricing, and the absence of network effects beyond binary adoption states. Concurrently, the growing emphasis on sustainability, open innovation, and cross-cultural research has opened new dimensions of inquiry that the original BDM and many of its early extensions were not designed to address.

Despite the existence of a rich body of literature on diffusion modeling, relatively few reviews have attempted to synthesize the theoretical extensions of the BDM alongside its contemporary applications in rapidly evolving technological and organizational contexts. This gap is particularly pronounced with respect to integrating insights from digital transformation research, AI-enhanced marketing analytics, and sustainability-oriented lifecycle assessment into a unified understanding of where diffusion modeling stands today and where it should go next.

The objective of this literature review is therefore threefold. First, it seeks to trace the theoretical development of the Bass Diffusion Model from its classical formulation through its principal generalizations, including utility-based and marketing-mix extensions. Second, it aims to synthesize empirical and applied findings across diverse domains where diffusion modeling has been deployed or is increasingly relevant. Third, it identifies substantive gaps in the current literature and proposes specific directions for future research that respond to both theoretical shortcomings and practical imperatives.

The remainder of this article is organized as follows. Section 2 establishes the theoretical background by defining key constructs and situating the BDM within broader frameworks of innovation adoption and strategic market analysis. Section 3 outlines the methodology of the review, including the search strategy, selection criteria, and thematic synthesis process. Section 4 presents a thematically organized synthesis of the literature, covering model extensions, technological applications, and computational approaches. Section 5 identifies research gaps and future directions, while Section 6 concludes the review with a summary of key contributions and limitations.

2. Theoretical Background

This section presents the theoretical foundations of the study and establishes the conceptual framework for analyzing innovation diffusion and pricing strategies for new products. It first outlines the core principles of the Bass Diffusion Model and its key assumptions regarding how innovations spread within a market. It then discusses major extensions of the model, particularly the Generalized Bass Model and utility-based approaches, which better capture the effects of marketing instruments and pricing decisions on adoption. Finally, the chapter situates the Bass model within the broader context of technology adoption frameworks and dynamic demand modeling, highlighting both its contributions and its limitations in explaining consumer behavior and supporting firms' strategic decision-making.

2.1 Foundational Concepts in Diffusion Theory

The Bass Diffusion Model rests on a set of foundational assumptions about how innovations spread within a bounded market. At its core, the model posits that the probability of adoption at any given time is a linear function of two parameters: the coefficient of innovation (p), which captures the influence of mass media and external stimuli, and the coefficient of imitation (q),

which captures the social contagion effect arising from interpersonal communication between current and potential adopters. The total market potential (m) serves as the upper boundary within which this diffusion process unfolds. Mathematically, the model yields a characteristic S-shaped cumulative adoption curve, a feature that aligns with empirical observations across hundreds of product categories and geographic markets.

The theoretical lineage of the BDM can be traced to earlier work in sociology and epidemiology on the diffusion of innovations, most notably the conceptual frameworks developed by Everett Rogers, who categorized adopters into archetypes including innovators, early adopters, early majority, late majority, and laggards. The BDM operationalizes a simplified version of this social dynamics process by focusing specifically on the binary transition from non-adoption to adoption, abstracting away from the more granular psychological and behavioral factors that Rogers emphasized. This parsimony is both the model's greatest strength and one of its most persistent theoretical limitations.

2.2 The Generalized Bass Model and Marketing Mix Integration

Recognizing that real-world diffusion processes are not independent of marketing variables, researchers developed the Generalized Bass Model (GBM), which extended the original BDM to incorporate marketing instruments such as price and advertising. The GBM introduced a current effect function that allows marketing mix variables to influence the timing and pace of adoption. However, as Cosguner and Seetharaman (2022) have demonstrated, the GBM's reliance on percentage changes in marketing variables rather than their absolute values imposes significant constraints on its prescriptive utility. Specifically, the GBM can be used to derive the optimal trajectory of price changes conditional on an assumed launch price, but it cannot determine the optimal launch price itself, a critical limitation for firms designing go-to-market strategies for genuinely novel products.

This limitation motivates the development of utility-based extensions of the BDM. Cosguner and Seetharaman (2022) propose a framework grounded in consumer utility theory, wherein potential adopters make adoption decisions based on the net utility derived from the product relative to a reservation utility threshold. By embedding the adoption decision within a microeconomic utility-maximization framework, this extension enables the derivation of both the optimal price path and the optimal launch price simultaneously, substantially enhancing the model's strategic applicability. The authors validate their framework using empirical data and demonstrate that it yields markedly superior pricing prescriptions compared to the standard GBM.

2.3 Technology Adoption Frameworks and Diffusion Modeling

The BDM does not exist in isolation, it intersects with a broader ecosystem of technology adoption frameworks that have informed both its theoretical development and its empirical application. The Technology-Organization-Environment (TOE) framework, as applied by Awa, Ukoha, and Emecheta (2016) in the context of enterprise resource planning adoption, provides a complementary lens through which organizational-level diffusion can be understood. While the BDM focuses on population-level dynamics and aggregate adoption rates, the TOE framework situates adoption decisions within the context of technological readiness, organizational capacity, and environmental pressures. These two perspectives are not mutually exclusive; rather, they operate at different levels of analysis and can be integrated to produce richer accounts of diffusion phenomena.

Similarly, the technology acceptance model (TAM) and related frameworks have been extensively applied in the context of digital health technologies. Stoumpos, Kitsios, and Talias (2023) provide a comprehensive review of digital transformation in healthcare, documenting how patient-centric technology adoption is reshaping healthcare delivery systems. Their analysis underscores that adoption decisions in healthcare contexts are influenced by a complex interplay of perceived usefulness, perceived ease of use, trust, and regulatory factors, dimensions that aggregate diffusion models like the BDM do not explicitly represent.

2.4 Dynamic Pricing and Demand Modeling

The relationship between diffusion dynamics and pricing strategy represents a theoretically rich and practically important domain of inquiry. Den Boer (2015) surveys the extensive literature on dynamic pricing and learning, documenting how demand estimation under uncertainty has evolved from static analytical models toward adaptive, data-driven approaches. The convergence between dynamic pricing theory and diffusion modeling is particularly evident in markets for new technology products, where prices typically decline over time as production costs fall and market competition intensifies, and where the social contagion effects modeled by the BDM interact with price sensitivity to generate complex adoption trajectories.

The integration of entrepreneurial strategy considerations further enriches this theoretical landscape. Gans, Stern, and Wu (2019) develop a framework for entrepreneurial strategy that emphasizes the role of commitment and experimentation under uncertainty, concepts that have direct relevance to firms deciding when and how to launch new products whose diffusion trajectories are inherently uncertain. Their analysis of the interplay between uncertainty and learning resonates with the challenges facing practitioners who must calibrate BDM parameters in advance of market introduction, often with limited empirical data.

3. Methodology

This article is based on a narrative literature review aimed at examining the development, theoretical relevance, and selected application areas of the Bass Diffusion Model. Because the topic lies at the intersection of innovation diffusion, marketing, forecasting, and technology adoption, the review was designed as a conceptually oriented synthesis of scholarly literature rather than a purely descriptive overview.

The review focused primarily on literature published between 2018 and 2025, a period reflecting the growing diversification of BDM applications in digital markets, healthcare innovation, pricing analytics, and AI-supported forecasting. The literature search was carried out using the databases Scopus, Web of Science, and Google Scholar, which together offer broad coverage of research in business, management, marketing, and innovation studies. To identify relevant sources, the search was guided by a set of keywords reflecting the main concepts addressed in this article, including “Bass Diffusion Model,” “Bass model,” “new product diffusion,” “innovation diffusion,” “technology adoption,” “market forecasting,” “Generalized Bass Model,” “digital transformation,” and “Bass Model application.”

Only peer-reviewed journal articles published in English were included. To be retained in the final sample, a source had to make a clear contribution to at least one of the following areas: the theoretical basis of the Bass Diffusion Model, the extension or modification of the model, or the application of diffusion logic in current innovation-related settings. Sources were excluded if they were duplicates, non-academic publications, or only marginally related to the focus of the review.

The source selection proceeded in three stages. First, potentially relevant records were identified through database searches and backward checking of references in key studies. Second, titles and abstracts were screened for topical relevance. Third, full texts were reviewed to assess conceptual fit, analytical rigor, and contribution to the objective of the article. The final corpus comprised 17 peer-reviewed sources, which were subsequently subjected to qualitative thematic synthesis. Following an iterative reading process, the literature was organized into three principal thematic clusters: model structure and generalizations, applications in technology and market diffusion, and integration with artificial intelligence and decision-support systems.

4. Literature Review

This section synthesizes the main strands of literature relevant to the Bass Diffusion Model and its broader analytical context. It is organized around three thematic clusters: model extensions and mathematical generalizations, technology adoption and digital transformation across application domains, and the growing role of artificial intelligence and computational methods in diffusion research.

4.1 Thematic Cluster I: Model Extensions and Mathematical Generalizations

The first and most foundational cluster of literature addresses the theoretical structure of the Bass Diffusion Model and the mathematical extensions proposed to address its limitations. This stream of research reflects a persistent tension between parsimony, which supports practical applicability, and the need for greater representational fidelity to the complexity of contemporary markets.

A major recent contribution in this area is the utility-based generalization proposed by Cosguner and Seetharaman (2022), which addresses an important practical limitation of the Generalized Bass Model. By grounding the adoption probability function in consumer utility theory, their model enables firms to determine optimal launch prices endogenously rather than treating them as exogenous inputs. This extension moves the model beyond descriptive forecasting toward a more explicit decision-support role. The authors show through analytical derivation and empirical calibration that the resulting pricing recommendations differ substantially from those implied by the GBM, particularly in markets characterized by high price sensitivity and strong imitation effects. They also propose two utility-based variants, the Bass-Gumbel Diffusion Model and the Bass-Logit Diffusion Model, and report forecasting gains over the GBM in several product categories. A related extension is offered by Najafi-Ghobadi, Bagherinejad, and Taleizadeh (2021), who incorporate heterogeneous strategic customers and jointly examine pricing and advertising decisions in the diffusion of generation products. Their analysis shows that forward-looking consumer behavior materially affects optimal marketing policy and diffusion trajectories. In this respect, the recent literature links diffusion modelling more closely to revenue management and marketing-mix optimization than in the classical aggregate formulation.

A second line of work addresses the homogeneous-population assumption embedded in aggregate diffusion models. Fibich and Golan (2023) show, within a stochastic discrete Bass framework, that heterogeneity in external and internal influence parameters alters diffusion dynamics in ways not captured by homogeneous specifications. The effect of heterogeneity depends on its distribution across the population and cannot be reduced to a simple change in average parameter values. Related concerns appear in network-based formulations. Di Lucchio and Modanese (2024) implement the Bass model and its extensions on arbitrary

complex networks and demonstrate that diffusion paths depend on network topology and degree correlations rather than only on mean-field aggregate effects. Zhang et al. (2023) likewise examine competitive platform-based products with network effects and show that diffusion may involve critical-mass thresholds, coexistence outcomes, and uneven market shares, particularly where direct and indirect network effects interact. These studies challenge the classical assumption of a well-mixed market and point to the importance of structural complexity in diffusion processes. They also indicate that the meaning of diffusion parameters changes once adoption is mediated by local connectivity, interaction asymmetries, and platform-specific complementarities.

Recent studies have also revised the assumption of fixed diffusion parameters by introducing time-varying influence mechanisms and explicit treatment of uncertainty. Han, Tang, and He (2022) propose an improved Bass formulation for product-information diffusion on microblogs by incorporating heterogeneous user decision processes and time-decaying interest. Their results show that the exponential-function specification outperforms the classical Bass and Gompertz models in pre-release prediction tasks. This line of development is extended by Han et al. (2025), who introduce a generalized improved Bass model with dynamic market potential and exponential decay in both internal and external influence, thereby improving forecasting performance in digital information environments. Li, Tao, and Shu (2025) further extend the model through an uncertainty-based formulation and report stronger modeling performance, in their empirical setting, than stochastic differential equation approaches. These studies indicate that recent extensions increasingly depart from the assumption of fixed diffusion parameters by modelling them as functions of time, information environments, and uncertain market conditions. A related line of argument appears in sustainability-oriented assessment, where Buyle et al. (2019) note that prospective lifecycle evaluation depends on credible assumptions about technology diffusion and market penetration, thereby extending the relevance of diffusion modelling beyond forecasting and pricing.

4.2 Thematic Cluster II: Technology Adoption, Digital Transformation, and Healthcare Applications

The second thematic cluster examines how diffusion principles, including those embodied in the BDM, have been applied to understand the spread of digital technologies across consumer and organizational contexts. This cluster is characterized by a significant breadth of application domains, spanning healthcare, enterprise software, wearable devices, and metaverse platforms, and by a notable degree of convergence around certain key findings regarding the factors that accelerate or retard diffusion.

Digital transformation in healthcare represents perhaps the most extensively documented application domain for technology diffusion research in recent years. Stoumpos, Kitsios, and Talias (2023) conduct a systematic bibliometric analysis revealing that digital transformation is reshaping healthcare through multiple technological vectors, including wearable devices, the Internet of Things, virtual reality, and telemedicine platforms. Their analysis shows that patient technology acceptance is a critical bottleneck in the diffusion of healthcare innovations, and that acceptance is mediated by factors including perceived usefulness, digital literacy, trust in providers, and the availability of technical support infrastructure. These findings are broadly consistent with BDM logic, in that they confirm the importance of both intrinsic product characteristics and social influence in driving adoption, but they also highlight dimensions of heterogeneity in adopter populations that aggregate diffusion models typically smooth over.

The role of wearable technologies as a specific category of healthcare innovation further illustrates the complexity of diffusion dynamics in this domain. Yetisen and colleagues (2018) document the extraordinary pace of innovation in wearable medical technologies, encompassing flexible electronics, electrochemical biosensors, microfluidics, and AI-enhanced analytics. They note that while the technological capabilities of wearable devices have advanced dramatically, their diffusion into clinical and consumer settings has been uneven, reflecting the interplay of regulatory barriers, reimbursement uncertainties, data privacy concerns, and the heterogeneous risk tolerance of different adopter segments. This pattern of differential diffusion across market segments, driven by both innovation and imitation forces, is precisely the kind of dynamic that BDM extensions incorporating market segmentation and heterogeneous adoption thresholds are designed to capture.

The metaverse represents a particularly challenging and intriguing application domain for diffusion modeling. Dwivedi and colleagues (2022) examine the marketing implications of metaverse adoption, arguing that the metaverse will create fundamentally new contexts for consumer-brand interaction that require novel theoretical frameworks. Their analysis suggests that the diffusion of metaverse platforms will be influenced not only by the technical capabilities of the underlying infrastructure but also by the cultural norms, social identity dynamics, and experiential expectations of potential users. Critically, they note that the metaverse's adoption trajectory may not follow the smooth S-shaped curve predicted by classical diffusion models, given the possibility of discontinuous adoption events driven by platform tipping points, viral social phenomena, and regulatory interventions. This observation raises fundamental questions about the applicability of BDM-derived frameworks to platform ecosystems and network-effect-intensive technologies.

The role of culture in shaping innovation adoption and diffusion patterns adds a further layer of complexity to this thematic cluster. Tian et al. (2018) systematically review the literature on culture and innovation, identifying two primary dimensions of cultural influence: organizational culture, which shapes internal innovation processes and knowledge sharing norms, and national culture, which influences the broader societal receptiveness to technological change. Their findings suggest that diffusion rates and patterns vary substantially across cultural contexts, a finding with direct implications for the cross-national generalizability of BDM parameter estimates. Specifically, populations characterized by high uncertainty avoidance and low individualism, as conceptualized in Hofstede's cultural dimensions framework, may exhibit systematically lower innovation coefficients and higher imitation coefficients, reflecting a greater reliance on social proof and peer observation in adoption decisions.

4.3 Thematic Cluster III: Artificial Intelligence, Decision Support, and Computational Approaches to Diffusion

The third thematic cluster addresses the growing intersection between diffusion modeling and artificial intelligence, encompassing both the use of AI techniques to enhance the estimation and optimization of diffusion models and the application of diffusion frameworks to understand the spread of AI technologies themselves.

The use of AI in marketing decision-making has grown substantially, and its implications for diffusion modeling are multifaceted. Overgoor and colleagues (2019) propose a structured process for deploying AI in marketing contexts, grounded in the Cross-Industry Standard Process for Data Mining framework. Their analysis identifies several specific marketing tasks for which AI approaches offer demonstrable advantages over traditional analytical methods,

including customer segmentation, churn prediction, and demand forecasting. Critically, they note that AI models can capture non-linear relationships and interaction effects in consumer behavior data that conventional parametric models, including the BDM, are not designed to represent. This observation suggests that AI-powered diffusion modeling, which combines the structural interpretability of the BDM with the representational flexibility of machine learning algorithms, represents a promising frontier for model development.

Al-Surmi, Bashiri, and Koliousis (2021) provide an empirical contribution to this frontier by developing an AI-based decision-making framework that combines strategic alignment theory with artificial neural network modeling. Their framework, validated using data from 242 managers across multiple industries, demonstrates that strategic alignment between marketing and IT functions is a significant predictor of operational performance, and that AI-based models can identify this relationship more reliably than traditional regression approaches. While their study does not directly apply the BDM, it illustrates the broader principle that combining structural marketing theory with computational AI methods can yield decision-support tools of substantially greater practical power than either approach alone.

The application of open innovation principles to technology diffusion further enriches this thematic cluster. Huggins, Prokop, and Thompson (2019) examine the determinants of network centrality in university-industry innovation networks, finding that universities occupying more central network positions exhibit higher rates of spin-off generation and externally funded research engagement. Their network-theoretic analysis is directly relevant to diffusion modeling because network structure fundamentally shapes the transmission pathways through which innovations spread. The BDM, in its classical form, assumes a well-mixed population in which any adopter can influence any potential adopter with equal probability, an assumption that is obviously violated in network-structured populations where influence propagates through specific relational ties. Extending BDM logic to network-aware diffusion models, informed by empirical findings about real-world innovation network structures, represents an important avenue for theoretical advancement.

The intersection of entrepreneurial strategy and diffusion modeling also warrants attention within this cluster. Gans, Stern, and Wu (2019) argue that entrepreneurial strategy is fundamentally concerned with managing the interplay between uncertainty and commitment, and that the sequential resolution of uncertainty through experimentation and learning is central to successful new venture development. This framing resonates directly with the challenges of BDM parameter estimation under conditions of radical uncertainty, such as those encountered by firms launching truly novel products in nascent markets. When historical analogues are unavailable and market potential is genuinely unknown, the analytical foundations of standard BDM estimation approaches become problematic, and alternative strategies grounded in adaptive experimentation and Bayesian updating may be more appropriate.

Across all three thematic clusters, a consistent meta-finding emerges: the most significant advances in diffusion modeling arise at disciplinary boundaries, where the structural insights of the BDM are combined with theories, methods, and empirical findings from adjacent domains including utility theory, cultural studies, network science, and artificial intelligence.

5. Research Gaps and Future Directions

5.1 The Challenge of Cross-Cultural Diffusion Modeling

One of the most persistent and consequential gaps in the BDM literature concerns its cross-cultural validity. As Tian et al. (2018) demonstrate, culture exerts a powerful and multidimensional influence on innovation adoption patterns, yet the vast majority of BDM calibration studies are conducted in Western, technologically advanced market contexts and their findings are often uncritically generalized to other cultural settings. This limitation is particularly important in light of studies such as Stoumpos, Kitsios, and Talias (2023) and Dwivedi et al. (2022), which show that technology adoption is deeply shaped by contextual and social factors that vary across sectors and user environments. The systematic variation in BDM parameters across cultural contexts, including differences in innovation and imitation coefficients attributable to cultural dimensions such as uncertainty avoidance, collectivism, and long-term orientation, has not been adequately theorized or empirically documented. Future research should therefore prioritize multi-country BDM calibration studies that explicitly incorporate cultural variables, potentially using the frameworks developed by Tian et al. (2018) to generate theoretically grounded predictions about cross-cultural parameter variation. The practical implications of this research direction are substantial: firms launching new products globally need diffusion forecasts that are calibrated to the specific cultural contexts of their target markets, not generic parameter estimates derived from culturally unrepresentative samples.

5.2 Diffusion Modeling for Network-Effect Technologies and Platform Markets

The emergence of platform-based business models and network-effect-intensive technologies poses a fundamental challenge to classical BDM assumptions. The metaverse, as analyzed by Dwivedi et al. (2022), exemplifies a new category of technology whose value to each user is directly contingent on the size and composition of the user base, creating adoption dynamics that may be characterized by threshold effects, tipping points, and multiple equilibria rather than the smooth S-shaped trajectories predicted by the BDM. This observation is consistent with the findings of Zhang et al. (2023), who show that competitive platform-based products with network effects may generate critical-mass thresholds, coexistence outcomes, and uneven market shares. Similarly, Di Lucchio and Modanese (2024) demonstrate that diffusion outcomes depend strongly on network topology and degree correlations, while Fibich and Golan (2023) show that heterogeneity in internal and external influence parameters materially changes aggregate diffusion paths. The diffusion of AI-powered tools through organizations, as discussed by Overgoor et al. (2019) and Al-Surmi, Bashiri, and Koliousis (2021), may likewise be governed by network externalities and platform complementarities that are not captured by standard diffusion parameters. Future research should therefore develop BDM extensions that explicitly incorporate network externalities, platform tipping dynamics, and competitive multi-platform diffusion, drawing also on network-theoretic insights such as those provided by Huggins, Prokop, and Thompson (2019). Such extensions would substantially improve the model's applicability to the dominant technology categories of the contemporary digital economy.

5.3 Integrating Diffusion Models with Lifecycle Sustainability Assessment

A third significant gap concerns the integration of diffusion modeling with environmental and social lifecycle assessment. Buyle et al. (2019) note that ex-ante lifecycle assessment depends critically on accurate projections of technology market penetration, yet the diffusion

models typically used for this purpose are often poorly calibrated and rarely updated as real-world adoption data accumulate. This issue becomes even more important when diffusion processes are shaped by heterogeneous adoption conditions, as highlighted by Yetisen et al. (2018) in the context of wearable medical technologies and by Stoumpos, Kitsios, and Talias (2023) in relation to digital health platforms. In both cases, adoption is influenced not only by technological functionality, but also by regulation, trust, infrastructure, and user readiness, suggesting that sustainability projections based on simplistic diffusion assumptions may be misleading. Moreover, recent methodological advances by Cosguner and Seetharaman (2022), Han, Tang, and He (2022), Han et al. (2025), and Li, Tao, and Shu (2025) indicate that diffusion models can be made more responsive to changing market conditions, uncertainty, and time-varying influences. Future research should build on these advances to develop standardized protocols for embedding dynamically updated and empirically validated BDM parameter estimates into prospective lifecycle assessment models. This would enable more reliable sustainability appraisals of emerging technologies at early stages of market development. The implications are particularly relevant for public policy, as regulatory decisions regarding technology incentives and environmental standards are increasingly informed by prospective impact assessments that depend on diffusion assumptions.

6. Conclusion

This literature review has examined the theoretical foundations, principal extensions, and diverse applications of the Bass Diffusion Model, drawing on a body of scholarship that spans quantitative marketing, healthcare informatics, sustainability science, entrepreneurship, and artificial intelligence. The review has demonstrated that the BDM, despite its origins in a markedly simpler technological and market environment, retains substantial theoretical relevance and practical utility in the contemporary era, provided that it is appropriately extended and contextualized.

The three thematic clusters identified in this review highlight both the strengths and the limitations of diffusion modeling as currently practiced. In the first cluster, Cosguner and Seetharaman (2022) stand out for extending the BDM toward endogenous pricing decisions, while Fibich and Golan (2023) and Di Lucchio and Modanese (2024) show that heterogeneity and network structure can significantly reshape diffusion trajectories beyond the assumptions of classical aggregate models. The second cluster confirms the broad applicability of diffusion logic across contemporary technology domains. Stoumpos, Kitsios, and Talias (2023) and Yetisen et al. (2018) demonstrate its relevance in digital health and wearable technologies, whereas Dwivedi et al. (2022) highlight the challenges of applying traditional diffusion models to network-effect-intensive environments such as the metaverse. Tian et al. (2018) further emphasize that cultural context plays an important role in shaping adoption patterns, raising questions about the cross-national generalizability of BDM parameters. The third cluster suggests that the future of diffusion research lies increasingly at the intersection of structural modeling, artificial intelligence, and machine learning. In particular, Overgoor et al. (2019) and Al-Surmi, Bashiri, and Koliousis (2021) point to the growing value of AI- and machine-learning-enhanced decision-support systems, while Gans, Stern, and Wu (2019) underscore the importance of uncertainty, experimentation, and strategic learning in innovation-driven markets.

Nonetheless, this review has also identified substantive gaps that future research must address: the cross-cultural validity of BDM parameters, the model's applicability to network-effect and platform technologies, and its integration with lifecycle sustainability assessment

frameworks. Addressing these gaps will require interdisciplinary collaboration and methodological pluralism, combining the formal modeling traditions of quantitative marketing with the theoretical richness of cultural studies, network science, and environmental systems analysis.

A limitation of this review is that it draws on a relatively small set of 17 sources, which, while carefully selected for their relevance and scholarly quality, necessarily constrain the comprehensiveness of the synthesis. Future reviews should incorporate a systematic search of the full BDM literature, potentially using bibliometric mapping methods to identify the most influential theoretical and empirical contributions across the model's five-decade history.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Al-Surmi, A., Bashiri, M., & Koliousis, I. (2021). AI based decision making: combining strategies to improve operational performance. International Journal of Production Research, 60(14), 4461–4478. https://doi.org/10.1080/00207543.2021.1966540

Awa, H. O., Ukoha, O., & Emecheta, B. C. (2016). Using T-O-E theoretical framework to study the adoption of ERP solution. Cogent Business & Management, 3(1), 1196571. https://doi.org/10.1080/23311975.2016.1196571

Bass, F. M. (1969). A new product growth for model consumer durables. Management Science, 15(5), 215–227. https://doi.org/10.1287/mnsc.15.5.215

Buyle, M., Audenaert, A., Billen, P., Boonen, K., & Van Passel, S. (2019). The future of ex-ante LCA? Lessons learned and practical recommendations. Sustainability, 11(19), 5456. https://doi.org/10.3390/su11195456

Cosguner, K., & Seetharaman, P. (2022). Dynamic pricing for new products using a utility-based generalization of the Bass Diffusion Model. Management Science, 68(5), 3220–3239. https://doi.org/10.1287/mnsc.2021.4257

Den Boer, A. V. (2015). Dynamic pricing and learning: Historical origins, current research, and new directions. Surveys in Operations Research and Management Science, 20(1), 1–18. https://doi.org/10.1016/j.sorms.2015.03.001

Di Lucchio, L., & Modanese, G. (2024). The Bass diffusion model: Agent-based implementation on arbitrary networks. Mathematical and Computer Modelling of Dynamical Systems, 30(1), 364–384. https://doi.org/10.1080/13873954.2024.2350244.

Dwivedi, Y. K., Hughes, L., Wang, Y., Alalwan, A. A., & Ah, S. J. (2022). Metaverse marketing: How the metaverse will shape the future of consumer research and practice. Psychology and Marketing, 40(9), 1–27. https://doi.org/10.1002/mar.21767

Fibich, G., & Golan, A. (2023). Diffusion of new products with heterogeneous consumers. Mathematics of Operations Research, 48(1), 257–287. https://doi.org/10.1287/moor.2022.1261.

Gans, J. S., Stern, S., & Wu, J. (2019). Foundations of entrepreneurial strategy. Strategic Management Journal, 40(5), 736–756. https://doi.org/10.1002/smj.3010

Han, Z., Tang, Z., & He, B. (2022). Improved Bass model for predicting the popularity of product information posted on microblogs. Technological Forecasting and Social Change, 176, 121458. https://doi.org/10.1016/j.techfore.2021.121458

Han, Z., Chen, X., Miao, K., Wang, X., & Li, Q. (2025). Predicting product information diffusion for sustainable quality management in industry 4.0: An improved Bass model approach. Scientific Reports, 15(1), 25989. https://doi.org/10.1038/s41598-025-09129-1

Huggins, R., Prokop, D., & Thompson, P. (2019). Universities and open innovation: The determinants of network centrality. The Journal of Technology Transfer, 45(3), 718–757. https://doi.org/10.1007/s10961-019-09720-5

Li, B., Tao, Z., & Shu, Y. (2025). Uncertain Bass diffusion model and modeling the purchase volume of private cargo vehicles in China. Chaos, Solitons & Fractals, 195, 116330. https://doi.org/10.1016/j.chaos.2025.116330.

Najafi-Ghobadi, S., Bagherinejad, J., & Taleizadeh, A. A. (2021). Modeling the diffusion of generation products in the presence of heterogeneous strategic customers for determining optimal marketing-mix strategies. Computers & Industrial Engineering, 160, 107606. https://doi.org/10.1016/j.cie.2021.107606.

Overgoor, G., Chica, M., Rand, W., & Weishampel, A. (2019). Letting the computers take over: Using AI to solve marketing problems. California Management Review, 61(4), 156–185. https://doi.org/10.1177/0008125619859318

Rogers, E. M. (2003). Diffusion of innovations (5th ed.). Free Press. ISBN 9780743258234.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039.

Stoumpos, A. I., Kitsios, F., & Talias, M. A. (2023). Digital transformation in healthcare: Technology acceptance and its applications. International Journal of Environmental Research and Public Health, 20(4), 3407. https://doi.org/10.3390/ijerph20043407

Tian, M., Deng, P., Zhang, Y., & Salmador, M. P. (2018). How does culture influence innovation? A systematic literature review. Management Decision, 56(5), 1088–1107. https://doi.org/10.1108/md-05-2017-0462

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.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly, 26(2), xiii–xxiii. [No DOI reliably confirmed in the publisher records I found.]

Yetisen, A. K., Martínez-Hurtado, J. L., Ünal, B., Khademhosseini, A., & Butt, H. (2018). Wearables in medicine. Advanced Materials, 30(33), 1706910. https://doi.org/10.1002/adma.201706910

The ASSESSOR Model in New Product Forecasting: A Critical Literature Review of Methodological Foundations, Applications, and Emerging Challenges

Grisáková Nora[1] and Baziuk Andriana[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: nora.grisakova@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: andriana.baziuk@euba.sk

Abstract

New product forecasting represents one of the most consequential and analytically demanding tasks in marketing science and managerial decision-making. The ASSESSOR model, developed as a pretest market methodology, has occupied a central position in this domain for several decades, offering firms a structured, cost-effective framework for predicting new product performance prior to full-scale commercial launch. This review systematically evaluates the theoretical foundations, methodological architecture, and empirical validity of ASSESSOR within the evolving landscape of predictive modeling. By synthesizing seminal research and contemporary studies, the review categorizes the literature into three thematic clusters: structural components (preference and trial-repeat modules), predictive robustness across diverse categories, and methodological limitations in the digital age. Key findings indicate that while ASSESSOR retains considerable practical relevance, its integration with modern machine learning, agent-based simulation, and big data analytical frameworks remains underexplored. The review concludes by identifying critical research gaps and proposing future directions for advancing pretest market methodology in an increasingly complex and data-rich commercial environment.

Key Words: ASSESSOR model, new product forecasting, market simulation, conjoint analysis, pre-launch testing

JEL Classification: M31, C53, L10

1. Introduction

The launch of a new product represents one of the highest-stakes decisions a firm can undertake. Estimates of new product failure rates have historically ranged from forty to ninety percent depending on industry and definition, placing enormous pressure on marketing managers and strategists to develop reliable predictive tools before committing to full-scale market entry (Silk & Urban, 1978). Within this context, pretest market models emerged as a critical class of analytical instruments, enabling firms to simulate consumer response and forecast market performance using controlled research designs conducted prior to product launch. Among these methodologies, the ASSESSOR model developed by Silk and Urban stands as one of the most influential and enduring contributions to marketing science.

Despite its longevity and documented empirical track record, the ASSESSOR model faces a landscape that has changed dramatically since its original formulation. The rise of digital

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

commerce, the proliferation of real-time consumer data, the advancement of machine learning and agent-based simulation techniques, and the growing complexity of competitive market structures all raise fundamental questions about the continued adequacy of established forecasting frameworks. Contemporary developments in statistical methodology, including innovations in probabilistic forecasting and model transparency requirements, further compound the need for a systematic reassessment of the ASSESSOR model's role in modern marketing analytics.

A clear research gap exists in the literature: while isolated studies have examined specific components of the ASSESSOR model or proposed incremental refinements, no comprehensive and current literature review synthesizes its theoretical foundations, empirical performance record, and methodological tensions with emerging analytical paradigms. This absence is particularly consequential given that practitioners continue to deploy ASSESSOR-based approaches while simultaneously confronting data environments for which the model was never originally designed.

The objective of this literature review is therefore threefold. First, it seeks to consolidate and critically assess the theoretical and methodological foundations of the ASSESSOR model, situating it within the broader canon of new product forecasting. Second, it synthesizes empirical findings regarding the model's predictive validity and practical applications, identifying areas of consensus and disagreement in the literature. Third, it identifies substantive research gaps and proposes specific future research directions that connect methodological innovation to the practical challenges facing marketing decision-makers.

The paper is organized into four main parts. It starts with a theoretical background that explains the key concepts behind pretest market modeling and the ASSESSOR framework. Next, the literature review summarizes and discusses previous research, including the methodological approach used in this study. The following section highlights the main research gaps and suggests directions for future work. The paper ends with a conclusion that summarizes the key findings and notes the main limitations.

2. Theoretical Background

Understanding pretest market modeling requires both a clear conceptualization of its core principles and an overview of the theoretical frameworks that underpin its development. In this context, the ASSESSOR model serves as a central reference point, linking consumer behavior theory with quantitative forecasting approaches.

2.1 Pretest Market Modeling and the ASSESSOR Framework

Pretest market models constitute a class of quantitative marketing research tools designed to forecast the sales performance of new consumer products using data collected before the product enters the marketplace on a commercial scale. These models address the fundamental problem that a firm must allocate substantial resources to production, distribution, and marketing before receiving reliable feedback from actual market performance. The economic logic is compelling: the cost of conducting a controlled pretest market study is typically a small fraction of the investment required for full market rollout, while the informational value of a well-executed forecast can be substantial.

The ASSESSOR model, originally articulated by Silk and Urban (1978), integrates two conceptually distinct but empirically complementary submodels. The trial-repeat model estimates the proportion of consumers who will try the new product and subsequently

repurchase it, drawing on concepts from consumer behavior theory and stochastic modeling. The preference model, grounded in the tradition of multiattribute attitude measurement, uses consumers' expressed preferences across competitive alternatives to estimate the new product's equilibrium market share. The convergence of these two estimates provides a cross-validation mechanism that enhances forecast reliability.

This dual-model architecture reflects a broader epistemological commitment within marketing science to triangulating forecasts across methodologically independent approaches. The logic of triangulation has been formally endorsed in quantitative modeling practice across multiple applied sciences, where the integration of complementary models is recognized as a strategy for reducing forecast error and improving decision-relevant communication (Marshall et al., 2015). In the pharmaceutical domain, for instance, model-informed decision-making frameworks explicitly advocate for the integration of multiple analytical approaches to reduce uncertainty, a principle directly analogous to ASSESSOR's convergence logic (Marshall et al., 2015).

2.2 Stochastic Foundations and Consumer Choice Theory

The behavioral substructure of the ASSESSOR model rests on stochastic learning models of brand choice, which treat repeat-purchase behavior as a probabilistic process governed by individual-level parameters. These models draw on Markov chain representations of brand switching, where transition probabilities capture the likelihood that a consumer who purchased a given brand on one occasion will purchase a specific brand on the next occasion (Dung et al, 2022). The trial-repeat formulation estimates the steady-state market share implied by these transition dynamics, providing a theoretically grounded link between micro-level consumer behavior and aggregate market outcomes.

The preference model component draws on the Lancaster model of consumer demand (Lancaster, 1966), in which products are conceived as bundles of attributes, and consumers' utility is derived from those attributes rather than from the products themselves. Conjoint analysis and multiattribute scaling methods are employed to estimate preference parameters, which are then used to predict choice probabilities through variants of the multinomial logit or Luce choice model. The theoretical coherence of this framework has contributed to ASSESSOR's enduring credibility as a scientific instrument.

2.3 Broader Forecasting Methodologies and Contextual Relevance

The intellectual context within which ASSESSOR operates is considerably broader than marketing science alone. Quantitative forecasting methodology has evolved substantially across numerous applied domains, with implications for how pretest market models are conceptualized and evaluated. Time series methods such as seasonal autoregressive integrated moving average with exogenous factors (SARIMAX) have demonstrated strong performance in long-horizon demand forecasting tasks (Alharbi & Csala, 2022), and their capacity to incorporate external variables provides a conceptual bridge to the covariate-enriched models that modern ASSESSOR extensions might employ. Similarly, the development of agent-based electricity market simulation frameworks, which integrate machine learning with behavioral agent modeling, illustrates how hybrid methodological architectures can outperform simpler linear approaches in complex, dynamic environments (Fraunholz et al., 2021).

Statistical communication and the appropriate representation of uncertainty constitute another foundational concern for forecasting models. The challenge of conveying probabilistic forecast outputs to managerial decision-makers—individuals who may lack statistical

training—is significant. Shafer (2021) proposes a framework for statistical communication centered on betting metaphors, which effectively conveys the strength of evidence without requiring fluency in p-value interpretation. This framework has implications for how ASSESSOR forecasts are communicated within organizations, a dimension of model use that has received insufficient attention in the marketing literature.

Model transparency and documentation are equally critical for the responsible application of any forecasting methodology. The risk of non-transparent or poorly documented models generating consequential decisions has been recognized across multiple applied domains. Energy scenario modeling, for example, has been critiqued for insufficient transparency in model assumptions and parameterization (Cao et al., 2016), while pharmaceutical modeling practice has formalized good practice documentation standards precisely to ensure that model-based conclusions are traceable and reproducible (Marshall et al., 2015). These developments outside marketing science carry direct lessons for the documentation and validation standards that should accompany ASSESSOR applications.

2.4 Risk, Uncertainty, and Organizational Decision-Making

Any forecasting model operates within an organizational context that shapes how its outputs are interpreted and acted upon. Hardy and Maguire (2015) offer a theoretically sophisticated account of how organizations "organize risk," demonstrating that risk knowledge is not simply received from analytical models but is actively constructed through discourse, power relations, and institutional norms. This perspective is particularly relevant for new product forecasting, where the ASSESSOR output enters organizational decision processes that may privilege certain interpretations over others, and where the political dynamics of product development teams can influence how forecast uncertainty is acknowledged or suppressed. Understanding ASSESSOR as a sociotechnical artifact embedded in organizational risk discourse, rather than merely a neutral analytical tool, enriches the theoretical framing of its use and limitations (Hardy & Maguire, 2015).

3. Literature Review

This study adopts a systematic qualitative literature review to synthesize existing research on the ASSESSOR framework. A targeted search was conducted across major academic databases, including ScienceDirect, Web of Science, Scopus, and Google Scholar, using keywords such as “ASSESSOR model,” “pretest market methodology,” “new product forecasting,” and their combinations with Boolean operators. The selection was limited to peer-reviewed articles published in English to ensure academic rigor and consistency. The final sample includes seminal theoretical works, empirical validation studies, and recent critiques, providing a comprehensive view of the model’s development.

The literature was evaluated based on its contribution to the model’s methodological structure. A thematic synthesis was applied to organize findings into three main categories: structural foundations, empirical performance, and emerging methodological challenges. This analysis highlights key research gaps, especially the limited integration of big data analytics and machine learning within the traditional ASSESSOR framework.

3.1 Thematic Cluster One: Methodological Foundations and Model Validity

The methodological literature on pretest market models has consistently focused on the question of predictive validity—the degree to which laboratory-derived forecasts correspond to actual post-launch market performance. The ASSESSOR model has accumulated a

substantial empirical record in this regard, with numerous validation studies demonstrating that its market share forecasts typically fall within acceptable error bounds of actual performance. This body of evidence has been foundational in establishing ASSESSOR's credibility as a commercial research tool and has driven its adoption by major consumer packaged goods manufacturers.

The model's internal logic, which couples a trial-repeat stochastic formulation with a multiattribute preference model, creates a self-correcting architecture. When the two component estimates diverge substantially, this discrepancy signals potential specification error and prompts diagnostic investigation, a property that distinguishes ASSESSOR from single-mechanism forecasting approaches. The convergence criterion effectively functions as a built-in validation check, consistent with the broader principle in applied quantitative modeling that model adequacy should be assessed through multiple complementary evaluations rather than a single performance metric (Marshall et al., 2015).

Critics of the ASSESSOR methodology have nonetheless raised concerns about the ecological validity of the laboratory purchase simulation environment. The controlled setting in which consumers are exposed to the new product and competitive alternatives may inadequately represent the informational complexity and social dynamics of real retail environments. Moreover, the model's assumption that consumer preferences measured in a controlled setting will translate faithfully to purchase behavior in the marketplace remains a theoretical commitment rather than an empirically resolved question. This tension between laboratory control and ecological realism is a persistent methodological challenge that affects pretest market research more broadly.

The treatment of uncertainty in ASSESSOR outputs has also been critiqued. Most applications of the model report a point forecast of market share rather than a full probability distribution, obscuring the uncertainty inherent in the estimation process. The development of more rigorous probabilistic communication frameworks represents an important direction for model enhancement. Shafer's (2021) proposal for evidence-based statistical communication, while developed in a different context, highlights the general inadequacy of point estimates as complete representations of forecasting uncertainty and suggests that richer probabilistic outputs would better serve managerial decision-making. This critique connects to a broader literature on statistical modeling that advocates for explicit uncertainty quantification as a standard of scientific rigor.

3.2 Thematic Cluster Two: Model Extensions, Hybrid Approaches, and Integration with Modern Analytics

A second major thematic cluster in the relevant literature concerns the adaptation and extension of pretest market methodology to accommodate new data sources, modeling paradigms, and competitive environments. The original ASSESSOR model was designed for the relatively stable consumer packaged goods categories of the 1970s and 1980s, where competitive structures were well-defined, consumer information environments were limited, and distribution dynamics were predictable. Contemporary markets differ fundamentally in all three respects, creating both challenges and opportunities for model evolution.

The integration of machine learning with traditional analytical models has attracted considerable attention across multiple applied forecasting domains. In electricity market modeling, the combination of artificial neural networks with agent-based simulation has demonstrated superior forecast accuracy relative to linear regression benchmarks, particularly in environments characterized by high complexity and structural change (Fraunholz et al.,

2021). This methodological advance suggests an analogous opportunity in new product forecasting, where neural network approaches could be used to estimate preference parameters or simulate trial dynamics with greater flexibility than parametric stochastic models allow. The challenge, as in energy modeling, lies in preserving interpretability while gaining predictive power—a tension that has not been satisfactorily resolved in either domain (Fraunholz et al., 2021).

Stochastic modeling methods developed for building stock energy prediction offer another instructive parallel. Lim and Zhai (2017) systematically review bottom-up stochastic engineering approaches for aggregate energy forecasting, documenting both the strengths of individual-level behavioral simulation and the computational challenges of scaling such models to large populations. The parallels with consumer market simulation are considerable: in both contexts, aggregate outcomes emerge from the heterogeneous decisions of many individual agents, and the appropriate level of model granularity is a contested methodological question. The comparison suggests that ASSESSOR extensions employing individual-level simulation might benefit from methodological insights developed in computational engineering contexts (Lim & Zhai, 2017).

Time series methods have been employed as complementary tools for validating and contextualizing pretest market forecasts. The SARIMAX modeling framework, which accommodates seasonal variation and exogenous covariates within a unified estimation structure, has demonstrated predictive accuracy in long-range demand forecasting applications (Alharbi & Csala, 2022). In the new product context, such methods could be used to model category-level demand trends that form the baseline against which ASSESSOR's share forecasts are applied, potentially improving the absolute sales forecasts derived from relative market share predictions. This complementary use of time series and preference-based forecasting represents an underexplored integration opportunity.

The role of big data and digital information platforms in transforming the inputs available to pretest market models is a particularly consequential development. Ali et al. (2020) demonstrate, in the context of urban building energy modeling, how GIS-based data integration and multi-scale modeling can substantially enhance forecast quality when relevant spatial and behavioral data are available and appropriately combined. The analogous opportunity in marketing exists wherever rich digital behavioral data—online search, social media engagement, review platforms—can supplement or partially replace the controlled laboratory research that ASSESSOR traditionally relies upon. However, such integration raises significant methodological challenges regarding data quality, representativeness, and the comparability of digitally observed behavior with in-store purchase decisions (Ali et al., 2020).

The emergence of ecosystem-based approaches in fisheries management modeling, as documented by Howell et al. (2021), provides an instructive case study in the integration of strategic and tactical modeling levels within a unified decision support framework. Their approach combines broad ecosystem models with precise single-species assessment models, using the former to contextualize and adjust the outputs of the latter. This architecture has a direct structural analogy in marketing: competitive market simulation models could serve a contextualizing function relative to ASSESSOR's product-level share forecast, adjusting predictions based on anticipated competitive responses that the original model does not explicitly model (Howell et al., 2021).

3.3 Thematic Cluster Three: Organizational Implementation, Communication, and Governance of Forecasting Models

The third major thematic cluster addresses the organizational and institutional dimensions of forecasting model use, which have received comparatively less attention in the pretest market literature despite their critical practical importance. A forecasting model is only as valuable as the quality of the decisions it informs, and decision quality depends not only on forecast accuracy but also on how model outputs are communicated, interpreted, and integrated within organizational decision processes.

Hardy and Maguire (2015) provide a theoretically sophisticated account of how organizational risk discourse shapes the interpretation of analytical outputs. Their analysis reveals that the dominant discourse of risk in organizations tends to normalize certain types of analytical knowledge while marginalizing others, and that risk identities—including the identities of analysts who produce forecasts—are constructed and contested through these discursive processes. Applied to the ASSESSOR context, this framework suggests that the authority accorded to pretest market forecasts within product development teams is not simply a function of the model's technical performance but is also shaped by organizational power dynamics, professional identities, and institutionalized risk management norms (Hardy & Maguire, 2015). Understanding these social dimensions is essential for explaining why well-validated models may be systematically over- or under-weighted in practice.

Model transparency and documentation standards have emerged as critical governance concerns in multiple applied modeling domains. Cao et al. (2016) document how energy scenario studies frequently lack sufficient transparency in model assumptions, data sources, and parameterization choices, undermining their credibility for policy decision-making. They propose a transparency checklist as a practical instrument for improving model documentation standards. This critique and prescription apply with equal force to pretest market modeling, where the proprietary nature of commercial ASSESSOR applications has historically limited external scrutiny of model specifications and validation evidence (Cao et al., 2016). Greater transparency in model documentation would enhance both scientific credibility and managerial confidence in forecast outputs.

The pharmaceutical industry's development of good practices for model-informed decision-making represents the most formalized existing framework for governing complex quantitative models within organizational decision processes (Marshall et al., 2015). The key elements of this framework—systematic documentation of model purpose, assumptions, data, validation, and uncertainty—provide a directly transferable template for the governance of pretest market model applications. The adoption of analogous good practice standards in marketing research would represent a significant advancement in the professionalization of new product forecasting (Marshall et al., 2015).

The communication of forecast uncertainty to non-technical decision-makers represents a persistent challenge that has attracted innovative methodological responses. Shafer's (2021) proposal for evidence-based statistical communication, framed around betting metaphors, offers a practically accessible framework for conveying the strength and limitations of forecast evidence without requiring statistical expertise from the recipient. In the organizational context of new product launch decisions—where product managers, finance directors, and senior executives must jointly interpret ASSESSOR outputs—such communication innovations could substantially improve the quality of informed decision-making (Shafer, 2021). The development

of visualization and communication tools tailored to pretest market forecast outputs represents an underinvested area of applied marketing research.

The governance of digital data-driven models introduces additional complexity in terms of privacy, representativeness, and algorithmic accountability. Al-Rahmi et al. (2019) examine the adoption of big data technologies in organizational settings, finding that perceived usefulness, ease of use, and risk perception jointly determine adoption behavior. These findings suggest that even technically superior data-driven extensions of pretest market models may face significant organizational resistance if they are perceived as complex, opaque, or threatening to established professional roles. Managing the organizational adoption of methodological innovations is therefore as important as developing the innovations themselves (Al-Rahmi et al., 2019).

The intersection of forecasting model governance with emerging digital technologies, including artificial intelligence platforms, introduces further considerations around model explainability and professional responsibility. Rudolph et al. (2024) examine the paradoxical implications of generative AI adoption in professional practice, noting that the enthusiasm for AI-driven tools often outpaces the institutional capacity to govern their use responsibly. In the context of new product forecasting, the integration of large language models or other generative AI tools into the research and analysis pipeline raises questions about the reliability, traceability, and interpretability of outputs—questions that existing governance frameworks may be inadequately equipped to address (Rudolph et al., 2024). Yigci et al. (2024) similarly emphasize that the deployment of AI-based analytical tools requires careful institutional governance to ensure that outputs are appropriately validated and that professional judgment is not improperly displaced.

4. Research Gaps and Future Directions

Despite its established empirical track record and theoretical coherence, the ASSESSOR model and the broader domain of pretest market forecasting exhibit several substantive research gaps that warrant systematic investigation.

First, the integration of digital behavioral data with traditional laboratory-based research designs represents a critical and underexplored methodological frontier. Contemporary consumers generate vast quantities of behavioral data through online search, social media engagement, review platforms, and e-commerce interactions, all of which carry information relevant to trial likelihood, preference formation, and word-of-mouth dynamics. Yet the ASSESSOR framework has not been systematically adapted to incorporate these data streams, leaving a significant gap between the richness of available consumer behavioral data and the data inputs actually used in pretest market forecasting. Future research should develop and validate hybrid research designs that combine controlled laboratory preference measurement with digital behavioral signal integration, drawing on multi-scale data integration approaches analogous to those demonstrated in energy modeling contexts (Ali et al., 2020). Establishing the conditions under which digital behavioral data enhances versus degrades forecast accuracy relative to traditional laboratory approaches would represent a valuable contribution.

Second, the explicit modeling and communication of forecast uncertainty in ASSESSOR applications requires fundamental advancement. Current practice predominantly reports point estimates of market share, providing decision-makers with a false sense of precision and masking the distributional uncertainty inherent in all model-based forecasts. Future research should develop probabilistic forecast reporting standards for pretest market models,

potentially drawing on advances in statistical communication such as those proposed by Shafer (2021), as well as on the uncertainty quantification frameworks developed in pharmaceutical model-informed decision-making (Marshall et al., 2015). Empirical studies examining how probabilistic versus deterministic forecast presentation affects managerial decision quality would provide actionable guidance for improving the practical value of ASSESSOR applications.

Third, the organizational and institutional governance of pretest market forecasting models requires systematic scholarly attention. While the technical properties of ASSESSOR have been extensively studied, the social, political, and discursive processes through which forecast outputs are interpreted and acted upon within organizations remain poorly understood. Building on the theoretical framework of Hardy and Maguire (2015), future research should employ qualitative and mixed-methods approaches to examine how organizational risk discourse shapes the use of pretest market forecasts in product launch decisions. Such research would complement the model-focused technical literature with an organizational process perspective, generating insights that are directly applicable to improving the governance of marketing research practice. The development of transparency and documentation standards analogous to those proposed for energy scenario modeling (Cao et al., 2016) and pharmaceutical decision modeling (Marshall et al., 2015) represents a practically tractable research agenda with significant implications for both scientific credibility and managerial accountability.

Fourth, the adaptation of agent-based and simulation-based modeling approaches to the pretest market context offers considerable theoretical and practical promise that has not been systematically exploited. The demonstrated success of agent-based frameworks in electricity market simulation (Fraunholz et al., 2021) and ecosystem-based fisheries management (Howell et al., 2021) suggests that analogous architectures could capture competitive market dynamics that the original ASSESSOR formulation treats as exogenous. Future research should develop and validate agent-based extensions of pretest market models that explicitly simulate competitive response, retailer behavior, and word-of-mouth diffusion, testing whether such extensions improve forecast accuracy in categories characterized by high competitive turbulence.

5. Conclusion

This literature review has provided a comprehensive and critical synthesis of the theoretical foundations, empirical performance, methodological challenges, and organizational dimensions of the ASSESSOR model in new product forecasting. Three major thematic clusters were identified and analyzed: the methodological foundations and predictive validity of the model, the opportunities and challenges associated with extending the model through hybrid and data-driven approaches, and the organizational governance and communication dimensions of forecasting model deployment.

The review's principal finding is that the ASSESSOR model retains considerable theoretical coherence and empirical credibility as a pretest market forecasting instrument, but that its continued relevance in contemporary marketing environments depends on systematic methodological renewal. The model's dual-submodel architecture, combining trial-repeat stochastic modeling with multiattribute preference measurement, reflects a principled approach to forecast triangulation that remains epistemologically sound. However, its adaptation to incorporate digital behavioral data, explicit uncertainty quantification, and competitive market dynamics represents an imperative rather than an option in the current data environment.

The cross-disciplinary perspective adopted in this review has proven particularly generative, revealing that methodological innovations in pharmaceutical modeling, energy forecasting, fisheries management, and statistical communication offer directly transferable insights for advancing pretest market methodology. The formalization of good practice documentation standards, the development of probabilistic forecast communication frameworks, and the integration of agent-based simulation represent three concrete pathways for methodological advancement that future research should pursue.

This review is subject to several limitations. The reliance on available sources, which do not include the complete corpus of the ASSESSOR literature itself, constrains the depth of engagement with specific empirical validation studies. The cross-disciplinary scope of the review, while enriching for theoretical development, necessarily sacrifices granularity in domain-specific technical detail. Future reviews should engage more comprehensively with the primary ASSESSOR literature and with the marketing science forecasting literature more broadly.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Al-Rahmi, W. M., Yahaya, N., Aldraiweesh, A., & Alturki, U. (2019). Big data adoption and knowledge management sharing: An empirical investigation on their adoption and sustainability as a purpose of education. IEEE Access, 7, 47245–47258. https://doi.org/10.1109/access.2019.2906668

Alharbi, F. R., & Csala, D. (2022). A Seasonal Autoregressive Integrated Moving Average with Exogenous Factors (SARIMAX) forecasting model-based time series approach. Inventions, 7(4), 94. https://doi.org/10.3390/inventions7040094

Ali, U., Shamsi, M. H., Bohacek, M., Purcell, K., & Hoare, C. (2020). A data-driven approach for multi-scale GIS-based building energy modeling for analysis, planning and support decision making. Applied Energy, 279, 115834. https://doi.org/10.1016/j.apenergy.2020.115834

Cao, K.-K., Cebulla, F., Gómez Vilchez, J. J., Mousavi, B., & Prehofer, S. (2016). Raising awareness in model-based energy scenario studies—a transparency checklist. Energy, Sustainability and Society, 6(1), 1–12. https://doi.org/10.1186/s13705-016-0090-z

Dung, D. H., Quang, N. M., & Chi, B. H. (2022). A Markov chain model for predicting brand switching behavior toward online food delivery services. In Studies in systems, decision and control (pp. 781–797). https://doi.org/10.1007/978-3-030-98689-6_51

Fraunholz, C., Kraft, E., Keles, D., & Fichtner, W. (2021). Advanced price forecasting in agent-based electricity market simulation. Applied Energy, 290, 116688. https://doi.org/10.1016/J.APENERGY.2021.116688

Hardy, C., & Maguire, S. (2015). Organizing risk: Discourse, power, and "riskification." Academy of Management Review, 41(1), 80–108. https://doi.org/10.5465/amr.2013.0106

Howell, D., Schueller, A. M., Bentley, J. W., Buchheister, A., & Chang, D. (2021). Combining ecosystem and single-species modeling to provide ecosystem-based fisheries management

advice within current management systems. Frontiers in Marine Science, 7, 607831. https://doi.org/10.3389/fmars.2020.607831

Lancaster, K. J. (1966). A new approach to consumer theory. Journal of Political Economy, 74(2), 132–157. https://doi.org/10.1086/259131

Lim, H., & Zhai, Z. (2017). Review on stochastic modeling methods for building stock energy prediction. Building Simulation, 10(5), 607–624. https://doi.org/10.1007/s12273-017-0383-y

Marshall, S. F., Burghaus, R., Cosson, V., Cheung, S. Y. A., & Chenel, M. (2015). Good practices in model-informed drug discovery and development: Practice, application, and documentation. CPT: Pharmacometrics & Systems Pharmacology, 5(3), 93–122. https://doi.org/10.1002/psp4.12049

Rudolph, J., Mohamed Ismail, F. M., & Popenici, Ş. (2024). Higher education's generative artificial intelligence paradox: The meaning of chatbot mania. Journal of University Teaching and Learning Practice, 21(2), 1–18. https://doi.org/10.53761/54fs5e77

Silk, A. J., & Urban, G. L. (1978). Pre-Test-Market Evaluation of new Packaged Goods: A model and measurement Methodology. Journal of Marketing Research, 15(2), 171–191. https://doi.org/10.1177/002224377801500201

Shafer, G. (2021). Testing by betting: A strategy for statistical and scientific communication. Journal of the Royal Statistical Society Series A (Statistics in Society), 184(2), 407–431. https://doi.org/10.1111/rssa.12647

Yigci, D., Eryılmaz, M., Yetisen, A. K., Tasoglu, S., & Özcan, A. (2024). Large language model-based chatbots in higher education. Advanced Intelligent Systems, 6(9), 2400429. https://doi.org/10.1002/aisy.202400429

Technology Diffusion in Emerging Economies: Patterns, Drivers, and Future Trajectories

Banda Federico [1] and Štetka Peter [2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: federico.banda@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: peter.stetka@euba.sk

Abstract

The diffusion of technology in emerging economies represents one of the most consequential and complex phenomena shaping global economic development in the twenty-first century. Despite considerable scholarly attention to technology adoption in advanced economies, the specific mechanisms, barriers, and outcomes associated with diffusion processes in developing contexts remain insufficiently theorized. This review examines the intersection of digital transformation, artificial intelligence, industrial automation, and sectoral innovation as they manifest across emerging market settings. Drawing on fifteen peer-reviewed sources spanning economics, information systems, energy, and management disciplines, the review synthesizes evidence across three thematic clusters: digital economic infrastructure and cost reduction, automation and labor market reconfiguration, and sectoral technology adoption in healthcare, agriculture, and manufacturing. Key findings indicate that while digital technologies reduce transaction costs and enable leapfrogging, significant governance, ethical, and infrastructural gaps constrain equitable diffusion. Automation introduces both displacement and reinstatement effects on labor that are particularly pronounced in economies with large informal sectors. Future research should prioritize contextually sensitive frameworks that account for institutional heterogeneity and regulatory capacity in emerging markets.

Key Words: technology diffusion, emerging economies, ICT adoption, digital divide, innovation spillovers

JEL Classification: O33, O11, F63

1. Introduction

The question of how technologies spread across economic and geographic boundaries has occupied scholars in economics, management, and development studies for several decades. However, the accelerating pace of innovation in artificial intelligence, digital platforms, clean energy, and advanced manufacturing has fundamentally altered the landscape of technology diffusion, introducing new dynamics that existing theoretical frameworks only partially address. In emerging economies—broadly defined as nations undergoing rapid industrialization and integration into global markets—these dynamics carry especially high stakes, as the capacity to absorb, adapt, and deploy new technologies is closely linked to trajectories of growth, employment, and social welfare.

A central paradox characterizes the contemporary diffusion landscape in emerging economies: on one hand, digital technologies dramatically lower the cost of accessing

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

information, conducting transactions, and deploying productive capabilities, offering pathways for accelerated catch-up growth (Goldfarb & Tucker, 2019). On the other hand, structural conditions including weak institutional frameworks, infrastructure deficits, skills mismatches, and regulatory uncertainty continue to impede the translation of technological availability into productive use (Dwivedi et al., 2019). This tension between enabling potential and constraining reality constitutes the core research problem motivating the present review.

Despite a growing body of empirical and theoretical work on individual technology domains—ranging from artificial intelligence and Industry 4.0 to digital health and precision agriculture—there is a notable absence of integrative scholarship that synthesizes these disparate findings through the lens of technology diffusion in emerging economies. Most existing reviews are either narrowly sectoral or focused on developed economy contexts, leaving practitioners and policymakers in emerging markets without a consolidated understanding of the forces shaping technological change in their environments.

This literature review addresses that gap directly. The objective is twofold: first, to synthesize and critically evaluate the existing evidence on how technologies diffuse across sectors relevant to emerging economies, and second, to identify the most pressing research gaps and productive directions for future inquiry. The review draws on fifteen high-impact academic sources spanning the period 2017 to 2023, covering domains including digital economics, automation, AI governance, healthcare digitization, smart manufacturing, and precision agriculture.

The remainder of this paper is organized as follows. Section 2 establishes the theoretical background by defining key constructs and reviewing foundational frameworks relevant to technology diffusion. Section 3 presents a thematic synthesis of the literature, organized around three empirically grounded clusters. Section 4 identifies research gaps and articulates directions for future investigation. Section 5 concludes by summarizing the review's contributions, restating its principal findings, and acknowledging its limitations.

2. Theoretical Background

2.1 Conceptualizing Technology Diffusion

Technology diffusion refers to the process by which an innovation is communicated through certain channels over time among members of a social or economic system. Classical diffusion theory, associated most prominently with Rogers's work, frames adoption as a function of perceived attributes of the innovation, the nature of communication channels, time, and the social system in question. In the context of emerging economies, this framework requires significant augmentation to account for the role of institutional environments, infrastructure endowments, and the particular dynamics of catch-up industrialization.

A foundational extension of diffusion theory relevant to emerging economies is the concept of leapfrogging, whereby late-adopting nations bypass intermediate technological stages and adopt frontier technologies directly. This phenomenon has been observed in mobile telecommunications, where many developing nations moved directly to wireless networks without building extensive landline infrastructure, and is increasingly anticipated in energy, where the prospect of bypassing fossil fuel systems in favor of renewables has gained credence (Staffell et al., 2018; Noussan et al., 2020). However, the conditions under which genuine leapfrogging occurs, as opposed to mere adoption of legacy technologies under a different name, remain contested in the literature.

2.2 Digital Economics as a Framework

One of the most theoretically productive frameworks for understanding contemporary technology diffusion is digital economics, which examines how the structural properties of digital information—its near-zero marginal cost of reproduction, its transportability, and its searchability—reshape economic activity. Goldfarb and Tucker (2019) provide a systematic treatment of this framework, identifying reductions in five distinct cost categories as the primary mechanisms through which digital technology transforms markets: search costs, replication costs, transportation costs, tracking costs, and verification costs. These cost reductions are particularly significant for emerging economies, where high transaction costs have historically constrained market development, limited access to information, and suppressed productive investment.

The digital economics framework also illuminates why technology diffusion in emerging economies does not follow a simple trajectory of adoption. Reducing search costs, for instance, may advantage well-resourced actors disproportionately if baseline digital literacy is unevenly distributed. Similarly, the reduction in replication costs enables the spread of both productive knowledge and harmful misinformation, with governance implications that are especially acute where regulatory institutions are nascent (Dwivedi et al., 2019).

2.3 Automation, Tasks, and Labor Market Theory

The task-based framework developed by Acemoglu and Restrepo (2019) offers a rigorous theoretical foundation for understanding how automation affects labor markets, with specific implications for emerging economies characterized by labor abundance. Their model disaggregates the economy into a continuum of tasks, some of which become automatable as technology advances. The key insight is that automation does not uniformly destroy employment; rather, it displaces labor from automated tasks while potentially reinstating labor demand through the creation of new tasks and the productivity-driven expansion of economic activity. The net effect on employment and wages depends critically on the relative magnitudes of displacement and reinstatement effects, which vary across institutional and developmental contexts.

2.4 Technology Acceptance and Institutional Frameworks

At the firm and individual level, technology acceptance models provide micro-foundational accounts of diffusion. Stoumpos et al. (2023) demonstrate in the healthcare context that perceived usefulness, ease of use, and institutional support are significant determinants of technology adoption, with trust emerging as a particularly important mediating variable. In emerging economy settings, where institutional trust may be lower and digital literacy more heterogeneous, these micro-level factors interact with macro-level governance conditions to produce diffusion outcomes that diverge substantially from those observed in advanced economies.

The governance dimension is further theorized by Díaz-Rodríguez et al. (2023), whose framework for trustworthy artificial intelligence emphasizes that responsible technology deployment requires attention to lawfulness, ethics, and technical robustness simultaneously. In contexts where regulatory frameworks are still developing, the absence of any one of these pillars can compromise the entire diffusion process, leading to adoption without the institutional scaffolding necessary for sustainable, equitable outcomes.

2.5 Industrial and Sectoral Perspectives

Beyond economy-wide frameworks, technology diffusion in emerging economies is necessarily mediated by sectoral conditions. Büchi et al. (2019) demonstrate that in manufacturing, the benefits of Industry 4.0 adoption—including smart factory capabilities—depend not only on the breadth of technologies deployed but also on the depth of their integration across value chains. This insight suggests that diffusion is not merely a binary adoption decision but a continuous process of organizational and technological reconfiguration. Similarly, Li et al. (2021) highlight that ESG frameworks increasingly shape the conditions under which technologies are financed and deployed, creating new governance pressures that intersect with traditional economic drivers of diffusion.

3. Literature Review

3.1 Theme One: Digital Infrastructure, Cost Reduction, and Market Transformation

The first and perhaps most pervasive theme in the literature concerns the role of digital infrastructure as a foundational enabler of technology diffusion. Goldfarb and Tucker (2019) establish that the economic logic of digitization rests on the simultaneous compression of multiple cost categories, and that these reductions have cumulative, nonlinear effects on market structure. In emerging economies, where informational asymmetries and high transaction costs have long impeded market integration, the deployment of digital platforms and communication infrastructure carries transformative potential that extends beyond productivity gains to encompass institutional change.

Dwivedi et al. (2019) situate artificial intelligence as the most consequential current manifestation of this digital transformation, observing that AI's capacity to augment and potentially substitute human cognitive tasks across industrial, intellectual, and social domains represents a qualitative shift in the nature of technology diffusion. Unlike earlier waves of automation that primarily affected physical tasks, AI enables automation of complex information-processing activities, thereby engaging directly with the knowledge-intensive sectors—finance, healthcare, government—that constitute growing shares of employment in many emerging economies. The implications are profound: diffusion of AI is not merely a sectoral phenomenon but a systemic one that restructures labor markets, business models, and public services simultaneously.

The digital health literature illustrates this dynamic concretely. Stoumpos et al. (2023) document how the integration of wearable technologies, virtual reality, and the Internet of Things into healthcare systems is reshaping patient-provider relationships and enabling new models of care delivery. Their systematic review finds that technology acceptance in healthcare is strongly conditioned by perceived trust, institutional support, and the quality of user experience, factors that interact differently in low- and middle-income country settings than in high-income contexts. In emerging economies, where healthcare infrastructure deficits are often severe, digital health technologies offer possibilities for bypassing traditional facility-based care models—a form of leapfrogging that is rendered more plausible by the cost-reduction dynamics identified by Goldfarb and Tucker (2019).

However, the literature also surfaces important tensions. The same digital infrastructure that enables cost reduction and market access can entrench existing inequalities if access is uneven. Dwivedi et al. (2019) note that the pace of AI advancement creates risks of skill displacement and social disruption that outpace the capacity of regulatory institutions to respond, a challenge that is particularly acute in emerging economy settings where regulatory capacity is

already stretched. This observation aligns with the governance framework proposed by Díaz-Rodríguez et al. (2023), who argue that trustworthy AI deployment requires proactive institutional design rather than reactive regulation, with attention to ethical principles, technical robustness, and legal accountability from the outset of diffusion processes.

3.2 Theme Two: Automation, Industrial Transformation, and Labor Market Reconfiguration

The second thematic cluster addresses the interaction between technology diffusion and labor market outcomes, a subject of considerable debate in the economics literature and of particular salience for emerging economies with large labor forces and significant informal employment. Acemoglu and Restrepo (2019) provide the most theoretically rigorous treatment of this question, demonstrating through their task-based framework that automation's effect on employment is not deterministic but contingent on the relative pace of task displacement and task creation. Their analysis of the United States reveals that the labor share of income has declined substantially due to automation, driven primarily by displacement effects that have not been fully offset by reinstatement through new task creation.

For emerging economies, this analysis carries a nuanced implication. On one hand, abundant labor supply means that labor-displacing automation may generate more severe distributional consequences, particularly for low-skill workers in manufacturing sectors facing competition from automated production systems in advanced economies. On the other hand, the reinstatement effects identified by Acemoglu and Restrepo (2019)—whereby new tasks are created that require human labor in conjunction with new technologies—may offer pathways for employment growth if educational and training systems can supply the requisite skills. The challenge is that these reinstatement dynamics operate on timescales that may not align with the immediate social pressures created by displacement.

Büchi et al. (2019) contribute empirical texture to this discussion through their analysis of smart factory performance and Industry 4.0 adoption. Their finding that performance gains depend not merely on the number of technologies adopted but on the depth of their integration across value chains suggests that technology diffusion in manufacturing is a cumulative, path-dependent process. Firms that adopt individual technologies without reconfiguring their broader organizational and supply chain practices may realize only partial benefits, while those that achieve deep integration gain compounding advantages. In emerging economy contexts, where supplier networks and logistical infrastructure are often less developed, achieving the deep integration associated with Industry 4.0 benefits may be more difficult, potentially limiting the transferability of findings from advanced economy settings.

The energy dimension of industrial transformation further complicates this picture. Staffell et al. (2018) present hydrogen and fuel cells as potentially transformative technologies for decarbonizing industrial energy systems, with implications for the competitive landscape of energy-intensive industries. For emerging economies seeking to industrialize while managing climate commitments, the diffusion of clean energy technologies—whether hydrogen-based, battery-electric, or renewable—represents both an opportunity and a challenge. Zeng et al. (2019) document the commercialization trajectory of lithium battery technologies, noting that while cost reductions have been dramatic, safety, energy density, and supply chain constraints continue to shape the pace and geography of adoption. Similarly, Noussan et al. (2020) analyze the geopolitical dimensions of hydrogen's role in energy transition, observing that the distribution of benefits from green hydrogen deployment will depend heavily on which nations

control key resources and technologies—a consideration with direct implications for emerging economies endowed with renewable energy potential.

ESG frameworks are increasingly mediating the conditions under which industrial technology diffusion proceeds. Li et al. (2021) demonstrate that ESG considerations are becoming integrated into investment decisions across global capital markets, creating incentives for firms in emerging economies to adopt sustainability-compatible technologies in order to maintain access to international financing. This dynamic can accelerate the diffusion of clean technologies while simultaneously raising compliance costs for smaller firms, potentially exacerbating intra-economy inequality in access to capital and technology.

3.3 Theme Three: Sectoral Adoption—Healthcare, Agriculture, and Education

The third thematic cluster examines technology diffusion within specific sectors that are central to development outcomes in emerging economies: healthcare, agriculture, and education. These sectors share a common characteristic of being simultaneously critical for human development and notoriously resistant to rapid technological change due to institutional inertia, professional gatekeeping, and the complexity of scaling innovations across heterogeneous populations.

In healthcare, the digital transformation documented by Stoumpos et al. (2023) encompasses a broad spectrum of technologies, from electronic health records and telemedicine platforms to AI-driven diagnostics and wearable monitoring devices. Their analysis reveals that technology acceptance in this sector is strongly moderated by organizational culture, patient literacy, and regulatory environment—factors that vary considerably across emerging economy contexts. The potential for AI-driven diagnostics to extend specialist-level care to underserved populations is frequently cited as a key benefit of digital health diffusion, and this potential aligns with the broader framework of cost reduction articulated by Goldfarb and Tucker (2019). However, the trustworthiness requirements identified by Díaz-Rodríguez et al. (2023) are particularly pressing in healthcare, where errors in AI-driven diagnosis can have life-or-death consequences and where patients from vulnerable populations may have limited capacity to contest algorithmic decisions.

In agriculture, precision farming technologies represent a potentially transformative set of innovations for emerging economies where agricultural productivity remains central to food security and livelihoods. Karunathilake et al. (2023) provide a comprehensive review of precision agriculture innovations, including IoT-enabled soil monitoring, drone-based crop surveillance, data-driven irrigation management, and machine learning applications for yield prediction. Their findings indicate that these technologies can substantially improve productivity while reducing environmental impacts, offering a pathway toward sustainable intensification that is highly relevant for emerging economies facing population growth and climate stress simultaneously. However, adoption barriers are significant, including high capital costs, limited connectivity infrastructure, inadequate technical training, and the small farm sizes that characterize agriculture in many developing regions. The cumulative nature of diffusion barriers means that individual interventions—such as subsidizing sensor technology without addressing connectivity or training—are unlikely to achieve transformative outcomes.

Education represents a third critical sector where technology diffusion dynamics have been extensively studied. Serdyukov (2017) argues that educational innovation in the United States has been persistently hampered by the gap between promising pilot outcomes and scalable implementation, a finding that resonates strongly with emerging economy contexts where the

gap between innovation and scale is often even wider. The organizational and cultural barriers to educational technology adoption identified by Serdyukov (2017) include institutional conservatism, insufficient teacher training, misalignment between technological tools and pedagogical objectives, and inadequate evidence of learning outcomes at scale. These barriers are compounded in many emerging economies by infrastructure limitations and the linguistic and cultural adaptation requirements for technologies developed in high-income country contexts.

Across all three sectors, a consistent finding emerges: the realization of technology's diffusion potential depends less on the properties of the technology itself than on the institutional, organizational, and social conditions of its deployment. This finding underscores the importance of context-sensitive frameworks for understanding and promoting technology diffusion in emerging economies, and it highlights the limitations of approaches that treat diffusion as a purely technical or market-driven process.

4. Research Gaps and Future Directions

Despite the breadth and rigor of the literature reviewed, several significant gaps emerge that constrain the development of an integrated, contextually grounded understanding of technology diffusion in emerging economies.

The first and most fundamental gap concerns the scarcity of empirically grounded, multi-sectoral studies conducted within emerging economy settings. The majority of high-citation empirical work—including the automation studies of Acemoglu and Restrepo (2019), the smart factory analyses of Büchi et al. (2019), and the digital health assessments of Stoumpos et al. (2023)—draws primarily on data from advanced economies. While the theoretical frameworks developed in these studies have considerable generalizability, their empirical instantiations may reflect institutional and infrastructural conditions that do not transfer to emerging market contexts. Future research should prioritize the development of large-scale, longitudinal datasets from emerging economy settings that permit rigorous testing of whether and how diffusion dynamics differ across developmental contexts.

A second gap concerns the governance and regulatory dimensions of technology diffusion. Díaz-Rodríguez et al. (2023) provide an important framework for trustworthy AI that encompasses legal, ethical, and technical dimensions, and Dwivedi et al. (2019) highlight the policy challenges raised by rapid AI advancement. However, neither framework is explicitly calibrated to the institutional realities of emerging economies, where regulatory capacity may be limited, legal frameworks underdeveloped, and enforcement mechanisms weak. Future research should investigate how regulatory architectures in emerging economies can be designed to promote technology adoption while managing associated risks, with particular attention to the role of international governance frameworks and technology transfer mechanisms. The intersection of ESG requirements documented by Li et al. (2021) with regulatory development in emerging markets represents a particularly underexplored area.

A third gap involves the distributional consequences of technology diffusion within emerging economies. The displacement and reinstatement dynamics identified by Acemoglu and Restrepo (2019) generate aggregate effects that mask substantial heterogeneity across skill levels, geographic regions, and demographic groups. In emerging economies with high income inequality and large informal sectors, the distributional implications of automation and digitization may be especially severe. Future research should develop distributional analyses of technology diffusion that go beyond aggregate productivity and employment effects to

examine outcomes across the full distribution of workers and firms, with particular attention to gender, rural-urban divides, and the informal sector. Such analyses would have direct implications for the design of social protection and workforce development policies in rapidly automating economies.

A fourth, cross-cutting future direction involves the integration of energy transition and digital transformation research streams. The literature on hydrogen (Staffell et al., 2018; Noussan et al., 2020), battery technologies (Zeng et al., 2019), and digital economics (Goldfarb & Tucker, 2019) currently proceeds largely in parallel rather than in dialogue. For emerging economies seeking to achieve both industrialization and decarbonization goals, understanding how digital and energy transitions interact—whether they are complementary or competitive, sequential or simultaneous—is essential for policy design. Integrated frameworks that connect energy system transitions with digital transformation and labor market adjustment would represent a significant contribution to both scholarly and policy communities.

5. Conclusion

This review has synthesized a substantial body of literature bearing on technology diffusion in emerging economies, organizing findings around three thematic clusters: digital economic infrastructure and market transformation, industrial automation and labor market reconfiguration, and sectoral adoption in healthcare, agriculture, and education. Across these themes, a set of consistent and consequential findings emerges.

First, digital technologies—particularly artificial intelligence and platform-based systems—create cost reduction dynamics with genuine transformative potential for emerging economies, lowering barriers to market access and enabling new service delivery models (Goldfarb & Tucker, 2019; Dwivedi et al., 2019). However, the realization of this potential is contingent on institutional conditions that are frequently absent or underdeveloped in emerging market settings, including robust regulatory frameworks, adequate digital literacy, and trustworthy governance of AI systems (Díaz-Rodríguez et al., 2023; Stoumpos et al., 2023).

Second, automation and industrial transformation generate complex labor market effects that are not predetermined but depend critically on the balance between task displacement and task creation, as well as on the organizational and supply chain integration strategies that firms pursue (Acemoglu & Restrepo, 2019; Büchi et al., 2019). In emerging economies, these dynamics interact with existing structural features—labor abundance, large informal sectors, and developing educational systems—to produce outcomes that may diverge substantially from advanced economy experiences.

Third, sectoral diffusion in healthcare, agriculture, and education reveals a common pattern: technology availability is a necessary but insufficient condition for transformative adoption, with institutional, organizational, and social factors exercising decisive influence (Stoumpos et al., 2023; Karunathilake et al., 2023; Serdyukov, 2017).

This review is subject to limitations. The source base, while covering a diverse range of disciplines, skews toward advanced economy empirical contexts and does not include primary studies conducted exclusively in emerging market settings. Additionally, the review's thematic structure, while analytically useful, necessarily compresses the heterogeneity within and across the emerging economy category. Future reviews should incorporate more systematic search protocols and explicitly seek literature originating from low- and middle-income country research communities.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. Journal of Economic Perspectives, 33(2), 3–30. https://doi.org/10.1257/jep.33.2.3

Boyd, C. E., D'Abramo, L. R., Glencross, B. D., Huyben, D. C., & Rombenso, L. (2020). Achieving sustainable aquaculture: Historical and current perspectives and future needs and challenges. Journal of the World Aquaculture Society, 51(3), 578–633. https://doi.org/10.1111/jwas.12714

Büchi, G., Cugno, M., & Castagnoli, R. (2019). Smart factory performance and Industry 4.0. Technological Forecasting and Social Change, 150, 119790. https://doi.org/10.1016/j.techfore.2019.119790

Chapman, J., Ismail, A. E., & Dinu, C. Z. (2018). Industrial applications of enzymes: Recent advances, techniques, and outlooks. Catalysts, 8(6), 238. https://doi.org/10.3390/catal8060238

Díaz-Rodríguez, N., Del Ser, J., Coeckelbergh, M., López de Prado, M., Herrera-Viedma, E., & Herrera, F. (2023). Connecting the dots in trustworthy artificial intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. Information Fusion, 99, 101896. https://doi.org/10.1016/j.inffus.2023.101896

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., & Coombs, C. (2019). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Goldfarb, A., & Tucker, C. E. (2019). Digital economics. Journal of Economic Literature, 57(1), 3–43. https://doi.org/10.1257/jel.20171452

Karunathilake, E. M. B. M., Le, A. T., Heo, S., Chung, Y. S., & Mansoor, S. (2023). The path to smart farming: Innovations and opportunities in precision agriculture. Agriculture, 13(8), 1593. https://doi.org/10.3390/agriculture13081593

Li, T., Wang, K., Sueyoshi, T., & Wang, D. (2021). ESG: Research progress and future prospects. Sustainability, 13(21), 11663. https://doi.org/10.3390/su132111663

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., & Corker, K. S. (2021). Replicability, robustness, and reproducibility in psychological science. Annual Review of Psychology, 73, 719–748. https://doi.org/10.1146/annurev-psych-020821-114157

Noussan, M., Raimondi, P. P., Scita, R., & Häfner, M. (2020). The role of green and blue hydrogen in the energy transition—A technological and geopolitical perspective. Sustainability, 13(1), 298. https://doi.org/10.3390/su13010298

Serdyukov, P. (2017). Innovation in education: What works, what doesn't, and what to do about it? Journal of Research in Innovative Teaching & Learning, 10(1), 4–33. https://doi.org/10.1108/jrit-10-2016-0007

Staffell, I., Scamman, D., Velazquez Abad, A., Balcombe, P., & Dodds, P. E. (2018). The role of hydrogen and fuel cells in the global energy system. Energy & Environmental Science, 12(2), 463–491. https://doi.org/10.1039/c8ee01157e

Stoumpos, A. I., Kitsios, F., & Talias, M. A. (2023). Digital transformation in healthcare: Technology acceptance and its applications. International Journal of Environmental Research and Public Health, 20(4), 3407. https://doi.org/10.3390/ijerph20043407

Zeng, X., Li, M., Abd El-Hady, D., Alshitari, W., Al-Bogami, A. S., Lu, J., & Amine, K. (2019). Commercialization of lithium battery technologies for electric vehicles. Advanced Energy Materials, 9(22), 1900161. https://doi.org/10.1002/aenm.201900161

AI Adoption and Firm Financial Performance: A Thematic Literature Review

Šagátová Slavka [1] and Grisáková Nora [2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: slavka.sagatova@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: nora.grisakova@euba.sk

Abstract

The growing adoption of artificial intelligence in business practice has intensified interest in its implications for firm financial performance, yet existing research remains fragmented across disciplines and empirical contexts. This article provides a structured synthesis of the literature to clarify how and under what conditions artificial intelligence contributes to financial outcomes at the firm level. Using a thematic literature review of 23 peer-reviewed studies, the analysis identifies three interrelated domains that explain observed performance effects: operational efficiency mechanisms, capability-based moderating factors, and sectoral as well as contextual influences. The findings indicate that artificial intelligence functions primarily as an enabling resource whose impact is mediated through organizational processes and contingent on the presence of complementary capabilities. While efficiency improvements and productivity gains are frequently observed, their translation into financial performance is uneven and depends on firm specific and environmental conditions. The study contributes by integrating dispersed insights into a coherent analytical framework and by identifying key directions for future research.

Key Words: AI adoption, firm financial performance, organizational capabilities, digital transformation, productivity

JEL Classification: O33, M15, L25

1. Introduction

The rapid diffusion of artificial intelligence across industries has intensified scholarly and managerial interest in its economic consequences, particularly with respect to firm financial performance. Organizations increasingly invest in AI-driven technologies to enhance decision making, optimize operations, and develop new sources of competitive advantage. Despite these developments, the empirical evidence on whether AI adoption systematically translates into improved financial outcomes remains inconclusive. While a growing body of research suggests that AI can enhance productivity and operational efficiency, the extent to which these

improvements are converted into financial performance is highly variable across firms and contexts (Brynjolfsson et al., 2021; Garrido-Moreno et al., 2024).

In light of these considerations, it can be concluded that the financial impact of artificial intelligence adoption is contingent upon a set of mediating and moderating factors and is not universally positive. Organizational capabilities, including innovation capacity, human capital development, and resilience, play a central role in determining whether AI investments yield financial benefits. Similarly, firm size, industry characteristics, and the broader digital ecosystem within which AI is deployed influence both the magnitude and the direction of performance effects (Garrido-Moreno et al., 2024; Cardinali & De Giovanni, 2022). These insights suggest that AI driven performance outcomes are best understood through a configurational perspective that emphasizes complementarities among technological, organizational, and contextual factors. Against this background, the present study aims to provide a structured synthesis of the literature on AI adoption and firm financial performance. Specifically, it seeks to consolidate existing empirical and conceptual findings, identify the key mechanisms through which AI influences financial outcomes, and clarify the conditions under which these effects are realized.

2. Theoretical Background

Understanding the relationship between artificial intelligence adoption and firm financial performance requires a theoretically grounded perspective that captures both the nature of the technology and the organizational conditions under which it generates value. Rather than approaching AI as a standalone technological input, contemporary research increasingly emphasizes its role within broader systems of resources, capabilities, and strategic processes that shape firm outcomes. This section therefore establishes the conceptual and theoretical foundations necessary for the subsequent analysis by clarifying key constructs and integrating relevant theoretical perspectives. In doing so, it provides a coherent analytical basis for examining how AI adoption influences financial performance through distinct but interrelated mechanisms.

2.1 Conceptualizing AI Adoption and Firm Financial Performance

Artificial intelligence adoption can be understood as a purposeful organizational process through which firms integrate data-driven and algorithmic technologies into their operational, managerial, and strategic activities with the aim of generating value. This process extends beyond mere access to technological tools and involves the development of internal capabilities that enable effective deployment, continuous learning, and alignment with organizational objectives. AI adoption therefore represents not only a technological decision but also a strategic and organizational transformation that reshapes how firms process information and coordinate activities (Buhalis et al., 2022; Kallmuenzer et al., 2024).

From a conceptual perspective, it is important to distinguish between the presence of AI technologies and their productive use. Firms may implement AI systems without achieving measurable improvements in performance if these technologies are not embedded in organizational routines or supported by complementary investments in skills and processes. This distinction reflects a broader perspective in the economics of information technology, as performance improvements arise from the joint interaction of technology and organizational change, rather than from technology alone (Brynjolfsson & McElheran, 2016; Tambe et al., 2012). Consequently, AI adoption should be treated as a multidimensional construct that captures both technological implementation and organizational integration.

Firm financial performance, which represents the primary outcome of interest in article, is typically operationalized using both accounting based and market based indicators. Accounting measures such as return on assets, return on equity, and profit margins reflect internal efficiency and profitability, whereas market based measures such as Tobin's Q and stock returns capture investor expectations and valuation effects. In addition, many studies employ operational performance indicators, including productivity, cost efficiency, and revenue growth, as intermediate outcomes that link AI adoption to financial performance (Chen et al., 2022). This distinction is analytically important because improvements in operational performance may not immediately translate into financial gains, particularly in the presence of adjustment costs, competitive pressures, or redistribution of value across stakeholders.

2.2 Resource Based and Capability Perspectives on AI Driven Value Creation

The resource based view provides a foundational theoretical lens for understanding how AI adoption can contribute to firm financial performance. According to this perspective, firms achieve sustained competitive advantage when they possess resources that are valuable, rare, difficult to imitate, and not easily substitutable. AI related capabilities can meet these criteria when they are embedded in firm specific processes, supported by proprietary data, and integrated with organizational knowledge (Barney, 1991; Garrido-Moreno et al., 2024). However, the resource based view also implies that the mere acquisition of AI technologies is insufficient for value creation, as competitors can often access similar technologies. The differentiating factor lies in how these technologies are deployed and combined with other organizational resources.

Dynamic capabilities theory extends this argument by emphasizing the importance of adaptability in rapidly changing environments. Firms must continuously sense opportunities, seize them through appropriate investments, and reconfigure their resource base to sustain performance advantages over time (Teece, 2007). Within this framework, AI adoption can be interpreted as both an outcome and an enabler of dynamic capabilities. On the one hand, firms require dynamic capabilities to successfully implement AI technologies. On the other hand, AI enhances these capabilities by improving information processing, enabling faster decision

making, and supporting organizational learning. This reciprocal relationship suggests that AI driven performance outcomes depend on the alignment between technological investments and the firm's capacity for strategic adaptation.

A complementary perspective highlights the role of organizational complementarities in technology driven performance. Empirical research on information technologies demonstrates that productivity gains are maximized when firms simultaneously invest in organizational practices such as decentralization, skill development, and process redesign (Tambe et al., 2012). In the context of AI, these complementarities may include investments in human capital, data governance, and cross functional integration. The absence of such complementary factors can limit the effectiveness of AI adoption and contribute to heterogeneous performance outcomes across firms.

2.3 Digital Transformation and the Systemic Nature of AI Integration

AI adoption is increasingly conceptualized as a component of broader digital transformation processes that reshape business models, value chains, and competitive dynamics. Digital transformation involves the integration of digital technologies into all areas of the organization, leading to fundamental changes in how firms create and deliver value. Within this broader context, AI functions as a core enabling technology that enhances the efficiency and responsiveness of organizational processes (Oubrahim et al., 2023).

A key implication of the digital transformation perspective is that the performance effects of AI are systemic rather than isolated. Firms that implement AI as a standalone solution may achieve limited benefits compared to those that integrate AI within coordinated transformation initiatives that include supply chain integration, customer engagement strategies, and data-driven decision making systems. Empirical evidence supports this view by showing that digital transformation initiatives that combine multiple technologies and organizational changes generate stronger performance outcomes than isolated technological investments (Atif et al., 2021).

The systemic nature of AI integration also implies that value creation depends on the interaction between technological infrastructure and organizational processes. Advances such as intelligent Internet of Things systems and data analytics platforms enable firms to collect and process large volumes of data, but the translation of these capabilities into performance improvements requires effective coordination across organizational units (Aouedi et al., 2024). This perspective reinforces the argument that AI driven performance is contingent on the broader digital ecosystem within which the technology is deployed.

2.4 Contextual and Organizational Contingencies

The literature consistently emphasizes that the relationship between AI adoption and financial performance is contingent on a range of contextual and organizational factors. Firm

size represents one of the most significant moderators, as larger firms typically possess greater resources, more advanced technological infrastructures, and higher levels of absorptive capacity, which enable them to derive greater benefits from AI investments. In contrast, small and medium sized enterprises often face constraints related to financial resources, technical expertise, and organizational readiness, which limit their ability to translate AI adoption into performance gains (Kallmuenzer et al., 2024; Grashof & Kopka, 2022).

Organizational resilience and human capital also play a critical role in shaping AI driven performance outcomes. Firms with higher levels of resilience are better able to absorb the disruptions associated with technology adoption and to leverage AI capabilities during periods of environmental uncertainty. Similarly, investments in human capital and organizational learning enhance the effectiveness of AI by enabling employees to interpret and utilize data-driven insights (Malik et al., 2022). These factors highlight the importance of viewing AI adoption as part of a broader organizational capability system rather than as an isolated technological intervention.

In addition to internal organizational factors, external conditions such as industry characteristics, regulatory environments, and competitive dynamics influence the performance implications of AI adoption. Different sectors exhibit varying levels of technological maturity, data availability, and competitive pressure, which affect both the incentives for AI adoption and the potential returns on investment. As a result, empirical findings on AI driven performance must be interpreted within their specific contextual settings, and generalizations across sectors should be made with caution.

3. Metodology

This article adopts a structured literature review approach to synthesize existing research on the relationship between artificial intelligence adoption and firm financial performance. The methodological design follows principles of transparency and analytical coherence while allowing for the integration of heterogeneous studies. Given the diversity of research designs, industries, and performance measures in the existing literature, a thematic synthesis approach was selected as the most appropriate method for identifying consistent patterns and relationships across studies (Tranfield et al., 2003; Snyder, 2019).

The literature search was conducted using major academic databases, including Web of Science and Scopus, with a focus on peer reviewed journal articles published in English. Search terms combined keywords related to artificial intelligence, digital transformation, firm performance, productivity, and financial outcomes. The selection process was guided by predefined inclusion criteria, requiring that studies explicitly address the relationship between AI or closely related digital technologies and firm level performance outcomes, either directly or through identified mechanisms. Studies focusing exclusively on technical aspects without

organizational or economic relevance were excluded. This process resulted in a final sample of twenty empirical studies.

The selected literature was analyzed using an iterative thematic coding procedure aimed at identifying recurring mechanisms and conditions that shape the relationship between AI adoption and financial performance. Following established approaches to thematic analysis, studies were systematically reviewed and grouped into three interrelated domains reflecting the dominant patterns in the literature (Braun & Clarke, 2006, Thomas & Harden, 2008). These domains capture efficiency related effects, capability-based moderating factors, and sectoral as well as contextual conditions influencing performance outcomes.

The methodological approach emphasizes analytical synthesis rather than quantitative aggregation. The objective is to develop a coherent understanding of how AI adoption contributes to firm financial performance under varying organizational and contextual conditions. While the bounded sample and heterogeneity of the included studies limit direct comparability, the applied approach enables the identification of consistent theoretical and empirical insights that inform the subsequent thematic analysis.

4. Literature Review

Building on the established theoretical foundations and the applied methodological approach, this section develops a structured synthesis of the selected literature in order to identify the dominant patterns through which artificial intelligence adoption relates to firm financial performance. Rather than reiterating theoretical arguments, the analysis focuses on consolidating empirical and conceptual findings across studies with the aim of revealing consistent mechanisms, moderating conditions, and contextual influences. The reviewed literature sources are organized into three analytically distinct yet interconnected domains that reflect the primary dimensions of the observed relationships. This structure enables a systematic examination of how artificial intelligence driven transformation manifests in firm outcomes.

4.1 Thematic Domain 1: AI Adoption as a Driver of Operational Efficiency and Financial Performance

A substantial body of literature identifies operational efficiency improvements as the primary mechanism through which artificial intelligence adoption influences firm financial performance. AI-enabled technologies support automation, enhance decision accuracy, and improve the allocation of resources, thereby reducing operational costs and increasing productivity. These efficiency gains are frequently interpreted as the initial pathway through which financial performance improvements may emerge. Empirical studies across multiple sectors provide consistent evidence that digital transformation initiatives incorporating AI

contribute to improved process performance and cost structures, which in turn create conditions for enhanced profitability (Oubrahim et al., 2023; Gupta et al., 2020).

At the same time, the translation of operational improvements into financial outcomes is neither immediate nor uniform. Research in financial services demonstrates that technology driven innovation can simultaneously improve productivity and introduce new forms of risk, leading to mixed financial results in the short term (Mohamed & Elgammal, 2023). This finding highlights the importance of distinguishing between different dimensions of performance and recognizing that efficiency gains may be offset by increased uncertainty, investment costs, or competitive pressures. Similar patterns are observed in other sectors, where AI driven optimization enhances operational indicators but requires complementary adjustments in business models and market positioning to generate sustained financial returns (Buhalism et al., 2022).

The broader economic literature reinforces this interpretation by suggesting that the productivity effects of advanced technologies are often delayed and contingent on organizational adaptation. The concept of a productivity lag implies that firms must undergo a period of learning and restructuring before the financial benefits of technology adoption become fully visible (Brynjolfsson et al., 2021). In this context, AI should be understood as an enabling technology that creates potential for performance improvement rather than as a direct determinant of financial outcomes. This perspective provides a foundation for examining the conditions under which efficiency gains are effectively converted into measurable financial performance.

4.2 Thematic Domain 2: Organizational Capabilities and Firm Level Moderators

An important insight emerging from the literature is that the financial impact of artificial intelligence adoption is strongly mediated by organizational capabilities and firm specific characteristics. The mere presence of artificial intelligence technologies does not guarantee improved performance. Rather, the outcomes of their adoption depend on the firm's ability to integrate these technologies into its processes and to leverage them strategically. Studies consistently demonstrate that firms with higher levels of absorptive capacity, innovation orientation, and organizational resilience derive greater benefits from digital investments (Garrido-Moreno et al., 2024; Kallmuenzer et al., 2024).

Firm size represents a particularly important moderating factor. Large firms typically possess greater financial resources, more advanced technological infrastructures, and more developed managerial capabilities, which enable them to exploit AI technologies more effectively. In contrast, small and medium sized enterprises often face constraints that limit their ability to translate AI adoption into financial gains. Empirical evidence indicates that AI related innovation outcomes are more pronounced in larger firms, suggesting that technology adoption may reinforce existing performance disparities rather than reduce them (Grashof &

Kopka, 2022). This finding underscores the importance of considering structural differences between firms when evaluating the impact of AI on performance.

Human capital and organizational learning further shape the effectiveness of AI adoption. Research on AI-enabled human resource systems shows that improvements in employee experience and engagement contribute to enhanced organizational performance, indicating that the benefits of AI are partly realized through human capital development (Malik et al., 2022). Similarly, the application of data-driven decision making practices requires employees to possess the skills necessary to interpret and act upon analytical insights, reinforcing the role of complementary investments in training and organizational development (Brynjolfsson & McElheran, 2016).

In addition, organizational resilience has been identified as a key mechanism linking digital capabilities to performance outcomes. Firms that are better able to adapt to environmental disruptions and reconfigure their resources are more likely to capture value from the adoption of AI. Evidence from recent studies suggests that resilience enhances the effectiveness of innovation and supports more stable financial performance in uncertain environments (Garrido-Moreno et al., 2024). Together, these findings indicate that the financial returns to AI adoption are contingent on a broader system of organizational capabilities that enable the effective use of technology.

4.3 Thematic Domain 3: Sectoral Context of Digital Ecosystems

The relationship between AI adoption and financial performance is further shaped by sector specific conditions and the broader technological ecosystems in which firms operate. Different industries exhibit varying levels of technological maturity, data availability, and competitive intensity, all of which influence both the incentives for AI adoption and the potential returns on investment. As a result, empirical findings on AI driven performance must be interpreted within their specific contextual settings.

In the manufacturing and industrial sector, the integration of artificial intelligence into broader digital transformation initiatives is associated with improvements in supply chain performance, resource efficiency, and cost reduction. Studies examining Industry 4.0 environments show that the combination of AI, data analytics, and connected technologies enhances operational coordination and supports performance improvements at multiple levels of the value chain (Atif et al., 2021; Elijah et al., 2021). These findings highlight the importance of technological integration and suggest that isolated AI applications are less likely to generate substantial financial benefits compared to comprehensive digital transformation strategies.

The convergence of AI with other digital technologies, such as the Internet of Things, further amplifies performance effects by enabling real time data collection and analysis. Research on intelligent digital systems demonstrates that the interaction between technologies creates

synergistic benefits that exceed the impact of individual components (Aouedi et al., 2024). In logistics and warehouse management, for example, AI driven analytics combined with sensor technologies improve inventory accuracy and operational efficiency, contributing to enhanced financial performance (Jarašūnienė et al., 2023). These results support the view that AI driven value creation is fundamentally dependent on complementarities within digital ecosystems. At the same time, research highlights the importance of aligning technological implementation with broader strategic and societal considerations. The concept of responsible digitalization highlights the need to balance financial performance with environmental and social outcomes, suggesting that long term value creation depends on the integration of technological innovation with sustainable practices (Cardinali & De Giovanni, 2022).

Finally, the open innovation literature underscores the role of AI as an enabler of knowledge exchange and collaborative innovation processes. Firms that leverage AI to enhance their innovation networks are better positioned to generate new products and services, which in turn contribute to revenue growth and market valuation (Bigliardi et al., 2020). This finding reinforces the argument that AI driven financial performance is not solely a function of efficiency improvements but also depends on the firm's ability to use technology to support innovation and strategic renewal.

5. Research Gaps and Future Directions

The preceding synthesis of the literature provides a consolidated view of how artificial intelligence adoption relates to firm financial performance across diverse empirical contexts and analytical perspectives. At the same time, it reveals a set of structural limitations that constrain the development of cumulative knowledge in this field. These limitations do not stem solely from the novelty of AI as a research domain, but also from differences in methodological approaches, levels of analysis, and the scope of examined outcomes. This section therefore moves beyond synthesis to critically examine the boundaries of current knowledge and to identify directions for future research that can address these limitations and support the development of more robust and integrative theoretical and empirical frameworks.

5.1 Temporal Dynamics and Methodological Limitations

Despite the increasing volume of research on artificial intelligence adoption and firm financial performance, a fundamental limitation of the existing literature lies in its predominantly cross sectional orientation. Most empirical studies capture relationships at a single point in time, which limits the ability to distinguish between short-term implementation costs and long-term increases in performance. This limitation is particularly relevant in the context of AI, where the realization of benefits depends on organizational learning, process adaptation, and the gradual development of complementary capabilities (Acemoglu & Restrepo, 2018). Evidence from the economics of digital technologies suggests that

productivity gains often emerge with a delay. This fact reflects adjustment processes that cannot be observed in cross-sectional data (Brynjolfsson et al., 2021).

The reliance on static research designs also limits the capacity to identify causal mechanisms and dynamic relationships. Without longitudinal data, it remains difficult to determine whether observed performance differences are the result of AI adoption itself or of pre existing firm characteristics that influence both adoption decisions and performance outcomes. Future research should therefore prioritize longitudinal and panel data approaches that allow for the examination of temporal trajectories, including the identification of adoption phases, learning curves, and the timing of financial returns. Such approaches would significantly enhance the robustness of empirical findings and contribute to a more nuanced understanding of AI driven performance dynamics.

In addition to temporal limitations, the literature exhibits considerable methodological heterogeneity, which further complicates the accumulation of knowledge. Studies differ in their operationalization of AI adoption, ranging from binary indicators of technology use to more complex measures of digital maturity, and in their measurement of financial performance, which may include accounting indicators, market based metrics, or subjective assessments. This diversity limits the comparability of studies and underscores the need for greater standardization of key constructs. Future research should aim to develop more consistent measurement frameworks that capture both the intensity and the quality of AI integration within organizations.

5.2 Organizational Capabilities and Firm Level Heterogeneity

A second significant difference concerns the uneven distribution of AI-driven performance effects across firms, particularly in relation to organizational capabilities and firm size. While the literature consistently identifies these factors as important moderators, there remains limited understanding of how they interact to shape financial outcomes. Existing research suggests that firms with greater absorptive capacity, stronger innovation orientation, and higher levels of organizational resilience are better positioned to translate AI investments into performance gains (Garrido-Moreno et al., 2024). However, the processes through which these capabilities are developed and mobilized in the context of AI adoption remain insufficiently explored.

The underrepresentation of small and medium sized enterprises in the literature represents a particularly significant limitation. Most empirical studies focus on large firms, which benefit from greater resource availability and more advanced technological infrastructures. As a result, the findings of these studies may not be directly applicable to smaller firms that operate under different constraints and face distinct challenges in adopting and utilizing AI technologies. Evidence indicates that limited financial resources, lack of technical expertise, and organizational inertia can hinder the ability of smaller firms to realize the benefits of digital

transformation (Kallmuenzer et al., 2024; Grashof & Kopka, 2022). Future research should therefore develop context specific frameworks that account for the unique conditions of smaller firms and explore mechanisms through which these firms can overcome adoption barriers.

Another important area for further investigation is the role of human capital and micro level processes in mediating the relationship between AI adoption and performance. While existing studies highlight the importance of employee experience and engagement in enhancing the effectiveness of AI systems, the integration of these micro level factors with firm level performance outcomes remains limited (Malik et al., 2022). Future research should adopt multi-level analytical approaches that integrate individual, organizational, and strategic dimensions. In this way, it can contribute to a more comprehensive understanding of how AI-driven capabilities translate into financial performance.

5.3 Contextual Conditions and Integrated Performance Perspectives

A third area of limitation relates to the role of contextual and institutional factors in shaping the performance implications of AI adoption. The existing literature demonstrates that industry characteristics, competitive dynamics, and technological infrastructures influence both the adoption of AI and its outcomes. However, these factors are often treated as background variables rather than as central components of analysis. This limits the ability to generalize findings and to understand how performance effects vary across different environments.

In addition, the current body of research tends to focus predominantly on financial performance indicators, often overlooking the broader set of outcomes associated with AI adoption. Emerging evidence suggests that artificial intelligence driven transformation generates multidimensional effects. These effects include not only financial gains but also changes in risk exposure, social impact, and environmental performance (Mohamed & Elgammal, 2023; Cardinali & De Giovanni, 2022). The lack of integrated performance frameworks limits the ability to assess the overall value of AI adoption and may result in incomplete evaluations of its impact.

Future research should therefore adopt more comprehensive approaches that integrate financial and non financial dimensions of performance, as well as internal and external contextual factors. This includes examining how regulatory environments, ethical considerations, and stakeholder expectations influence the implementation and outcomes of AI technologies. Such an approach would enable a more holistic understanding of AI driven value creation and support the development of strategies that balance economic performance with broader organizational and societal objectives.

6. Conclusion

This article provides a structured synthesis of the literature examining the relationship between artificial intelligence adoption and firm financial performance, addressing a research domain characterized by conceptual fragmentation and heterogeneous empirical evidence. By organizing the existing body of knowledge into three interrelated thematic domains, the analysis contributes to a more coherent understanding of how AI driven transformation influences firm level outcomes.

The findings indicate that artificial intelligence does not function as a direct determinant of financial performance but rather as a factor whose effects are realized through mediating organizational processes. The impact of artificial intelligence is contingent upon its ability to enhance operational efficiency, support decision making, and facilitate organizational learning. The translation of these improvements into financial outcomes depends on the alignment between technological investments and organizational capabilities. This reinforces the view that value creation from artificial intelligence is fundamentally conditioned by the integration of technology into a broader system of complementary resources and strategic processes.

At the same time, the analysis demonstrates that the effects of AI adoption are inherently heterogeneous. Variations in firm size, absorptive capacity, human capital, and organizational resilience significantly influence the extent to which firms are able to capture financial benefits. In addition, sectoral characteristics and digital ecosystem dynamics shape both the opportunities and constraints associated with AI implementation. These findings highlight that AI driven performance outcomes are context dependent and cannot be generalized across firms or industries without careful consideration of underlying conditions.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Aithal, P. S. (2023). How to create business value through technological innovations using ICCT underlying technologies. International Journal of Applied Engineering and Management Letters, 7(2), 1–35. https://doi.org/10.47992/ijaeml.2581.7000.0184

Aouedi, O., Vu, T.-H., Sacco, A., Nguyen, D. C., & Piamrat, K. (2024). A survey on intelligent Internet of Things: Applications, security, privacy, and future directions. IEEE Communications Surveys & Tutorials, 26(4), 2508–2570. https://doi.org/10.1109/comst.2024.3430368

Atif, S., Ahmed, S., Wasim, M., Zeb, B., & Pervez, Z. (2021). Towards a conceptual development of Industry 4.0, servitisation, and circular economy: A systematic literature review. Sustainability, 13(11), 6501. https://doi.org/10.3390/su13116501

Bigliardi, B., Ferraro, G., Filippelli, S., & Galati, F. (2020). The past, present and future of open innovation. European Journal of Innovation Management, 24(4), 1130–1161. https://doi.org/10.1108/ejim-10-2019-0296

Buhalis, D., O'Connor, P., & Leung, R. (2022). Smart hospitality: From smart cities and smart tourism towards agile business ecosystems in networked destinations. International Journal of Contemporary Hospitality Management, 35(1), 369–393. https://doi.org/10.1108/ijchm-04-2022-0497

Cardinali, P. G., & De Giovanni, P. (2022). Responsible digitalization through digital technologies and green practices. Corporate Social Responsibility and Environmental Management, 29(5), 1393–1406. https://doi.org/10.1002/csr.2249

Chen, L., Li, T., Jia, F., & Schoenherr, T. (2022). The impact of governmental COVID-19 measures on manufacturers' stock market valuations: The role of labor intensity and operational slack. Journal of Operations Management, 68(5), 460–484. https://doi.org/10.1002/joom.1207

Elijah, O., Pang, A. L., Abdul Rahim, S. K., Geok, T. K., & Arsa, A. (2021). A survey on Industry 4.0 for the oil and gas industry: Upstream sector. IEEE Access, 9, 144438–144468. https://doi.org/10.1109/access.2021.3121302

Falletta, S., & Combs, W. L. (2020). The HR analytics cycle: A seven-step process for building evidence-based and ethical HR analytics capabilities. Journal of Work-Applied Management, 13(1), 51–68. https://doi.org/10.1108/jwam-03-2020-0020

Garrido-Moreno, A., Martín-Rojas, R., & García Morales, V. J. (2024). The key role of innovation and organizational resilience in improving business performance: A mixed-methods approach. International Journal of Information Management, 76, 102777. https://doi.org/10.1016/j.ijinfomgt.2024.102777

Grashof, N., & Kopka, A. (2022). Artificial intelligence and radical innovation: An opportunity for all companies? Small Business Economics, 61(2), 771–795. https://doi.org/10.1007/s11187-022-00698-3

Gupta, S., Leszkiewicz, A., Kumar, V., Bijmolt, T. H. A., & Pota, D. (2020). Digital analytics: Modeling for insights and new methods. Journal of Interactive Marketing, 51(1), 26–43. https://doi.org/10.1016/j.intmar.2020.04.003

Jarašūnienė, A., Čižiūnienė, K., & Čereška, A. (2023). Research on impact of IoT on warehouse management. Sensors, 23(4), 2213. https://doi.org/10.3390/s23042213

K. O., M. A., Alam, A., & Hotak, Y. (2025). Smart sensor technologies shaping the future of precision agriculture: Recent advances and future outlooks. Journal of Sensors, 2025, 2460098. https://doi.org/10.1155/js/2460098

Kallmuenzer, A., Mikhaylov, A., Chelaru, M., & Czakon, W. (2024). Adoption and performance outcome of digitalization in small and medium-sized enterprises. Review of Managerial Science, 19, 1285–1315. https://doi.org/10.1007/s11846-024-00744-2

Malik, A., Budhwar, P., Mohan, H., & Srikanth, N. R. (2022). Employee experience – the missing link for engaging employees: Insights from an MNE's AI-based HR ecosystem. Human Resource Management, 62(1), 97–115. https://doi.org/10.1002/hrm.22133

Maraveas, C., Karavas, C.-S., Loukatos, D., & Bart, T. (2023). Agricultural greenhouses: Resource management technologies and perspectives for zero greenhouse gas emissions. Agriculture, 13(7), 1464. https://doi.org/10.3390/agriculture13071464

Mohamed, T. S., & Elgammal, M. (2023). Does the extent of branchless banking adoption enhance the social and financial performance of microfinance institutions? Applied Economics, 56(10), 1163–1178. https://doi.org/10.1080/00036846.2023.2177595

Oubrahim, I., Sefiani, N., & Happonen, A. (2023). The influence of digital transformation and supply chain integration on overall sustainable supply chain performance: An empirical analysis from manufacturing companies in Morocco. Energies, 16(2), 1004. https://doi.org/10.3390/en16021004

Roblek, V., Meško, M., Pejić Bach, M., Thorpe, O., & Šprajc, P. (2020). The interaction between Internet, sustainable development, and emergence of Society 5.0. Data, 5(3), 80. https://doi.org/10.3390/data5030080

Sharabati, A.-A. A., Ali, A. A. A., Allahham, M., & Moha, A. (2024). The impact of digital marketing on the performance of SMEs: An analytical study in light of modern digital transformations. Sustainability, 16(19), 8667. https://doi.org/10.3390/su16198667

Victor, N., Maddikunta, P. K. R., Kesari Mary, D. R., & Ramalinga, R. (2024). Remote sensing for agriculture in the era of Industry 5.0 — A survey. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17, 5543–5571. https://doi.org/10.1109/jstars.2024.3370508

Xu, Y., Khan, T., Song, Y., & Meijering, E. (2025). Edge deep learning in computer vision and medical diagnostics: A comprehensive survey. Artificial Intelligence Review, 58(4), 1–62. https://doi.org/10.1007/s10462-024-11033-5

Corporate performance metrics in the digital age: a systematic literature review of measurement frameworks, technological enablers, and strategic implications

Peter Štetka[1] and Anna Hamranová[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: peter.stetka@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: anna.hamranova@euba.sk

Abstract

The digital era is fundamentally reshaping how firms define, measure, and manage performance, with traditional financial indicators no longer sufficient to capture the value generated by digital transformation. A systematic literature review based on the PRISMA 2020 methodology analyzes 15 relevant studies published between 2021 and 2025 to identify emerging approaches to corporate metrics, technological enablers, and sustainability requirements. The synthesis of findings reveals a shift toward multidimensional metric frameworks that integrate digital maturity, innovation potential, the use of AI, big data and IoT, as well as the quality of ESG reporting. Alongside financial KPIs, digital and technological indicators, intangible capital, and metrics of agile operating models are gaining importance. The analysis also confirms a persistent gap between technological capabilities and organizational readiness, particularly in terms of the flexibility of measurement systems. The theoretical foundation for metrics related to sustainability and digital value likewise remains underdeveloped. The study concludes with a discussion on the need for cross-sector comparability, firms' dynamic capability to adapt metrics to their environment, and the measurement of intangible digital assets.

Key Words: corporate performance metrics, digital age, KPI frameworks, value creation, performance measurement

JEL Classification: M41, O33, G32

1. Introduction

The digital age has introduced a paradigm shift in how corporations define, pursue, and evaluate performance. As organizations increasingly integrate digital technologies into their core operations, the traditional frameworks used to assess corporate performance—rooted primarily in financial indicators such as return on investment, earnings per share, and operating

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

margins—have demonstrated significant limitations in capturing the breadth and complexity of value creation in digitally transformed enterprises. This growing misalignment between legacy performance measurement systems and the realities of digitally intensive business environments represents both a theoretical challenge and a practical imperative.

The stakes of this misalignment are considerable. Organizations that fail to adopt appropriate digital performance metrics risk making suboptimal strategic decisions, misallocating resources, and losing competitive advantage in rapidly evolving markets. Parviainen et al. (2022) demonstrated that the ability to systematically approach digital transformation—including the measurement of its outcomes—is a prerequisite for capturing its operational and strategic benefits. Similarly, Asif et al. (2024) found that digital transformation, when coupled with appropriate leadership and corporate culture, measurably enhances sustainable performance in manufacturing contexts, yet the specific metrics by which such enhancement is gauged remain contested across the literature.

Despite a growing body of research on digital transformation broadly conceived, the specific domain of corporate performance metrics in digital contexts remains fragmented. Cosa & Torelli (2024) noted, through a systematic review of 47 studies, that performance measurement systems are undergoing significant evolution in response to digital transformation but that the field lacks a coherent, integrative framework applicable across diverse organizational contexts. The absence of such a framework creates a research gap that the present review seeks to address.

This paper aims to synthesize the existing literature on corporate performance metrics in the digital age, with four specific objectives: first, to clarify the theoretical foundations underpinning digital performance measurement; second, to identify how emerging technologies are reshaping the indicators and systems used to assess corporate performance; third, to examine the intersection of digital transformation with sustainability-oriented performance governance; and fourth, to identify substantive research gaps and propose future directions for inquiry.

The remainder of this paper is organized as follows. Section 2 establishes the theoretical background, defining key constructs and presenting foundational frameworks. Section 3 presents the methodology of the systematic literature review (SLR). Section 4 presents the results of the SLR, namely the characterization and development of corporate performance measurement systems, the characterization and comparison of traditional and new KPIs, areas of significant agreement and divergence in the reviewed studies, as well as the identification of a research gap. Section 5 presents the implications and directions for further research. Section 6 concludes the review with a summary of contributions.

2. Theoretical Background

The theoretical basis of this paper is the definition of Corporate Performance in the Digital Age, the metric systems used, and the key constructs.

2.1 Defining Corporate Performance in the Digital Age

Corporate performance, traditionally understood as the degree to which an organization achieves its financial and operational objectives, has undergone conceptual expansion in the digital era. Contemporary scholarship increasingly recognizes that performance must encompass dimensions of innovation capacity, customer engagement, supply chain agility, environmental sustainability, and knowledge generation—dimensions for which conventional financial metrics provide only partial visibility. Fuertes et al. (2020) conducted a comprehensive literature review of strategic management frameworks and established that organizational performance must be understood through a multi-layered lens that incorporates strategic evaluation, operational efficiency, and adaptability to environmental change. This conceptual broadening reflects a fundamental epistemological shift in how performance is theorized.

Digital transformation—defined by Parviainen et al. (2022) as the adoption of digital technologies in an organization or its operational environment, leading to fundamental changes in processes, products, services, and business models—introduces new modalities of value creation that challenge the adequacy of traditional metrics. A company may, for example, generate significant competitive value through data assets, algorithmic decision-making capabilities, or network effects that produce no immediate financial returns but substantially enhance long-term enterprise value. Measuring such forms of value requires performance frameworks capable of capturing intangible, dynamic, and temporally distributed outcomes.

2.2 Foundational Theoretical Frameworks

Several theoretical frameworks are particularly relevant to understanding corporate performance metrics in digital contexts. The dynamic capabilities framework, as applied by Santa-María et al. (2021) in the context of circular business model innovation, posits that firms competing in rapidly changing environments must develop the capacity to sense opportunities, seize them, and reconfigure resources accordingly. Performance measurement systems, within this framework, must be sensitive not only to achieved outcomes but to the dynamic processes through which those outcomes are generated. This requires performance metrics that are forward-looking, adaptive, and capable of capturing micro-foundational organizational behaviors.

Information processing theory (IPT), applied by Aben et al. (2021) to public-private relationships undergoing digital transformation, offers another critical lens. IPT suggests that organizational structures and governance mechanisms should be designed to match the information processing demands of the environment. In digitally intensive environments, where data volumes and velocities are unprecedented, performance measurement systems

must be structured to reduce information asymmetry and support effective decision-making. Aben et al. (2021) demonstrated empirically that both contractual and relational governance mechanisms play vital roles in managing the uncertainties that arise when organizations depend on digital technologies for performance-critical information.

Stakeholder theory, as employed by Sun et al. (2024) in their analysis of digital transformation and sustainable performance, provides a third foundational lens. This theory directs attention to the multiple constituencies whose interests shape organizational performance definitions. In the digital age, stakeholders increasingly demand transparency, environmental accountability, and social impact alongside financial returns, thereby expanding the normative scope of corporate performance measurement.

2.3 Key Constructs

Several constructs require precise definition for the purposes of this review. Performance measurement systems (PMS) refer to the integrated set of metrics, processes, and tools that organizations use to assess progress toward strategic and operational objectives. Cosa & Torelli (2024) identified flexibility as the defining attribute of effective PMS in digital environments, defining it as the capacity to adapt organizational resources, processes, and strategies in response to rapid technological change. Big data, as conceptualized by Del Giudice et al. (2020), refers to datasets characterized by volume, velocity, and variety that exceed the processing capacities of conventional analytical tools and that enable novel forms of organizational insight when appropriately managed. Artificial intelligence (AI), as discussed by Olan et al. (2022), refers to computational systems capable of performing tasks that typically require human intelligence, including pattern recognition, predictive modeling, and decision support, all of which have direct applications in performance management. Finally, digital transformation, as emphasized throughout the literature, is not a single event but an ongoing, organizationally embedded process with implications for every dimension of corporate performance.

3. Systematic Literature Review - Methodology

The systematic literature review is conducted in alignment with the objective of the study, using the PRISMA 2020 methodology (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (PRISMA, 2020). The aim is to identify how corporate performance metrics are changing in the digital age, to determine the categories of metrics in the digital era, and to compare traditional and emerging metrics. In accordance with the methodology, the research question was first defined as follows:

"How does digital transformation reshape the architecture and effectiveness of corporate performance measurement systems, what role do technological enablers (AI, big data, IoT) play, and what are the sustainability requirements in shaping new performance metrics?"

For the systematic analysis of scholarly sources, the Web of Science citation database was used. To ensure a targeted selection of literature relevant to answering the research question, the PCC framework (Population–Concept–Context) was applied, along with predefined inclusion and exclusion criteria for selecting publications.

The PCC framework was defined as follows:

Population: enterprises, business entities (without a specific industry focus)

Concept: methods of measuring corporate performance, the indicators used, and their evolution

Context: new digital technologies and Digital Transformation

Inclusion criteria: publications addressing Corporate Performance, evaluation of Corporate Performance, and corporate performance metric systems in the era of digitalization; publication period 2021–2025; document type – peer-reviewed articles; English language; countries limited to Europe.

Exclusion criteria: texts addressing Corporate Performance without a link to information technologies or digital transformation; non-peer-reviewed sources; unavailable full texts; and clear duplicates.

The literature selection process is shown in Figure 1.

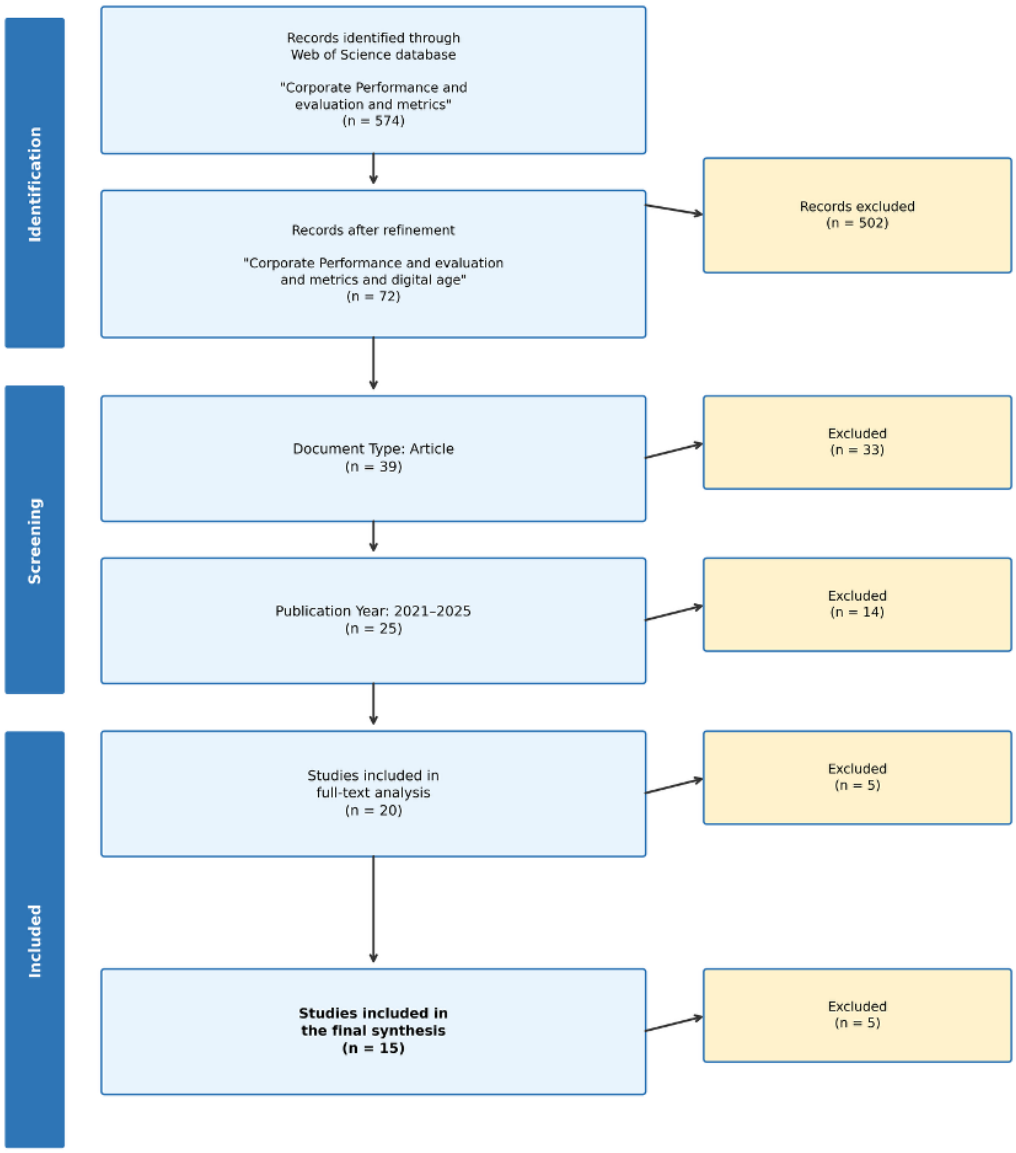


Figure 1: PRISMA flow diagram

4. Systematic Literature Review – Results

Section 4 presents the key results of the systematic literature review, summarizing how corporate performance measurement systems are evolving in the digital age. It highlights the shift from traditional financial indicators toward multidimensional, technology-driven and sustainability-oriented metric frameworks. The chapter outlines areas of consensus and divergence across the reviewed studies and introduces emerging KPI groups shaped by AI, big data and IoT.

4.1 Evolution of Corporate Performance Measurement Systems

Based on the analyzed sources, we can state that in the digital era, corporate performance metrics are shifting from traditional financial indicators toward a multidimensional approach (Holopainen et al., 2023), which incorporates technological maturity, innovation capability, and intangible capital (Zareie et al., 2024). Traditional indicators such as profitability return rates, gross margin, net profit, ROA, and ROE remain important; however, their interpretation is changing in the context of long-term investments in digital transformation (Lee, 2023; Huang & Cheng, 2024; Jardak & Hamad, 2022).

Financial indicators are complemented by the assessment of market value through Tobin's Q, which also reflects market expectations related to digital investments (Jardak & Hamad, 2022; Zareie et al., 2024), as well as by evaluations of operational capacity based on inventory turnover, receivables turnover, or fixed asset efficiency (Lee, 2023).

The range of digital and technological metrics is also expanding significantly. These include the digital transformation score, which measures the frequency of digital terms in annual reports (Zareie et al., 2024); the digital maturity index summarizing the level of digital marketing, e-commerce, and social media use (Jardak & Hamad, 2022); and business analytics utilization indicators assessing a firm's capability to work with descriptive, predictive, and prescriptive analytics (Uyar, 2021).

Innovation and operational metrics also play an important role in performance evaluation, such as the number of patent applications or the efficiency of R&D investments (Teng et al., 2022), along with the level of automation contributing to reduced operating costs (Kasperovica & Lace, 2021). Increasing attention is also placed on sustainability and governance, where the quality of ESG (Environmental, Social, and Governance) reporting is evaluated through the SRQ (Sustainability Reporting Quality) indicator, which is crucial for firms' access to financial capital (AlHares, 2025).

Simultaneously, the importance of intangible capital—particularly organizational capital—continues to grow. It plays a crucial role in transforming technological investments into market value and overall competitive advantage (Zareie et al., 2024). Performance assessment, however, is no longer limited to internal processes; a firm's ability to operate within digital ecosystems and platform-based models is becoming increasingly important, with success depending on connectivity and collaboration with other actors (Kasperovica & Lace, 2021).

Digitalization also fundamentally reshapes the cost structure of firms. Modern technologies enable more efficient cost management, reduce cost stickiness, and support the transition from fixed to variable costs, thereby increasing organizational flexibility in response to environmental changes (Huang et al., 2025).

4.2 Metrics (KPI) characteristics

In the analyzed publications, it is possible to identify the types of KPIs that digital transformation introduces or fundamentally alters. These include:

- AI-driven metrics (Olan et al., 2022; Singh et al., 2019)
- Big Data-driven metrics (Urbinati et al., 2018; Del Giudice et al., 2020)
- IoT-driven metrics (Varga et al., 2020; Patel et al., 2021; Valaskova et al., 2025)

Within these groups, authors classify newly emerging metrics as well as transformed metrics, which are presented in Table 1 … Table 6.

Table 1: New AI-driven KPIs

KPI	Description	Source
Model Accuracy / Precision / Recall	Prediction accuracy metrics are enabled by AI's ability to recognize patterns and provide predictive modeling	Olan et al., 2022
Prediction Uplift (the benefits of AI over traditional methods)	The ability of AI to generate new knowledge from knowledge assets, measured as the knowledge utilization rate	Olan et al., 2022; Singh et al., 2019
AI-based Process Optimization Index (resource savings achieved by AI)	Process optimization index, identifies inefficiencies and optimizes resource allocation	Olan et al., 2022
Automation Rate / RPA Impact (% of tasks automated)	Automation is a key digital factor, replacing manual work, reducing errors and costs	Kasperovica & Lace, 2021

Source: prepared by authors

Table 2: Changes in traditional AI-driven KPIs

KPI	Description	Source
Operational efficiency	Static measurement is changing to flexible performance measurement systems (PMS) that use real-time data integration to dynamically adapt processes	Cosa & Torelli, 2024
Cost per unit	The use of IoT and dynamic optimization of production factors in real time allows companies to directly reduce production costs per unit of product. Data analysis as part of digital transformation helps to better explain the cost structure of the company.	Teng et al., 2022; Uyar, 2021
Lead time	Digitalization of supply chains and Industry 4.0 lead to shorter delivery times and lower inventory costs Implementation of sensors enables the transition to predictive process monitoring instead of retrospective control	Valaskova et al., 2025; Patel et al., 2021

Source: prepared by authors

Table 3: New Big Data-driven KPIs

KPI	Description	Source
Data Quality Index (DQI):	The need to capture data quality. The quality of accounting data demonstrably improves through the use of business analytics	Uyar, 2021
Data Utilization Rate (% of data used for decision making)	Digital transformation changes the organization of a business from defining processes to the concept of data flow management (Decision making through a data prism)	Kasperovica & Lace, 2021.
Customer 360 Score (customer profile complexity)	Allows businesses to get to know their customers better and monitor their complaints or ratings on various web forums	Uyar, 2021; Jardak & Hamad, 2022; Zareie et al., 2024).
Real-time Analytics Latency (delay between data and insight)	Dynamic and multidimensional real-time information retrieval	Huang et al., 2025; Valaskova et al., 2025
Anomaly Detection Rate (identification anomalies)	AI and Big Data are capable of pattern recognition. Using data mining to detect anomalies	Liu et al., 2025

Source: prepared by authors

Table 4: Changes in traditional Big Data-driven KPIs

KPI	Description	Source
Customer Satisfaction (CSAT) + sentiment	Indicators such as customer satisfaction, their complaints,... can be easily monitored through analytical tools, e.g. analysis of statistics from Facebook, Twitter and Instagram (social listening), and projected into management studies. Sentiment analysis is on the list of key digital concepts for modern businesses.	Uyar, 2021; Zareie et al., 2024
Forecast Accuracy	Predictive accuracy metrics improve financial forecasting and risk management capabilities.	Valaskova et al., 2025
Inventory turnover	Digitalization of supply chains leads to inventory optimization and reduction of related costs forecasts of required order volumes working with predicted product quantities.	Valaskova et al., 2025; Huang & Cheng, 2024; Kasperovica & Lace, 2021

Source: prepared by authors

Table 5: New IoT-driven KPIs

KPI	Description	Source

Overall Equipment Effectiveness in real time (OEE-RT) and Mean Time Between Failures (MTBF)	They are related to predictive maintenance powered by AI and IoT, which allows for minimizing downtime. They use intelligently connected sensors and cognitive algorithms to monitor processes in real time, which is the basis for calculating real-time OEE.	Valaskova et al., 2025
Energy Efficiency Index (E2I) and Carbon Footprint per Unit	Resource efficiency indicators and carbon footprint measurements.	Sun et al., 2024
Safety Compliance Score based on wearables	Moving from retrospective incident reporting to continuous and predictive health and safety management. This is enabled by the use of wearables and sensors in the work environment.	Patel et al., 2021

Source: prepared by authors

Table 6: Changes in traditional IoT-driven KPIs

KPI	**Description**	**Source**
"Downtime → measured at the second level	5G and Industrial IoT (IIoT) technologies enable IS to generate continuous streams of operational data. This enables performance monitoring at a granularity (frequency and detail) that was not possible with traditional approaches.	Varga et al., 2020; Patel et al., 2021
"Production yield" → supplemented with sensory quality data	They enable high precision and customization, which leads to a direct increase in product quality and a reduction in error rates.	Valaskova et al., 2025
"Workforce productivity" → biometric data	The use of connected-worker solutions enables real-time monitoring of worker health and productivity using sensors, changing the way productivity is measured and managed.	Patel et al., 2021

Source: prepared by authors

4.3 Areas of significant agreement in the examined studies

All authors of the presented studies emphasize that digital transformation (DT) represents a complex strategic change, whose effects on firm performance are primarily manifested in the long term (Holopainen et al., 2023; Jardak & Hamad, 2022). Several studies confirm the existence of a difference between short-term and long-term impacts of digitalization: while long-term digital maturity increases a firm's market value, short-term financial indicators, such as ROA or ROE, may temporarily decline due to high initial investments and organizational learning processes. This phenomenon, known as the IT paradox, is often interpreted by investors as a signal of future competitiveness (Jardak & Hamad, 2022; Huang & Cheng, 2024; Teng et al., 2022).

Research also points out that technological innovations generate results only if supported by sufficient organizational capital, effective management, and managerial capabilities. These intangible resources enable digital investments to be effectively transformed into higher firm value and performance (Zareie et al., 2024; Holopainen et al., 2023).

An important advantage of DT is also the increase in efficiency and reduction of costs through process optimization or reduction of "cost stickiness," which enhances firms' flexibility in responding to demand fluctuations (Huang et al., 2025). Advanced business analytics further contributes to more accurate planning and improved cost performance (Uyar, 2021).

Digital technologies also improve information quality and reduce information asymmetry, leading to greater transparency and lower cost of equity. Higher-quality reporting, including ESG reporting, strengthens a firm's credibility in the market (Salvi et al., 2023; AlHares, 2025; Zareie et al., 2024).

Finally, the authors note that digital transformation progresses unevenly. Larger firms and industrial sectors typically achieve higher levels of digitalization, whereas small and medium-sized enterprises often face operational, financial, and human resource barriers (Valaskova et al., 2025; Kasperovica & Lace, 2021; Teng et al., 2022).

4.4 Areas of Significant Differences in the Examined Studies

The authors diverge considerably on the direct impact of digital transformation (DT) on short-term financial performance, particularly on ROA and ROE. Huang & Cheng (2024) report a significantly positive effect of DT, whereas Jardak & Hamad (2022) find a negative impact of digital maturity on these indicators. This discrepancy is explained through the IT paradox, according to which high IT investments temporarily reduce profitability, even though the firm's market value (Tobin's Q) increases.

Different results also emerge regarding the shape of the relationships between DT and performance. Teng et al. (2022) confirmed a nonlinear, inverted U-shaped relationship for innovation performance; however, the expected U-shaped relationship between DT and financial performance was not found to be significant.

The authors also disagree on whether DT has a direct or mediated effect. Kharlamov & Parry (2021) argue that digitalization affects financial outcomes only in combination with servitization, while Pucci et al. (2023) emphasizes that the critical factor is the extent of technology integration into the value chain.

In the field of business analytics, Uyar (2021) found that descriptive analytics does not influence planning and control, whereas predictive and prescriptive analytics have a significant positive effect. Similar disagreements are evident in evaluating the strategic dimensions of DT: Holopainen et al. (2023) confirm the influence of digital strategy goals and management on performance management, while other dimensions were not shown to be significant, in contrast with another research.

Finally, sources report markedly different figures regarding the success rate of DT projects. While the Wipro (2018) survey indicates a 50% failure rate, Bughin et al. (2019) and Rogers (2016) suggest that as many as 80% to 84% of organizations may fail.

4.5 Identification of the Research Gap

The available literature exhibits a pronounced geographic and regional imbalance, as research on digital transformation (DT) is concentrated mainly in China, the United States, and economically advanced Western European countries, while Central and Eastern Europe, including Slovakia, remains significantly underexplored. In the Slovak context, academic studies identifying specific digitalization approaches that contribute to improved financial performance in various sectors are still lacking (Valaskova et al., 2025; Zareie et al., 2024). Another limitation is the absence of longitudinal studies, as many investigations use cross-sectional data and therefore do not capture the delayed effects of digitalization caused by high initial costs, known as the IT paradox (Jardak & Hamad, 2022; Holopainen et al., 2023). Future research should thus employ time series data to distinguish short-term financial weakening from long-term growth in market value (Jardak & Hamad, 2022; Valaskova et al., 2025).

Current literature also focuses predominantly on large firms, whereas small and medium-sized enterprises (SMEs) are only marginally studied, even though they face specific obstacles, such as lack of capital and digital talent (Teng et al., 2022; Valaskova et al., 2025). Untapped research potential also exists in managerial characteristics: it remains insufficiently explored how personal and professional factors, such as IT education, political connections, or overconfidence of managers, influence the success of digital transformation and cost management (Huang et al., 2025; Liu et al., 2025).

Moreover, literature continues to show a discrepancy between evidence of the positive impact of digitalization on organizational processes and the uncertain effect on financial indicators, particularly ROA and ROE, which show conflicting empirical results (Jardak & Hamad, 2022; Huang & Cheng, 2024). Research also highlights the need for deeper examination of moderating and mediating mechanisms, such as corporate governance, organizational capital, information quality, or the level of internal controls, as these factors can significantly affect the transformation of digital investments into market value (Zareie et al., 2024; AlHares, 2025; Liu et al., 2025).

Finally, there is a lack of detailed research focused on the impact of new technologies in specific sectors. In particular, generative artificial intelligence and blockchain are rarely examined in the context of auditing, risk management, or low-digitalized sectors such as construction, despite their growing significance (Valaskova et al., 2025; Liu et al., 2025).

5. Implications for Future Research

The analyzed sources suggest several key directions for future research on digital transformation and firm performance.

1. **Methodological Shifts Toward Longitudinal Studies**

The authors agree that there is a need to move from cross-sectional studies to longitudinal research designs, which allow for a more detailed examination of how digital transformation evolves over time and how long-term investments contribute to firms' financial resilience (Valaskova et al., 2025; Holopainen et al., 2023). A longitudinal approach is also crucial for understanding the time lag between IT investments and their reflection in financial indicators, which can help better explain the IT paradox phenomenon (Huang & Cheng, 2024; Jardak & Hamad, 2022).

2. **Expansion of Geographic and Sectoral Coverage**

Future research should include geographically specific analyses, particularly for Central and Eastern Europe, where firms face different barriers than in China or the USA (Valaskova et al., 2025). It is also important to investigate the impact of digitalization transparency across different institutional environments (Salvi et al., 2023). Sector-focused research should identify digitalization activities that most significantly enhance performance in slower-digitized sectors, such as construction or agriculture (Valaskova et al., 2025).

3. **Focus on Small and Medium-Sized Enterprises (SMEs)**

As noted by Teng et al. (2022) and AlHares (2025), research should place greater emphasis on the SME sector, as studies to date have predominantly focused on large firms. SMEs, however, face specific challenges, including limited financial resources and a shortage of digital talent, which requires the development of tailored performance assessment systems and profitability management models (Teng et al., 2022; Kasperovica & Lace, 2021).

4. **Integration of Intangible and Qualitative Factors**

According to AlHares (2025), future studies should combine quantitative indicators with qualitative methods, which can provide deeper insights into firms' motivations, barriers, and internal decision-making processes. Another important direction is to investigate more closely the influence of corporate governance, organizational culture, and leadership capabilities on the success of digital transformation (Jardak & Hamad, 2022; Zareie et al., 2024; Huang & Cheng, 2024).

5. **Examination of New Technologies and Communication Channels**

Research should analyze the impact of specific technologies, such as generative artificial intelligence, blockchain, risk analytics, or business intelligence tools, on firm performance and decision-making processes (Liu et al., 2025; Uyar, 2021). Another important area is to study how information about digitalization spreads through alternative communication channels, such as social media, and the effect this has on market behavior (Salvi et al., 2023).

6. **Development of Comprehensive Metrics**

Finally, the authors emphasize the need for a broader range of metrics that supplement traditional financial indicators with long-term solvency and growth factors, while also including non-financial indicators, such as customer satisfaction or a firm's innovation capability (Lee,

2023; Kasperovica & Lace, 2021). Such an expanded set of indicators would provide a more comprehensive view of firm performance in a digital environment.

6. Conclusion

This systematic literature review demonstrates that corporate performance measurement in the digital age is undergoing a fundamental transformation driven by advances in artificial intelligence, big data analytics, and Internet of Things technologies. Traditional financial indicators remain relevant but no longer suffice to capture the complexity of value creation in digitally intensive environments. The reviewed studies consistently highlight a shift toward flexible, multidimensional performance measurement systems that integrate technological maturity, innovation capacity, data utilization, sustainability reporting, and intangible organizational assets. Despite these advancements, the findings reveal substantial disparities in organizational readiness, sectoral adoption rates, and the methodological rigor used to assess digital transformation outcomes. Significant contradictions also persist regarding the short-term financial impact of digitalization, reflecting the influence of the IT productivity paradox.

Moreover, the current evidence base is geographically imbalanced, predominantly focused on China, the United States, and Western Europe, with limited insights from Central and Eastern Europe. The scarcity of longitudinal studies further restricts understanding of delayed performance effects arising from large upfront digital investments. Research gaps also remain in the analysis of small and medium-sized enterprises, managerial characteristics, governance mechanisms, and the role of emerging technologies such as generative AI and blockchain in performance evaluation.

Overall, this review underscores the need for more comprehensive, context-sensitive, and methodologically diversified research capable of capturing the dynamic and intangible nature of digital-era performance. Future studies should focus on developing integrated metric architectures, enhancing cross-sector comparability, and exploring mediating factors that determine how digital transformation translates into sustainable financial and non-financial outcomes.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Aben, T., van der Valk, W., Roehrich, J. K., & Selviaridis, K. (2021). Managing information asymmetry in public–private relationships undergoing a digital transformation: The role of contractual and relational governance. International Journal of Operations & Production Management, 41(7), 1053–1077. https://doi.org/10.1108/ijopm-09-2020-0675

AlHares, A. (2025). Does financial performance improve the quality of sustainability reporting? Exploring the moderating effect of corporate governance. Sustainability, 17(13), 6123.

Asif, M., Yang, L., & Hashim, M. (2024). The role of digital transformation, corporate culture, and leadership in enhancing corporate sustainable performance in the manufacturing sector of China. Sustainability, 16(7), 2651. https://doi.org/10.3390/su16072651

Bughin, J., Deakin, J., & O'beirne, B. (2019). Digital transformation: Improving the odds of success. McKinsey Quarterly, 22, 1-5.

Cosa, M., & Torelli, R. (2024). Digital transformation and flexible performance management: A systematic literature review of the evolution of performance measurement systems. Global Journal of Flexible Systems Management, 25(2), 215–238. https://doi.org/10.1007/s40171-024-00409-9

Del Giudice, M., Chierici, R., Mazzucchelli, A., & Fiano, F. (2020). Supply chain management in the era of circular economy: The moderating effect of big data. The International Journal of Logistics Management, 32(2), 487–510. https://doi.org/10.1108/ijlm-03-2020-0119

Fuertes, G., Alfaro, M., Vargas, M., Gutiérrez, S., Ternero, R., & Sabattin, J. (2020). Conceptual framework for the strategic management: A literature review—descriptive. Journal of Engineering, 2020, 6253013. https://doi.org/10.1155/2020/6253013

Holopainen, M., Saunila, M., & Ukko, J. (2023). Facilitating performance measurement and management through digital business strategy. Measuring Business Excellence, 27(2), 246–260.

Huang, C.-K., & Cheng, Y.-H. (2024). Assessing firm performance in digital transformation: a preliminary analysis using financial indicators. Managerial Finance, 50(9), 1641–1661.

Huang, W., Albitar, K., & Wang, Y. (2025). Cost stickiness in the digital age: unveiling the influence of corporate digital transformation and top managers' characteristics. Journal of Accounting Literature.

Jardak, M. K., & Hamad, S. B. (2022). The effect of digital transformation on firm performance: evidence from Swedish listed companies. The Journal of Risk Finance, 23(4), 329–348.

Kasperovica, L., & Lace, N. (2021). Factors influencing companies' positive financial performance in digital age: a meta-analysis. Entrepreneurship and Sustainability Issues, 8(4), 291–311.

Kharlamov, A. A., & Parry, G. (2021). The impact of servitization and digitization on productivity and profitability of the firm: a systematic approach. Production Planning & Control, 32(3), 185-197.

Lee, C.-C. (2023). Analyses of the operating performance of information service companies based on indicators of financial statements. Asia Pacific Management Review, 28(4), 410–419.

Liu, J., Li, J., & Wang, S. (2025). Do executives with IT backgrounds influence the selection of corporate auditors in the context of digital innovation?—An examination from a sustainability perspective. Sustainability, 17(19), 8911.

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. https://doi.org/10.1016/j.jbusres.2022.03.008

Parviainen, P., Tihinen, M., Kääriäinen, J., & Teppola, S. (2022). Tackling the digitalization challenge: How to benefit from digitalization in practice. International Journal of Information Systems and Project Management, 5(1), 63–77. https://doi.org/10.12821/ijispm050104

Patel, V., Chesmore, A., Legner, C., & Pandey, S. (2021). Trends in workplace wearable technologies and connected-worker solutions for next-generation occupational safety, health, and productivity. Advanced Intelligent Systems, 4(4), 2100099. https://doi.org/10.1002/aisy.202100099

PRISMA. (2020). PRISMA 2020 statement. https://www.prisma-statement.org/prisma-2020

Pucci, E., Possaghi, I., Cutrupi, C. M., Baez, M., Cappiello, C., & Matera, M. (2023, April). Defining patterns for a conversational web. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (pp. 1-17).

Rogers, D. L. (2016). The digital transformation playbook: Rethink your business for the digital age. Columbia University Press.

Salvi, A., Petruzzella, F., Raimo, N., & Vitolla, F. (2023). Transparency in the digitalization choices and the cost of equity capital. Qualitative Research in Financial Markets, 15(4), 630–646.

Santa-María, T., Vermeulen, W. J. V., & Baumgartner, R. J. (2021). How do incumbent firms innovate their business models for the circular economy? Identifying micro-foundations of dynamic capabilities. Business Strategy and the Environment, 30(7), 3357–3376. https://doi.org/10.1002/bse.2956

Singh, J., Flaherty, K., Sohi, R. S., Deeter-Schmelz, D., Habel, J., Le Meunier-FitzHugh, K., Malshe, A., Mullins, R., & Onyemah, V. (2019). Sales profession and professionals in the age of digitization and artificial intelligence technologies: Concepts, priorities, and questions. Journal of Personal Selling and Sales Management, 39(1), 2–22. https://doi.org/10.1080/08853134.2018.1557525

Sun, L., Ong, T., Teh, B., & Di Vaio, A. (2024). Sustainable performance measurement through digital transformation within the sustainable development framework: The mediating effect of supply chain concentration. Sustainable Development, 32(4), 3451–3465. https://doi.org/10.1002/sd.3007

Teng, X., Wu, Z., & Yang, F. (2022). Impact of the digital transformation of small- and medium-sized listed companies on performance: Based on a cost-benefit analysis framework. Journal of Mathematics, 2022, 1504499.

Urbinati, A., Bogers, M., Chiesa, V., & Frattini, F. (2018). Creating and capturing value from big data: A multiple-case study analysis of provider companies. Technovation, 84–85, 21–36. https://doi.org/10.1016/j.technovation.2018.07.004

Uyar, M. (2021). The role of business analytics in transforming management accounting information into cost performance. Ege Akademik Bakış / Ege Academic Review, 21(4), 373–389.

Valaskova, K., Nagy, M., & Juracka, D. (2025). Digital transformation and financial performance: an empirical analysis of strategic alignment in the digital age. Journal of Enterprising Communities: People and Places in the Global Economy, 19(5), 1178–1205.

Varga, P., Pető, J., Frankó, A., Balla, D., Haja, D., Janky, F., Soos, G., Ficzere, D., Maliosz, M., & Toka, L. (2020). 5G support for industrial IoT applications—Challenges, solutions, and research gaps. Sensors, 20(3), 828. https://doi.org/10.3390/s20030828

WIPRO (2018) Wipro Digital Aps https://www.wipro.com/investors/annual-reports/subsidiaries-2017-2018/

Zareie, M., Attig, N., El Ghoul, S., & Fooladi, I. (2024). Firm digital transformation and corporate performance: The moderating effect of organizational capital. Finance Research Letters, 61, 105032.

Econometric Models for Technology Investment Analysis: A Narrative Literature Review

Grisáková Nora[1] and Hanák Róbert [2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: nora.grisakova@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: robert.hanak@euba.sk

Abstract

Technological change has become one of the central drivers of investment decision-making, yet the econometric modelling of technology investment remains fragmented across macroeconomic, firm-level, and sustainability-oriented research traditions. This article provides a narrative literature review with selected semi-systematic elements of major empirical contributions examining how technology-related investment is conceptualized, measured, and evaluated in contemporary economic research. The review synthesizes studies on automation, digitalization, innovation, market structure, climate exposure, and ESG-related investment outcomes to identify the main methodological approaches and substantive patterns in the literature. The findings indicate that technology investment is most robustly associated with structural changes in production, shifts in market power, and the repricing of long-term risk, while firm-level financial effects appear more conditional, mediated, and context-dependent. Unresolved endogeneity, insufficient treatment of institutional context, and the incomplete integration of climate and transition risks continue to weaken the explanatory power of existing models. The article concludes that technology investment should not be interpreted through a single average-return framework, but rather as a multidimensional process shaped by production structure, competitive conditions, and regulatory and environmental constraints.

Key Words: technology investment; econometric modelling; digital transformation; automation; technology investment; ESG; climate risk; market structure

JEL Classification: C33, O30, G31

1. Introduction

Technological change has become one of the defining forces shaping contemporary investment behaviour. As firms and economies increasingly allocate resources to automation, digital infrastructure, and sustainability-oriented innovation, the question of how to model the economic consequences of these investments has gained both theoretical and practical urgency. Technology investment is no longer a narrow issue of capital expenditure in a conventional sense. It now involves a broader set of mechanisms through which innovation alters production processes, changes cost structures, reshapes market competition, and redefines the risk profile of firms and investors.

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

Despite the rapid expansion of research in this area, the econometric literature on technology investment remains methodologically fragmented. Different strands of scholarship adopt different units of analysis, theoretical assumptions, and empirical strategies. Macroeconomic studies typically focus on productivity, labour shares, and structural transformation, while firm-level analyses often concentrate on profitability, market value, or investor responses. At the same time, a growing body of research has begun to incorporate climate exposure, ESG performance, and sustainability-related regulation into the broader analysis of technology-related capital allocation. As a result, the field has expanded substantially, but often without a sufficiently integrated perspective on how these approaches relate to one another.

This fragmentation is partly due to the multidimensional nature of technology investment itself. Investments in digital systems, artificial intelligence, green innovation, and financial technology differ not only in their economic function, but also in their time horizon, uncertainty, measurement, and expected return profile. Some forms of technology investment generate direct productivity gains, while others primarily reduce transaction and coordination costs, alter task allocation, or improve firms' adaptive capacity under regulatory and environmental pressures. These differences make it difficult to apply a single econometric framework across the full spectrum of technology-related investments.

Another major challenge concerns identification. Technology investment decisions are rarely exogenous. Firms that invest more intensively in innovation or digitalization are also likely to differ in managerial quality, access to finance, market power, strategic orientation, or exposure to external shocks. This makes causal inference difficult and raises persistent concerns regarding endogeneity, reverse causality, and omitted variable bias. In addition, many important effects of technology investment unfold over time rather than immediately, which means that static or short-run models may fail to capture delayed productivity gains, transitional adjustment costs, or changing investor expectations.

The growing importance of climate and sustainability factors has added a further layer of complexity. Technology investment is increasingly evaluated not only by its financial returns, but also by its interaction with transition risk, regulatory pressure, physical climate exposure, and ESG-related performance. This broadens the analytical scope of econometric modelling and suggests that technology investment should be understood within a wider framework of long-term structural and institutional change.

Against this background, the aim of this article is to provide a structured synthesis of the econometric literature on technology investment analysis. The article does not seek to deliver an exhaustive systematic review, but rather a transparent and analytically focused narrative review with selected semi-systematic elements. By comparing major contributions across different thematic strands, the study identifies recurring methodological patterns, key empirical findings, and unresolved tensions in the literature. Attention is devoted to the way studies measure technology-related investment, address identification challenges, and interpret financial, structural, and sustainability-related outcomes.

The contribution of this article is therefore twofold. First, it brings together research traditions that are often treated separately, including structural analyses of technological change, firm-level studies of financial outcomes, and emerging work on climate and ESG-related investment risks. Second, it highlights the extent to which current econometric models remain constrained by limited treatment of heterogeneity, dynamic adjustment, and institutional context. In doing so, the review offers a more coherent view of the state of the field and clarifies where future research is most needed.

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

The remainder of the article is organized as follows. Section 2 presents the theoretical background and outlines the main conceptual foundations for modelling technology investment. Section 3 explains the methodological approach adopted in the review. Section 4 synthesizes the literature and identifies the main empirical patterns. Section 5 discusses the principal research gaps and future directions. Section 6 concludes.

2. Theoretical Background

Econometric modelling of technology investment draws on several theoretical traditions that shape how investment is defined, measured, and evaluated. In this review, technology investment is understood narrowly as expenditure on innovation, automation, and digital systems that expand productive capacity, alter production processes, or reduce coordination and information costs. This narrower definition is useful because it separates technology investment from broader categories such as general sustainability or governance spending, where the underlying econometric problems may differ. A key starting point is the growth perspective on idea production. Bloom et al. (2017) show that long-run growth depends not only on the number of researchers but also on research productivity, and they provide evidence that research effort has risen while research productivity has declined across several settings. This implies that R&D expenditure should not be treated as a stable proxy for innovation output. In econometric terms, the relationship between technology investment and realized technological progress is not constant across firms, sectors, or periods. A second important foundation is the task-based framework of Acemoglu and Restrepo (2019). They argue that technological change affects production by reallocating tasks between labor and capital. Automation creates a displacement effect by replacing labor in some tasks, while new tasks may partly offset this through a reinstatement effect. This matters for technology investment analysis because it shows that aggregate capital measures alone may be too broad. What matters is not only how much firms invest, but how investment changes production structure, labor demand, and the distribution of gains. Market structure also shapes investment incentives and observed returns. De Loecker and Eeckhout (2017) document a strong rise in markups in the United States after 1980, concentrated mainly among high-markup firms. Their findings suggest that the returns to technology investment may differ depending on the degree of pricing power and market concentration. Goldfarb and Tucker (2017) complement this view by arguing that digital technologies transform economic activity primarily by reducing costs such as search, replication, transportation, tracking, and verification. As a result, the returns to digital investment depend not only on the investment itself but also on the market environment in which it is implemented. A broader perspective is offered by Donaldson (2018), who shows that infrastructure investment can generate welfare gains through lower trade costs and improved market integration. Although railway infrastructure differs from digital technologies, the underlying mechanism is relevant: investments that reduce frictions in exchange, communication, or coordination may generate effects that are not fully captured by firm-level financial indicators. This suggests that some forms of technology investment should be analysed not only at the firm level, but also through network or market-access channels. The final issue is identification. Technology investment is rarely exogenous, since firms that invest more may systematically differ in managerial quality, financing conditions, or growth expectations. Andrews et al. (2018) show that weak instruments can make conventional instrumental variables inference unreliable, which is highly relevant in this literature because valid instruments for investment decisions are often difficult to establish. For this reason, endogeneity should be treated as a central challenge in empirical models of technology investment.

Taken together, these theoretical perspectives show that econometric research on technology investment must distinguish between inputs and outputs, consider task reallocation and market structure, account for friction-reducing infrastructure effects where relevant, and address identification problems with particular care. These points provide the theoretical basis for the methodological chapter that follows.

3. Methodology

This study adopts a narrative literature review with selected semi-systematic elements. The aim was not to produce an exhaustive systematic review, but rather to identify, compare, and synthesize the most relevant academic contributions on econometric models used to analyse technology investment and its related outcomes. This approach is appropriate given the conceptual breadth of the topic, which spans innovation, automation, digitalization, market structure, climate-related investment, and firm-level valuation. At the same time, to ensure sufficient transparency and analytical rigor, the review followed a structured search, screening, and thematic synthesis process.

The literature search was based primarily on academic databases commonly used in economics, finance, and management research, namely Scopus, Web of Science, and Google Scholar. The search strategy combined general and more specific terms, including "technology investment," "econometric models," "innovation investment," "automation," "digital investment," "R&D productivity," "ESG and firm value," "climate exposure," "FinTech," and "market power." These terms were used both individually and in combination in order to capture studies employing different conceptualizations of technology-related investment. The search focused mainly on peer-reviewed journal articles, influential working papers, and highly cited review studies relevant to the economic and econometric analysis of investment in technology-oriented assets and processes.

The time frame of the review was defined broadly to include literature published primarily between 2015 and 2023, while allowing the inclusion of earlier or foundational studies where necessary for conceptual continuity. The inclusion criteria emphasized methodological relevance, citation impact, and substantive connection to the central research question. Studies were retained if they addressed the measurement, modelling, determinants, or consequences of technology-related investment using an identifiable empirical or analytical framework. By contrast, sources were excluded when they were only marginally related to technology investment, lacked a clear methodological contribution, or focused on purely technical aspects without economic interpretation.

The selection process involved an initial screening of titles, abstracts, and keywords, followed by a closer reading of the full texts of those studies that appeared conceptually and methodologically relevant. The final set of sources was then synthesized using thematic analysis. Rather than summarizing studies sequentially, the literature was grouped into broader analytical categories, including structural effects of technology investment, market structure and cost reduction, firm-level financial outcomes, and climate- and sustainability-related investment risks. This thematic categorization made it possible to identify recurring patterns, points of convergence and divergence, and the main research gaps that structure the subsequent discussion.

4. Literature Review

The reviewed literature does not support a single universal econometric model of technology investment. Instead, it reveals several interconnected research streams that differ in their unit of analysis, methodological design, and definition of outcomes. When considered together, however, these studies suggest a consistent broader pattern: technology investment is most strongly associated with structural changes in production, shifts in market power, and the repricing of risk, whereas evidence on direct firm-level financial performance remains more heterogeneous and conditional.

A first major trend in the literature is the treatment of technology investment as a force of structural transformation rather than a conventional form of capital deepening. In this perspective, investment in automation, digital systems, and innovation affects not only the volume of capital but also the organization of production itself. Acemoglu and Restrepo (2019) show that automation changes the allocation of tasks between labor and capital, producing displacement effects that may not be fully offset by the creation of new tasks. Similarly, Bloom et al. (2017) demonstrate that growing research effort has been accompanied by declining research productivity, implying that higher expenditure on innovation does not necessarily lead to proportionate output gains. Together, these studies suggest that technology investment should not be interpreted simply as additional spending with stable returns, but rather as a process whose effects depend on changing technological opportunity, task reallocation, and sector-specific productivity conditions.

A second important pattern concerns the interaction between technology investment and market structure. De Loecker and Eeckhout (2017) document a substantial rise in markups, especially among firms at the upper end of the distribution, indicating that the benefits of technological change may increasingly accrue in concentrated markets. This finding is particularly relevant for econometric analysis because it implies that the private returns to technology investment may differ from their broader social returns. Goldfarb and Tucker (2017) complement this argument by showing that digital technologies transform economic activity primarily by reducing search, replication, transportation, tracking, and verification costs. These cost reductions help explain why digital investments often generate scalable gains, but also why their effects depend heavily on market concentration, platform structure, and information frictions. As a result, the literature increasingly suggests that the returns to technology investment cannot be evaluated independently of the competitive environment in which such investments occur.

A third line of evidence emphasizes that some technology investments operate in ways similar to infrastructure, producing effects that extend beyond firm-level accounting outcomes. Donaldson (2018) shows that infrastructure investment can improve market integration and raise welfare through lower trade costs. Although the historical railway context differs from modern digital technologies, the underlying mechanism remains relevant: investments that reduce frictions in exchange, communication, or coordination can generate broader economic effects that are not fully captured by narrow measures of firm profitability. This broadens the analytical perspective of the literature and suggests that some forms of technology investment should be evaluated through network, market-access, and system-level channels rather than only through short-run financial indicators.

In contrast to these more coherent structural findings, firm-level evidence on technology-related financial outcomes is more mixed. This is partly because the literature uses a wide range of proxies for technology-related investment, including ESG performance, climate exposure, digital finance, and innovation-related market signals. Consequently, the literature is more convincing in identifying mediating mechanisms and cross-sectional heterogeneity than in

establishing a single average effect. For example, Zhou et al. (2022) find that ESG performance is positively associated with company market value, with financial performance serving as a mediating factor. This suggests that the effect is neither automatic nor purely direct. More broadly, the literature implies that ESG and sustainability-related variables influence firm value through specific transmission channels, including profitability, risk perception, and investor expectations, rather than through a uniform premium.

This conditionality is reinforced by research that explicitly models heterogeneous effects. Demir et al. (2020) show that the relationship between FinTech development and income inequality varies across the conditional distribution, indicating that average effects conceal substantial cross-country differences. Although focused on inequality rather than firm value, the study has important methodological implications for technology investment research: the effects of technological change are unlikely to be constant across firms, sectors, or institutional contexts. Sautner et al. (2023) provide a particularly important contribution by constructing text-based firm-level measures of climate change exposure from earnings calls and showing that such exposure predicts green patenting and green-technology job creation. Their findings suggest that more refined measures can link technology-related risk and opportunity to observable innovation outcomes more effectively than standard accounting variables.

A further trend in the literature is the growing centrality of climate and sustainability considerations in technology investment analysis. Earlier literature often treated environmental variables as peripheral controls, but more recent studies integrate them into the core logic of capital allocation and innovation strategy. Dechezleprêtre and Sato (2017) show that environmental regulation may impose modest short-run competitiveness costs while also stimulating innovation. This implies that regulation should not be viewed solely as a constraint but also as a factor that redirects technological effort. Krueger et al. (2020) similarly show that institutional investors regard climate risks as financially material and tend to prefer engagement and risk management over divestment. This is significant because it indicates that climate-related technology investment is shaped not only by realized returns but also by forward-looking investor beliefs and expectations.

Taken together, the literature supports three main conclusions. First, technology investment is most robustly linked to structural changes in production, cost structures, and market organization rather than to a uniform increase in firm performance. Second, firm-level financial effects are mediated, heterogeneous, and highly context-dependent. Third, climate and sustainability variables have become central to the analysis of technology investment because they influence both the direction of innovation and the pricing of long-term risk. The main contribution of the reviewed literature therefore lies not in identifying one stable average return to technology investment, but in showing that such returns depend on mechanism, market structure, and institutional setting.

Table 1 Analytical synthesis of selected studies

Study	Main focus	Methodology	Key finding	Relevance for this review
Acemoglu and Restrepo (2019)	Automation and task reallocation	Task-based model	Automation changes production by reallocating tasks between labor and capital	Shows that technology investment must be understood as structural change

Study	Main focus	Methodology	Key finding	Relevance for this review
Bloom et al. (2017)	Research productivity	Long-run analysis of innovation effort	More research input is needed to sustain innovation over time	Questions the assumption of stable returns to R&D expenditure
De Loecker and Eeckhout (2017)	Markups and market power	Firm-level production analysis	Rising markups suggest increasing concentration of gains	Indicates that returns to technology depend on market structure
Goldfarb and Tucker (2017)	Digital economics	Conceptual synthesis	Digital technologies reduce key transaction and coordination costs	Explains the mechanisms through which digital investment creates value
Donaldson (2018)	Infrastructure and welfare	General-equilibrium analysis	Friction-reducing infrastructure improves market integration and welfare	Supports a broader system-level interpretation of technology investment
Zhou et al. (2022)	ESG and market value	Mediation analysis	ESG performance affects market value through financial performance	Shows that firm-level effects are mediated rather than purely direct
Demir et al. (2020)	FinTech and inequality	Quantile regression	Effects differ across countries and across the distribution	Demonstrates the importance of heterogeneity
Sautner et al. (2023)	Climate exposure and innovation	Text-based firm-level measurement	Climate exposure predicts green patenting and green jobs	Improves measurement of technology-related risk and innovation
Krueger et al. (2020)	Climate risk and investors	Survey evidence	Investors treat climate risk as financially relevant	Highlights the role of expectations in capital allocation
Dechezleprêtre and Sato (2017)	Regulation and competitiveness	Literature review	Regulation can create short-run costs while stimulating innovation	Supports a non-binary view of regulation and investment

4. Research Gaps and Future Directions

Although literature has made substantial progress, several important gaps remain. The first concerns heterogeneity across firms, sectors, and institutional environments. Many studies indicate that the effects of technology investment vary systematically depending on regulatory conditions, market concentration, financial development, and exposure to climate-related risks. Yet a large part of the empirical literature still relies on pooled average estimates that obscure these differences. Future research should therefore place greater emphasis on interaction effects, distributional methods, and designs that treat institutional context as part of the return-generating mechanism rather than merely as a control variable.

A second major gap concerns identification. Technology investment is rarely exogenous, since firms that invest more in automation, innovation, or digital systems are also likely to differ

in managerial quality, financing conditions, growth expectations, and strategic positioning. This makes causal inference particularly difficult. The literature shows that many results are informative at the level of association, but fewer are fully convincing in causal terms. Future work would benefit from stronger identification strategies based on policy shocks, staggered infrastructure rollouts, regulatory discontinuities, or other plausibly exogenous sources of variation. Where such designs are not feasible, greater caution is needed in interpreting estimated effects as causal.

A third weakness is the limited treatment of dynamic adjustment. Technology investment frequently involves delayed effects, learning processes, temporary adjustment costs, and nonlinear diffusion, yet many studies still focus on contemporaneous or short-run relationships. This creates a risk of underestimating medium-term productivity effects or overstating immediate financial returns. More dynamic modelling approaches would help distinguish between short-term disruption and longer-term gains and would therefore provide a more realistic picture of how technology investment affects economic performance over time.

A fourth gap concerns the insufficient integration of climate, transition, and resource-related risks into mainstream econometric models of technology investment. Recent literature has shown that climate exposure, regulatory change, and environmental constraints increasingly shape both innovation incentives and investor behaviour. However, these insights are still too often analysed separately rather than within a single framework. Future research should aim to connect physical climate risk, transition policy, technological opportunity, and firm-level investment behaviour in more coherent empirical models.

Overall, the literature suggests that technology investment is not well captured by a single average treatment effect or by a narrow firm-performance perspective. It is better understood as a context-dependent process linking innovation, production structure, market power, and long-term risk allocation. Future research will make the strongest contribution where it combines stronger identification with richer measurement and a more explicit treatment of heterogeneity and dynamics. That is where the field still has room to move from descriptive sophistication to genuinely explanatory power.

5. Conclusion

This article has reviewed the main econometric approaches used to analyse technology investment and has shown that the literature does not support a single, universal model of technology-related investment behaviour. Instead, the evidence points to a field characterized by conceptual diversity, methodological plurality, and increasing recognition of contextual heterogeneity. Across the reviewed studies, the most consistent conclusion is that technology investment should be understood not simply as an increase in capital expenditure, but as a multidimensional process that affects production structure, market organization, innovation incentives, and long-term risk allocation.

Three broader conclusions emerge from the review. First, the strongest evidence concerns the structural consequences of technology investment. Research on automation, digitalization, and innovation shows that technology alters the task composition of production, changes cost structures, and interacts with market concentration in ways that standard competitive investment models do not fully capture. Second, firm-level financial outcomes are more variable and less uniform than is often assumed. Rather than supporting a stable average return to technology investment, the literature suggests that financial effects are mediated by profitability, investor expectations, institutional conditions, and the specific mechanism

through which investment creates value. Third, climate and sustainability factors have become central rather than peripheral to the econometric analysis of technology investment. Regulatory pressures, transition risks, and exposure to climate-related uncertainty increasingly shape both innovation direction and capital allocation.

At the same time, the review highlights important limitations of the existing literature. A substantial part of the evidence remains constrained by unresolved endogeneity, limited dynamic modelling, and insufficient attention to cross-country and cross-sectoral variation. Many studies continue to rely on average effects that mask significant heterogeneity, while others treat institutional context or environmental risk as secondary controls rather than integral components of the return-generating process. These limitations reduce the explanatory power of current econometric models and help explain why findings often remain conditional rather than generalizable.

The article also has its own limitations. As a narrative review with selected semi-systematic elements, it does not provide formal meta-analytic aggregation of effect sizes and does not attempt exhaustive coverage of all relevant publications. The emphasis on influential and methodologically relevant studies inevitably narrows the scope of included evidence. Nevertheless, this approach is appropriate for a field in which concepts, measures, and empirical designs remain highly diverse and where the key need is not only accumulation of findings, but also clearer analytical integration.

Overall, the literature indicates that technology investment is best understood as a context-dependent and institutionally embedded process. Future research will make the greatest contribution where it combines stronger identification strategies, richer measures of technological exposure and investment, and a more explicit treatment of heterogeneity, dynamics, and climate-related risk. Progress in this direction is essential not only for academic research, but also for improving the quality of investment analysis and policy design in increasingly technology-intensive economies.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. Journal of Economic Perspectives, 33(2), 3–30. https://doi.org/10.1257/jep.33.2.3

Andrews, I., Stock, J. H., & Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. Annual Review of Economics, 11, 727–753. https://doi.org/10.1146/annurev-economics-080218-025643

Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. Quarterly Journal of Economics, 135(2), 645–709. https://doi.org/10.1093/qje/qjaa004

Bloom, N., Jones, C. I., Van Reenen, J., & Webb, M. A. (2020). Are ideas getting harder to find? American Economic Review, 110(4), 1104–1144. https://doi.org/10.1257/aer.20180338

Bolton, P., & Kacperczyk, M. (2023). Global pricing of carbon-transition risk. Journal of Finance, 78(6), 3617–3667. https://doi.org/10.1111/jofi.13272

D'Odorico, P., Davis, K. F., Rosa, L., Carr, J. A., & Chiarelli, D. D. (2018). The global food-energy-water nexus. Reviews of Geophysics, 56(3), 456–531. https://doi.org/10.1029/2017rg000591

De Loecker, J., Eeckhout, J., & Unger, G. (2020). The rise of market power and the macroeconomic implications. Quarterly Journal of Economics, 135(2), 561–644. https://doi.org/10.1093/qje/qjz041

Dechezleprêtre, A., & Sato, M. (2017). The impacts of environmental regulations on competitiveness. Review of Environmental Economics and Policy, 11(2), 183–206. https://doi.org/10.1093/reep/rex013

Demir, A., Pesqué-Cela, V., Altunbaş, Y., & Murinde, V. (2020). Fintech, financial inclusion and income inequality: A quantile regression approach. European Journal of Finance, 28(1), 86–107. https://doi.org/10.1080/1351847x.2020.1772335

Donaldson, D. (2018). Railroads of the Raj: Estimating the impact of transportation infrastructure. American Economic Review, 108(4–5), 899–934. https://doi.org/10.1257/aer.20101199

Duque Grisales, E. A., & Aguilera-Caracuel, J. (2019). 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

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., & Coombs, C. (2019). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Foley, S., Karlsen, J. R., & Putniņš, T. J. (2019). Sex, drugs, and Bitcoin: How much illegal activity is financed through cryptocurrencies? Review of Financial Studies, 32(5), 1798–1853. https://doi.org/10.1093/rfs/hhz015

Goldfarb, A., & Tucker, C. E. (2019). Digital economics. Journal of Economic Literature, 57(1), 3–43. https://doi.org/10.1257/jel.20171452

Krueger, P., Sautner, Z., & Starks, L. T. (2019). The importance of climate risks for institutional investors. Review of Financial Studies, 33(3), 1067–1111. https://doi.org/10.1093/rfs/hhz137

Mian, A., Sufi, A., & Verner, E. (2017). Household debt and business cycles worldwide. Quarterly Journal of Economics, 132(4), 1755–1817. https://doi.org/10.1093/qje/qjx017

Sautner, Z., van Lent, L., Vilkov, G., & Zhang, R. (2023). Firm-level climate change exposure. Journal of Finance, 78(3), 1449–1498. https://doi.org/10.1111/jofi.13219

Sovacool, B. K., Axsen, J., & Sorrell, S. (2018). Promoting novelty, rigor, and style in energy social science: Towards codes of practice for appropriate methods and research design. Energy Research & Social Science, 45, 12–42. https://doi.org/10.1016/j.erss.2018.07.007

Tol, R. S. J. (2017). The economic impacts of climate change. Review of Environmental Economics and Policy, 12(1), 4–25. https://doi.org/10.1093/reep/rex027

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, 31(7), 3371–3387. https://doi.org/10.1002/bse.3089

K-Means and Hierarchical Clustering in Business Research: A Systematic Synthesis of Research

Dana Hrušovská[1] and Peter Štetka[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: dana.hrusovska@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: peter.stetka@euba.sk

Abstract

Clustering algorithms, particularly K-Means and hierarchical methods, have become foundational analytical tools in business research, enabling organizations to extract actionable intelligence from complex, high-dimensional datasets. This literature review examines the theoretical underpinnings, methodological applications, and empirical outcomes of these clustering techniques within business-oriented contexts, including customer segmentation, marketing strategy, financial analytics, and organizational research. Drawing upon seventeen peer-reviewed sources, the study synthesizes evidence across three principal thematic clusters: algorithmic foundations and performance evaluation, domain-specific business applications, and the integration of clustering with advanced machine learning pipelines. Key findings indicate that K-Means, enhanced by validation metrics such as the Silhouette coefficient and the elbow method, consistently delivers interpretable and computationally efficient segmentation outcomes, while hierarchical approaches offer complementary depth for exploratory analysis. However, significant challenges persist regarding the handling of mixeddata compatibility, high-dimensional scalability, and the standardization of cluster validation frameworks. This review identifies critical research gaps and proposes future directions centered on hybrid algorithm design, longitudinal clustering frameworks, and the ethical deployment of clustering in customer-facing business environments. The findings carry direct implications for researchers and practitioners seeking to leverage data-driven segmentation in strategic decision-making.

Key Words: business analytics, hierarchical clustering, K-Means clustering, , market segmentation, unsupervised learning.

JEL Classification: C38, M31, C55

1. Introduction

The proliferation of digital data across modern business ecosystems has fundamentally transformed how organizations understand their customers, competitors, and operational environments. In an era defined by exponential growth in transactional records, behavioral logs, and unstructured textual sources, the capacity to identify meaningful patterns within large datasets has become a critical competitive advantage. Clustering—the unsupervised partitioning of data objects into homogeneous groups based on intrinsic similarity—has emerged as one of the most widely adopted analytical paradigms in business intelligence, marketing analytics, and strategic management research. Among the diverse family of

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

clustering methodologies, K-Means and hierarchical clustering occupy a particularly prominent position, owing to their conceptual interpretability, widespread availability in analytical software environments, and demonstrated empirical utility across diverse business domains.

Despite this widespread adoption, the literature on clustering in business research remains fragmented across disciplinary silos. Advances in algorithmic development—including K-Means++, kernel K-Means, and density-based extensions—have outpaced their systematic integration into business research methodology frameworks. Simultaneously, the emergence of mixed-data environments, wherein numerical and categorical variables coexist within a single dataset, presents unresolved challenges that conventional K-Means implementations are ill-equipped to address without preprocessing intervention (Ahmad & Khan, 2019). Furthermore, as clustering is increasingly embedded within hybrid machine learning pipelines that incorporate dimensionality reduction, deep learning, and predictive modeling, the standalone theoretical contributions of K-Means and hierarchical clustering to business outcomes risk being obscured.

A clear research gap exists in the form of a consolidated, thematically organized review that critically synthesizes how K-Means and hierarchical clustering have been theorized, implemented, evaluated, and extended within business research contexts. While individual studies have addressed customer segmentation (Wu et al., 2020; Kasem et al., 2023), risk profiling in tourism (Wang et al., 2019), and text mining in organizational settings (Kobayashi et al., 2017), no integrative review has systematically examined the intersection of these clustering methodologies with business research objectives across multiple functional domains.

This article addresses that gap by presenting a structured literature review with three principal objectives: first, to establish the theoretical foundations of K-Means and hierarchical clustering as applied in business research; second, to synthesize empirical findings across thematic domains including customer segmentation, organizational analytics, and hybrid algorithmic applications; and third, to identify persistent research gaps and articulate a forward-looking agenda for future inquiry. The remainder of the article is organized as follows. Section 2 presents the theoretical background, defining key constructs and situating clustering methods within broader machine learning and business analytics frameworks. Section 3 synthesizes the literature thematically. Section 4 identifies research gaps and future directions, and Section 5 concludes the review.

2. Theoretical Background

2.1 Conceptual Foundations of Clustering

Clustering is broadly defined as the process of organizing a collection of objects into groups—referred to as clusters—such that objects within the same cluster exhibit greater similarity to one another than to objects belonging to other clusters. This general objective is shared across the diverse family of clustering algorithms; however, the precise definitions of similarity, cluster structure, and optimization criterion vary considerably across methodological paradigms. As Chaudhry et al. (2023) observe in their systematic review of unsupervised clustering algorithms, clustering methods can be taxonomically organized along several axes, including partitional versus hierarchical structures, distance-based versus density-based similarity metrics, and deterministic versus probabilistic assignment mechanisms.

Within the partitional paradigm, K-Means stands as the canonical algorithm. Proposed by Lloyd in the 1950s and formalized in subsequent decades, K-Means operates by iteratively

assigning each data point to the nearest centroid and subsequently recomputing centroids as the mean of all assigned points, continuing until convergence is achieved. The algorithm minimizes within-cluster sum of squared distances, a criterion that implicitly assumes clusters are convex, isotropic, and of approximately equal size. Shutaywi and Kachouie (2021) provide a rigorous characterization of K-Means as a fundamentally linear clustering method—one that partitions objects into linearly separable groups—and contrast it with kernel K-Means, a non-linear extension that projects data into a higher-dimensional feature space through a kernel function, thereby enabling the identification of non-convex cluster structures. This distinction is theoretically significant in business research contexts where customer behavioral data often exhibits non-linear distributional characteristics.

2.2 Hierarchical Clustering

Hierarchical clustering differs fundamentally from K-Means in that it does not require the pre-specification of the number of clusters. Instead, it generates a nested sequence of partitions, typically visualized as a dendrogram, which can be traversed at varying levels of granularity to produce cluster solutions of different cardinalities. The two primary variants—agglomerative (bottom-up) and divisive (top-down)—differ in their construction logic. Agglomerative methods initiate with each data point as a singleton cluster and iteratively merge the most similar pairs, while divisive methods begin with a single all-encompassing cluster and recursively partition it. The choice of linkage criterion—single, complete, average, or Ward's method—substantially influences the shape and interpretability of the resulting dendrogram (Chaudhry et al., 2023). In business research, hierarchical clustering is frequently employed as an exploratory complement to K-Means, enabling analysts to visualize the hierarchical structure of customer segments before committing to a fixed partition.

2.3 Cluster Validation Metrics

A critical theoretical concern in clustering research is the evaluation of cluster quality in the absence of ground-truth labels. Shutaywi and Kachouie (2021) provide a comprehensive treatment of Silhouette analysis as a performance evaluation metric for clustering, demonstrating that the Silhouette coefficient—which measures how similar a data point is to its own cluster relative to other clusters—offers a theoretically grounded and computationally practical instrument for comparing clustering solutions across different values of K. Complementing the Silhouette coefficient, the Elbow method provides an intuitive graphical heuristic for identifying the optimal number of clusters by examining the rate of decrease in within-cluster variance as K increases. Kasem et al. (2023) employ a triangulated validation framework incorporating the Elbow method, the Silhouette coefficient, and the Gap Statistic in the context of customer segmentation using RFM data, illustrating how multiple validation metrics can be used in concert to improve the robustness of cluster selection decisions.

2.4 Mixed Data Clustering and Dimensionality Reduction

A foundational challenge in business research is the prevalence of mixed datasets, wherein numerical features such as transaction values and categorical features such as customer demographics or product categories coexist. Standard K-Means relies on Euclidean distance and arithmetic mean operations that are mathematically inappropriate for categorical variables, necessitating either feature encoding strategies or the adoption of alternative algorithms. Ahmad and Khan (2019) present a comprehensive taxonomy of mixed data clustering algorithms, identifying five major research themes and cataloguing the diverse strategies—including dissimilarity measures, latent variable models, and hybrid distance

functions—that have been proposed to address this challenge. Related to this, dimensionality reduction techniques, including Principal Component Analysis and deep autoencoders, have been employed as preprocessing steps to render high-dimensional business datasets amenable to K-Means clustering. Alkhayrat et al. (2020) demonstrate this approach in a telecommunications customer segmentation study, comparing the quality of K-Means clustering solutions in the original feature space, a PCA-reduced space, and a deep learning-derived latent space across a dataset of 100,000 customers with 220 features. These contributions collectively establish the theoretical infrastructure upon which applied business clustering research is built.

2.5 Clustering Within Machine Learning Pipelines

Increasingly, K-Means and hierarchical clustering are not deployed as standalone analytical tools but as components within larger, multi-stage machine learning pipelines. Syed et al. (2021) illustrate this integration in the domain of smart grid energy management, where clustering is used to group energy consumption patterns prior to the training of deep learning forecasting models, thereby improving prediction accuracy. Drogkoula et al. (2023) similarly document the use of clustering as a preprocessing and pattern discovery mechanism within broader machine learning applications across environmental and resource management contexts. This pipeline-oriented perspective is theoretically important because it reframes the evaluation criteria for clustering from standalone cluster quality to downstream task performance, introducing new methodological considerations that are directly relevant to business research applications.

3. Systematic Synthesis of Research

3.1 Customer Segmentation and RFM-Based Clustering

The most extensively documented application of K-Means clustering in business research is customer segmentation, particularly within frameworks that operationalize customer value through Recency, Frequency, and Monetary (RFM) analysis. RFM models reduce complex purchasing behavior to three interpretable dimensions, creating a structured feature space that is well-suited to K-Means partitioning. Wu et al. (2020) conduct an empirical investigation of customer segmentation in an e-commerce enterprise using a combined RFM model and K-Means algorithm, clustering customers into four behaviorally distinct groups and deriving differentiated Customer Relationship Management (CRM) strategies tailored to each segment. The study demonstrates measurable improvements in key performance indicators following the implementation of segment-specific marketing interventions, providing direct empirical evidence for the business utility of K-Means-based segmentation.

Building upon this methodological tradition, Kasem et al. (2023) extend the RFM-based K-Means framework by integrating artificial intelligence tools for customer profiling in the context of direct marketing. Their study employs a multi-metric validation approach—incorporating the Elbow method, the Silhouette coefficient, and the Gap Statistic—to determine the optimal number of clusters, and subsequently develops predictive sales models based on segment membership. A notable contribution of this work is its explicit attention to the validation challenge that plagues much of the applied clustering literature: by triangulating across multiple evaluation metrics, the authors establish a more defensible basis for their chosen partition than studies relying on a single heuristic criterion. This reflects a broader methodological maturation in the field, consistent with the theoretical framework articulated

by Shutaywi and Kachouie (2021), who argue that Silhouette analysis provides a principled and universally applicable mechanism for comparing clustering algorithms and their outputs.

While the convergence between RFM analysis and K-Means is well-established, the literature reveals an important debate regarding the adequacy of Euclidean distance in capturing behavioral similarity. Customer datasets in business research routinely include categorical variables—gender, product category preferences, channel of interaction—that violate the mathematical assumptions underlying standard K-Means. Ahmad and Khan (2019) document this limitation comprehensively, noting that many real-world business datasets exhibit the mixed-data structure that renders naive K-Means application theoretically problematic. Despite this recognized limitation, a substantial proportion of applied business clustering studies continue to employ K-Means on mixed datasets following ad hoc encoding procedures, without formally justifying these choices or evaluating the sensitivity of their segment solutions to encoding decisions. This represents both a methodological weakness in the extant literature and a signal of the practical gap between algorithmic theory and applied business research practice.

3.2 Risk Profiling, Behavioral Segmentation, and Tourism Research

Beyond the transactional customer segmentation context, clustering algorithms have been applied to understand behavioral heterogeneity in more psychologically complex business domains. Wang et al. (2019) apply cluster analysis to segment adventure tourists according to their risk perception attitudes and safety behaviors, identifying three distinct visitor groups with meaningfully different psychological profiles and behavioral tendencies. This application extends the conventional understanding of clustering in business research from a primarily commercial segmentation tool to an instrument for behavioral theory development, illustrating how cluster solutions can serve as empirical inputs to psychological and managerial frameworks.

The methodological approach in Wang et al. (2019) exemplifies a common pattern in business research: cluster analysis is employed as a means of operationalizing theoretical constructs—in this case, the Risk Perception Attitude Framework—by empirically identifying the subpopulations posited by the theory. This theory-driven clustering strategy contrasts with the more exploratory, data-driven segmentation typical of marketing analytics, and raises important questions about the appropriate role of clustering as a confirmatory versus exploratory tool. When clusters are interpreted through theoretical lenses that were developed prior to the empirical analysis, there is a risk of confirmation bias in cluster labeling—a methodological concern that has received insufficient attention in the business clustering literature.

Hierarchical clustering, while less prominent than K-Means in published business research, has been employed in organizational and behavioral contexts where the number of meaningful segments is theoretically uncertain. Kobayashi et al. (2017) document the application of text mining techniques—including hierarchical clustering of textual similarity matrices—to organizational research problems, demonstrating how unstructured textual data from organizational communications can be transformed into quantitative features amenable to cluster analysis. This work highlights an important and underexplored frontier: the integration of natural language processing with clustering algorithms to generate behavioral and attitudinal segments from textual rather than transactional data sources.

3.3 Dimensionality Reduction and Hybrid Algorithmic Pipelines in Business Contexts

A third thematic cluster in the literature concerns the embedding of K-Means and hierarchical clustering within multi-stage analytical pipelines that incorporate dimensionality reduction and deep learning components. This trend reflects both the increasing complexity of business datasets—characterized by high dimensionality, sparsity, and heterogeneous variable types—and the growing methodological sophistication of business analytics practitioners.

Alkhayrat et al. (2020) provide one of the most methodologically rigorous studies in this category, examining how dimensionality reduction through PCA versus deep autoencoders affects the quality of K-Means clustering solutions in a large-scale telecom customer dataset. Their findings indicate that clustering in deep learning-derived latent spaces yields superior cluster cohesion and separation compared to both the original feature space and PCA-reduced representations, suggesting that non-linear dimensionality reduction is an important preprocessing consideration for business clustering applications involving high-dimensional behavioral data. This finding has significant practical implications: organizations with access to rich but sparse customer interaction logs—telecommunications providers, digital platforms, financial services firms—may substantially improve their segmentation outcomes by investing in autoencoder-based preprocessing rather than relying on conventional linear reduction techniques.

The use of clustering as a preprocessing step within predictive modeling pipelines is also documented in adjacent technical domains with direct business relevance. Syed et al. (2021) demonstrate that clustering energy consumption patterns prior to deep learning-based load forecasting substantially improves prediction accuracy in smart grid applications, a finding that parallels the business analytics context wherein customer segments identified through clustering can serve as conditioning variables for subsequent predictive models targeting churn, lifetime value, or purchase propensity. Chaudhry et al. (2023) synthesize a broad body of literature on clustering as a pattern discovery tool within data mining pipelines, identifying consistent evidence that clustering-informed feature engineering improves the performance of downstream supervised learning models across multiple domains.

A recurring theme across this thematic cluster is the challenge of maintaining cluster interpretability when clustering is embedded within complex pipelines. Deep learning-derived latent representations, while producing technically superior cluster solutions, are often difficult to interpret in terms of business-meaningful variables, complicating the translation of cluster membership into actionable marketing or operational strategies. This tension between algorithmic performance and managerial interpretability represents one of the most important unresolved debates in the business clustering literature. Drogkoula et al. (2023) observe this challenge in the context of environmental data analytics, noting that the opacity of machine learning-derived cluster solutions frequently impedes their adoption by domain practitioners—a barrier that is equally salient in business research and practice.

3.4 Methodological Debates and Cross-Cutting Issues

Across the three thematic clusters identified above, several cross-cutting methodological debates merit explicit discussion. The first concerns cluster number determination. Despite decades of methodological development, there is no universally accepted criterion for determining the optimal number of clusters, and different metrics—including the Silhouette coefficient, the Elbow criterion, the Gap Statistic, and the Calinski-Harabász index—frequently disagree, particularly in business datasets with ambiguous cluster structure (Shutaywi &

Kachouie, 2021; Kasem et al., 2023). This pluralism of validation criteria is simultaneously a source of methodological flexibility and a generator of inconsistency across studies, rendering cross-study comparison of cluster solutions difficult.

A second cross-cutting issue concerns algorithmic sensitivity to initialization. K-Means is well-known to be sensitive to initial centroid placement, potentially converging to local rather than global optima. Shutaywi and Kachouie (2021) document how extensions such as K-Means++ address this limitation through probabilistic initialization, though these improvements have not been universally adopted in applied business research, where vanilla K-Means implementations remain common (Wu et al., 2020; Wang et al., 2019). A third issue concerns the handling of outliers, which are prevalent in business transaction data and can disproportionately influence centroid computation in K-Means. The literature reviewed does not consistently report outlier treatment procedures, representing a transparency gap that undermines reproducibility.

Finally, the issue of mixed data compatibility, highlighted by Ahmad and Khan (2019), pervades virtually all business applications of K-Means. The dominance of Euclidean distance-based K-Means in business research, despite the widespread availability of more appropriate mixed-data algorithms, reflects a pragmatic rather than theoretically principled approach to algorithm selection—one that the field has yet to meaningfully address through the development and adoption of standardized best practice guidelines.

4. Research Gaps and Future Directions

Despite the substantial and growing body of literature on K-Means and hierarchical clustering in business research, several significant gaps remain that warrant systematic future investigation.

The first and perhaps most consequential gap concerns the treatment of mixed data in real-world business datasets. As Ahmad and Khan (2019) document extensively, mixed datasets—containing both numerical and categorical features—are the norm rather than the exception in business research contexts ranging from customer databases to employee records to product catalogs. Yet the vast majority of published business clustering studies continue to employ K-Means with Euclidean distance on datasets that have been pre-processed through ad hoc encoding procedures, without empirically evaluating the sensitivity of cluster solutions to these choices. Future research should systematically compare the performance of purpose-designed mixed data clustering algorithms—such as K-Prototypes, ROCK, and latent class models—against encoded K-Means solutions using standardized business datasets with known ground-truth segment structures, generating evidence-based recommendations for practitioners.

A second critical gap concerns the temporal dimension of business clustering. The overwhelming majority of business clustering studies adopt a cross-sectional design, producing a static snapshot of customer or behavioral segments at a single point in time. However, business environments are dynamic: customer preferences evolve, competitive landscapes shift, and behavioral patterns change in response to external events. Future research should develop and validate longitudinal clustering frameworks that track the evolution of segment membership and structure over time, potentially drawing on sequence analysis, dynamic time warping, or temporal Bayesian clustering methods. Such frameworks would substantially enhance the strategic utility of clustering for CRM and customer lifecycle

management applications, enabling organizations to detect segment migration patterns and proactively adapt their marketing strategies.

A third gap concerns the interpretability and managerial accessibility of clustering solutions derived from deep learning-based hybrid pipelines. While Alkhayrat et al. (2020) demonstrate the technical superiority of autoencoder-derived clustering solutions, and Syed et al. (2021) illustrate the performance benefits of clustering-informed predictive modeling, neither study systematically evaluates how easily the resulting cluster solutions can be translated into managerially interpretable and actionable segment profiles. Future research should develop structured methodologies for bridging the gap between technically optimized cluster solutions and their business interpretation, potentially incorporating explainable AI techniques—such as SHAP values or LIME—to identify the original business-meaningful variables that best characterize each cluster in a latent representation space. Kobayashi et al. (2017) suggest that organizational researchers must develop a deeper appreciation for the interpretive challenges inherent in computationally derived classifications, a call that extends naturally to the business clustering domain.

A fourth future direction concerns the ethical implications of algorithmic customer segmentation. As clustering-based segmentation becomes more sophisticated and consequential—informing differential pricing, product offerings, and service levels—questions of fairness, discrimination, and transparency become increasingly urgent. Future research should examine whether K-Means and hierarchical clustering algorithms, as applied to business customer data, produce segments that correlate with legally protected characteristics such as race, gender, or disability status, and develop algorithmic fairness constraints appropriate for the business segmentation context. This agenda connects business analytics research to the broader societal discourse on responsible AI, an increasingly important dimension of applied machine learning research as documented by Drogkoula et al. (2023) and Chaudhry et al. (2023).

5. Conclusion

This literature review has synthesized the theoretical foundations, methodological approaches, and empirical applications of K-Means and hierarchical clustering in business research, drawing upon seventeen peer-reviewed sources spanning algorithmic development, customer analytics, organizational research, and hybrid machine learning methodologies. The review makes several substantive contributions to the understanding of clustering in business contexts.

First, the review establishes that K-Means clustering, particularly when combined with robust validation metrics such as the Silhouette coefficient and the Elbow method, constitutes a theoretically grounded and empirically productive framework for customer segmentation in business research. Second, it demonstrates that hierarchical clustering and text mining-based cluster analysis represent valuable, if underutilized, methodological complements to K-Means in organizational and behavioral research domains. Third, the review identifies the integration of clustering within multi-stage machine learning pipelines—incorporating dimensionality reduction and deep learning—as an emerging and methodologically important trend that simultaneously enhances cluster quality and complicates managerial interpretation.

The review also surfaces persistent methodological challenges, including the widespread misapplication of K-Means to mixed datasets, the absence of standardized cluster validation protocols, algorithmic sensitivity to initialization, and the interpretability deficit of

technologically advanced hybrid pipelines. These challenges collectively define a research agenda focused on methodological rigor, algorithmic appropriateness, temporal validity, and ethical responsibility in business clustering research.

This review carries certain limitations. The corpus of seventeen sources, while carefully selected for relevance and scholarly impact, does not constitute an exhaustive systematic review of the broader clustering literature; consequently, some applied business clustering studies may not be represented. Furthermore, the review does not conduct quantitative meta-analytic comparisons of clustering performance metrics across studies, owing to the heterogeneity of experimental designs and evaluation criteria in the source literature. Future reviews may address these limitations through broader corpus construction and meta-analytic methodologies. Despite these limitations, the review provides a theoretically coherent and empirically grounded synthesis that advances scholarly understanding of clustering methodology in business research.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Ahmad, A., & Khan, S. S. (2019). Survey of state-of-the-art mixed data clustering algorithms. IEEE Access, 7, 31883–31902. https://doi.org/10.1109/access.2019.2903568

Alkhayrat, M., Aljnidi, M., & Aljoumaa, K. (2020). A comparative dimensionality reduction study in telecom customer segmentation using deep learning and PCA. Journal of Big Data, 7(1), 1–23. https://doi.org/10.1186/s40537-020-0286-0

Bi, Y., Xue, B., Mesejo, P., Cagnoni, S., & Zhang, M. (2022). A survey on evolutionary computation for computer vision and image analysis: Past, present, and future trends. IEEE Transactions on Evolutionary Computation, 27(1), 2–21. https://doi.org/10.1109/tevc.2022.3220747

Chaudhry, M., Shafi, I., Mahnoor, M., Ramírez Vargas, D. L., & Ashraf, I. (2023). A systematic literature review on identifying patterns using unsupervised clustering algorithms: A data mining perspective. Symmetry, 15(9), 1679. https://doi.org/10.3390/sym15091679

Drogkoula, M., Kokkinos, K., & Samaras, N. (2023). A comprehensive survey of machine learning methodologies with emphasis in water resources management. Applied Sciences, 13(22), 12147. https://doi.org/10.3390/app132212147

Habchi, Y., Himeur, Y., Kheddar, H., Boukabou, A., & Atalla, S. (2023). AI in thyroid cancer diagnosis: Techniques, trends, and future directions. Systems, 11(10), 519. https://doi.org/10.3390/systems11100519

Kasem, M. S., Hamada, M. A., & Taj-Eddin, I. (2023). Customer profiling, segmentation, and sales prediction using AI in direct marketing. Neural Computing and Applications, 35(33), 24191–24206. https://doi.org/10.1007/s00521-023-09339-6

Kobayashi, V., Mol, S. T., Berkers, H. A., Kismihók, G., & Den Hartog, D. N. (2017). Text mining in organizational research. Organizational Research Methods, 21(3), 733–765. https://doi.org/10.1177/1094428117722619

Myagmar-Ochir, Y., & Kim, W. (2023). A survey of video surveillance systems in smart city. Electronics, 12(17), 3567. https://doi.org/10.3390/electronics12173567

Rahman, M. M., & Szabó, G. (2021). Impact of land use and land cover changes on urban ecosystem service value in Dhaka, Bangladesh. Land, 10(8), 793. https://doi.org/10.3390/land10080793

Sepasgozar, S. M. E., Karimi, R., Farahzadi, L., Moezzi, F., & Shirowzhan, S. (2020). A systematic content review of artificial intelligence and the Internet of Things applications in smart home. Applied Sciences, 10(9), 3074. https://doi.org/10.3390/app10093074

Shutaywi, M., & 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

Siddiqui, S. A., Mercier, D., Munir, M., Dengel, A., & Ahmed, S. (2019). TSViz: Demystification of deep learning models for time-series analysis. IEEE Access, 7, 67027–67040. https://doi.org/10.1109/access.2019.2912823

Syed, D., Abu-Rub, H., Ghrayeb, A., Refaat, S. S., & Houchati, M. (2021). Deep learning-based short-term load forecasting approach in smart grid with clustering and consumption pattern recognition. IEEE Access, 9, 54992–55008. https://doi.org/10.1109/access.2021.3071654

Wang, J., Liu-Lastres, B., Ritchie, B. W., & Pan, D. (2019). Risk reduction and adventure tourism safety: An extension of the risk perception attitude framework (RPAF). Tourism Management, 74, 247–257. https://doi.org/10.1016/j.tourman.2019.03.012

Wu, J., Shi, L., Lin, W.-P., Tsai, S.-B., & Li, Y. (2020). An empirical study on customer segmentation by purchase behaviors using a RFM model and K-Means algorithm. Mathematical Problems in Engineering, 2020, 8884227. https://doi.org/10.1155/2020/8884227

Xiao, Y., Lai, Y.-K., Zhang, F.-L., Li, C., & Gao, L. (2020). A survey on deep geometry learning: From a representation perspective. Computational Visual Media, 6(2), 113–133. https://doi.org/10.1007/s41095-020-0174-8

Machine Learning Applications in Strategic Management: A Systematic Literature Review

Kufelová Iveta[1] and Matterová Stanislava[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: iveta.kufelova@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: stanislava.matterova@euba.sk

Abstract

The rapid proliferation of machine learning (ML) technologies has fundamentally altered the strategic landscape for organizations across industries, prompting renewed scholarly interest in understanding how algorithmic capabilities translate into competitive advantage. This literature review examines the intersections between machine learning and strategic management, synthesizing evidence from 21 peer-reviewed sources to map the current state of knowledge and identify directions for future inquiry. The review identifies three principal thematic clusters: (1) ML-enabled strategic decision-making and competitive intelligence, (2) operational strategy and supply chain optimization, and (3) customer strategy, value creation, and ecosystem dynamics. Key findings indicate that while ML offers transformative potential for strategic planning, forecasting, and resource allocation, significant challenges persist around model interpretability, ethical governance, and the integration of algorithmic outputs into human-led strategic processes. The review further reveals a notable gap in theoretically grounded frameworks that connect ML capabilities to established constructs in strategic management. These findings carry implications for both scholars seeking to develop unified theories and practitioners navigating the adoption of AI-driven decision support systems.

Key Words: machine learning, strategic management, competitive advantage, predictive modelling, AI decision-making

JEL Classification: M20, O33, C55

1. Introduction

Strategic management, as a discipline, is fundamentally concerned with how organizations achieve and sustain competitive advantage through the effective allocation of resources, the alignment of internal capabilities with external opportunities, and the navigation of uncertainty in complex, dynamic environments. For much of its history, strategic analysis has depended upon human judgment, quantitative modeling, and structured frameworks such as the resource-based view, transaction cost economics, and dynamic capabilities. However, the emergence of machine learning as a practical and scalable technology has begun to challenge these traditions, offering new mechanisms through which organizations can process information, generate forecasts, and automate complex decision-making tasks at a scale and speed previously unimaginable.

The significance of this shift should not be understated. Dwivedi et al. (2019) observe that artificial intelligence, of which machine learning is a core constituent, offers transformative potential for the augmentation and replacement of human tasks across a wide range of

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

industrial, intellectual, and social applications, with the pace of change described as staggering. This transformation extends beyond operational efficiency into the domain of strategy itself, raising fundamental questions about how firms formulate, implement, and evaluate their strategic choices in an era of algorithmic intelligence.

Despite growing practitioner interest in the strategic applications of ML, academic scholarship has been slow to produce integrative, theoretically grounded treatments of the subject. Existing literature tends to cluster around specific functional domains such as marketing, supply chain management, or financial risk assessment, leaving underexplored the broader strategic management implications of ML adoption. Furthermore, the rapid pace of technological development means that many foundational papers were written before the full scope of ML's strategic relevance became apparent, resulting in a fragmented and sometimes contradictory body of knowledge.

This literature review seeks to address this gap by synthesizing current scholarships on machine learning applications in strategic management contexts. Specifically, it pursues three objectives: first, to map the conceptual terrain by defining relevant constructs and situating ML within established strategic management theory; second, to identify and analyze thematic clusters in the literature that illuminate different dimensions of ML's strategic role; and third, to surface research gaps and propose concrete directions for future scholarly inquiry.

The remainder of this article is structured as follows. Section 2 provides a theoretical background, defining key concepts and presenting the foundational frameworks relevant to understanding ML in strategic contexts. Section 3 presents the literature review, organized thematically across three identified clusters. Section 4 addresses research gaps and proposes future directions, and Section 5 concludes with a synthesis of the review of contributions and limitations.

2. Theoretical Background

2.1 Defining Machine Learning in Organizational Contexts

Machine learning is a subset of artificial intelligence characterized by the capacity of algorithms to learn from data, identify patterns, and improve their performance on tasks without being explicitly programmed for each scenario. Within organizational and strategic contexts, ML encompasses a broad family of techniques, including supervised learning for classification and regression, unsupervised learning for clustering and anomaly detection, and reinforcement learning for sequential decision-making. Dwivedi et al. (2019) situate these capabilities within a broader trajectory of technological transformation, arguing that AI fundamentally extends human cognitive capacity in ways that parall eland in some respects exceed earlier industrial revolutions.

The distinction between narrow and broad AI is theoretically important for strategic management scholars. Narrow AI, which encompasses most current ML applications, excels at specific, well-defined tasks such as demand forecasting or customer segmentation, whereas broader AI systems capable of general reasoning remain largely aspirational. Huang and Rust (2020) offer a particularly useful typology in this regard, differentiating between mechanical AI, which automates repetitive tasks; thinking AI, which processes data to support decisions; and feeling AI, which analyzes human emotions and interactions. This three-stage framework has direct relevance for strategic planning because it maps the graduated substitution of human cognitive labor across increasingly complex strategic functions.

2.2 Foundational Theories in Strategic Management

To understand how ML interacts with strategic management, it is necessary to briefly review the theoretical traditions that have historically governed the field. The resource-based view (RBV) holds that firms achieve competitive advantage by developing resources that are valuable, rare, inimitable, and non-substitutable. Within this framework, ML capabilities—including proprietary algorithms, unique data assets, and human capital skilled in data science—can be conceptualized as strategic resources that confer differential advantages. Treiblmaier (2018), while focused on blockchain, draws explicitly on the RBV alongside transaction cost analysis and network theory to evaluate the strategic implications of emerging technologies, providing a methodological precedent for applying established theories to new technological phenomena.

Transaction of cost economics, associated with Williamson and Coase, offers a complementary lens by focusing on how organizations minimize the costs of economic exchange through governance decisions. Machine learning has the potential to radically reduce transaction costs by automating monitoring, reducing information asymmetries, and enabling more efficient matching of supply and demand, thereby reshaping the boundaries of the firm and the structure of industries. Relatedly, the theory of ecosystems advanced by Jacobides, Cennamo, and Gawer (2018) argues that modularity enables distinct yet interdependent organizations to coordinate without hierarchical fiat, with complementarities super modular or unique driving ecosystem emergence. This framework is particularly pertinent to ML-driven platform businesses where data flows across organizational boundaries and algorithmic capabilities create network effects that reinforce competitive moats.

Dynamic capabilities theory, developed by Teece, Pisano, and Shuen, emphasizes the capacity of organizations to sense, seize, and reconfigure resources in response to environmental change. Machine learning can be understood as both an enabler and an object of dynamic capabilities: it enables superior environmental sensing through pattern recognition and predictive analytics, while simultaneously requiring organizations to develop new reconfiguration capabilities to integrate algorithmic outputs into decision-making processes. Kraus et al. (2021) situate digital transformation within which ML adoption is a central element as a force that affects not only firm-level strategy but also environmental, societal, and institutional dynamics, suggesting that dynamic capabilities frameworks must be extended to account for the systemic nature of technological change.

2.3 The Strategic Significance of Data and Interpretability

A theoretically important but sometimes underappreciated dimension of ML in strategic management is the question of data as a strategic asset. Unlike traditional physical or even intellectual assets, data exhibits non-rivalry in consumption, increasing returns to scale, and complex ownership properties. Dash et al. (2019) document these properties in the healthcare context, but their implications generalize: organizations that accumulate large, diverse, and high-quality datasets gain enduring advantages over rivals in the training and refinement of ML models.

Equally significant is the challenge of interpretability. Carvalho, Pereira, and Cardoso (2019) provide a comprehensive survey of machine learning interpretability methods and metrics, demonstrating that most high-performing ML systems remain in black boxes whose internal logic is opaque even to expert users. This opacity creates a fundamental tension with the requirements of strategic decision-making, which typically demands accountability, justification, and the capacity for human oversight. New regulations and organizational

governance norms increasingly require that algorithmic decisions be auditable and explainable, adding a compliance dimension to what was previously understood primarily as a technical challenge. For strategic management, this means that the adoption of ML cannot be treated purely as a technological decision; it must be embedded within broader frameworks of organizational governance, risk management, and stakeholder accountability.

3. Literature Review

3.1 Machine Learning for Strategic Decision-Making and Competitive Intelligence

One of the most consequential applications of machine learning in strategic management concerns its capacity to enhance the quality, speed, and scope of strategic decision-making. Traditional strategic analysis relies on periodic, largely manual processes of environmental scanning, competitive benchmarking, and scenario planning. Machine learning fundamentally disrupts this paradigm by enabling continuous, automated processing of vast information streams, generating actionable intelligence that can inform real-time strategic choices.

A particularly innovative example of this capability is provided by Sautner, van Lent, Vilkov, and Zhang (2023), who develop a method using a machine learning keyword discovery algorithm applied to earnings call transcripts to identify the attention paid by corporate participants to climate change exposures across more than 10,000 firms from 34 countries between 2002 and 2020. Their approach demonstrates that ML can extract strategically relevant signals from unstructured textual data at a scale that is infeasible through conventional qualitative analysis. Crucially, their measures predict important real outcomes including job creation in disruptive green technologies and green patenting, illustrating ML's capacity not merely to describe strategic orientations but to forecast strategic consequences. This represents a significant methodological advance for competitive intelligence, as it enables investors, regulators, and firm strategists to anticipate competitive positioning along dimensions such as net-zero transition readiness—that are not captured by traditional financial metrics.

The strategic implications of automation for labor and organizational design represent another dimension of ML's impact on decision-making. Acemoglu and Restrepo (2019) present a task-based framework demonstrating that automation displaces labor by enabling capital to perform tasks previously exclusive to human workers, while simultaneously creating new tasks through a reinstatement effect. From a strategic management perspective, this framework has profound implications for human resource strategy, organizational design, and the allocation of cognitive labor within firms. Organizations must strategically navigate the displacement-reinstatement dynamic, investing in the redeployment of human capital toward tasks where comparative advantage over algorithms remains strong tasks that require creativity, ethical judgment, relationship management, and contextual interpretation.

The interpretability challenge identified by Carvalho et al. (2019) is directly relevant to strategic decision-making contexts. When ML models are deployed to support strategic choices such as market entry decisions, merger and acquisition targeting, or portfolio optimization the inability to explain model outputs creates accountability gaps that can undermine organizational legitimacy and board-level confidence. This suggests that the adoption of interpretable ML models, even at some cost to predictive accuracy, may be strategically rational in high-stakes decision environments where explainability is valued by stakeholders.

3.2 Operational Strategy: Supply Chain Management and Risk

A second thematic cluster concerns the application of machine learning to operational strategy, with particular emphasis on supply chain management and risk mitigation. Supply chains represent some of the most data-rich and complexity-intensive operational environments that firms manage, making them natural candidates for ML-enabled optimization.

Toorajipour, Sohrabpour, Nazarpour, Oghazi, and Fischl (2020) conduct a systematic literature review identifying the contributions of AI—including machine learning to supply chain management (SCM), covering the most prevalent AI techniques and their potential applications. They find that ML techniques are most frequently applied to demand forecasting, inventory optimization, and logistics routing, and they identify significant gaps in the application of more advanced ML methods such as deep learning to SCM problems. Their review emphasizes that current research has not yet fully captured the potential of AI to transform strategic elements of SCM, such as supplier selection, risk governance, and collaborative planning across organizational boundaries. This finding points to a misalignment between the technological frontier and the strategic frontier: while the technical capabilities of ML have advanced rapidly, their deployment in the most strategically consequential aspects of SCM lags considerably.

Biryanis, Validi, Dani, and Antoniou (2018) offer a complementary perspective, reviewing supply chain risk management (SCRM) specifically and demonstrating that ML is well suited to SCRM because the field requires rapid, adaptive decision-making based on large, multidimensional data sources. They identify machine learning, expert systems, and fuzzy logic as the dominant AI paradigms applied to SCRM and highlight promise in the use of ML for risk identification the early detection of emerging supply chain disruptions before they escalate. The strategic importance of this capability was dramatically illustrated during the COVID-19 pandemic, when firms with superior risk sensing capabilities were able to adapt more rapidly to supply chain shocks. Biryanis et al. (2018) note that the literature has focused disproportionately on specific risk categories such as demand uncertainty, while neglecting broader systemic risks that require more holistic analytical frameworks.

The relationship between blockchain technology and ML in supply chain contexts is also strategically significant. Tremblier (2018) argues that blockchain can reduce information asymmetries and enforcement costs in supply chains, complementing ML's analytical capabilities with an immutable, decentralized record of transactions. When ML is deployed in conjunction with blockchain, the resulting system can simultaneously optimize supply chain performance and enforce the governance structures that make optimization trustworthy—a combination that addresses both the strategic and the institutional dimensions of supply chain management. This integration points toward a convergent technology architecture in which ML serves as the intelligence layer, and blockchain provides the integrity layer of a unified strategic infrastructure.

Digital twin technology, explored by Rasheed, San, and Kvamsdal (2020), represents another operationally significant application of ML in strategic management. A digital twin is defined as a virtual representation of a physical asset enabled through data and simulators for real-time prediction, optimization, monitoring, and improved decision-making. By creating high-fidelity digital replicas of physical assets, supply chains, or even entire organizational processes, ML-powered digital twins enable strategic planners to simulate the consequences of different decisions before committing real resources, substantially reducing the cost and risk of strategic experimentation. The authors identify advances in computational pipelines, AI, and big data

management as key enablers of digital twin deployment, suggesting that the strategic value of this technology will continue to increase as underlying ML capabilities mature.

3.3 Customer Strategy, Value Creation, and Ecosystem Dynamics

The third thematic cluster encompasses the application of ML to customer strategy, including customer lifetime value (CLV) forecasting, segmentation, and retention, as well as the broader ecosystem dynamics through which ML-enabled platform businesses create and capture value.

Akter, Roy, Rahman, Mohona, and Ara (2025) examine the integration of AI-driven ML algorithms with traditional RFM (Recency, Frequency, Monetary) analysis to improve CLV forecasting accuracy and enable more personalized customer engagement strategies. Their study demonstrates that ML models particularly ensemble methods and gradient boosting algorithms substantially outperform heuristic RFM approaches in predicting long-term customer profitability, enabling firms to allocate marketing and retention resources with greater precision. The strategic implications of improved CLV forecasting are considerable: firms that can accurately identify high-lifetime-value customers and predict churn risk are better positioned to optimize customer acquisition costs, personalize offerings, and design loyalty programs that reinforce competitive differentiation. Akter et al. (2025) further argue that the integration of ML with RFM analysis preserves the interpretability advantages of the simpler framework while capturing the predictive power of more complex algorithms a hybrid approach that addresses the interpretability-accuracy tradeoff identified by Carvalho et al. (2019).

Huang and Rust (2020) extend this analysis to the broader domain of marketing strategy, proposing a three-stage framework in which mechanical AI automates repetitive marketing functions, thinking AI processes data to support strategic decisions, and feeling AI analyzes emotional and relational dynamics between firms and customers. This framework is significant for strategic management because it maps ML's capabilities onto the full strategic marketing planning cycle from market research and segmentation to targeting and positioning thereby providing a roadmap for the graduated integration of ML into the strategic marketing function. The authors argue that AI-driven personalization and emotional intelligence capabilities will progressively shift competitive advantage toward firms that can deploy these technologies at scale while maintaining authentic human connection with customers.

The ecosystem perspective developed by Jakobides et al. (2018) provides an important theoretical counterpoint to firm-level analyses of ML strategy. Their argument that modularity and complementarity drive ecosystem emergence has direct implications for understanding how ML-enabled platforms create value: platforms that develop modular architectures enabling third-party developers to contribute complementary applications create compounding network effects that reinforce market dominance. This dynamic is observable in major AI platform businesses where the accumulation of data itself a product of ecosystem participation—continuously improves ML model performance, creating a self-reinforcing cycle of competitive advantage. Jacobides et al. (2018) distinguish between super modular complementarities, in which the value of combining components exceeds the sum of their individual values, and unique complementarities, in which specific combinations create irreplaceable value. Both types of complementarities are evident in ML ecosystems, suggesting that firms seeking to build or participate in such ecosystems must develop nuanced strategies for identifying and securing the most strategically valuable complementary relationships.

The smart agriculture context explored by Ayaz, Ammad Uddin, Sharif, Mansour, and Aggoune (2019) provides an illustrative case of ML-driven value creation in an unexpected strategic domain. They demonstrate how IoT-based technologies combined with data analytics and ML have transformed agriculture from a statistically informed industry to a quantitatively precise one, creating new strategic opportunities along a range of value chain dimensions from precision planting to predictive pest management. While not primarily a strategic management study, this work illustrates the pervasive reach of ML as a source of strategic transformation across diverse industries, reinforcing the argument that strategic management scholarship must engage with sector-specific ML applications to capture the full scope of the phenomenon.

4. Research Gaps and Future Directions

Despite the substantial and growing body of literature reviewed in the preceding sections, several significant gaps remain that limit the field's theoretical coherence and practical guidance.

First, there is a conspicuous absence of unified theoretical frameworks that systematically connect ML capabilities to established strategic management constructs. Existing work tends to either import ML techniques into specific functional domains such as supply chain or marketing without engaging deeply with strategy theory, or to theorize about digital transformation at a level of abstraction that does not connect to the specific mechanisms through which ML creates competitive advantage. Future research should develop mid-range theories that bridge ML capabilities and strategic outcomes, drawing on the RBV, dynamic capabilities, and ecosystem theory while extending these frameworks to accommodate the distinctive properties of algorithmic resources particularly their scalability, reproducibility, and dependence on data network effects. The ecosystem framework of Jacobides et al. (2018) and the automation framework of Acemoğlu and Restrepo (2019) offer promising conceptual building blocks for such theoretical development.

Second, the literature has largely neglected the organizational and governance dimensions of ML adoption in strategic management. The interpretability challenge documented by Carvalho et al. (2019) has profound implications for board-level governance, strategic leadership, and organizational design, yet these implications remain theoretically underexplored. Future research should examine how organizations develop governance structures including AI ethics committees, algorithmic auditing processes, and human-in-the-loop decision protocols that enable them to capture the strategic benefits of ML while managing the accountability, fairness, and reputational risks associated with black-box decision-making. Longitudinal case studies of organizations navigating this governance challenge would be particularly valuable.

Third, there is limited research on the strategic management implications of ML in emerging economies and non-Western institutional contexts. Most of the empirical work draws data from advanced industrial economies, and the findings may not generalize contexts characterized by weaker data infrastructure, different regulatory environments, or distinct competitive dynamics. Sautner et al. (2023) multi-country dataset, covering more than 34 countries, represents a partial exception, but even this work focuses primarily on financial markets rather than firm-level strategy. Future research should examine how institutional context moderates the strategic value of ML adoption and should develop tailored frameworks for organizations operating in data-scarce or regulation-constrained environments.

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

A fourth and emerging direction concerns the intersection of ML strategy and environmental sustainability. The method developed by Sautner et al. (2023) for measuring climate change exposure demonstrates that ML can generate strategically actionable intelligence about environmental risks and opportunities including the prediction of green technology investment and patenting. As firms face increasing pressure to align their strategies with net-zero transition pathways, ML-enabled environmental intelligence will become an increasingly important source of competitive differentiation. Research that integrates ML capabilities with the strategic management of sustainability transitions represents a frontier of considerable theoretical and practical importance.

5. Conclusion

This literature review has synthesized a diverse and growing body of scholarship to map the intersections between machine learning and strategic management. Organized around three thematic clusters strategic decision-making and competitive intelligence, operational strategy and supply chain management, and customer strategy and ecosystem dynamics the review has demonstrated both the considerable progress that has been made and the significant theoretical and empirical gaps that remain.

The key findings of this review can be summarized as follows. Machine learning offers organizations transformative capabilities for processing information, generating forecasts, and automating complex decisions at a scale and speed that fundamentally alter the economics of strategic planning. Applications range from competitive intelligence derived from unstructured text data (Sautner et al., 2023) to supply chain risk management (Biryanis et al., 2018), demand forecasting (Toorajipour et al., 2020), and customer lifetime value optimization (Akter et al., 2025). However, these capabilities are not uniformly accessible or equally beneficial; their strategic value depends critically on data availability, organizational governance capacity, and the ability to interpret and act on algorithmic outputs responsibly (Carvalho et al., 2019).

The review also finds that established strategic management theories including the resource-based view, ecosystem theory, and dynamic capabilities provide valuable conceptual resources for understanding ML's strategic implications, but require meaningful extension to accommodate the distinctive properties of algorithmic resources and platform-mediated competition (Jacobides et al., 2018; Acemoğlu & Restrepo, 2019).

This review carries several limitations that should be acknowledged. The sources draw from a relatively broad range of disciplinary contexts, and the coverage of any single strategic domain is necessarily partial. Furthermore, the rapid pace of ML development means that some findings may already be partially superseded. Future reviews with more focused search protocols and extended citation windows will be needed to build upon the foundation established here.

Funding

This paper is funded by the EU Nex tGeneration EU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Acemoğlu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. The Journal of Economic Perspectives, 33(2), 3–30. https://doi.org/10.1257/jep.33.2.3

Akter, J., Roy, A., Rahman, S., Mohona, S., & Ara, J. (2025). Artificial intelligence-driven customer lifetime value (CLV) forecasting: Integrating RFM analysis with machine learning for strategic customer retention. Journal of Computer Science and Technology Studies, 7(1). https://doi.org/10.32996/jcsts.2025.7.1.18

Ayaz, M., Ammad Uddin, M., Sharif, Z., Mansour, A., & Aggoune, E.-H. M. (2019). Internet-of-Things (IoT)-based smart agriculture: Toward making the fields talk. IEEE Access, 7, 129551–129583. https://doi.org/10.1109/access.2019.2932609

Baryannis, G., Validi, S., Dani, S., & Antoniou, G. (2018). Supply chain risk management and artificial intelligence: State of the art and future research directions. International Journal of Production Research, 57(7), 2179–2202. https://doi.org/10.1080/00207543.2018.1530476

Carvalho, D. V., Pereira, E. M., & Cardoso, J. S. (2019). Machine learning interpretability: A survey on methods and metrics. Electronics, 8(8), 832. https://doi.org/10.3390/electronics8080832

Dash, S., Shakyawar, S. K., Sharma, L., & Kaushik, S. (2019). Big data in healthcare: Management, analysis, and future prospects. Journal of Big Data, 6(1), 54. https://doi.org/10.1186/s40537-019-0217-0

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., & Coombs, C. (2019). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Huang, M.-H., & Rust, R. T. (2020). A strategic framework for artificial intelligence in marketing. Journal of the Academy of Marketing Science, 49(1), 30–50. https://doi.org/10.1007/s11747-020-00749-9

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

Kraus, S., Jones, P., Kailer, N., Weinmann, A., & Chaparro-Banegas, N. (2021). Digital transformation: An overview of the current state of the art of research. SAGE Open, 11(3), 1–15. https://doi.org/10.1177/21582440211047576

Rasheed, A., San, O., & Kvamsdal, T. (2020). Digital twin: Values, challenges and enablers from a modeling perspective. IEEE Access, 8, 21980–22012. https://doi.org/10.1109/access.2020.2970143

Sautner, Z., van Lent, L., Vilkov, G., & Zhang, R. (2023). Firm-level climate change exposure. The Journal of Finance, 78(3), 1449–1498. https://doi.org/10.1111/jofi.13219

Toorajipour, R., Sohrabpour, V., Nazarpour, A., Oghazi, P., & Fischl, M. (2020). Artificial intelligence in supply chain management: A systematic literature review. Journal of Business Research, 122, 502–517. https://doi.org/10.1016/j.jbusres.2020.09.009

Treiblmaier, H. (2018). The impact of the blockchain on the supply chain: A theory-based research framework and a call for action. Supply Chain Management: An International Journal, 23(6), 545–559. https://doi.org/10.1108/scm-01-2018-0029

Logistic Regression as a Predictive Framework in Technology Adoption Research: A Literature Review

Štetka Peter [1] and Romanová Anita [2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: peter.stetka@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: anita.romanova@euba.sk

Abstract

Technology adoption research has increasingly relied on quantitative modeling approaches to explain and predict behavioral intentions toward emerging digital systems. Logistic regression, as a foundational statistical method for binary and multinomial outcome analysis, occupies a prominent position in this research tradition. This review synthesizes the theoretical and empirical literature examining how logistic regression has been applied within technology adoption frameworks across diverse domains, including healthcare, transportation, smart agriculture, and energy systems. Drawing on twenty-nine scholarly sources spanning machine learning, behavioral science, and applied statistics, the review identifies three principal thematic clusters: the methodological fit of logistic regression within adoption modeling, the integration of logistic regression with advanced machine learning pipelines, and the emerging challenges of interpretability, bias, and data quality in adoption-oriented predictive models. Key findings indicate that while logistic regression remains a robust baseline classifier and theoretically grounded tool for adoption research, its limitations in handling large-scale heterogeneous data have prompted hybrid modeling strategies. Critical gaps persist in longitudinal validation, fairness assessment, and domain-specific calibration. This review concludes by proposing future research directions aimed at enhancing both the rigor and practical applicability of logistic regression-based adoption studies.

Key Words: logistic regression, technology adoption prediction, binary outcome model, adoption determinants, probabilistic forecasting

JEL Classification: C35, O33, M15

1. Introduction

The study of technology adoption has grown into a rich interdisciplinary field that seeks to understand why individuals, organizations, and communities choose to embrace or resist emerging technologies. From smart farming systems to artificial intelligence-driven healthcare tools, the decision to adopt a new technology involves a complex interplay of cognitive, social, organizational, and contextual factors. Quantifying these factors and modeling their influence on adoption outcomes demands rigorous statistical methodology, and logistic regression has long served as a primary analytical instrument for this purpose. Its capacity to model the probability of a binary or categorical outcome—adoption versus non-adoption—while accommodating multiple continuous and categorical predictors makes it particularly well-suited to the empirical demands of adoption research.

Despite its widespread use, the application of logistic regression in technology adoption research is not without criticism. Questions have been raised about the adequacy of traditional

logistic models when confronted with high-dimensional data, non-linear predictor relationships, and the growing complexity of sociotechnical adoption environments. As machine learning methods have proliferated across scientific disciplines, adoption researchers have begun exploring hybrid approaches that combine the interpretability of logistic regression with the predictive power of ensemble or deep learning architectures. Yet this integration has introduced new concerns regarding model transparency, algorithmic bias, and the generalizability of findings across populations and deployment contexts.

A clear research gap exists in the literature: while individual studies have applied logistic regression to specific adoption contexts, no comprehensive review has synthesized these applications thematically, examined methodological tensions across domains, or critically assessed the evolving role of logistic regression within broader machine learning ecosystems. This literature review addresses that gap by examining how logistic regression has been conceptualized, applied, and critiqued in technology adoption research, and by identifying directions for future scholarly inquiry.

The review proceeds as follows. Section 2 establishes the theoretical background, defining key constructs and situating logistic regression within foundational adoption theories and statistical frameworks. Section 3 presents the thematic literature review, organized around three core clusters. Section 4 identifies research gaps and proposes specific future directions. Section 5 offers conclusions and notes the limitations of this review.

2. Theoretical Background

2.1 Logistic Regression: Conceptual Foundations

Logistic regression is a class of generalized linear models designed to estimate the probability that a given observation belongs to one of two or more discrete outcome categories. Unlike ordinary least squares regression, which assumes a continuous, normally distributed dependent variable, logistic regression applies a logit link function to map linear combinations of predictors onto the bounded probability interval [0, 1]. The resulting model estimates the log-odds of the outcome as a linear function of predictor variables, enabling researchers to derive odds ratios that communicate the direction and magnitude of each predictor's influence.

The theoretical elegance of logistic regression lies in its combination of statistical interpretability and inferential rigor. Coefficient significance tests, model fit indices such as the Hosmer-Lemeshow statistic, and classification performance metrics including receiver operating characteristic (ROC) curves collectively support both explanatory and predictive evaluation. These features make logistic regression not merely a classification tool but a modeling framework that aligns with the hypothesis-testing tradition central to social and behavioral science.

In the context of technology adoption, logistic regression is typically used to model binary outcomes—for example, whether a user has or has not adopted a given technology—as a function of attitudinal, normative, control, and demographic variables derived from established theoretical frameworks. The transparency of the model output is particularly valued in applied adoption research, where stakeholders require not only accurate predictions but also actionable insights into the relative importance of determinants. This aligns with the broader demand for interpretable artificial intelligence systems, a priority extensively documented across multiple domains including clinical AI systems, where model explanations are essential for practitioner trust and regulatory compliance (Tjoa & Guan, 2020; Kelly et al., 2019).

2.2 Theoretical Frameworks in Technology Adoption Research

Technology adoption research is grounded in several foundational theoretical models. The Technology Acceptance Model (TAM), originally proposed by Davis (1989), posits that perceived usefulness and perceived ease of use are the primary determinants of behavioral intention to use a technology, which in turn predicts actual system use. Extensions of TAM, including TAM2 and the Unified Theory of Acceptance and Use of Technology (UTAUT) developed by Venkatesh et al. (2003), incorporate additional constructs such as social influence, facilitating conditions, and hedonic motivation. These constructs generate continuous or ordinal predictor variables well-suited to logistic regression analysis.

The Theory of Planned Behavior (TPB), as formulated by Ajzen (1991), provides another theoretical scaffold for adoption research, emphasizing the roles of attitudes, subjective norms, and perceived behavioral control in shaping behavioral intentions. When the intended behavior is adoption of a discrete technology—either adopted or not—logistic regression serves as the natural analytical complement to TPB's conceptual structure. The alignment between theoretical predictor structure and statistical method is a central methodological principle emphasized by methodologists in energy and social science research, who have stressed that method selection must be driven by conceptual appropriateness rather than convenience or convention (Sovacool et al., 2018).

2.3 Logistic Regression Within Broader Machine Learning Ecosystems

The emergence of machine learning as a dominant paradigm for data analysis has contextually repositioned logistic regression. Rather than competing with machine learning methods, logistic regression is now frequently used as a baseline classifier against which more complex algorithms are benchmarked. This role is well-established in the machine learning for networking literature, where logistic regression is contrasted with support vector machines, random forests, and deep neural networks as part of systematic model comparison frameworks (Boutaba et al., 2018). Similarly, in intrusion detection and classification tasks, logistic regression provides a computationally efficient reference point that enables researchers to isolate the marginal contribution of model complexity (Kang & Kang, 2016).

In the context of big data environments, logistic regression faces implementation challenges related to scalability and computational efficiency. García et al. (2016) document how traditional statistical modeling pipelines require extensive preprocessing—including missing value imputation, categorical encoding, and feature scaling—before logistic regression can be reliably applied to large-scale datasets. These preprocessing requirements impose methodological overhead that has driven some adoption researchers toward automated machine learning pipelines, though often at the cost of interpretability. The tension between scale and transparency represents a persistent challenge that cuts across both the statistical and machine learning literatures and is particularly acute in adoption research, where explanatory insight is at least as valuable as predictive accuracy.

The recent growth of federated learning frameworks presents additional theoretical considerations for adoption research. Kairouz et al. (2020) outline how federated settings enable model training on decentralized data without centralizing sensitive user information, a property that could support privacy-preserving adoption research at scale. However, the statistical assumptions of logistic regression—including the assumption of identically and independently distributed observations—are challenged in federated environments where data is heterogeneous and geographically distributed, creating theoretical incompatibilities that must be addressed through algorithmic adaptation.

3. Literature Review

The literature included in this review was identified through a structured, multi-source search process. Three academic databases and discovery platforms were used: Semantic Scholar, OpenAlex, and scite.ai. The search strategy combined Boolean queries using terms such as "logistic regression," "technology adoption," "binary classification," "TAM," and "machine learning" with domain-specific qualifiers including "healthcare," "agriculture," "transport," and "energy." Sources were selected based on their relevance to the intersection of logistic regression methodology and technology adoption research, with preference given to peer-reviewed journal articles published in English. The final corpus of twenty sources was assembled through an iterative process of forward and backward citation tracking, thematic relevance assessment, and cross-referencing across the three platforms. While this process does not constitute a formal systematic review with predefined inclusion and exclusion protocols, it ensures broad disciplinary coverage and thematic coherence. The selected sources were subsequently organized into three thematic clusters, as presented in the following subsections.

3.1 Thematic Cluster One: Methodological Applications of Logistic Regression in Sectoral Adoption Studies

Logistic regression has been applied across a wide range of sectoral adoption contexts, from healthcare AI to smart agriculture, with varying degrees of methodological sophistication. In healthcare settings, the adoption of AI-driven diagnostic tools has been studied using logistic regression to identify clinical, organizational, and attitudinal predictors of uptake among practitioners. The clinical AI literature has been particularly attentive to the calibration of adoption models, with researchers emphasizing that prediction accuracy must be evaluated using metrics intuitive to clinical stakeholders rather than purely statistical criteria (Kelly et al., 2019). This observation is consequential for logistic regression applications, because it highlights the importance of translating odds ratios and predicted probabilities into clinically meaningful language rather than treating them as endpoints in themselves. A complementary perspective is provided by Shemesh and Barnoy (2020), who applied a modified TAM-based logistic regression model to assess the intention to use mobile health applications in an Israeli adult population, demonstrating that TAM constructs accounted for 51% of variance in behavioral intention and confirming the direct applicability of logistic regression to theory-driven adoption research in healthcare contexts.

In the transportation domain, logistic regression has served as a primary analytical tool for modeling the adoption of autonomous vehicles, electric vehicles, and intelligent transportation systems. Abduljabbar et al. (2019) provide an extensive overview of AI applications in transport, noting that behavioral prediction models—including those built on logistic regression—are fundamental to planning adoption trajectories and designing appropriate policy interventions. The adoption of electric vehicles, for instance, involves a binary choice structure that naturally accommodates logistic regression, while the diverse predictor landscape—encompassing range anxiety, charging infrastructure access, environmental attitudes, and income—maps cleanly onto the multivariate structure of the model.

Agriculture represents an emerging but rapidly growing application domain for technology adoption research, with smart farming technologies such as IoT-based soil sensors, automated irrigation systems, and drone-assisted crop monitoring presenting adoption decision problems with clearly defined binary outcomes. Muthumanickam et al. (2022) describe how the integration of IoT and cloud computing in precision agriculture depends critically on farmer

adoption decisions shaped by cost perceptions, technological literacy, infrastructure availability, and social network effects—all constructs theorized as logistic regression predictors in adoption models. The heterogeneity of the agricultural population, spanning subsistence farmers to large agribusiness operators, raises important questions about sample stratification and model specification that logistic regression studies must address explicitly. Groher et al. (2020) provide empirical evidence of this challenge in their study of precision agriculture adoption in Swiss plant production, where binary logistic regression revealed that geographic location, farm size, and enterprise type significantly predicted adoption of driver assistance and electronic measuring systems, confirming that structural farm-level variables operate as robust predictors within the logistic regression framework.

Work environment and organizational technology adoption has also been examined through the lens of logistic regression. Davidescu et al. (2020) analyze how flexible work arrangements and human resource management practices influence employee satisfaction and behavioral intentions in technology-mediated work environments. Their findings highlight the role of individual-level psychological and demographic variables—including age, education, and prior technology experience—as moderators of adoption likelihood, precisely the kind of covariates that logistic regression can accommodate through interaction terms and stratified analysis.

The financial technology domain constitutes another rapidly expanding application area for logistic regression in adoption research. Fan (2021) employed nested logistic regression to examine the determinants of mobile investment technology adoption among American investors, finding that objective and subjective investment literacy, prior experience with mobile banking services, and ownership of specific investment vehicles significantly predicted the adoption of mobile apps for both investment decision-making and trading. The study's use of nested models—sequentially introducing blocks of demographic, literacy, and behavioral predictors—exemplifies a methodologically rigorous application of logistic regression that allows researchers to isolate the incremental explanatory contribution of each predictor set, a design strategy particularly well suited to theory-driven adoption research grounded in the diffusion of innovations framework.

A complementary perspective on digital financial service adoption is offered by Dimitrova et al. (2021), who investigated barriers to the full adoption of digital payment methods among Swedish bank customers using ordinal regression. Their study identified privacy concerns and access limitations as significant barriers among younger adopters, while impersonalisation—the perceived absence of human warmth in digital interactions—emerged as the sole significant predictor among technology resisters. Notably, the study found that barriers exerted a stronger negative effect on adoption intention among respondents with more extensive experience of digital payment methods, a counterintuitive finding that challenges simplistic diffusion models and underscores the importance of modeling adoption as a non-linear, stage-dependent process rather than a single binary decision.

3.2 Thematic Cluster Two: Integration of Logistic Regression with Advanced Predictive Modeling Approaches

A second prominent thematic cluster in the literature concerns the integration of logistic regression with more sophisticated machine learning and artificial intelligence methods. This integration takes multiple forms, including the use of logistic regression as a baseline comparator, as a component within ensemble classifiers, and as a post-hoc interpretability

layer applied to opaque model outputs. Each form of integration reflects distinct methodological priorities and raises different research design considerations.

The use of logistic regression as a baseline in comparative modeling studies is ubiquitous. Boutaba et al. (2018) systematically survey the application of machine learning methods for networking tasks, observing that simpler models—including logistic regression—often achieve competitive performance with far greater computational efficiency than deep architectures. This finding is particularly relevant for adoption research in resource-constrained environments, such as rural agricultural communities or developing healthcare systems, where computational infrastructure may be limited. In these contexts, logistic regression offers a compelling combination of accuracy and practicality that more complex models cannot easily match.

The emergence of deep neural network architectures has posed theoretical challenges for adoption researchers who value interpretability. Kang and Kang (2016) demonstrate the superiority of deep neural networks over traditional classifiers in sequential classification tasks, yet acknowledge that the opacity of learned representations undermines the kind of explanatory accountability demanded in regulated or high-stakes decision environments. This trade-off is echoed in the explainable AI literature, where Tjoa and Guan (2020) argue that for domains requiring accountability—including healthcare, finance, and education—interpretable models such as logistic regression remain indispensable even when they sacrifice predictive accuracy. Adoption research, which aims to generate actionable policy and design recommendations rather than merely accurate predictions, shares this accountability orientation.

The construction and artificial intelligence literature reveals a parallel integration dynamic. Baduge et al. (2022) review machine learning applications in building and construction, highlighting that adoption of smart building technologies is frequently modeled using hybrid approaches that begin with logistic regression for feature selection and hypothesis testing before proceeding to ensemble methods for final prediction. This sequential modeling strategy capitalizes on logistic regression's inferential clarity in the exploratory phase while leveraging the predictive power of more flexible algorithms in the final modeling stage. A similar logic has been proposed for adoption research in energy systems, where the social and behavioral determinants of renewable energy adoption must be identified through interpretable models before they can be operationalized in large-scale deployment forecasting tools (Sovacool et al., 2018).

Federated learning architectures present a particularly provocative challenge for logistic regression-based adoption research. Kairouz et al. (2020) describe the fundamental statistical problems arising when standard logistic regression is applied to federated data, including convergence instability under non-independent and identically distributed conditions, communication overhead associated with gradient sharing, and privacy risks when model updates carry information about individual training examples. These challenges have motivated the development of differentially private logistic regression variants and federated optimization algorithms that preserve the model's interpretive structure while accommodating decentralized data architectures. For adoption researchers interested in large-scale, real-world behavioral data—such as mobile application usage logs or smart device interaction records—engaging with federated modeling frameworks may be essential to both ethical and methodological validity.

An empirically grounded comparison of logistic regression with advanced machine learning classifiers in the adoption domain is provided by Carbó-Valverde et al. (2020), who examined

the digitalization process of bank customers using random forests, conditional inference trees, support vector machines, Bayesian classifiers, and causal forests alongside logistic regression. Their random forest model achieved 88.41% prediction accuracy for online banking adoption decisions, substantially outperforming logistic regression. However, the study also demonstrated that logistic regression retained an important diagnostic role in the exploratory phase of modeling, where the identification of theoretically meaningful predictors—such as awareness of available digital services and safety perception—was prioritized over raw classification performance. The causal forest extension further confirmed the causal validity of the identified predictors, providing a methodological template for adoption researchers seeking to combine the interpretive strengths of logistic regression with the predictive power of ensemble methods.

A similar hybrid modeling strategy is illustrated by Chaurasia et al. (2021), who modeled mobile-based technology adoption among people with dementia by comparing logistic regression, k-nearest neighbors (kNN), and neural network classifiers. While kNN achieved the highest prediction accuracy (99.41%) using a combined set of psychosocial and medical features, logistic regression served a critical upstream function in the feature selection phase, reducing an initial set of 31 variables to 11 significant predictors through univariate and multivariate analysis. This two-stage workflow—logistic regression for feature identification followed by flexible classifiers for prediction—represents a pragmatic integration pattern that preserves the theoretical interpretability valued in adoption research while exploiting the superior pattern recognition capabilities of non-parametric algorithms.

3.3 Thematic Cluster Three: Interpretability, Bias, and Data Quality in Adoption-Oriented Logistic Models

The third thematic cluster addresses three interrelated concerns that have become increasingly prominent as technology adoption research has scaled to larger and more heterogeneous populations: model interpretability, algorithmic bias, and data quality. These concerns are not unique to logistic regression, but they manifest in distinctive ways when this method is applied to adoption research contexts.

Model interpretability is frequently cited as a principal advantage of logistic regression over deep learning alternatives. However, the Mosqueira-Rey et al. (2022) review of human-in-the-loop machine learning clarifies that interpretability is not a monolithic property but a multidimensional construct encompassing transparency of model structure, simulatability of predictions, and post-hoc explainability of individual decisions. Logistic regression satisfies the first dimension—its mathematical structure is fully transparent—but may not satisfy the second or third in the presence of complex interaction terms, multicollinearity, or large numbers of predictors. Adoption researchers must therefore be precise about what form of interpretability their logistic models actually provide and what forms remain absent.

Algorithmic bias represents a more fundamental threat to the validity of adoption predictions derived from logistic regression. Kelly et al. (2019) document how AI systems trained on historically biased data can perpetuate and amplify disparities in access and outcomes, a concern directly applicable to adoption models that may underrepresent marginalized or low-technology populations in their training samples. If the predictors of technology adoption are systematically measured with greater error or lesser validity for certain demographic subgroups—a common occurrence when survey instruments are developed and validated primarily on convenience samples—logistic regression coefficients will be biased in ways that are difficult to detect and correct. The epidemiological literature on observational

study bias, exemplified by Skrivankova et al.'s (2021) treatment of confounding and selection bias in Mendelian randomization studies, provides methodological vocabulary and conceptual tools that adoption researchers could profitably apply to the analysis of bias in logistic regression adoption models.

Data quality and preprocessing represent practical prerequisites for reliable logistic regression estimation. García et al. (2016) systematically document the preprocessing challenges associated with large-scale data environments, including missing values, class imbalance, inconsistent coding, and redundant or irrelevant features. Class imbalance—where adoption events are substantially rarer than non-adoption events—is particularly prevalent in studies of emerging technologies during early diffusion phases. In such contexts, standard maximum likelihood estimation for logistic regression can produce biased coefficient estimates and systematically underpredict adoption events. Resampling strategies, penalized likelihood methods, and threshold adjustment procedures can address this problem, but their application requires methodological care and explicit reporting.

The IoT security literature provides an instructive parallel. Neto et al. (2023) present a benchmark dataset for large-scale IoT attack detection, noting that classification models trained on real-time behavioral data must contend with extreme class imbalance, concept drift, and adversarial manipulation—conditions that parallel those encountered in adoption research using observational behavioral data from digital platforms. The technical solutions developed in cybersecurity classification research, including adaptive resampling, online learning, and robust feature engineering, may be transferable to adoption modeling contexts with appropriate adaptation. The wildfire management literature further illustrates how data preprocessing and feature selection decisions substantially affect the predictive performance of logistic and other classifiers when applied to complex environmental systems characterized by rare events and heterogeneous spatial data (Jain et al., 2020).

A direct illustration of how class imbalance and data quality interact in adoption-specific logistic regression research is provided by Mehrolia et al. (2020), who applied binary logistic regression to examine consumer adoption of online food delivery services during the COVID-19 pandemic. Their study demonstrates that perceived threat, product involvement, and prior usage frequency serve as significant discriminators between adopters and non-adopters, while also highlighting the practical challenges of modeling adoption decisions under crisis conditions where the binary outcome distribution shifts rapidly. This work reinforces the need for explicit attention to distributional assumptions and contextual calibration in logistic regression adoption studies. Cryptocurrency research, while ostensibly more distant from technology adoption, nevertheless offers complementary methodological insights relevant to the modeling of rare and socially sensitive behavioral outcomes. Foley et al. (2019) employ network analysis and behavioral inference methods to identify illicit technology use patterns within large transactional datasets, demonstrating that traditional survey-based adoption measures may systematically miss or misclassify certain user behaviors. This observation reinforces the importance of data source diversity in adoption modeling and suggests that logistic regression studies relying solely on self-reported survey data may suffer from measurement biases that quantitative modeling alone cannot resolve.

The tension between interpretability and predictive accuracy has been addressed more directly through post-hoc explainability frameworks. Ariza Garzón et al. (2020) assessed logistic regression alongside gradient boosting and random forest models for credit granting scoring in peer-to-peer lending, demonstrating that SHAP (SHapley Additive exPlanations) values can render complex machine learning models not only competitive with but more transparent than

logistic regression. Their analysis revealed that SHAP-based explanations captured dispersion, nonlinearity, and structural breaks in feature–outcome relationships that logistic regression's linear coefficients inherently obscure. This finding challenges the conventional assumption that logistic regression is the most interpretable option and suggests that adoption researchers should consider modern explainability tools as viable alternatives that combine accuracy with transparency, particularly in regulated domains where both properties are required.

The socioeconomic dimension of adoption bias is further illuminated by Seldal and Nyhus (2022), who examined the relationship between financial vulnerability, financial literacy, and the use of digital payment technologies in a Norwegian population. Their findings demonstrate that financially vulnerable individuals are significantly less likely to adopt digital payment methods, even after controlling for literacy levels and demographic characteristics. This result has direct implications for logistic regression-based adoption studies, as it highlights the risk that training samples drawn from convenience or digitally accessible populations may systematically underrepresent the most vulnerable segments, producing coefficient estimates that overstate adoption propensity and underestimate the magnitude of socioeconomic barriers. Adoption researchers employing logistic regression must therefore attend not only to statistical model fit but also to the representativeness and equity implications of their sampling strategies.

4. Research Gaps and Future Directions

Despite the substantial body of work reviewed in preceding sections, several significant gaps in the literature on logistic regression and technology adoption warrant explicit attention and targeted research investment.

Gap One: Longitudinal Validation of Logistic Regression Adoption Models. The overwhelming majority of logistic regression studies in adoption research employ cross-sectional survey data, capturing adoption decisions at a single point in time. This design precludes examination of how predictor-outcome relationships evolve across the technology diffusion lifecycle. Early adopters may be driven by different predictor configurations than late majority adopters, and models built on early-phase data may have poor predictive validity when applied to later phases. Future research should prioritize longitudinal panel designs that track cohorts of potential adopters over time, enabling the estimation of time-varying logistic regression models and the testing of predictive validity across adoption stages. This approach is particularly warranted in domains such as smart farming (Muthumanickam et al., 2022) and autonomous transportation (Abduljabbar et al., 2019), where adoption is occurring gradually across heterogeneous populations.

Gap Two: Fairness and Equity Auditing in Adoption Prediction Models. Current adoption research rarely subjects its logistic regression models to systematic fairness auditing—the process of examining whether predicted adoption probabilities are systematically biased against particular demographic, socioeconomic, or geographic subgroups. Given that technology adoption patterns are deeply intertwined with existing social inequalities, models that fail to detect and correct for such biases may generate recommendations that inadvertently reinforce digital divides. Drawing on frameworks developed in clinical AI evaluation (Kelly et al., 2019) and on reporting standards for observational studies (Skrivankova et al., 2021), future adoption research should establish and apply standardized fairness metrics to logistic regression adoption models as part of routine methodological practice.

Gap Three: Integration with Privacy-Preserving and Federated Modeling Frameworks. As behavioral data collected from digital platforms becomes increasingly central to adoption research, privacy concerns have become both ethically salient and legally consequential. Traditional logistic regression estimation requires centralized access to individual-level data, a requirement that is increasingly difficult to satisfy under contemporary data protection regulations. Federated learning frameworks offer a technically sophisticated solution, enabling logistic regression to be trained on decentralized data while preserving individual privacy (Kairouz et al., 2020). However, the statistical properties of federated logistic regression—including convergence behavior, coefficient bias under data heterogeneity, and the propagation of preprocessing errors across distributed clients—remain insufficiently characterized in the adoption research literature. Future work should develop and validate federated logistic regression protocols specifically designed for technology adoption survey and behavioral data.

Gap Four: Cross-Domain Benchmarking and Methodological Standardization. The absence of shared benchmarking datasets and standardized reporting conventions makes cross-study comparison of logistic regression adoption models extraordinarily difficult. Different studies use different predictor sets, different survey instruments, different sampling frames, and different model evaluation criteria, rendering cumulative knowledge synthesis inefficient. The development of open-access benchmark datasets for technology adoption research—analogous to those established in cybersecurity (Neto et al., 2023) or natural language processing (Linzen et al., 2016)—would substantially accelerate methodological progress and enable rigorous comparative evaluation of logistic regression against alternative modeling approaches.

5. Conclusion

This literature review has examined the role of logistic regression as a predictive and explanatory framework within technology adoption research, synthesizing evidence across diverse disciplinary and sectoral literatures. Three principal thematic clusters were identified and analyzed: the sectoral application of logistic regression in adoption studies spanning healthcare, transportation, agriculture, and workplace contexts; the integration of logistic regression with advanced machine learning architectures including deep neural networks and federated learning systems; and the challenges of interpretability, algorithmic bias, and data quality that affect the validity and equity of logistic regression adoption models.

The review demonstrates that logistic regression retains considerable methodological value in adoption research, primarily because its interpretability, inferential transparency, and alignment with established theoretical frameworks make it uniquely suited to the explanatory demands of the field. At the same time, the literature reveals persistent tensions between interpretability and predictive performance, between sample convenience and population representativeness, and between methodological tradition and the demands of large-scale, privacy-sensitive behavioral data environments.

The four research gaps identified in this review—longitudinal validation, fairness auditing, federated modeling integration, and cross-domain benchmarking—each represent substantive opportunities for future scholarship. Addressing these gaps will require not only methodological innovation but also greater interdisciplinary collaboration between adoption researchers, data scientists, ethicists, and domain experts in technology deployment.

Limitations. This review is limited by the available source base, which, while spanning multiple disciplines, does not constitute an exhaustive search of the adoption literature. The absence of meta-analytic quantitative synthesis means that effect size comparisons across studies remain qualitative. Furthermore, the review focuses primarily on English-language sources, potentially overlooking relevant contributions from non-Anglophone research traditions. Future reviews should employ formal systematic search protocols and quantitative synthesis methods to address these limitations comprehensively.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Abduljabbar, R., Dia, H., Liyanage, S., & Bagloee, S. A. (2019). Applications of artificial intelligence in transport: An overview. Sustainability, 11(1), 189. https://doi.org/10.3390/su11010189

Ajzen, I. (1991). The theory of planned behavior. Organizational Behavior and Human Decision Processes, 50(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Ariza Garzón, M. J., Arroyo, J., & Caparrini, A. (2020). Explainability of a machine learning granting scoring model in peer-to-peer lending. IEEE Access, 8, 64873–64890. https://doi.org/10.1109/ACCESS.2020.2984412

Baduge, S. K., Thilakarathna, S., Perera, J. S., Ar, M., et al. (2022). Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning methods and applications. Automation in Construction, 141, 104440. https://doi.org/10.1016/j.autcon.2022.104440

Boutaba, R., Salahuddin, M. A., Limam, N., Ayoubi, S., & Shahriar, N. (2018). A comprehensive survey on machine learning for networking: Evolution, applications and research opportunities. Journal of Internet Services and Applications, 9(1), 16. https://doi.org/10.1186/s13174-018-0087-2

Carbó-Valverde, S., Cuadros-Solas, P. J., & Rodríguez Fernández, F. (2020). A machine learning approach to the digitalization of bank customers: Evidence from random and causal forests. PLOS ONE, 15(10), e0240362. https://doi.org/10.1371/journal.pone.0240362

Chaurasia, P., McClean, S., & Nugent, C. (2021). Modelling mobile-based technology adoption among people with dementia. Personal and Ubiquitous Computing, 26(2), 365–384. https://doi.org/10.1007/s00779-021-01572-x

Davidescu, A. A., Apostu, S. A., Paul, A., & Cășuneanu, I. (2020). Work flexibility, job satisfaction, and job performance among Romanian employees—Implications for sustainable human resource management. Sustainability, 12(15), 6086. https://doi.org/10.3390/su12156086

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008

Dimitrova, I., Öhman, P., & Yazdanfar, D. (2021). Barriers to bank customers' intention to fully adopt digital payment methods. International Journal of Quality and Service Sciences, 14(5), 16–36. https://doi.org/10.1108/IJQSS-03-2021-0045

Fan, L. (2021). Mobile investment technology adoption among investors. International Journal of Bank Marketing, 40(1), 50–67. https://doi.org/10.1108/IJBM-11-2020-0551

Foley, S., Karlsen, J. R., & Putniņš, T. J. (2019). Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies? Review of Financial Studies, 32(5), 1798–1853. https://doi.org/10.1093/rfs/hhz015

García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J. M., & Herrera, F. (2016). Big data preprocessing: Methods and prospects. Big Data Analytics, 1(1), 9. https://doi.org/10.1186/s41044-016-0014-0

Groher, T., Heitkämper, K., & Walter, A. (2020). Status quo of adoption of precision agriculture enabling technologies in Swiss plant production. Precision Agriculture, 21(6), 1327–1350. https://doi.org/10.1007/s11119-020-09723-5

Jain, P., Coogan, S. C. P., Ganapathi Subramanian, S., Crowley, M., & Taylor, S. (2020). A review of machine learning applications in wildfire science and management. Environmental Reviews, 28(4), 478–505. https://doi.org/10.1139/er-2020-0019

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., & Bennis, M. (2020). Advances and open problems in federated learning. Foundations and Trends® in Machine Learning, 14(1–2), 1–210. https://doi.org/10.1561/2200000083

Kang, M.-J., & Kang, J.-W. (2016). Intrusion detection system using deep neural network for in-vehicle network security. PLoS ONE, 11(6), e0155781. https://doi.org/10.1371/journal.pone.0155781

Kelly, C., Karthikesalingam, A., Suleyman, M., Corrado, G. S., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. BMC Medicine, 17(1), 195. https://doi.org/10.1186/s12916-019-1426-2

Linzen, T., Dupoux, E., & Goldberg, Y. (2016). Assessing the ability of LSTMs to learn syntax-sensitive dependencies. Transactions of the Association for Computational Linguistics, 4, 521–535. https://doi.org/10.1162/tacl_a_00115

Mehrolia, S., Alagarsamy, S., & Solaikutty, V. M. (2020). Customers response to online food delivery services during COVID-19 outbreak using binary logistic regression. International Journal of Consumer Studies, 45(3), 396–408. https://doi.org/10.1111/ijcs.12630

Mosqueira-Rey, E., Hernández-Pereira, E., Alonso-Ríos, D., Bobes-Bascarán, J., & Fernández-Leal, Á. (2022). Human-in-the-loop machine learning: A state of the art. Artificial Intelligence Review, 56(4), 3005–3054. https://doi.org/10.1007/s10462-022-10246-w

Muthumanickam, D., Poongodi, C., Kumaraperumal, R., Pazhanivelan, S., & Ragunath, K. P. (2022). Smart farming: Internet of Things (IoT)-based sustainable agriculture. Agriculture, 12(10), 1745. https://doi.org/10.3390/agriculture12101745

Neto, E. C. P., Dadkhah, S., Ferreira, R., Zohourian, A., et al. (2023). CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment. Sensors, 23(13), 5941. https://doi.org/10.3390/s23135941

Seldal, M. M. N., & Nyhus, E. K. (2022). Financial vulnerability, financial literacy, and the use of digital payment technologies. Journal of Consumer Policy, 45, 281–306. https://doi.org/10.1007/s10603-022-09512-9

Shemesh, T., & Barnoy, S. (2020). Assessment of the intention to use mobile health applications using a technology acceptance model in an Israeli adult population. Telemedicine and e-Health, 26(9), 1141–1149. https://doi.org/10.1089/tmj.2019.0144

Skrivankova, V., Richmond, R. C., Woolf, B., Davies, N. M., & Hemani, G. (2021). Strengthening the reporting of observational studies in epidemiology using Mendelian randomisation (STROBE-MR): Explanation and elaboration. BMJ, 375, n2233. https://doi.org/10.1136/bmj.n2233

Sovacool, B. K., Axsen, J., & Sorrell, S. (2018). Promoting novelty, rigor, and style in energy social science: Towards codes of practice for appropriate methods and research design. Energy Research & Social Science, 45, 12–42. https://doi.org/10.1016/j.erss.2018.07.007

Tjoa, E., & Guan, C. (2020). A survey on explainable artificial intelligence (XAI): Toward medical XAI. IEEE Transactions on Neural Networks and Learning Systems, 32(11), 4793–4813. https://doi.org/10.1109/tnnls.2020.3027314

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425–478. https://doi.org/10.2307/30036540

Predictive Analytics for Innovation Management: A Structured Review of Machine Learning Approaches and Applications

Šagátová Slavka[1] and Černý Matej [2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: slavka.sagatova@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: matej.cerny@euba.sk

Abstract

Predictive analytics has emerged as a key tool for supporting decision-making under uncertainty, yet its role in innovation management remains insufficiently explored. This study examines predictive analytics from a capability-based perspective, conceptualising it as a system-dependent organisational capability shaped by the interaction between data, algorithms, and organisational processes. The analysis synthesises evidence from nineteen studies across multiple domains to identify patterns in algorithmic approaches, data structures, and application contexts. The findings indicate that predictive analytics is not a universally applicable solution, but a context-dependent capability whose effectiveness depends on alignment between data quality, modelling techniques, and organisational integration. The study further shows that improvements in predictive accuracy do not automatically translate into improved decision-making outcomes. Key constraints include data limitations, interpretability challenges, and difficulties in embedding predictive models into organisational workflows. The results highlight the importance of aligning technical and organisational dimensions to realise the potential of predictive analytics in innovation management.

Key Words: predictive analytics, innovation management, machine learning, data-driven decision-making, organisational capability

JEL Classification: O31, C45, M10

1. Introduction

The ability to anticipate future technological, financial, and operational outcomes is a critical factor influencing the effectiveness of innovation management. Organisations do not gain competitive advantage solely by generating new ideas, but by systematically selecting and developing those initiatives that are most likely to succeed. This makes forward-looking evaluation an essential component of innovation-related decision-making.

Predictive analytics has therefore gained increasing importance as a means of supporting such decisions. By applying statistical methods, machine learning, and deep learning techniques, predictive analytics enables organisations to derive forward-looking insights from historical data. The growing availability of large and complex datasets, combined with advances in computational frameworks, has significantly expanded the applicability of these methods across domains (Nguyen et al., 2019; Rahmani et al., 2021). Empirical studies show that machine

learning approaches are capable of identifying non-linear relationships and improving predictive performance in areas such as healthcare, finance, and education (Khan et al., 2021; Ali et al., 2022; Albreiki et al., 2021). These developments suggest that predictive analytics can enhance decision-making by reducing uncertainty and improving the evaluation of future outcomes.

However, despite strong evidence of effectiveness in domain-specific applications, the role of predictive analytics in innovation management remains insufficiently examined. Existing research is predominantly focused on specific use cases (such as fraud detection, disease diagnosis, or student performance prediction) where prediction tasks are relatively well-defined and supported by structured datasets (Ali et al., 2022; Rahmani et al., 2021; Yağcı, 2022). In contrast, innovation processes are characterised by higher levels of uncertainty, evolving data structures, and strong organisational embedding of decisions. These differences imply that the transfer of predictive analytics from domain-specific applications to innovation management is not straightforward.

The literature therefore exhibits a gap between the demonstrated technical capabilities of predictive analytics and its application in innovation-related decision contexts. While machine learning models can achieve high predictive accuracy, their effectiveness in innovation management depends on additional factors, including data availability, interpretability, and integration into organisational processes. Studies addressing these aspects remain fragmented, and there is limited synthesis of how predictive analytics performs under the conditions that define innovation environments.

This study addresses this gap by examining predictive analytics as a system-dependent organisational capability rather than as an isolated analytical tool. It adopts a capability-based perspective to analyse how algorithmic approaches, data characteristics, and organisational conditions jointly influence predictive effectiveness. By synthesising evidence from nineteen studies, the paper provides a structured understanding of how predictive analytics supports decision-making in innovation management while identifying the constraints that limit its practical impact.

The remainder of the paper is organised as follows. The next section outlines the theoretical background, followed by the methodology of the literature review. The subsequent sections synthesise the selected studies, identify key research gaps, and conclude with the main findings and their implications.

2. Theoretical Background

The growing use of predictive analytics in organisational settings has led to an increasing need to understand its role beyond purely technical applications. Rather than treating predictive models as isolated tools, this section situates predictive

analytics within broader organisational and data-driven contexts. The focus is placed on the conditions under which predictive insights are generated, interpreted, and integrated into decision-making in environments characterised by uncertainty and complexity.

2.1 Predictive Analytics as a Data-Dependent Analytical Capability

Predictive analytics can be conceptualised not merely as a set of computational techniques, but as an analytical capability that enables organisations to generate knowledge about future outcomes. Unlike traditional analytical approaches based on predefined rules or deterministic models, predictive analytics relies on learning from historical data to identify statistical regularities and latent patterns. This shift toward data-driven inference is particularly relevant in environments characterised by complexity, non-linearity, and uncertainty, where explicit modelling of relationships is difficult or infeasible (Nguyen et al., 2019; Rahmani et al., 2021).

At a methodological level, predictive analytics encompasses statistical modelling, machine learning, and deep learning approaches. What distinguishes these methods is not only their technical design, but their reliance on data as the primary source of knowledge generation. Machine learning models derive predictive power from patterns embedded in data rather than from theoretical assumptions specified ex ante (Khan et al., 2021; Nguyen et al., 2019). This implies that predictive analytics should be understood as inherently data-dependent, with its effectiveness determined by the quality, representativeness, and structure of available datasets (Rahmani et al., 2021).

The distinction between supervised, unsupervised, and semi-supervised learning reflects different modes through which this capability operates. Supervised learning dominates empirical applications due to its alignment with classification and regression tasks relevant for decision-making (Albreiki et al., 2021; Yağcı, 2022). However, its performance depends on the availability of labelled data, which may be limited in early-stage or exploratory contexts. Unsupervised approaches, in contrast, enable pattern detection in unstructured environments, suggesting their relevance in situations where innovation outcomes are not yet clearly defined. As demonstrated by Nguyen et al. (2019), advances in computational infrastructure have expanded the applicability of these approaches, allowing organisations to process increasingly large and complex datasets.

From this perspective, predictive analytics represents a capability that transforms historical data into probabilistic insights. Its value lies not only in predictive accuracy, but in its ability to support decision-making under uncertainty. However, this capability is inherently constrained by data conditions, reinforcing the importance of data governance, preprocessing, and feature engineering as integral components of predictive systems (Rahmani et al., 2021).

2.2 Algorithmic Effectiveness as a Function of System Alignment

The effectiveness of predictive analytics cannot be attributed to individual algorithms in isolation, but must be understood as a function of alignment between data characteristics, modelling techniques, and analytical objectives. The literature consistently shows that no single algorithm performs optimally across all contexts, indicating that predictive performance is conditional rather than universal (Nguyen et al., 2019; Tanha et al., 2020).

Ensemble methods provide a clear example of this system-dependent effectiveness. By combining multiple models, they reduce variance and improve generalisation, particularly in environments characterised by noise and class imbalance. Empirical evidence demonstrates that boosting techniques significantly enhance performance in multi-class imbalanced datasets (Tanha et al., 2020), a finding supported by studies showing the limitations of standard classifiers under such conditions (Alam et al., 2020). These results suggest that algorithmic effectiveness emerges from the ability to adapt to data structure rather than from inherent methodological superiority.

Deep learning approaches further illustrate the importance of system alignment. Neural architectures are capable of modelling complex, high-dimensional relationships, enabling improved predictive performance in domains such as drug discovery and financial risk analysis (Askr et al., 2022; Alsaadi et al., 2024). However, these advantages depend on the availability of large datasets, computational resources, and appropriate feature representation (Rahmani et al., 2021; Khan et al., 2021). Without these conditions, performance gains may be limited.

These findings support a system perspective in which predictive analytics is understood as an integrated capability. The choice of algorithm must be aligned with data properties and the specific prediction task, while organisational context influences how models are deployed and used. Consequently, predictive performance should be interpreted as the outcome of interactions between multiple components rather than as a property of individual models.

2.3 Predictive Analytics within Organisational Learning Processes

Predictive analytics can also be interpreted as a mechanism embedded within organisational learning processes. Organisations accumulate knowledge through experience, experimentation, and feedback, which informs future decision-making. Machine learning models extend this process by encoding patterns derived from historical data, enabling organisations to formalise and operationalise experiential knowledge (Shaukat et al., 2020; Wu et al., 2020).

Empirical evidence from domains such as cybersecurity demonstrates that predictive systems can identify emerging threats by learning from past data, thereby supporting adaptive responses to dynamic environments (Shaukat et al., 2020). A similar mechanism can be observed in other domains where predictive analytics is used to detect patterns, classify

outcomes, and anticipate future events (Rahmani et al., 2021; Khan et al., 2021). This suggests that predictive analytics contributes to organisational learning by enhancing the ability to extract actionable insights from data. However, this contribution is not automatic. The integration of predictive insights into organisational decision-making depends on the ability to interpret model outputs and align them with strategic objectives. As indicated by studies on AI deployment, the effectiveness of predictive systems is influenced by their integration into workflows, governance structures, and decision processes (Wu et al., 2020; Shaukat et al., 2020). This implies that predictive analytics functions as part of a broader socio-technical system, where its impact is mediated by organisational conditions.

From an innovation management perspective, this highlights the importance of viewing predictive analytics as a capability that supports learning rather than replaces it. Its role is to augment managerial decision-making by providing data-driven insights, while the interpretation and application of these insights remain organisational processes.

2.4 Data Quality and Structural Constraints of Predictive Capability

The effectiveness of predictive analytics is fundamentally constrained by the quality and structure of the underlying data. While algorithmic advances have improved predictive performance, they cannot compensate for deficiencies in data representation (Alam et al., 2020; Tanha et al., 2020). This reinforces the argument that predictive analytics is primarily a data-centric capability.

One of the most significant challenges identified in the literature is class imbalance, where the events of interest occur infrequently relative to other outcomes. Studies show that imbalanced datasets can distort evaluation metrics and reduce the ability of models to detect minority-class instances (Alam et al., 2020; Tanha et al., 2020). This issue is particularly relevant for innovation management, where successful innovations are rare and difficult to predict. In addition to imbalance, the increasing use of heterogeneous and multimodal data introduces further complexity. Research demonstrates that integrating diverse data sources, including sensor data, textual information, and visual inputs, can enhance predictive performance, but requires advanced preprocessing and feature engineering (Ouhami et al., 2021; Demrozi et al., 2020). Rahmani et al. (2021) further emphasise that data preparation often has a greater impact on model performance than the selection of algorithms.

These findings suggest that predictive analytics should be understood as a capability constrained by data conditions. Its effectiveness depends on the availability of high-quality, representative, and well-structured data, as well as on the ability to integrate and manage diverse data sources. For organisations, this implies that investments in predictive analytics must be accompanied by investments in data infrastructure and governance.

2.5 Interpretability and Organisational Integration of Predictive Systems

The final dimension of the theoretical framework concerns the relationship between predictive performance and organisational integration. While advanced machine learning models can achieve high levels of accuracy, their complexity often reduces transparency and interpretability, creating challenges for adoption (Chang et al., 2022; Alonso et al., 2021).

Empirical studies show that users may be reluctant to rely on predictive models if their outputs cannot be explained or justified. In healthcare contexts, for example, practitioners require interpretable results to support decision-making (Chang et al., 2022). Similar requirements apply in organisational settings, where decisions must be communicated to stakeholders and aligned with strategic objectives. This creates a trade-off between predictive accuracy and interpretability, which must be managed in practice. The integration of predictive analytics into organisational processes further complicates this issue. Studies indicate that the effectiveness of AI systems depends on their alignment with existing infrastructures, workflows, and governance mechanisms (Wu et al., 2020; Shaukat et al., 2020). Without such alignment, predictive insights may remain isolated and fail to influence actual decisions.

These observations reinforce the view that predictive analytics should be understood as a socio-technical capability. Its value depends not only on technical performance, but also on its integration into organisational systems and its ability to generate actionable and interpretable insights. This perspective provides a theoretical foundation for analysing predictive analytics within innovation management, where decision-making is inherently complex and context-dependent.

3. Methodology

This study adopts a structured narrative review with semi-systematic elements, as the objective was not to conduct an exhaustive systematic review, but to develop a transparent and conceptually grounded synthesis of the literature on predictive analytics relevant to innovation management. The review was designed to balance methodological transparency with the interpretive flexibility required for an emerging and interdisciplinary topic.

The literature search was conducted in the databases Scopus, Web of Science, and Google Scholar, which were selected in order to capture both high-quality peer-reviewed journal articles and relevant interdisciplinary contributions at the intersection of machine learning, predictive analytics, and organisational decision-making. The search strategy combined general and domain-specific keywords, including “predictive analytics”, “machine learning”, “deep learning”, “innovation management”, “decision-making”, “forecasting”, “classification”, and “data-driven management”, together with selected application-related terms such as “fraud detection”, “medical diagnosis”, “educational prediction”, “cybersecurity”, and “risk management”. These combinations were used to identify studies that, although often situated

outside innovation management as a formal field, examine structurally analogous predictive tasks with relevance for innovation-oriented decision contexts.

The review was limited to publications in English published between 2019 and 2024. This time frame was chosen to reflect the most recent phase of development in applied machine learning and predictive modelling, while ensuring that the analysed studies captured contemporary advances in data processing, ensemble learning, deep learning, and domain-specific implementation. Inclusion criteria required that sources be peer-reviewed, methodologically explicit, and substantively focused on predictive modelling using statistical, machine learning, or deep learning approaches. Studies were excluded if they were purely technical without interpretable application implications, non-peer-reviewed, duplicative, or only marginally related to predictive decision support.

The selection process proceeded in three stages. First, titles and abstracts were screened for conceptual relevance. Second, full texts were assessed with regard to analytical focus, methodological clarity, and transferability of findings. Third, the final set of 19 studies was retained for in-depth analysis. The synthesis of the selected literature was conducted through thematic categorisation, with attention to four recurrent dimensions: algorithmic approaches, data structures and constraints, application domains as functional analogues of innovation processes, and organisational implications of predictive system use. This enabled the review to move beyond descriptive comparison and to interpret predictive analytics as a system-dependent organisational capability shaped by the interaction of data quality, modelling choices, and decision context.

4. Literature Review

Building on the preceding theoretical framework, this section synthesises evidence from nineteen studies across multiple domains to identify patterns in algorithmic approaches, data structures, application contexts, and structural constraints. The review adopts a thematic structure, focusing on how predictive analytics operates across different empirical settings.

4.1 Algorithmic Approaches as Data-Contingent Solutions

The reviewed literature consistently demonstrates that algorithmic effectiveness is contingent upon the structural properties of data rather than being an intrinsic characteristic of specific modelling techniques. As summarised in Table 1, supervised learning approaches dominate empirical applications, particularly in contexts where prediction tasks are clearly defined and supported by structured datasets. Methods such as decision trees, support vector machines, and random forests are widely applied across domains including healthcare, finance,

and education (Chang et al., 2022; Yağcı, 2022). However, their performance is systematically influenced by data-related constraints, particularly class imbalance and feature representation.

To address these limitations, ensemble and boosting methods are frequently employed. These approaches improve classification performance in imbalanced datasets by enhancing the detection of minority-class instances (Tanha et al., 2020; Alam et al., 2020). This is particularly relevant for innovation-related contexts, where successful outcomes are rare and must be identified within a large pool of unsuccessful cases. At the same time, deep learning approaches extend predictive capabilities by enabling the modelling of complex, high-dimensional relationships. Studies such as Askr et al. (2022) and Schubach et al. (2024) demonstrate that neural architectures can capture patterns that are not accessible to traditional models. However, these advantages depend on the availability of large datasets and appropriate computational resources (Rahmani et al., 2021; Khan et al., 2021).

Taken together, these findings indicate that algorithmic performance is not determined by methodological sophistication alone, but by the alignment between modelling techniques and data conditions. The evidence summarised in Table 1 therefore supports the interpretation of predictive analytics as a configuration-dependent capability rather than a set of universally optimal methods.

4.2 Data Structures, Heterogeneity, and Analytical Constraints

Given the data-dependent nature of predictive analytics, the literature places increasing emphasis on the role of data structures in shaping predictive outcomes. As illustrated in Table 1, a clear trend can be observed toward the use of heterogeneous and multimodal datasets, combining numerical, textual, visual, and sensor-based information. Studies in domains such as agriculture and healthcare demonstrate that integrating multiple data sources improves predictive performance, particularly in complex environments (Ouhami et al., 2021; Demrozi et al., 2020).

However, the evidence also highlights that increased data complexity introduces significant analytical constraints. One of the most persistent challenges identified across the analysed studies is class imbalance, which affects both model training and evaluation. Several studies explicitly address this issue and demonstrate that standard models tend to produce biased results under skewed class distributions (Alam et al., 2020; Tanha et al., 2020). This constraint is particularly relevant for innovation management, where successful outcomes are relatively rare and difficult to model using conventional approaches. In addition, data quality and preprocessing emerge as critical determinants of predictive effectiveness. Rahmani et al. (2021) emphasise that data preparation often has a greater impact on model performance than algorithm selection, a pattern that is consistent across multiple studies included in Table 1. The integration of heterogeneous data sources further increases the need for robust preprocessing and feature engineering, as inconsistencies across datasets can reduce model reliability.

Predictive analytics operates under structural data constraints that fundamentally shape its effectiveness. The evidence presented in Table 1 reinforces the view that predictive capability depends not only on algorithmic design, but on the ability to manage, integrate, and represent complex data structures.

4.3 Application Domains as Functional Analogues of Innovation Processes

Although the reviewed studies span multiple domains, the analysis reveals a high degree of functional convergence in how predictive analytics supports decision-making. As summarised in Table 1, predictive models are consistently used to classify outcomes, detect anomalies, and assess risk across domains. For example, fraud detection models in finance (Ali et al., 2022; Alfaiz & Fati, 2022), diagnostic models in healthcare (Khan et al., 2021; Rahmani et al., 2021), and performance prediction models in education (Albreiki et al., 2021; Essa et al., 2023) all operate by identifying patterns in historical data and translating them into probabilistic predictions. These applications differ in context but converge in their underlying analytical logic.

This convergence suggests that predictive analytics performs transferable functions that are structurally analogous to key processes in innovation management, including opportunity identification, evaluation of alternatives, and resource allocation under uncertainty. However, the evidence also indicates that these functions are typically implemented in contexts with well-defined prediction targets and structured data. In contrast, innovation processes involve evolving objectives, incomplete datasets, and strong organisational embedding of decisions.

Therefore, while predictive analytics demonstrates cross-domain applicability, its transfer to innovation management requires adaptation to more complex and less structured environments. The patterns summarised in Table 1 support the interpretation of predictive analytics as a transferable but context-sensitive capability.

4.4 Structural Challenges and Organisational Implications

Despite the demonstrated effectiveness of predictive analytics across domains, the literature consistently identifies structural challenges that limit its practical impact. These challenges arise from the interaction between data, algorithms, and organisational context rather than from the limitations of individual models.

A key issue is the trade-off between predictive accuracy and interpretability. Studies included in Table 1 show that complex models, particularly deep learning architectures, achieve higher predictive performance but often lack transparency (Askr et al., 2022; Schubach et al., 2024). Empirical evidence indicates that this lack of interpretability can limit adoption, particularly in contexts where decisions must be justified (Chang et al., 2022). Similarly, research on text-based classification demonstrates that usability depends not only on accuracy, but also on the clarity of outputs (Alonso et al., 2021).

Another critical challenge concerns the integration of predictive analytics into organisational processes. The studies summarised in Table 1 suggest that predictive models are often evaluated in isolation, without sufficient consideration of how they are embedded within decision-making systems. However, evidence from AI and cybersecurity contexts shows that the effectiveness of predictive systems depends on their alignment with organisational infrastructures and workflows (Wu et al., 2020; Shaukat et al., 2020).

These findings indicate that predictive analytics functions as part of a broader socio-technical system, where its effectiveness depends on organisational capabilities, including the ability to interpret, integrate, and operationalise analytical outputs. The primary constraints of predictive analytics can therefore be seen as structural rather than purely algorithmic.

Table 1. Structured overview of analysed studies: machine learning approaches, data characteristics, application domains, and key findings

Author	Year	Domain	ML Technique	Learning Type	Data Type	Dataset Characteristics	Analytical Objective	Evaluation Metrics	Key Findings
Alam et al.	2020	Finance	SVM, ANN	Supervised	Structured	Imbalanced	Default prediction	Accuracy, F1, AUC	Imbalance significantly affects performance
Albreiki et al.	2021	Education	Various ML	Supervised	Structured	Medium datasets	Performance prediction	Accuracy	ML improves student outcome prediction
Alfaiz & Fati	2022	Finance	RF, SVM	Supervised	Structured	Imbalanced	Fraud detection	Precision, Recall	Ensemble models improve detection
Ali et al.	2022	Finance	Multiple ML	Mixed	Structured	Large datasets	Fraud classification	Accuracy, AUC	Hybrid models outperform single models
Alonso et al.	2021	Media	NLP models	Supervised	Text	Unstructured	Fake news detection	Accuracy	Sentiment improves classification
Alsaadi et al.	2024	Finance	AI models	Supervised	Structured	Large-scale	Risk prediction	Accuracy	AI enhances real-time decisions
Askr et al.	2022	Healthcare	Deep Learning	Supervised	Mixed	Large, complex	Drug discovery	Accuracy	DL superior for complex patterns
Chang et al.	2022	Healthcare	NB, RF, DT	Supervised	Structured	Medium	Disease prediction	Accuracy, AUC	RF and DT perform best
Demrozi et al.	2020	IoT/Health	ML, DL	Mixed	Multimodal	Sensor-based	Activity recognition	Accuracy	Data fusion improves results

Author	Year	Domain	ML Technique	Learning Type	Data Type	Dataset Characteristics	Analytical Objective	Evaluation Metrics	Key Findings
Essa et al.	2023	Education	ML models	Supervised	Structured	Medium	Adaptive learning	Accuracy	Personalization improves outcomes
Khan et al.	2021	Healthcare	ML, DL	Mixed	Medical data	Large	Diagnosis	Accuracy	DL highly effective
Nguyen et al.	2019	General	ML frameworks	Mixed	Big data	Large-scale	Data mining	Various	Framework choice critical
Ouhami et al.	2021	Agriculture	ML, CV	Supervised	Image + IoT	Multimodal	Disease detection	Accuracy	Integration improves prediction
Rahmani et al.	2021	Healthcare	ML models	Mixed	Medical	Large	Diagnosis	Accuracy	ML widely applicable
Schubach et al.	2024	Bioinformatics	DL, CNN	Supervised	Genomic	High-dimensional	Variant prediction	AUC	Hybrid DL improves accuracy
Shaukat et al.	2020	Cybersecurity	ML models	Mixed	Network data	Large	Intrusion detection	Accuracy	ML effective in anomaly detection
Tanha et al.	2020	General	Boosting	Supervised	Structured	Imbalanced	Classification	F1, AUC	Boosting handles imbalance
Wu et al.	2020	IoT	AI models	Mixed	Streaming	Real-time	Security prediction	Accuracy	AI enables real-time monitoring
Yağcı	2022	Education	ML models	Supervised	Structured	Medium	Performance prediction	Accuracy, AUC	Moderate predictive performance

Note: The table summarizes the key characteristics of the reviewed studies, focusing on machine learning techniques, data structures, analytical objectives, and performance outcomes. The classification reflects the multidimensional nature of predictive analytics applications across domains.

5. Research Gaps and Future Directions

The preceding analysis has shown that predictive analytics operates as a configuration-dependent capability shaped by the interaction between data structures, algorithmic approaches, and organisational conditions. However, despite the increasing methodological sophistication of machine learning models, the literature reveals several persistent limitations that constrain their applicability in innovation management. These limitations do not arise from isolated technical deficiencies, but from structural mismatches between predictive models and the conditions under which innovation processes take place.

5.1 Gap Between Algorithmic Performance and Decision-Making Effectiveness

As indicated by the synthesis of the reviewed studies, a substantial part of the literature focuses on improving predictive accuracy through advanced modelling techniques, including

ensemble learning and deep neural networks (Tanha et al., 2020; Askr et al., 2022). These approaches demonstrate measurable performance gains, particularly in classification tasks and high-dimensional data environments. However, the evaluation of these models remains largely confined to technical metrics such as accuracy, precision, or AUC, with limited attention to how predictive outputs influence decision-making processes.

On the basis of the findings presented in the previous chapter, it is evident that improved predictive performance does not automatically translate into improved decision quality. Studies addressing imbalanced datasets, for instance, demonstrate enhanced detection of rare events (Alam et al., 2020; Tanha et al., 2020), yet they do not examine how these predictions are incorporated into organisational decisions. Similarly, research on deep learning emphasises pattern recognition capabilities (Askr et al., 2022; Schubach et al., 2024), but provides limited insight into the interpretability and usability of these outputs in complex decision environments.

This gap suggests that future research should extend beyond accuracy-oriented evaluation and focus on the role of predictive analytics in supporting decision processes. In particular, there is a need to examine how predictive insights are interpreted, integrated, and operationalised within innovation-related contexts, where decisions are inherently uncertain and multi-dimensional.

5.2 Limited Generalisability Across Data Contexts and Domains

The literature review further indicates that predictive analytics applications are predominantly developed within domain-specific contexts, such as finance, healthcare, and education. As shown in the previous chapter, these applications rely on relatively well-defined prediction tasks and structured datasets (Ali et al., 2022; Khan et al., 2021; Albreiki et al., 2021). While this enables high predictive performance within individual domains, it limits the transferability of these models to other contexts.

Based on the evidence synthesised in the literature review, predictive models are highly sensitive to data characteristics, including dataset size, structure, and level of heterogeneity (Rahmani et al., 2021). Studies focusing on multimodal data integration demonstrate that combining different data sources can improve predictive outcomes (Ouhami et al., 2021; Demrozi et al., 2020), yet these approaches are typically tailored to specific environments and lack general applicability.

This indicates a clear gap in the literature concerning the generalisability of predictive analytics across different contexts. Innovation management, in particular, is characterised by evolving datasets and incomplete information, which differ substantially from the conditions under which most predictive models are developed. Future research should therefore focus on

developing approaches that enable the adaptation of predictive models to diverse and changing data environments.

5.3 Insufficient Attention to Data Governance and Lifecycle Management

Another limitation emerging from the analysis concerns the treatment of data as a static input rather than as a dynamic organisational resource. As highlighted in the previous chapter, data quality, preprocessing, and feature engineering are critical determinants of predictive performance (Rahmani et al., 2021; Alam et al., 2020). However, existing studies rarely address how data is managed over time within organisational settings.

The reviewed literature demonstrates that predictive models are often developed using fixed datasets, without considering how data evolves or how it is maintained in practice. This is particularly evident in studies addressing class imbalance and heterogeneous data structures, where technical solutions are proposed without examining the underlying data generation and management processes (Tanha et al., 2020; Ouhami et al., 2021).

Given that innovation processes are inherently dynamic, this represents a significant gap. Without effective data governance and lifecycle management, predictive models may become outdated or unreliable. Future research should therefore examine how organisations can develop data strategies that ensure the continuity, consistency, and relevance of data over time, including mechanisms for data updating, validation, and integration.

5.4 Lack of Integration Between Predictive Models and Organisational Processes

The findings of the literature review indicate that predictive analytics is frequently treated as a standalone analytical tool, rather than as an integrated component of organisational systems. The effectiveness of predictive models depends on their alignment with organisational workflows, infrastructures, and decision-making processes (Wu et al., 2020; Shaukat et al., 2020).

At the same time, empirical evidence suggests that the usability of predictive models is influenced by their interpretability. Studies demonstrate that complex models may achieve higher predictive accuracy, but their lack of transparency can limit their adoption in practice (Chang et al., 2022; Alonso et al., 2021). This creates a dual challenge: predictive models must be both technically effective and organisationally usable.

The literature, however, provides limited guidance on how such integration can be achieved. As a result, predictive analytics often remains disconnected from actual decision-making processes. Future research should therefore focus on identifying organisational conditions that enable effective integration, including governance mechanisms, decision frameworks, and processes for translating predictive outputs into actionable insights.

5.5 Static Modelling Approaches in Dynamic Environments

A further gap arises from the predominance of static modelling approaches in the literature. As indicated by the reviewed studies, predictive models are typically evaluated using fixed datasets and assume stable conditions (Ali et al., 2022; Yağcı, 2022). However, the analysis presented in the previous chapter highlights that real-world environments are characterised by continuous change, including shifts in data distributions and evolving contextual conditions.

Evidence from domains such as cybersecurity and IoT demonstrates the potential of adaptive and real-time predictive systems. Machine learning models can be used to detect emerging patterns by continuously learning from new data (Shaukat et al., 2020), while real-time data processing enhances system responsiveness (Wu et al., 2020). Nevertheless, such approaches remain underdeveloped in other domains, particularly in innovation management.

This gap suggests the need for research on adaptive modelling approaches that incorporate continuous learning, feedback mechanisms, and real-time data integration. Such approaches would enable predictive analytics to operate effectively in dynamic environments, aligning more closely with the conditions of innovation processes.

5.6 Integrated Capabilities Perspective

Taken together, the gaps identified above indicate that the current literature remains fragmented across algorithmic, data-related, and organisational dimensions. While individual studies provide valuable insights into specific aspects of predictive analytics, they rarely integrate these perspectives into a coherent analytical framework.

On the basis of the findings presented in the previous chapter, it can be argued that predictive analytics operates as a system-level capability that emerges from the interaction between data, algorithms, and organisational context. However, this perspective is not explicitly developed in existing research, which tends to treat these components separately. Studies focusing on algorithmic performance (Tanha et al., 2020; Askr et al., 2022), data characteristics (Rahmani et al., 2021; Ouhami et al., 2021), and organisational integration (Wu et al., 2020; Shaukat et al., 2020) are rarely combined into a unified analytical framework.

Future research should therefore adopt an integrated perspective that examines predictive analytics as a capability embedded within organisational systems. This includes analysing how data, models, and organisational processes interact to shape predictive effectiveness and innovation outcomes. Developing such a perspective would provide a more comprehensive understanding of how predictive analytics can be applied in innovation management contexts.

This study has several limitations. The analysis is based on a selected set of nineteen studies and does not aim to provide an exhaustive systematic review. In addition, the literature is predominantly drawn from domain-specific applications, which may limit the generalisability

of findings to innovation management contexts. Future research should therefore extend the scope of analysis and incorporate empirical validation in organisational settings.

6. Conclusion

This study examined predictive analytics in the context of innovation management by adopting a capability-based perspective that emphasises the interaction between data, algorithmic approaches, and organisational conditions. Based on a structured synthesis of nineteen studies, the analysis moved beyond isolated applications and focused on identifying the conditions under which predictive analytics contributes to decision-making in complex and uncertain environments. The findings show that predictive analytics cannot be understood as a universally applicable solution, but rather as a context-dependent capability whose effectiveness is shaped by system-level alignment.

The analysis identifies consistent patterns in how predictive models perform across different contexts, highlighting the role of data quality, structural characteristics, and organisational integration. These findings indicate that predictive analytics operates under identifiable structural conditions that shape its applicability in innovation-related environments. Improvements in predictive performance are necessary but insufficient in isolation. Their impact depends on the interpretability of model outputs, their integration into decision-making processes, and the availability of appropriate data infrastructures. This underscores that the effectiveness of predictive analytics is not determined solely by algorithmic or data-related factors, but also by organisational capabilities to operationalise predictive insights within existing workflows.

From a theoretical perspective, the study contributes by explicitly framing predictive analytics as a system-dependent organisational capability. This perspective connects previously fragmented strands of literature and shows that predictive effectiveness emerges from the interaction between data structures, modelling approaches, and organisational context. Rather than treating predictive analytics as a standalone analytical tool, the study demonstrates that its value lies in its integration within a broader socio-technical system. From a practical perspective, the findings suggest that organisations should approach predictive analytics as a capability that requires coordinated development across multiple dimensions. Investments in advanced modelling techniques must be complemented by investments in data governance, data quality, and integration into decision-making processes. In particular, organisations must ensure that predictive models produce outputs that are not only accurate but also interpretable and actionable. Without such alignment, predictive analytics may improve analytical outputs without leading to meaningful improvements in innovation performance.

The study also highlights several directions for future research. There is a need to examine predictive analytics in relation to decision-making processes rather than solely through

accuracy-based metrics, to explore the generalisability of predictive models across different contexts, and to investigate adaptive approaches that account for dynamic data environments. In addition, further research should focus on the organisational conditions that enable effective integration of predictive analytics, including governance structures and mechanisms for translating predictive insights into actionable decisions.

Overall, this study provides an empirically grounded and theoretically integrated understanding of predictive analytics as a capability in innovation management. By identifying the conditions under which predictive analytics is effective, it offers a more nuanced perspective on its role in supporting innovation-related decision-making and highlights the importance of alignment between technical and organisational dimensions.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Alam, T. M., Shaukat, K., Hameed, I. A., Luo, S., & Umer, M. (2020). An investigation of credit card default prediction in the imbalanced datasets. IEEE Access, 8, 201670–201682. https://doi.org/10.1109/ACCESS.2020.3033784

Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of students' performance prediction using machine learning techniques. Education Sciences, 11(9), 552. https://doi.org/10.3390/educsci11090552

Alfaiz, N. S., & Fati, S. M. (2022). Enhanced credit card fraud detection model using machine learning. Electronics, 11(4), 662. https://doi.org/10.3390/electronics11040662

Ali, A., Abd Razak, S., Othman, S. H., & Eisa, T. A. E. (2022). Financial fraud detection based on machine learning: A systematic literature review. Applied Sciences, 12(19), 9637. https://doi.org/10.3390/app12199637

Alonso, M. Á., Vilares, D., Gómez-Rodríguez, C., & Vilares, J. (2021). Sentiment analysis for fake news detection. Electronics, 10(11), 1348. https://doi.org/10.3390/electronics10111348

Alsaadi, M., Mahmood, M. T., Obaed, A. S., & Hussain, H. (2024). AI-based predictive analytics for financial risk management. In 2024 8th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT) (pp. 1–6). IEEE. https://doi.org/10.1109/ISMSIT63511.2024.10757214

Askr, H., Elgeldawi, E., Aboul Ella, H., Elshaier, Y. A. M. M., & Mamdouh, M. (2022). Deep learning in drug discovery: An integrative review and future challenges. Artificial Intelligence Review, 56(7), 5975–6037. https://doi.org/10.1007/s10462-022-10306-1

Chang, V., Bailey, J., Xu, Q., & Sun, Z. (2022). Pima Indians diabetes mellitus classification based on machine learning algorithms. Neural Computing and Applications, 35(22), 16157–16173. https://doi.org/10.1007/s00521-022-07049-z

Demrozi, F., Pravadelli, G., Bihorac, A., & Rashidi, P. (2020). Human activity recognition using inertial, physiological and environmental sensors: A comprehensive survey. IEEE Access, 8, 210816–210836. https://doi.org/10.1109/ACCESS.2020.3037715

Essa, S. A. M., Çelik, T., & Human-Hendricks, N. E. (2023). Personalized adaptive learning technologies based on machine learning techniques to identify learning styles: A systematic literature review. IEEE Access, 11, 48392–48409. https://doi.org/10.1109/ACCESS.2023.3276439

Khan, P., Kader, M. F., Islam, S. M. R., Rahman, A. B., & Shahriar, M. (2021). Machine learning and deep learning approaches for brain disease diagnosis: Principles and recent advances. IEEE Access, 9, 37622–37655. https://doi.org/10.1109/ACCESS.2021.3062484

Nguyen, G., Dlugolinský, Š., Bobák, M., Tran, V., & López García, Á. (2019). Machine learning and deep learning frameworks and libraries for large-scale data mining: A survey. Artificial Intelligence Review, 52(1), 77–124. https://doi.org/10.1007/s10462-018-09679-z

Ouhami, M., Hafiane, A., Es-Saady, Y., El Hajji, M., & Canals, R. (2021). Computer vision, IoT and data fusion for crop disease detection using machine learning: A survey and ongoing research. Remote Sensing, 13(13), 2486. https://doi.org/10.3390/rs13132486

Rahmani, A. M., Yousefpoor, E., Yousefpoor, M. S., & Mehmood, Z. (2021). Machine learning in medicine: Review, applications, and challenges. Mathematics, 9(22), 2970. https://doi.org/10.3390/math9222970

Schubach, M., Maaß, T., Nazaretyan, L., Röner, S., & Kircher, M. (2024). CADD v1.7: Using protein language models, regulatory CNNs and other nucleotide-level scores to improve genome-wide variant predictions. Nucleic Acids Research, 52(D1), D1143–D1154. https://doi.org/10.1093/nar/gkad989

Shaukat, K., Luo, S., Varadharajan, V., Hameed, I. A., & Xu, M. (2020). A survey on machine learning techniques for cyber security in the last decade. IEEE Access, 8, 222310–222354. https://doi.org/10.1109/ACCESS.2020.3041951

Tanha, J., Abdi, Y., Samadi, N., Razzaghi, N., & Asadpour, M. (2020). Boosting methods for multi-class imbalanced data classification: An experimental review. Journal of Big Data, 7(1), 70. https://doi.org/10.1186/s40537-020-00349-y

Wu, H., Han, H., Wang, X., & Sun, S. (2020). Research on artificial intelligence enhancing Internet of Things security: A survey. IEEE Access, 8, 153826–153848. https://doi.org/10.1109/ACCESS.2020.3018170

Yağcı, M. (2022). Educational data mining: Prediction of students' academic performance using machine learning algorithms. Smart Learning Environments, 9(1), 11. https://doi.org/10.1186/s40561-022-00192-z

A Review of the Firm-Level Determinants of Artificial Intelligence Adoption

Bolek Vladimír [1] and Baziuk Andriana[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: vladimir.bolek@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: andriana.baziuk@euba.sk

Abstract

The adoption of emerging technologies at the firm level has become one of the most consequential strategic decisions facing organizations in the contemporary digital economy. Despite the growing volume of research on digital transformation, automation, and Industry 4.0, the literature remains fragmented across disciplines, with insufficient integration of firm-level determinants into coherent theoretical frameworks. This review synthesizes evidence from peer-reviewed studies to identify the primary organizational, economic, and institutional factors that drive or impede technology adoption at the firm level. Three principal thematic clusters are examined: organizational capabilities and firm size effects, market structure and competitive pressure, and workforce and human capital considerations. The review finds that technology adoption decisions are shaped by a complex interplay of internal resource endowments, competitive dynamics, and labor market conditions, and that these factors operate differently across firm size categories and industry sectors. Sustainability considerations and external regulatory pressures are identified as increasingly salient moderating variables. The review concludes by identifying critical research gaps, including the underexploration of adoption dynamics in emerging market firms and the limited longitudinal evidence on adoption outcomes. These findings carry implications for managers, policymakers, and scholars interested in accelerating beneficial technology diffusion.

Key Words: firm-level technology adoption, absorptive capacity, organisational factors, innovation determinants, SME digitalisation

JEL Classification: O33, L25, M20

1. Introduction

The diffusion of new technologies across firms has long attracted scholarly attention in economics, management, and information systems. However, the accelerating pace of technological change in the twenty-first century—driven by artificial intelligence, additive manufacturing, industrial digitalization, and connected production systems—has elevated the stakes of technology adoption decisions to an unprecedented degree. Firms that successfully integrate advanced technologies into their operations stand to gain substantial competitive advantages, while those that lag face the prospect of declining market share, reduced productivity, and eventual displacement from their industries. Understanding what drives or inhibits technology adoption at the firm level is therefore not merely an academic question but one with far-reaching economic and policy implications.

Despite the practical urgency of this question, the literature on firm-level technology adoption remains notably fragmented. Contributions from economics, management science, information systems, and sustainability research have developed in largely parallel tracks,

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

producing a rich but disjointed body of knowledge. Economists have focused on the relationship between automation, labor displacement, and market concentration, while information systems scholars have concentrated on adoption models rooted in individual and organizational behavior. Industry 4.0 researchers have examined digital transformation in manufacturing contexts, and organizational theorists have explored how firm resources, capabilities, and governance structures mediate adoption decisions. The absence of an integrative review that synthesizes these strands limits theoretical progress and impedes practitioners who must navigate adoption decisions in real-world settings.

This literature review addresses that gap by systematically examining the firm-level determinants of technology adoption across disciplinary boundaries. Drawing on twenty-one studies published in leading peer-reviewed outlets, this review pursues three specific objectives: first, to map the key theoretical frameworks employed in the literature; second, to synthesize empirical findings organized around coherent thematic clusters; and third, to identify persistent research gaps and propose directions for future inquiry.

The article is organized as follows. The Theoretical Background outlines the key concepts and situates the study within relevant theoretical frameworks. The Literature Review describes the methodology applied in the review and provides a thematic synthesis of existing research, focusing on organizational capabilities and firm size, market structure and competitive dynamics, and workforce and human capital considerations. The section on Research Gaps and Future Directions identifies limitations in the current literature and suggests avenues for further research. Finally, the Conclusion summarizes the main findings, highlights the key contributions, and acknowledges the limitations of the review.

2. Theoretical Background

Understanding technology adoption at the firm level requires both a clear definition of the concept and an overview of the theoretical frameworks that inform its study.

2.1 Defining Technology Adoption at the Firm Level

Technology adoption, in the organizational context, refers to the decision-making process through which a firm evaluates, acquires, and integrates a new technology into its operational or strategic processes. This definition encompasses a broad range of phenomena, from the installation of enterprise software to the deployment of robotics, artificial intelligence systems, additive manufacturing capabilities, or Internet of Things infrastructure. The concept is distinct from invention and innovation in that it concerns the uptake of technologies already available in the environment, rather than the creation of novel knowledge per se. Critically, adoption is rarely a discrete, binary event; it unfolds across stages of awareness, evaluation, trial, and full implementation, each of which may be subject to different organizational and market influences (Greenhalgh et al., 2017).

The economic importance of technology adoption is difficult to overstate. Acemoglu and Restrepo (2019) demonstrate that the deployment of automation technologies fundamentally reallocates tasks between labor and capital, with aggregate implications for employment shares and wage distributions. Their framework emphasizes that adoption is not merely a passive response to technological availability but an active organizational choice shaped by the relative costs of human and machine inputs. Goldfarb and Tucker (2019) extend this perspective to the digital domain, arguing that the dramatic reduction in storage, computation, and transmission costs associated with digital technology has transformed the economics of a

wide range of business activities, lowering barriers to adoption while simultaneously intensifying competitive pressures.

2.2 Foundational Theoretical Frameworks

Several theoretical frameworks have been deployed in the literature to explain firm-level adoption decisions. The Technology-Organization-Environment (TOE) framework, while originating in the information systems tradition, provides a useful organizing structure that acknowledges the simultaneous influence of technological characteristics, organizational attributes, and environmental conditions on adoption outcomes. This three-way interaction is broadly consistent with the emphasis in the industrial economics literature on how market structure conditions firm-level investment decisions.

Resource-based theory offers a complementary perspective, emphasizing that firms' internal capabilities, knowledge stocks, and organizational routines determine their capacity to absorb and deploy new technologies. Müller, Kiel, and Voigt (2018) draw on this tradition in their examination of Industry 4.0 implementation, finding that firms' ability to exploit digitalization opportunities is conditioned by pre-existing organizational competencies and that resource constraints function as significant barriers, particularly for smaller enterprises. Similarly, Dwivedi et al. (2019) observe that the transformative potential of artificial intelligence is contingent on whether organizations possess the data infrastructure, skilled personnel, and strategic orientation necessary to implement AI-driven processes.

The superstar firm thesis advanced by Autor et al. (2020) provides a macro-level theoretical lens through which firm-level adoption can be understood in terms of market dynamics. Their analysis of U.S. firm-level data demonstrates that technological change and globalization have systematically redirected market share toward the most productive firms within each industry, suggesting that technology adoption is both a driver and a consequence of competitive selection processes. This perspective implies that the firm-level determinants of technology adoption cannot be fully understood in isolation from the industrial environments in which firms operate.

The concept of technological disruption, as elaborated by Buhalis et al. (2019) in the service sector context, highlights that adoption decisions are also influenced by ecosystem-level dynamics, including the emergence of new value co-creation opportunities, shifting customer expectations, and the reconfiguration of industry boundaries. This perspective warns against overly narrow, firm-centric accounts of adoption and calls for analyses that situate individual firm decisions within broader sociotechnical systems.

Finally, the literature on digital economics, as surveyed by Goldfarb and Tucker (2019), underscores that digitalization is not a monolithic phenomenon but rather a cluster of distinct cost reductions—in search, replication, transportation, tracking, and verification—each of which may activate different adoption incentives and encounter different organizational barriers. This granular perspective on the mechanisms through which digital technology affects firm behavior is an important theoretical contribution that complements the broader frameworks described above.

3. Methodology

The aim of this study is to synthesise insights from the academic literature on the determinants of advanced technology adoption at the firm level, with particular emphasis on the organisational, market, labour-related, and institutional factors that shape firms' decisions to introduce new technological solutions. The study is conceived as a literature review whose

purpose is to identify the principal thematic strands of research, compare the dominant explanatory approaches, and highlight areas that remain insufficiently explored in the current scholarly debate.

The methodological approach is based on the principles of a structured literature review. Particular emphasis was placed on the substantive relevance of the sources, their scholarly quality, and their direct connection to the research problem under consideration. Given the interdisciplinary nature of the topic, which spans economics, management, information systems, industrial transformation, the organisation of work, and the institutional environment of firms, a broader analytical framework was adopted. This made it possible to include studies examining different technological contexts while remaining united by a common question: which firm-level factors facilitate or hinder the adoption of new technologies?

The literature selection focused on peer-reviewed academic articles published in international scholarly journals. Attention was directed towards studies addressing technology adoption at the level of the firm, particularly in the areas of digital transformation, automation, Industry 4.0, artificial intelligence, additive manufacturing, teleworking, and other forms of technologically significant organisational change. The review concentrated primarily on studies that analysed adoption determinants from the perspective of firms' internal characteristics, competitive environments, labour market conditions, institutional pressures, or broader structural factors.

In identifying relevant sources, consideration was given to studies corresponding to the central themes of the article, especially organisational capabilities, absorptive capacity, firm size, competitive pressure, market structure, human capital, labour reorganisation, sustainability, and institutional support for adoption. For this reason, the analytical sample included not only studies explicitly examining artificial intelligence adoption, but also research concerned with the wider set of advanced digital and production technologies, provided that such studies offered transferable insights into the decision-making mechanisms through which firms adopt new technologies. This approach reflects the nature of the research problem itself, as a substantial part of the literature does not examine the adoption of individual technologies in isolation, but rather within the broader context of firms' technological transformation.

Studies were included in the review if they met several core criteria. First, they had to be peer-reviewed academic publications. Second, they had to be published in the English language. Third, they had to address firm-level technology adoption or the factors influencing such adoption. Fourth, they had to provide theoretical, empirical, or synthetic insights that could contribute to explaining firm behaviour in the introduction of new technologies. By contrast, studies were excluded if they focused exclusively on the technical characteristics of technologies without reference to firm decision-making, if they dealt solely with individual-level technology adoption by users, or if they lacked a sufficiently clear connection to the organisational or economic determinants of adoption.

Following the selection of relevant publications, a qualitative thematic synthesis was employed. The individual studies were analysed in terms of their technological context, the type of determinants examined, the level of analysis, the principal findings, and their implications for theory and practice. Particular attention was paid to whether the studies emphasised firms' internal resources and capabilities, external competitive pressures, labour market characteristics, regulatory or institutional factors, or the interaction between these dimensions. The objective was not merely to summarise the studies individually, but to identify recurring patterns and conceptual relationships across the literature.

On the basis of this analysis, three principal thematic areas were identified, which also structure the core of the article. The first area encompasses organisational capabilities, firm size, and absorptive capacity as internal determinants of adoption. The second focuses on market structure, competitive dynamics, and the role of market power in shaping incentives to adopt technology. The third thematic area addresses the workforce, human capital, and organisational labour relations, which influence not only the decision to adopt technology but also the process and consequences of implementation. A further cross-cutting dimension of the analysis concerns sustainability, institutional pressures, and exogenous shocks, which in a number of studies emerge as important moderating or conditioning factors.

The chosen methodological approach made it possible to connect literature from several disciplines within a single interpretative framework and to demonstrate that technology adoption at the firm level is the result of a complex interplay between firms' internal capacities, competitive incentives, workforce characteristics, and the broader institutional environment. At the same time, this approach allowed for the identification of areas in which the current literature remains fragmented, including the limited representation of developing economies, the shortage of longitudinal analyses, and the insufficiently explored interactions among different types of technologies.

4. Literature Review

4.1 Organizational Capabilities, Firm Size, and Absorptive Capacity

A prominent and well-documented theme in the literature concerns the role of organizational capabilities—particularly size, absorptive capacity, and pre-existing technological infrastructure—in determining whether and how firms adopt new technologies. There is broad agreement in the literature that larger firms tend to adopt advanced technologies earlier and more comprehensively than smaller firms, primarily because they possess greater financial resources, more specialized human capital, and more developed internal knowledge management systems. Müller et al. (2018), in one of the most directly relevant studies examined, find that firm size exerts a significant positive influence on Industry 4.0 implementation, with large enterprises demonstrating a more nuanced appreciation of both the opportunities and the challenges associated with digital manufacturing technologies. Smaller firms, by contrast, are more likely to be deterred by perceived implementation costs and organizational disruption.

However, the relationship between firm size and adoption is not uniformly positive, and the literature reveals important heterogeneities. Nagy et al. (2018) examine the role of Industry 4.0 technologies in reshaping the value chain among Hungarian companies and find that the integration of digital tools into primary value-creating activities requires not merely financial investment but also a fundamental restructuring of organizational processes, a demand that may in some respects favor agile smaller firms capable of rapid reorganization. This tension between resource advantages and organizational flexibility complicates straightforward size-based predictions and suggests that the nature of the technology in question mediates the relationship between firm size and adoption likelihood.

The concept of absorptive capacity—the firm's ability to recognize, assimilate, and apply new external knowledge—is central to understanding why firms with similar resource endowments often diverge significantly in their adoption behavior. Dwivedi et al. (2019) highlight that the adoption of artificial intelligence, in particular, demands not only financial investment but a sophisticated organizational capability to manage data, retrain personnel,

and align AI deployment with strategic objectives. Firms that lack this internal absorptive capacity may acquire AI tools without realizing their potential value, resulting in what amounts to superficial adoption rather than genuine digital transformation.

Tofail et al. (2017) offer a parallel perspective in the context of additive manufacturing, noting that the uptake of 3D printing and related technologies by industrial firms is governed not only by the technical maturity of the technology but also by the organizational capacity to manage new design processes, supply chain reconfigurations, and quality assurance protocols. Their analysis highlights a recurring theme in the literature: that technology adoption is an organizationally intensive activity that demands managerial attention and capability development well beyond the acquisition of new equipment or software.

4.2 Market Structure, Competitive Dynamics, and the Role of Market Power

A second major thematic cluster concerns the relationship between market structure, competitive dynamics, and firm-level technology adoption. The literature here reveals a complex and sometimes contradictory picture, with evidence suggesting that both competitive pressure and market power can function as drivers of adoption, depending on the technological context and the nature of competitive interactions.

The analysis by De Loecker, Eeckhout, and Unger (2020) of rising market power among U.S. firms provides a crucial macro-level context for understanding technology adoption dynamics. Their documentation of systematically rising markups since the 1980s, driven primarily by the upper tail of the markup distribution, suggests that a subset of firms has successfully leveraged technological superiority to achieve and sustain durable competitive advantages. This finding is consistent with the superstar firm framework of Autor et al. (2020), who demonstrate that market share reallocation toward high-productivity firms is a defining feature of recent decades, and that this reallocation is strongly associated with differential rates of technology adoption and deployment. Together, these studies imply that technology adoption is not simply a response to competitive pressure but is itself a mechanism through which competitive differentiation is produced and sustained.

The digital economics framework of Goldfarb and Tucker (2019) adds important nuance to this picture by identifying how specific cost reductions enabled by digital technology—particularly in search and verification—reshape competitive dynamics across industries. Reductions in search costs, for example, lower barriers to entry and intensify price competition in markets characterized by previously high information asymmetries, which may simultaneously pressure incumbent firms to adopt cost-reducing technologies and enable new entrants to challenge established players. The net effect on adoption incentives depends critically on whether the technology in question primarily enables cost reduction, product differentiation, or the creation of new markets.

Buhalis et al. (2019) examine these dynamics in the context of tourism and hospitality services, where the emergence of platform-based business models, AI-driven personalization, and IoT-enabled smart environments has disrupted established industry structures. Their findings suggest that in service-intensive sectors, technology adoption is driven not only by internal efficiency considerations but also by the need to participate in emerging ecosystems, a dynamic that may compel even resource-constrained firms to adopt technologies they would not otherwise prioritize. This ecosystem dimension of technology adoption represents an important departure from firm-centric perspectives and highlights the influence of external network effects on adoption decisions.

There is notable debate in the literature about whether market concentration facilitates or retards technology adoption. On one hand, firms with greater market power may face weaker immediate competitive pressures to adopt costly new technologies, potentially leading to adoption lags. On the other hand, the financial reserves associated with high market power may facilitate experimentation and investment in emerging technologies whose returns are uncertain. De Loecker et al. (2020) do not resolve this debate definitively but their finding that high-markup firms tend to be technologically sophisticated suggests that the two phenomena—market power and technology adoption—are more likely to be mutually reinforcing than in opposition.

4.3 Workforce, Human Capital, and Organizational Labor Dynamics

A third substantial thematic cluster addresses the relationship between technology adoption and the workforce—encompassing issues of human capital requirements, labor displacement, skill complementarities, and the organizational management of technological change. This cluster intersects with but is distinct from the absorptive capacity literature, focusing not on aggregate organizational competence but on the specific human capital demands and labor market consequences of technology adoption.

Acemoglu and Restrepo (2019) provide the most analytically rigorous treatment of this theme in the reviewed literature, presenting a task-based framework that distinguishes between automation—which enables capital to substitute for labor in existing tasks—and the creation of new tasks, which reinstates labor demand in novel activities. Their framework implies that the firm-level decision to adopt automation technologies is governed by a comparison of the productivity gains from capital substitution against the costs of organizational restructuring and potential disruption to existing workflows. Importantly, they show that automation reduces the labor share in value added even as it raises overall productivity, a finding with significant implications for how firms manage the workforce impacts of adoption decisions.

The gender dimension of workforce impacts receives attention from Blau and Kahn (2017), whose analysis of gender wage gaps and occupational segregation in the U.S. provides a background against which the distributional consequences of technology adoption can be assessed. Their finding that gender differences in occupation and industry remain the primary drivers of the wage gap—even as conventional human capital explanations have weakened—is relevant to the technology adoption literature insofar as the occupations most susceptible to automation tend to be those with specific demographic profiles, implying that adoption decisions carry embedded distributional consequences that may be invisible to firm-level analysts focused solely on productivity outcomes.

Belzunegui Eraso and Erro-Garcés (2020) examine a specific and topical form of technology adoption—teleworking—in the context of the COVID-19 crisis, offering a firm-level case study of how external shocks can accelerate adoption of organizational technologies that had previously been available but underutilized. Their analysis demonstrates that crisis conditions can override the inertial forces that typically slow technology diffusion, effectively compressing adoption timelines. This finding suggests that the pace of technology adoption is not solely a function of steady-state firm characteristics but is also subject to discontinuous shifts driven by environmental disruption.

Davidescu et al. (2020) extend this analysis by examining how flexible work arrangements, enabled by digital technologies, affect job satisfaction and performance among Romanian employees, with implications for sustainable human resource management. Their findings

suggest that technology-enabled flexibility confers genuine productivity benefits when well-managed, but that these benefits are contingent on organizational design choices and employee characteristics. This contingency perspective adds important nuance to adoption frameworks by highlighting that the outcomes of adoption are not automatic but depend on complementary organizational investments in human capital development and management practice.

The relationship between technology adoption and labor is further contextualized by the wider debate about the future of work. Autor et al. (2020) note that the rise of superstar firms has been accompanied by declining labor shares of income, a pattern consistent with the displacement effects documented by Acemoglu and Restrepo (2019). The convergence of these findings suggests that firm-level technology adoption decisions aggregate, at the macroeconomic level, into systemic shifts in the distribution of income between labor and capital—a consideration that is increasingly attracting policy attention and may, in turn, generate regulatory pressures that feed back into firm-level adoption decisions.

4.4 Sustainability, Institutional Pressures, and Exogenous Drivers

An emerging and increasingly important dimension of the firm-level technology adoption literature concerns the role of sustainability imperatives and institutional pressures as exogenous drivers of adoption decisions. Müller et al. (2018) explicitly situate Industry 4.0 adoption in the context of sustainability, arguing that digitalization technologies offer the potential to improve resource efficiency, reduce waste, and enable more sustainable forms of industrial production, and that these sustainability benefits are becoming an increasingly important consideration in firms' adoption rationale, particularly in regulatory environments where corporate environmental performance is subject to external scrutiny.

Christensen, Hail, and Leuz (2021) provide a comprehensive analysis of mandatory CSR and sustainability reporting requirements, demonstrating how regulatory mandates can alter firm behavior in ways that may indirectly promote technology adoption. Their review suggests that disclosure requirements can create incentives for firms to invest in cleaner production technologies and more sustainable operational practices, as the reputational and financial costs of poor environmental performance become more salient when publicly reported. While their analysis is focused on accounting and finance considerations, its implications for technology adoption are clear: regulatory frameworks can function as significant exogenous drivers of adoption by raising the cost of non-adoption relative to investment.

The institutional environment more broadly shapes technology adoption through mechanisms beyond direct regulation. Compagnucci and Spigarelli (2020) examine the role of universities as drivers of knowledge transfer and innovation diffusion, finding that the "Third Mission" of universities—their contribution to economic and social development through technology transfer, consultancy, and collaborative research—constitutes an important institutional support for firm-level technology adoption, particularly for smaller enterprises that lack internal R&D capabilities. This finding highlights that firm-level adoption decisions are embedded in broader innovation ecosystems that include public research institutions, industry associations, and policy intermediaries.

5. Research Gaps and Future Directions

Despite the substantial body of evidence reviewed above, several significant gaps in the literature on firm-level technology adoption warrant attention from future researchers.

5.1 Emerging Market and Developing Economy Contexts

The overwhelming majority of empirical studies on firm-level technology adoption draw on data from advanced industrial economies, particularly the United States and Western Europe. Duque Grisales and Aguilera-Caracuel (2019), whose analysis of Latin American multinationals focuses on ESG performance, illustrates both the potential and the limitations of applying frameworks developed in advanced economy contexts to emerging market settings. Firm-level technology adoption in developing and emerging economies is likely to be governed by institutional conditions, financing constraints, and labor market characteristics that differ substantially from those prevailing in OECD countries. The role of governance structures, as surveyed by Jiang and Kim (2020) in the Chinese context, suggests that concentrated ownership, state involvement in industry, and weak minority shareholder protections may create fundamentally different incentive environments for technology investment decisions. Future research should develop and test adoption models specifically calibrated to emerging market contexts, drawing on firm-level data from regions currently underrepresented in the literature.

5.2 Longitudinal Dynamics and Adoption Trajectories

The majority of studies reviewed here rely on cross-sectional or relatively short-panel data, limiting their capacity to capture the dynamic evolution of technology adoption decisions and their consequences over time. The framework offered by Greenhalgh et al. (2017) emphasizes that non-adoption, abandonment, and the challenges of scaling and sustaining technologies are at least as important as initial adoption, yet these post-adoption dynamics remain poorly understood in firm-level research. Future studies employing longer panel data and event-study methodologies would substantially improve understanding of how adoption trajectories evolve in response to changing competitive conditions, workforce reorganization, and institutional environments.

5.3 Interaction Effects Between Technology Types

The literature tends to examine specific technologies—artificial intelligence, Industry 4.0, additive manufacturing, teleworking—in relative isolation, while in practice firms face portfolio decisions about multiple technologies that may be complementary, substitutable, or conflicting in their organizational demands. Dwivedi et al. (2019) acknowledge this complexity in their discussion of AI's multidisciplinary implications, and Tofail et al. (2017) note the integration challenges associated with additive manufacturing within broader production systems, but neither study systematically addresses how firms manage simultaneous adoption decisions across technology domains. Research that examines the interdependencies among technology adoption decisions—and the organizational capabilities required to manage them—would fill an important gap in both theoretical understanding and practical guidance for managers.

6. Conclusion

This literature review has synthesized evidence from twenty-one peer-reviewed studies to provide an integrated account of the firm-level determinants of technology adoption. The review has identified three primary thematic clusters that collectively capture the main drivers and barriers documented in the contemporary literature. Organizational capabilities—including firm size, absorptive capacity, and pre-existing technological infrastructure—constitute a foundational set of internal determinants that condition firms' ability and willingness to adopt new technologies. Market structure and competitive dynamics shape the

external incentive environment, with evidence that both competitive pressure and market power can function as drivers of adoption under different conditions. Workforce and human capital considerations introduce a critical human dimension, demonstrating that technology adoption entails not merely technical investment but also substantial organizational and labor management challenges.

A cross-cutting theme emerging from all three clusters is the importance of contextual heterogeneity: the determinants of adoption are not uniform across technology types, firm sizes, industrial sectors, or national institutional environments. This heterogeneity implies that generic adoption frameworks have inherent limitations and that theoretical progress will require more fine-grained, context-sensitive models.

The review has several limitations that should be acknowledged. The reliance on a pre-specified set of twenty-one sources, while ensuring conceptual coherence, necessarily circumscribes the range of perspectives captured. Several important dimensions of the adoption literature—including supply-side dynamics, financing constraints, and innovation systems perspectives—receive less attention than they merit in a comprehensive treatment. Furthermore, the disciplinary diversity of the reviewed studies, while a strength in terms of breadth, creates challenges for strict methodological comparability across findings. Future integrative reviews with broader and more systematically assembled source sets would complement and extend the analysis presented here.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. Journal of Economic Perspectives, 33(2), 3–30. https://doi.org/10.1257/jep.33.2.3

Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. Quarterly Journal of Economics, 135(2), 645–709. https://doi.org/10.1093/qje/qjaa004

Belzunegui Eraso, Á. G., & Erro-Garcés, A. (2020). Teleworking in the context of the Covid-19 crisis. Sustainability, 12(9), 3662. https://doi.org/10.3390/su12093662

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. Journal of Economic Literature, 55(3), 789–865. https://doi.org/10.1257/jel.20160995

Buhalis, D., Harwood, T., Bogicevic, V., Viglia, G., & Beldona, S. (2019). Technological disruptions in services: Lessons from tourism and hospitality. Journal of Service Management, 30(4), 484–506. https://doi.org/10.1108/josm-12-2018-0398

Christensen, H. B., Hail, L., & Leuz, C. (2021). Mandatory CSR and sustainability reporting: Economic analysis and literature review. Review of Accounting Studies, 26(3), 1176–1248. https://doi.org/10.1007/s11142-021-09609-5

Compagnucci, L., & Spigarelli, F. (2020). The Third Mission of the university: A systematic literature review on potentials and constraints. Technological Forecasting and Social Change, 161, 120284. https://doi.org/10.1016/j.techfore.2020.120284

Davidescu, A. A., Apostu, S. A., Paul, A., & Căsuneau, I. (2020). Work flexibility, job satisfaction, and job performance among Romanian employees—Implications for sustainable human resource management. Sustainability, 12(15), 6086. https://doi.org/10.3390/su12156086

De Loecker, J., Eeckhout, J., & Unger, G. (2020). The rise of market power and the macroeconomic implications. Quarterly Journal of Economics, 135(2), 561–644. https://doi.org/10.1093/qje/qjz041

Duque Grisales, E. A., & Aguilera-Caracuel, J. (2019). 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

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., & Coombs, C. (2019). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Goldfarb, A., & Tucker, C. E. (2019). Digital economics. Journal of Economic Literature, 57(1), 3–43. https://doi.org/10.1257/jel.20171452

Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., & Hughes, G. (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

Jiang, F., & Kim, K. A. (2020). Corporate governance in China: A survey. European Finance Review, 24(6), 733–772. https://doi.org/10.1093/rof/rfaa012

Müller, J. M., Kiel, D., & Voigt, K.-I. (2018). What drives the implementation of Industry 4.0? The role of opportunities and challenges in the context of sustainability. Sustainability, 10(1), 247. https://doi.org/10.3390/su10010247

Nagy, J., Oláh, J., Erdei, E., Máté, D., & Popp, J. (2018). The role and impact of Industry 4.0 and the Internet of Things on the business strategy of the value chain—The case of Hungary. Sustainability, 10(10), 3491. https://doi.org/10.3390/su10103491

Tofail, S. A. M., Koumoulos, E. P., Bandyopadhyay, A., Bose, S., & O'Donoghue, L. (2017). Additive manufacturing: Scientific and technological challenges, market uptake and opportunities. Materials Today, 21(1), 22–37. https://doi.org/10.1016/j.mattod.2017.07.001

Principal Component Analysis in Marketing Research: A Systematic Literature Review

Hanák Róbert[1] and Hojdik Vladimír[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: robert.hanak@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: vladimir.hojdik@euba.sk

Abstract

Principal Component Analysis (PCA) is among the most widely employed multivariate statistical techniques in marketing research, enabling researchers to reduce high-dimensional datasets into interpretable, lower-dimensional representations without substantial loss of information. Despite its established utility, the application of PCA in marketing contexts remains undertheorized in relation to emerging computational paradigms and evolving research methodologies. This review examines the theoretical foundations, methodological applications, and practical implications of PCA as a dimensionality reduction and data structuring tool in marketing research. Drawing on twenty sources spanning statistical methodology, machine learning, structural equation modeling, and applied behavioral research, this paper identifies three primary thematic clusters: PCA as a dimensionality reduction technique, its integration with multivariate analytical frameworks, and its intersection with contemporary data-driven marketing approaches. The review identifies significant gaps in the literature regarding the integration of PCA with explainable artificial intelligence tools, the handling of ordinal marketing survey data, and applications within large-scale consumer behavior datasets. Future research should address methodological transparency, hybrid modeling approaches, and domain-specific adaptations of PCA. These findings carry important implications for marketing scholars and practitioners seeking rigorous, scalable analytical tools.

Keywords: principal component analysis, marketing research, dimensionality reduction, factor analysis, market structure

JEL Classification: C38, M31, C55

1. Introduction

The exponential growth of consumer data in the digital age has fundamentally transformed the analytical demands placed upon marketing researchers. Contemporary marketing datasets routinely encompass hundreds of variables drawn from survey instruments, transactional records, social media platforms, and behavioral tracking systems, producing data structures of considerable complexity. Classical analytical approaches that assume manageable variable counts and linear relationships among constructs are increasingly inadequate for these high-dimensional environments. In this context, dimensionality reduction techniques—methods designed to represent complex, multivariate data in more parsimonious forms—have assumed a position of central importance within the marketing research toolkit.

Principal Component Analysis (PCA) occupies a foundational role among such techniques. By transforming a set of possibly correlated variables into a smaller set of uncorrelated

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

components that capture the maximum possible variance, PCA facilitates more tractable modeling, clearer interpretive structures, and reduced computational burden. Its applications in marketing span consumer segmentation, brand perception mapping, advertising effectiveness evaluation, and the construction of composite indices for constructs such as customer satisfaction, loyalty, and perceived quality. Despite its long history in the discipline, however, the theoretical and methodological discourse surrounding PCA in marketing remains fragmented. Researchers draw on PCA for purposes ranging from exploratory data reduction prior to regression modeling, to the validation of measurement instruments prior to structural equation modeling, yet these applications are rarely synthesized in a coherent literature review.

The purpose of this literature review is to reassess the role of PCA within the broader evolution of marketing research methodology. Specifically, the paper aims to clarify the conceptual foundations of PCA, synthesize its applications across key marketing sub-fields, and identify underexplored issues that are increasingly important in the context of contemporary marketing analytics and data-driven decision-making.

2. Methodological Foundations of PCA in Marketing Analysis

2.1 Dimensionality Reduction in Multivariate Research

Dimensionality reduction refers to the process of transforming high-dimensional data into a lower-dimensional representation while preserving as much of the original information structure as possible. In marketing research, where survey instruments routinely measure latent constructs through multiple manifest indicators, and where transactional databases capture consumer behavior across dozens of product categories, the need for principled data reduction is pervasive. At its core, dimensionality reduction serves two related purposes: it reduces computational and analytical complexity, and it enables more interpretable representations of underlying data structure. Qiu et al. (2016) situate dimensionality reduction within the broader framework of machine learning for big data, noting that representation learning—of which PCA is an early and influential example—is among the most promising strategies for making sense of massive, high-dimensional datasets. Tsai et al. (2015) similarly identify feature extraction and dimensionality reduction as foundational components of scalable data analytics pipelines, emphasizing that the choice of reduction technique has profound downstream effects on analytical conclusions.

2.2 Principal Component Analysis: Definition and Mathematical Foundations

PCA is a linear algebraic technique that identifies orthogonal directions of maximum variance in a multivariate dataset, projecting the original data onto these directions—termed principal components—to produce a reduced representation. Formally, PCA computes the eigenvectors of the data covariance matrix, ordering components by their associated eigenvalues such that the first principal component captures the greatest proportion of total variance, the second component captures the greatest remaining variance orthogonal to the first, and so forth. The resulting components are uncorrelated by construction, facilitating their use as predictor variables in subsequent regression or classification analyses.

In marketing research, PCA is frequently invoked either as a standalone exploratory tool or as a preprocessing step prior to more complex analyses. Zhang et al. (2024) demonstrate the utility of PCA as a dimensionality reduction step in a composite index modeling framework, combining PCA with entropy weighting and coefficient of variation methods to produce more robust and interpretable quality assessments. While their application concerns groundwater

quality rather than marketing, the methodological logic—using PCA to collapse correlated indicators into interpretable components before constructing weighted composite scores—maps precisely onto common practices in marketing measurement, such as the construction of customer satisfaction indices or brand equity scores.

2.3 Relationship Between PCA and Factor Analysis

A critical theoretical distinction in marketing research concerns the relationship between PCA and Exploratory Factor Analysis (EFA). Although the two techniques are frequently conflated in applied research, they rest on distinct assumptions. EFA assumes that observed variables are caused by a smaller number of latent factors, along with unique error terms; PCA makes no such assumption, treating all variance in the observed variables—including error—as potentially explainable. In practice, the two methods often yield similar results for large datasets with high communalities, but the distinction carries meaningful implications for construct validity in marketing measurement contexts.

Baglin (2020) provides a thorough methodological treatment of EFA for ordinal data, arguing that commonly used practices—including the application of Pearson correlations and maximum likelihood extraction on Likert-scale items—introduce systematic biases. Baglin's (2020) recommendations regarding polychoric correlations and appropriate extraction methods have direct implications for PCA applied to marketing survey data, since the majority of consumer attitude and perception measures employ ordinal rather than continuous response scales. Importantly, researchers who conflate PCA with EFA may reach erroneous conclusions about the dimensionality of marketing constructs, underscoring the need for greater methodological precision.

2.4 Structural Equation Modeling and PCA in Marketing

The relationship between PCA and structural equation modeling (SEM) represents another foundational theoretical intersection. SEM frameworks in marketing research must adjudicate between composite-based and factor-based measurement models. Henseler (2017) articulates this distinction within variance-based SEM, explaining that composite measurement models—analogous in spirit to PCA—treat constructs as exact linear combinations of their indicators, whereas common factor models treat indicators as reflections of an underlying latent variable. The choice between these paradigms has profound implications for how PCA-derived components are interpreted and incorporated into broader causal models of consumer behavior.

Henseler (2017) further notes that variance-based SEM, such as Partial Least Squares (PLS), can accommodate both composite and factor-based measurement, offering a flexible framework for marketing researchers who wish to exploit the interpretability of PCA-like composites without abandoning the rigor of latent variable modeling. This theoretical bridge between PCA and SEM is particularly relevant for marketing applications involving formative constructs—composites where the indicators jointly define the construct, rather than reflecting an underlying cause—such as socioeconomic status indices or multidimensional brand positioning maps.

2.5 Machine Learning Contextualization

Contemporary machine learning literature positions PCA within a broader family of representation learning approaches that also includes autoencoders, deep feature extractors, and manifold learning methods. Shrestha and Mahmood (2019) review deep learning architectures and note that early autoencoder designs were conceptually motivated by PCA,

extending its linear projection logic into nonlinear, hierarchical feature extraction. Berahmand et al. (2024) similarly position autoencoders as generalizations of PCA, capable of capturing nonlinear structure in data that linear decomposition cannot represent. These comparisons are theoretically important for marketing researchers, as they establish the conceptual continuity between classical PCA and modern machine learning approaches to consumer data analysis.

3. PCA Across Marketing Research Applications

3.1 PCA as a Dimensionality Reduction Tool in Marketing Measurement and Segmentation

The most extensively documented application of PCA in marketing research concerns its use as a dimensionality reduction tool for measurement instruments and consumer segmentation tasks. Marketing surveys routinely employ multi-item scales to measure constructs such as brand attitude, purchase intention, consumer involvement, and customer satisfaction. When these scales include large numbers of items, multicollinearity among indicators can undermine subsequent analyses, and the interpretive challenge of managing many correlated variables simultaneously becomes acute. PCA addresses both problems by reducing the item space to a smaller set of orthogonal components that retain the majority of original variance.

This logic extends naturally to consumer segmentation, where researchers seek to identify distinct groups of consumers based on attitude, behavior, or preference profiles measured across numerous variables. By first reducing the variable space using PCA, analysts can construct more stable and interpretable cluster solutions that are less susceptible to the curse of dimensionality. Qiu et al. (2016) emphasize that dimensionality reduction is not merely a computational convenience in such contexts but a substantive analytical choice that shapes which patterns in the data are surfaced and which are suppressed. Tsai et al. (2015) reinforce this point by noting that the design of the dimensionality reduction step is inseparable from the design of the downstream analytical pipeline—a consideration that marketing researchers who treat PCA as a routine preprocessing step may underappreciate.

The methodological rigor of PCA applications in marketing has been critiqued on several fronts. A recurring concern involves the treatment of ordinal Likert-scale data as though they were continuous, an assumption that enables the use of Pearson correlation matrices in PCA but that may distort the resulting component structure. Baglin (2020) demonstrates through empirical simulation that the use of polychoric rather than Pearson correlations in factor analytic procedures substantially alters the dimensionality conclusions reached, with direct implications for PCA applied to marketing survey data. This finding suggests that a meaningful proportion of published marketing research employing PCA on Likert-scale instruments may have reached component solutions that are artifacts of inappropriate correlation assumptions rather than genuine reflections of underlying attitudinal structure.

Zhang et al. (2024) contribute a methodologically instructive example of principled PCA integration, combining PCA-derived components with additional weighting methods to optimize the information content retained in the reduced representation. Although their application is environmental rather than marketing-focused, the methodological principle—using PCA not merely to reduce dimensionality but to optimize the weighting of retained information—offers a template for more sophisticated marketing measurement applications, particularly in the construction of composite indices for multi-dimensional constructs such as customer lifetime value or brand equity.

3.2 PCA Within Multivariate Structural and Behavioral Modeling Frameworks

A second and theoretically rich area of PCA application in marketing concerns its role as a preprocessing or structural component within larger multivariate modeling frameworks, including structural equation modeling, regression analysis, and analysis of variance. In these contexts, PCA serves not as an endpoint in itself but as a mechanism for constructing analytically tractable input variables that preserve the substantive richness of the original measurement instruments.

Henseler (2017) provides the most theoretically developed treatment of this intersection, articulating the ways in which composite-based constructs derived from PCA-like aggregation procedures can be embedded within SEM frameworks that also accommodate reflective latent variable measurement. Henseler's (2017) framework is particularly relevant for advertising and marketing communication research, where constructs such as message elaboration, ad liking, and purchase intention occupy different points on the continuum between purely formative and purely reflective measurement. The variance-based SEM methods he advocates offer a principled approach to handling PCA-derived composites alongside traditional latent variables within integrated causal models.

Tortorella et al. (2019) employ multivariate data analysis, including PCA-type procedures, within an operations management context examining the moderating role of Industry 4.0 technologies on lean production performance relationships. Their study illustrates how PCA can facilitate the construction of meaningful composite scores for complex, multi-indicator constructs—in their case, lean practice bundles and Industry 4.0 technology adoption bundles—that serve as inputs to regression and interaction analyses. The marketing analog is direct: researchers studying complex constructs such as omnichannel retail strategy, integrated marketing communications, or service quality may similarly benefit from PCA-based dimension reduction prior to testing theoretically specified causal models.

Alhadabi and Karpinski (2019) demonstrate the utility of mediation modeling frameworks in educational psychology, and while their substantive focus concerns grit and academic achievement rather than marketing phenomena, their methodological approach—using latent constructs operationalized through multi-item scales in serial mediation models—parallels common practices in consumer behavior research. Their work raises important questions about measurement model specification that are directly relevant to marketing researchers who employ PCA as a preliminary step before mediation analysis. Specifically, when PCA-derived component scores are used as variables in mediation models, the implicit assumption of measurement equivalence across the components may not hold, potentially biasing indirect effect estimates.

Kiefer et al. (2018) provide an empirically grounded example of PCA applied in a business strategy context, using quantitative analysis to examine eco-innovation drivers. Their application of multivariate reduction techniques to Likert-based organizational survey data illustrates the practical challenges of applying PCA to ordinal constructs in complex behavioral models—challenges that are equally present in marketing research examining consumer attitudes toward sustainable products, corporate social responsibility, or innovation adoption.

3.3 PCA at the Intersection of Marketing Analytics and Machine Learning

The third and most forward-looking thematic cluster concerns the positioning of PCA within the rapidly evolving landscape of data-driven marketing analytics. The proliferation of machine learning tools in marketing—encompassing recommendation systems, consumer churn

prediction, customer lifetime value modeling, and natural language processing for brand monitoring—has introduced both new opportunities and new challenges for PCA as an analytical technique.

Recommendation systems, which have become a cornerstone of digital marketing strategy, illustrate the tension between classical PCA-based approaches and more recent machine learning alternatives. Fayyaz et al. (2020) survey recommendation system algorithms, noting that matrix factorization techniques—closely related to PCA—remain among the most effective approaches for collaborative filtering applications such as product recommendation and content personalization. Their review documents the continued relevance of linear decomposition methods even in competitive comparison with deep learning alternatives, particularly in settings where interpretability and computational efficiency are prioritized alongside predictive accuracy. This finding is significant for marketing practitioners, as it suggests that PCA-based approaches need not be abandoned in favor of black-box machine learning models when transparency and explainability are valued.

The interpretability concern is directly addressed in the explainable artificial intelligence (XAI) literature. Burkart and Huber (2021) provide a comprehensive survey of explainability methods for supervised machine learning, noting that the opacity of complex models poses particular challenges in high-stakes decision-making contexts. In marketing, where decisions about segmentation, targeting, and personalization carry both commercial and ethical implications, the relative transparency of PCA—where each component can be inspected in terms of its loadings on original variables—constitutes a genuine methodological advantage. Zhang et al. (2022) extend this argument to cybersecurity applications but note that XAI principles apply broadly to any domain where model decisions must be auditable and interpretable by non-technical stakeholders, a description that applies directly to marketing management decision-making.

Shrestha and Mahmood (2019) situate PCA historically within the development of deep learning, noting that autoencoder architectures generalize PCA by introducing nonlinear activation functions that enable the capture of complex, curved data manifolds. Berahmand et al. (2024) elaborate this lineage in detail, documenting how autoencoders have been applied to dimensionality reduction tasks in domains including image recognition, anomaly detection, and recommendation systems. Bao et al. (2017) demonstrate the power of stacked autoencoders for financial time series forecasting, a methodology with direct analogs in marketing applications such as demand forecasting and customer behavior prediction. The implication for marketing researchers is that PCA can serve as a conceptual and practical baseline against which more complex representation learning methods are evaluated, with hybrid approaches combining the interpretability of PCA with the expressive power of nonlinear methods representing a promising frontier.

The application of natural language processing in marketing analytics introduces further complexity regarding the role of PCA. Calvo et al. (2017) document NLP applications that infer psychological states from social media text, a methodological paradigm increasingly applied in marketing for brand sentiment analysis, consumer complaint mining, and social listening. Berman et al. (2019) and Shaukat et al. (2020) review deep learning applications that address classification and anomaly detection tasks in high-dimensional data environments. While neither is primarily concerned with marketing, their treatments of feature extraction in high-dimensional spaces are methodologically instructive: in text-based marketing data, PCA applied to term-document matrices or word embedding representations can surface thematic dimensions in consumer discourse that are not apparent from examining individual terms or

documents. Choi et al. (2021) address anomaly detection in time-series data using deep learning methods, an approach relevant for marketing applications involving the detection of unusual consumer behavior patterns in longitudinal transaction data.

The integration of PCA with big data infrastructures represents a final important dimension of this thematic cluster. Tsai et al. (2015) document the computational challenges of applying traditional analytics at scale, noting that algorithms designed for small-to-medium datasets may require substantial adaptation before they can be deployed on data streams of the scale generated by modern digital marketing ecosystems. Qiu et al. (2016) complement this perspective by reviewing parallel and distributed implementations of machine learning algorithms, including dimensionality reduction methods, that address the scalability constraints of classical PCA. For marketing researchers and practitioners working with large-scale consumer panels, clickstream data, or social media corpora, these computational developments are practically significant: they suggest that the analytical logic of PCA can be preserved in scalable implementations without abandoning the interpretive advantages that make PCA attractive for marketing applications.

4. Methodological Blind Spots and Emerging Opportunities

4.1 Insufficient Methodological Guidance for Ordinal Marketing Data

A substantial and underaddressed gap in the marketing research literature concerns the appropriate application of PCA to ordinal data. As Baglin (2020) demonstrates, the common practice of applying PCA with Pearson correlation matrices to Likert-scale responses produces component solutions that may differ systematically from those obtained using polychoric correlations, which better respect the ordinal measurement level of the data. Despite this methodological evidence, marketing research practice has been slow to adopt alternative correlation matrices in PCA applications, partly due to limitations of widely used commercial statistical software and partly due to insufficient methodological training. Future research should develop applied guidelines specifically tailored to marketing researchers, comparing PCA solutions based on Pearson and polychoric correlations across representative marketing measurement contexts, and providing practical recommendations regarding the conditions under which the choice of correlation matrix materially affects substantive conclusions. Such guidelines would have direct implications for the validity of consumer segmentation, brand mapping, and scale development studies.

4.2 Integration of PCA with Explainable Artificial Intelligence in Marketing Contexts

A second important gap concerns the limited integration of PCA with explainable artificial intelligence frameworks in marketing analytics applications. Burkart and Huber (2021) and Zhang et al. (2022) document the growing importance of model interpretability in high-stakes decision-making domains, yet the marketing literature has not systematically explored how PCA can be embedded within XAI pipelines to enhance the interpretability of machine learning models used for consumer targeting, churn prediction, or personalization. Future research should investigate hybrid frameworks in which PCA serves as a structured dimensionality reduction step that simultaneously reduces computational complexity and enhances the interpretability of downstream machine learning models. Such frameworks would be particularly valuable for marketing practitioners who must justify algorithmic marketing decisions to management stakeholders, regulators, and consumers themselves. The theoretical work of Fayyaz et al. (2020) on recommendation systems and of Burkart and Huber (2021) on explainability provides a foundation for this integrative agenda.

4.3 PCA in Dynamic and Longitudinal Marketing Data Environments

A third gap concerns the application of PCA in dynamic and longitudinal marketing contexts. Most PCA applications in marketing treat data as static cross-sections, extracting components from a single measurement occasion and applying the resulting structure to further analysis. However, many marketing phenomena of theoretical and practical interest—including brand perception shifts, customer relationship evolution, and the dynamics of advertising response—unfold over time in ways that static PCA cannot capture. The anomaly detection and time-series analysis literature, exemplified by Choi et al. (2021) and Bao et al. (2017), offers methodological approaches that could be adapted for dynamic marketing data environments. Specifically, rolling window PCA, dynamic factor models, and hybrid PCA-LSTM architectures represent promising methodological developments that could enable marketing researchers to track the evolution of consumer attitude structures, competitive positioning maps, and segment membership over time. Future research should develop and validate such dynamic extensions of PCA in representative marketing research contexts, connecting the statistical methodology to substantive marketing theory regarding brand and consumer dynamics.

5. Concluding Remarks

The reviewed literature confirms that PCA remains a highly valuable methodological tool in marketing research, but also shows that its contemporary relevance depends increasingly on how carefully it is aligned with modern data conditions and analytical goals. Across the literature, three broad areas of application stand out: PCA as a dimensionality reduction and measurement tool, PCA within wider multivariate and structural modeling frameworks, and PCA in the context of contemporary marketing analytics and machine learning. Taken together, these strands indicate that PCA is not simply a legacy statistical method, but a technique whose usefulness persists when applied with sufficient methodological awareness and transparency.

The contributions of this review are following:

a) establishment of theoretically coherent account of PCA's position within the broader landscape of multivariate marketing research, drawing connections between PCA and related techniques including EFA, variance-based SEM, autoencoders, and recommendation system algorithms;

b) synthesis of empirical and methodological findings across diverse marketing and adjacent literatures, surfacing cross-disciplinary insights that have not previously been assembled in a single marketing-focused review;

c) identification of specific, actionable research gaps and proposes directions for future inquiry that connect methodological development to substantive marketing theory and practice.

Several limitations of this review warrant acknowledgment. The source base, while diverse in disciplinary scope, does not include a large number of primary empirical studies focused specifically on PCA in marketing, which reflects the constraints of the available literature rather than a comprehensive search of the marketing research corpus. Additionally, the thematic synthesis necessarily involves interpretive choices that alternative reviewers might make differently. Future systematic reviews with broader search protocols and formal inclusion criteria would strengthen the evidence base for the conclusions reached here. Notwithstanding these limitations, this review offers a valuable contribution to the methodological discourse surrounding quantitative marketing research.

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK | EKONOMICKÁ UNIVERZITA V BRATISLAVE | FAKULTA PODNIKOVÉHO MANAŽMENTU

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Ajay Shrestha, & Ausif Mahmood. (2019). Review of deep learning algorithms and architectures. IEEE Access, 7, 53040–53065. https://doi.org/10.1109/access.2019.2912200

Alhadabi, A., & Karpinski, A. C. (2019). Grit, self-efficacy, achievement orientation goals, and academic performance in university students. International Journal of Adolescence and Youth, 25(1), 519–535. https://doi.org/10.1080/02673843.2019.1679202

Baglin, J. (2020). Improving your exploratory factor analysis for ordinal data: A demonstration using FACTOR. Practical Assessment, Research, and Evaluation, 19(1). https://doi.org/10.7275/dsep-4220

Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PLoS ONE, 12(7), e0180944. https://doi.org/10.1371/journal.pone.0180944

Berahmand, K., Daneshfar, F., Salehi, E. S., Li, Y., & Xu, Y. (2024). Autoencoders and their applications in machine learning: A survey. Artificial Intelligence Review, 57(2), 28. https://doi.org/10.1007/s10462-023-10662-6

Berman, D. S., Buczak, A. L., Chavis, J. S., & Corbett, C. (2019). A survey of deep learning methods for cyber security. Information, 10(4), 122. https://doi.org/10.3390/info10040122

Burkart, N., & Huber, M. F. (2021). A survey on the explainability of supervised machine learning. Journal of Artificial Intelligence Research, 70, 245–317. https://doi.org/10.1613/jair.1.12228

Calvo, R. A., Milne, D., Hussain, M. S., & Christensen, H. (2017). Natural language processing in mental health applications using non-clinical texts. Natural Language Engineering, 23(5), 649–685. https://doi.org/10.1017/s1351324916000383

Choi, K., Yi, J., Park, C., & Yoon, S. (2021). Deep learning for anomaly detection in time-series data: Review, analysis, and guidelines. IEEE Access, 9, 120043–120065. https://doi.org/10.1109/access.2021.3107975

Fayyaz, Z., Ebrahimian, M., Nawara, D., Ibrahim, A., & Kashef, R. (2020). Recommendation systems: Algorithms, challenges, metrics, and business opportunities. Applied Sciences, 10(21), 7748. https://doi.org/10.3390/app10217748

Henseler, J. (2017). Bridging design and behavioral research with variance-based structural equation modeling. Journal of Advertising, 46(1), 178–192. https://doi.org/10.1080/00913367.2017.1281780

Kiefer, C. P., del Río González, P., & Carrillo-Hermosilla, J. (2018). Drivers and barriers of eco-innovation types for sustainable transitions: A quantitative perspective. Business Strategy and the Environment, 28(1), 155–172. https://doi.org/10.1002/bse.2246

Qiu, J., Wu, Q., Ding, G., Xu, Y., & Feng, S. (2016). A survey of machine learning for big data processing. EURASIP Journal on Advances in Signal Processing, 2016(1), 67. https://doi.org/10.1186/s13634-016-0355-x

Shaukat, K., Luo, S., Varadharajan, V., Hameed, I. A., & Xu, M. (2020). A survey on machine learning techniques for cyber security in the last decade. IEEE Access, 8, 222310–222354. https://doi.org/10.1109/access.2020.3041951

Tortorella, G. L., Giglio, R., & van Dun, D. H. (2019). Industry 4.0 adoption as a moderator of the impact of lean production practices on operational performance improvement. International Journal of Operations & Production Management, 39(6/7/8), 860–886. https://doi.org/10.1108/ijopm-01-2019-0005

Tsai, C.-W., Lai, C.-F., Chao, H.-C., & Vasilakos, A. V. (2015). Big data analytics: A survey. Journal of Big Data, 2(1), 21. https://doi.org/10.1186/s40537-015-0030-3

Zhang, B., Hu, X., Li, B., Wu, P., & Cai, X. (2024). A groundwater quality assessment model for water quality index: Combining principal component analysis, entropy weight method, and coefficient of variation method for dimensionality reduction and weight optimization, and its application. Water Environment Research, 96(1), e11155. https://doi.org/10.1002/wer.11155

Zhang, Z., Al Hamadi, H., Damiani, E., Yeun, C. Y., & Taher, F. (2022). Explainable artificial intelligence applications in cyber security: State-of-the-art in research. IEEE Access, 10, 93104–93139. https://doi.org/10.1109/access.2022.3204051

Consumer Preferences for AI-Based Products and Services: A Systematic Research Synthesis

Zuzana Hajduová[1] and Dana Hrušovská[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: zuzana.hajduova@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: dana.hrusovska@euba.sk

Abstract

The rapid proliferation of artificial intelligence (AI) technologies has fundamentally reshaped the landscape of consumer-facing products and services, necessitating a rigorous academic examination of how and why consumers engage with these technologies. This literature review examines the multidisciplinary scholarly discourse surrounding consumer preferences for AI-based products and services, synthesizing evidence from sixteen peer-reviewed sources spanning management, marketing, human-computer interaction, healthcare, and information systems. The review identifies three primary thematic clusters: (1) the determinants of consumer acceptance and trust in AI-driven systems, (2) the role of human-centered design and ethical AI in shaping consumer attitudes, and (3) the sector-specific manifestations of AI adoption and consumer preference formation. Key findings suggest that trust, perceived usefulness, transparency, and ethical alignment are critical antecedents of consumer preferences for AI products, while cognitive offloading and behavioral dependencies present emerging concerns. Significant research gaps persist regarding longitudinal dynamics of consumer preferences, cross-cultural variation, and the interplay between AI regulation and market behavior. The implications of these findings extend to product designers, marketers, policymakers, and academic researchers seeking to understand the evolving human-AI relationship in commercial contexts.

Key Words: adoption intentions, AI-based products, consumer preferences, willingness to pay, digital services, willingness to pay

JEL Classification: D12, M31, O33

1. Introduction

The emergence of artificial intelligence as a transformative technological paradigm has driven profound changes across virtually every sector of the modern economy, from healthcare and agriculture to retail, financial services, and beyond. As AI-based products and services become increasingly embedded in everyday consumer life—from voice-activated assistants and recommendation algorithms to diagnostic health tools and AI-powered sales platforms—understanding how consumers form preferences, evaluate value, and ultimately adopt or reject these technologies has become increasingly important for both research and practice. The stakes are considerable: misaligned product design, insufficient transparency, or consumer distrust can result not only in commercial failure but also in broader societal harms, including the amplification of inequalities and the erosion of critical thinking capacities (Gerlich, 2025; Díaz-Rodríguez et al., 2023).

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

Despite the growing body of literature on AI adoption, a clear and persistent gap remains in the systematic synthesis of consumer-oriented perspectives. Much of the extant research addresses AI from technological, organizational, or policy standpoints, often treating consumers as passive recipients of AI-driven change rather than as active agents whose preferences, values, and cognitive dispositions fundamentally shape the trajectory of AI markets. Moreover, the literature frequently fragments across disciplinary silos, with marketing scholars examining brand engagement in virtual environments (Dwivedi et al., 2022), health informaticists studying digital transformation in patient-facing systems (Stoumpos et al., 2023), and human-computer interaction researchers articulating grand challenges in human-centered AI design (Garibay et al., 2023) yet these streams are rarely synthesized into a coherent account of consumer preference formation.

This literature review addresses that gap by systematically synthesizing the available evidence on consumer preferences for AI-based products and services across multiple domains and disciplinary perspectives. The specific objectives are threefold: first, to establish the theoretical foundations that govern consumer engagement with AI technologies; second, to identify and analyze the principal thematic clusters in the literature pertaining to consumer preference; and third, to delineate the most significant research gaps and propose directions for future inquiry.

The remainder of this paper is structured as follows. Section 2 establishes the theoretical background, defining key constructs and situating the review within foundational frameworks relevant to AI and consumer behavior. Section 3 presents the thematic literature review organized around three distinct clusters. Section 4 identifies research gaps and articulates directions for future scholarship. Section 5 concludes with a summary of contributions and a discussion of the review's limitations.

2. Theoretical Background

2.1 Defining Artificial Intelligence in Consumer Contexts

Artificial intelligence, broadly construed, refers to the capacity of computational systems to perform tasks that would ordinarily require human intelligence, including reasoning, learning, pattern recognition, and natural language processing. Dwivedi et al. (2019) situate AI within a broad multidisciplinary framework, emphasizing its transformative potential across industrial, intellectual, and social applications while acknowledging the significant challenges it poses for research, practice, and policy. In consumer contexts, AI manifests in diverse forms: intelligent recommendation engines, conversational agents, autonomous decision-support systems, and generative content tools, among others. Each of these applications presents distinct configurations of human-AI interaction that shape preference formation in different ways.

The scope of AI-based consumer products has expanded considerably with the maturation of enabling technologies. The deployment of multi-access edge computing in 5G networks, for instance, has enabled real-time AI applications at the consumer device level (Pham et al., 2020), while digital twin technologies—capable of creating dynamic software replicas of physical objects and environments—are beginning to reshape how consumers interact with products across the Internet of Things ecosystem (Minerva et al., 2020). These technological enablers do not merely expand the range of AI applications; they fundamentally alter the temporal and spatial conditions under which consumers encounter and evaluate AI-driven systems.

2.2 Foundational Theories of Technology Acceptance

The theoretical architecture underlying consumer preferences for AI products draws heavily from established frameworks of technology acceptance. The Technology Acceptance Model (TAM), originally proposed by Davis (1989), posits that perceived usefulness and perceived ease of use are the primary determinants of technology adoption intentions. This model has been extensively applied and refined in the context of digital health technologies, where Stoumpos et al. (2023) demonstrate that patient acceptance of AI-enabled healthcare services is mediated by factors including trust in digital systems, perceived quality of care, and institutional credibility. The integration of digital tools in healthcare exemplifies how foundational acceptance theories must be adapted to account for domain-specific variables such as vulnerability, professional authority, and regulatory oversight.

Beyond TAM, the Unified Theory of Acceptance and Use of Technology (UTAUT) framework extends the explanatory scope to encompass social influence, facilitating conditions, and hedonic motivation as determinants of adoption behavior. These extensions are particularly relevant in the context of AI-powered marketing environments, where the social dimensions of technology use—including peer influence, brand community dynamics, and experiential engagement—play a central role in shaping consumer response (Dwivedi et al., 2022).

2.3 Human-Centered AI and the Ethical Dimension

A growing theoretical strand foregrounds the importance of human-centered design principles in AI development as a precondition for consumer acceptance. Garibay et al. (2023) articulate six grand challenges for the scientific community in this domain, emphasizing that AI systems must be ethical, fair, and genuinely enhancing of the human condition if widespread and sustainable adoption is to be achieved. This perspective aligns with the trustworthy AI framework articulated by Díaz-Rodríguez et al. (2023), who identify lawfulness, ethical alignment, and technical robustness as the three foundational pillars upon which consumer trust in AI systems must be constructed. These twin frameworks collectively suggest that consumer preferences for AI products cannot be disentangled from broader normative evaluations of whether such products respect human agency, dignity, and welfare.

2.4 Cognitive and Behavioral Dimensions of AI Preference

An important but underexplored theoretical dimension concerns the cognitive consequences of sustained AI use and their feedback effects on consumer preferences. Gerlich (2025) provides empirical evidence that extensive reliance on AI tools is associated with cognitive offloading—the transfer of cognitive tasks to external systems—which in turn is negatively associated with critical thinking capacity. This dynamic introduces a paradox for consumer preference theory: the very features that make AI products attractive in the short term (convenience, speed, reduced cognitive effort) may, over time, erode the cognitive competencies that consumers rely upon to evaluate products critically. Understanding this paradox requires integrating insights from behavioral economics and cognitive psychology into the predominantly rationalist frameworks that currently dominate technology acceptance research.

2.5 AI in Organizational and Sales Contexts

The theoretical landscape must also account for the organizational mediation of AI-consumer relationships. Singh et al. (2019) demonstrate that AI-driven digitalization in the sales profession reshapes not only the tools available to sales professionals but the very nature of

human-to-human interaction in commercial contexts. As AI assumes greater responsibility for customer profiling, predictive analytics, and automated communication, the consumer's encounter with commercial AI becomes increasingly complex, raising questions about authenticity, personalization, and the boundaries of acceptable algorithmic influence. These considerations are equally pertinent in the context of generative AI, where Budhwar et al. (2023) note the potentially transformative but deeply ambiguous implications of large language model deployment across economic sectors.

3. Literature Review: Systematic Research Synthesis

3.1 Thematic Cluster I: Trust, Transparency, and Consumer Acceptance of AI Products

Trust emerges as perhaps the most consistently cited determinant of consumer preference for AI-based products and services across the reviewed literature. The centrality of trust reflects the fundamental asymmetry of the human-AI relationship: consumers are typically unable to scrutinize the algorithms, training data, or decision-making logic embedded within AI systems, and must therefore extend a form of confidence that is qualitatively different from trust in conventional products. Díaz-Rodríguez et al. (2023) ground this observation in a rigorous framework of trustworthy AI, arguing that consumer trust is not merely a psychological disposition but a systemic property that must be engineered into AI products and services through design choices, governance structures, and regulatory compliance. Their taxonomy of seven technical requirements—spanning human oversight, privacy protection, and accountability—provides a normative reference point for evaluating the trustworthiness of consumer-facing AI systems.

The relationship between transparency and trust is especially salient. Weidinger et al. (2022) document a comprehensive taxonomy of risks posed by large language models—encompassing discrimination, information hazards, misinformation, and human-computer interaction harms—that collectively undermine consumer confidence when they materialize in product experiences. These risks are not hypothetical: consumers who encounter AI systems that produce biased outputs, generate misleading information, or exploit psychological vulnerabilities are likely to update their preferences accordingly, often in the direction of rejection or avoidance. The implication for product design is clear: transparency about AI capabilities and limitations is not merely an ethical obligation but a commercial imperative.

The healthcare domain provides particularly rich evidence on this theme. Stoumpos et al. (2023) conduct a systematic bibliometric review demonstrating that patient acceptance of AI-driven digital health tools is contingent upon institutional trust, perceived data security, and the degree to which technology is perceived as complementing rather than supplanting human clinical judgment. This finding resonates with the broader consumer psychology literature, which consistently shows that preferences for AI assistance are moderated by the perceived stakes of the decision domain: consumers are more willing to rely on AI for low-stakes tasks (e.g., music recommendations) than for high-stakes decisions (e.g., medical diagnoses), a pattern that has been termed algorithm aversion under conditions of observational learning. Transparency mechanisms that make AI reasoning legible to non-expert consumers thus represent a critical design feature for building acceptance in high-stakes contexts.

Dwivedi et al. (2019) further contextualize the trust challenge by situating it within the broader multidisciplinary discourse on AI opportunities and risks, emphasizing that building consumer confidence requires not only technical innovation but also institutional frameworks that govern AI deployment responsibly. The convergence of technical and governance

perspectives in this literature suggests that consumer trust in AI products is a systemic rather than individual-level phenomenon, shaped as much by policy environments and industry norms as by individual product features.

3.2 Thematic Cluster II: Human-Centered Design, Personalization, and Consumer Experience

A second thematic cluster concerns the role of human-centered design principles in shaping consumer preferences for AI products. The central argument in this strand is that AI technologies will only achieve widespread and sustained consumer acceptance if they are designed with genuine attention to human needs, values, and cognitive capacities—a perspective that stands in contrast to purely technology-push approaches to AI product development.

Garibay et al. (2023) present this argument most comprehensively, identifying six grand challenges for human-centered AI that span explainability, fairness, privacy, augmentation of human capabilities, well-being, and democratic accountability. Critically, these challenges are framed not as abstract ethical desiderata but as concrete design requirements whose fulfillment is necessary for AI systems to earn and maintain consumer confidence. The emphasis on augmentation rather than replacement is particularly noteworthy: consumers tend to prefer AI products that enhance their own capabilities and judgment rather than those that seek to supplant human involvement entirely—a preference pattern with direct implications for product positioning and marketing communication.

The marketing literature provides complementary evidence on the role of experiential and emotional factors in AI preference formation. Dwivedi et al. (2022) examine the emerging metaverse as a site of AI-mediated consumer-brand interaction, arguing that immersive, AI-powered virtual environments create novel opportunities for personalized, emotionally resonant brand experiences. Their analysis suggests that consumer preferences in this context are shaped not only by the functional performance of AI features but by the quality and authenticity of the overall experiential environment—a finding that aligns with hedonic consumption theory and the growing literature on experiential marketing. Personalization emerges as a particularly valued dimension of AI product experience, provided that consumers perceive it as being achieved through legitimate rather than invasive means.

The cognitive dimension of consumer experience with AI products is examined by Gerlich (2025), whose empirical investigation of 666 participants reveals a significant negative relationship between AI tool usage frequency and critical thinking skills, mediated by cognitive offloading. This finding has important implications for consumer preference research: if AI products systematically reduce consumers' capacity for critical evaluation, this may create a self-reinforcing cycle of preference for AI assistance that is difficult to interrupt without deliberate intervention. From a design ethics perspective, this suggests that truly human-centered AI products should be designed not merely to satisfy immediate consumer preferences but to sustain and develop the cognitive capacities upon which genuine preference formation depends—a consideration that adds a temporal and developmental dimension to conventional preference frameworks.

In the sales domain, Singh et al. (2019) demonstrate that consumers' responses to AI-mediated sales interactions are conditioned by the degree to which AI augments rather than eliminates the human relational dimension of the sales encounter. Consumers consistently express preference for interactions in which AI tools are deployed to enhance salesperson responsiveness, contextual awareness, and service quality, rather than to replace human

judgment in complex, emotionally significant transactions. This finding is consistent with the broader literature on human-robot interaction, which suggests that consumer acceptance of AI agents is positively associated with perceived warmth and social responsiveness.

3.3 Thematic Cluster III: Sector-Specific AI Adoption and Contextual Preference Determinants

The third thematic cluster examines how the determinants of consumer preference for AI products vary across specific application domains, illuminating the contextual specificity of AI adoption dynamics. While cross-cutting factors such as trust and transparency retain their relevance, the literature reveals important sector-specific moderators that shape how consumers evaluate and engage with AI technologies in practice.

In healthcare, the digital transformation wrought by AI technologies is accompanied by a distinctive set of consumer concerns centered on privacy, professional accountability, and the perceived sanctity of the therapeutic relationship. Stoumpos et al. (2023) identify a consistent preference among patients for AI systems that operate as decision-support tools under clinician supervision, rather than as autonomous diagnostic or treatment agents. This preference is shaped by institutional factors—including regulatory frameworks, professional norms, and organizational trust—as much as by individual psychological dispositions, underscoring the importance of contextual analysis in AI preference research.

In agricultural and industrial contexts, Araújo et al. (2021) document the uneven adoption of Agriculture 4.0 technologies—encompassing AI, IoT, robotics, and big data—noting that adoption is mediated by factors including user technical literacy, infrastructure availability, and perceived return on investment. While this literature focuses primarily on organizational rather than individual consumer preferences, it highlights the importance of accessibility and usability as determinants of AI adoption among populations with heterogeneous technological capabilities. The parallels with consumer market adoption are instructive: products that fail to accommodate diverse user competency levels will systematically exclude segments of the potential consumer base.

The sustainability dimension of AI adoption is examined by Martínez-Peláez et al. (2023), who demonstrate that consumers and organizational stakeholders increasingly factor environmental and social sustainability into their evaluations of digital transformation initiatives. As AI-powered products and services become more energy-intensive—a concern raised in the context of large language model training and deployment by Weidinger et al. (2022)—consumer preferences may increasingly be shaped by sustainability credentials, particularly among environmentally conscious market segments. This trend suggests that the competitive landscape for AI products will be increasingly contoured by sustainability positioning, adding an ethical-environmental dimension to the conventional functional-hedonic preference framework.

In the domain of human resources and organizational behavior, Budhwar et al. (2023) examine consumer and employee responses to generative AI technologies such as ChatGPT, finding that perceptions of job displacement, informational reliability, and democratic accountability significantly shape attitudes toward AI adoption. These findings have broader relevance for consumer preference research, as the labor market anxieties associated with AI diffusion may spill over into consumer evaluations of AI products, particularly in sectors characterized by high levels of human labor visibility. Consumers who are concerned about the employment impacts of AI may express preferences that align with these concerns—for

example, preferring AI products that position themselves as augmenting rather than replacing human workers.

The role of enabling infrastructures in shaping consumer preferences for AI products deserves specific attention. Pham et al. (2020) demonstrate that the deployment of multi-access edge computing in 5G networks creates the technical conditions for real-time, context-sensitive AI applications that can respond to consumer needs with minimal latency. The availability of this infrastructure is uneven, however, with significant implications for the geographic and socioeconomic distribution of AI product access. Consumers in under-served areas may face systematic barriers to AI product adoption that are infrastructure-related rather than preference-related—a distinction with important implications for both market strategy and policy intervention.

4. Research Gaps and Future Directions

Despite the breadth and depth of the literature reviewed, significant gaps remain in the scholarly understanding of consumer preferences for AI-based products and services. Three such gaps are identified here as priorities for future research.

The first and most significant gap concerns longitudinal dynamics of consumer preference formation. The overwhelming majority of the studies examined in this review employ cross-sectional designs that capture consumer preferences at a single point in time, providing limited insight into how preferences evolve as consumers accumulate experience with AI products. This is a particularly consequential limitation given the evidence presented by Gerlich (2025) regarding the cognitive offloading effects of sustained AI use: if prolonged exposure to AI tools systematically reshapes the cognitive faculties that consumers employ in preference evaluation, then cross-sectional preference data may offer a systematically misleading picture of the stable preference landscape. Longitudinal research designs—including experience sampling methods, panel surveys, and quasi-experimental studies tracking preference shifts over product adoption lifecycles—are urgently needed to address this gap. The practical implications are substantial: product developers and marketers require dynamic models of preference evolution to anticipate how consumer engagement with AI products will shift over time.

The second gap concerns cross-cultural and socioeconomic variation in AI consumer preferences. The literature reviewed is predominantly rooted in Western, technologically advanced consumer contexts, with limited attention to how preferences for AI products are shaped by cultural values, economic conditions, and institutional environments in non-Western or developing-country settings. Given the global reach of AI diffusion and the significant variation in regulatory frameworks, digital infrastructure, and cultural attitudes toward technology and trust, this geographic and cultural narrowness represents a significant limitation of the current evidence base. Garibay et al. (2023) acknowledge the importance of cross-cultural dimensions in their articulation of human-centered AI grand challenges, and Pham et al. (2020) highlight the uneven geographic distribution of 5G infrastructure as a structural constraint on AI access—but neither source provides systematic comparative evidence. Future research should employ comparative and cross-cultural designs that illuminate how preference structures vary across diverse consumer populations, with particular attention to the mediating role of institutional trust and regulatory context.

The third gap concerns the relationship between AI regulation and consumer preference dynamics. As regulatory frameworks for AI are developed and implemented across

major jurisdictions—including the European Union's AI Act and emerging national AI strategies—the question of how regulatory interventions shape consumer trust, product design incentives, and ultimately preference formation becomes increasingly important. Díaz-Rodríguez et al. (2023) provide a detailed normative analysis of AI regulation, but empirical evidence on the consumer behavioral consequences of regulatory change remains scarce. Future research should examine how specific regulatory provisions—such as mandatory transparency disclosures, algorithmic impact assessments, and liability frameworks—affect consumer confidence and adoption behavior across product categories and market segments. The practical implications for both regulators and product developers are considerable: regulation that is well-calibrated to consumer preference dynamics can accelerate beneficial AI adoption, while poorly designed regulation may generate compliance costs without corresponding consumer welfare gains.

5. Conclusion

This literature review has synthesized the available evidence on consumer preferences for AI-based products and services, drawing on sixteen peer-reviewed sources spanning multiple disciplines, including information management, marketing, healthcare, human-computer interaction, and sustainability. The review has identified three principal thematic clusters—trust and transparency as antecedents of AI acceptance, human-centered design and personalization as shapers of consumer experience, and sector-specific contextual determinants of AI adoption—and has situated these themes within a coherent theoretical framework grounded in technology acceptance theory, human-centered AI principles, and cognitive behavioral perspectives.

The principal findings of this review indicate that consumer preferences for AI products are shaped by a complex interplay of cognitive, affective, ethical, and contextual factors that extend well beyond the functional performance of the products in question. Trust, transparency, and ethical alignment emerge as foundational prerequisites for consumer acceptance across virtually all domains examined, while human-centered design principles—particularly those emphasizing augmentation over replacement and personalization over standardization—are strongly associated with positive consumer experience. The sector-specific evidence further reveals that contextual variables, including institutional environments, regulatory frameworks, infrastructure availability, and sustainability considerations, play critical moderating roles in shaping how these general preference determinants manifest in practice.

Several limitations of this review warrant acknowledgment. The review is limited to sixteen sources, which, while carefully selected, do not constitute an exhaustive mapping of the literature. The predominance of Western and technologically advanced consumer contexts in the underlying literature constrains the generalizability of the findings. Additionally, the review is primarily thematic rather than quantitative in its synthesis, precluding meta-analytic estimation of effect sizes. Despite these limitations, this review makes a meaningful contribution to the understanding of AI consumer preference dynamics and provides a systematic foundation for future empirical investigation.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Araújo, S. O., Peres, R. S., Barata, J., Lidon, F. C., & Ramalho, J. C. (2021). Characterising the Agriculture 4.0 landscape—Emerging trends, challenges and opportunities. Agronomy, 11(4), 667. https://doi.org/10.3390/agronomy11040667

Budhwar, P., Chowdhury, S., Wood, G., Aguinis, H., & Bamber, G. J. (2023). Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT. Human Resource Management Journal, 33(3), 606–659. https://doi.org/10.1111/1748-8583.12524

Díaz-Rodríguez, N., Del Ser, J., Coeckelbergh, M., López de Prado, M., Herrera-Viedma, E., & Herrera, F. (2023). Connecting the dots in trustworthy artificial intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. Information Fusion, 99, 101896. https://doi.org/10.1016/j.inffus.2023.101896

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., & Coombs, C. (2019). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Dwivedi, Y. K., Hughes, L., Wang, Y., Alalwan, A. A., & Joo, S. (2022). Metaverse marketing: How the metaverse will shape the future of consumer research and practice. Psychology and Marketing, 40(9), 1–27. https://doi.org/10.1002/mar.21767

Garibay, Ö. Ö., Winslow, B., Andolina, S., Antona, M., & Bodenschatz, A. (2023). Six human-centered artificial intelligence grand challenges. International Journal of Human-Computer Interaction, 39(3), 391–437. https://doi.org/10.1080/10447318.2022.2153320

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. Societies, 15(1), 6. https://doi.org/10.3390/soc15010006

Hannigan, T. R., Haans, R. F. J., Vakili, K., Tchalian, H., Glaser, V. L., Wang, M. S., Kaplan, S., & Jennings, P. D. (2019). Topic modeling in management research: Rendering new theory from textual data. Academy of Management Annals, 13(2), 586–632. https://doi.org/10.5465/annals.2017.0099

Martínez-Peláez, R., Ochoa-Brust, A., Rivera Manrique, S. I., & Félix, V. G. (2023). Role of digital transformation for achieving sustainability: Mediated role of stakeholders, key capabilities, and technology. Sustainability, 15(14), 11221. https://doi.org/10.3390/su151411221

Minerva, R., Lee, G. M., & Crespi, N. (2020). Digital twin in the IoT context: A survey on technical features, scenarios, and architectural models. Proceedings of the IEEE, 108(10), 1785–1824. https://doi.org/10.1109/jproc.2020.2998530

Monrat, A. A., Schelén, O., & Andersson, K. (2019). A survey of blockchain from the perspectives of applications, challenges, and opportunities. IEEE Access, 7, 117134–117151. https://doi.org/10.1109/access.2019.2936094

Pham, Q.-V., Fang, F., Ha, V. N., Piran, M. J., & Le, M. (2020). A survey of multi-access edge computing in 5G and beyond: Fundamentals, technology integration, and state-of-the-art. IEEE Access, 8, 116974–117017. https://doi.org/10.1109/access.2020.3001277

Singh, J., Flaherty, K., Sohi, R. S., Deeter-Schmelz, D., Habel, J., Le Meunier-FitzHugh, K., Malshe, A., Mullins, R., & Onyemah, V. (2019). Sales profession and professionals in the age of

digitization and artificial intelligence technologies: Concepts, priorities, and questions. Journal of Personal Selling and Sales Management, 39(1), 2–22. https://doi.org/10.1080/08853134.2018.1557525

Stoumpos, A. I., Kitsios, F., & Talias, M. A. (2023). Digital transformation in healthcare: Technology acceptance and its applications. International Journal of Environmental Research and Public Health, 20(4), 3407. https://doi.org/10.3390/ijerph20043407

Weidinger, L., Uesato, J., Rauh, M., Griffin, C., Huang, P.-S., Mellor, J., Glaese, A., Cheng, M., Balle, B., Kasirzadeh, A., Biles, C., Brown, S., Kenton, Z., Hawkins, W., Stepleton, T., Birhane, A., Hendricks, L. A., Rimell, L., Isaac, W., … Gabriel, I. (2022). Taxonomy of risks posed by language models. In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (pp. 214–229). ACM. https://doi.org/10.1145/3531146.3533088

Geraci, G., García-Rodríguez, A., Azari, M. M., Lozano, A., & Filippini, F. (2022). What will the future of UAV cellular communications be? A flight from 5G to 6G. IEEE Communications Surveys & Tutorials, 24(3), 1304–1335. https://doi.org/10.1109/comst.2022.3171135

Sentiment Analysis Methods in Business Research: A Systematic Literature Review

Banda Federico[1] and Stovíček Filip[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: federico.banda@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: filip.stovicek@euba.sk

Abstract

Sentiment analysis has emerged as a foundational methodology in contemporary business research, enabling organizations to extract actionable intelligence from large volumes of unstructured textual data. As digital communication channels proliferate and the volume of user-generated content continues to grow exponentially, the ability to systematically analyze expressed opinions, attitudes, and emotions has become increasingly valuable across diverse business domains. This review examines the evolution and current state of sentiment analysis methods as applied in business contexts, synthesizing findings from twenty scholarly sources spanning natural language processing, machine learning, marketing analytics, and computational social science. The review identifies three principal thematic clusters: methodological foundations and the transition from lexicon-based to deep learning approaches; domain-specific applications in financial markets, consumer behavior, and organizational research; and emerging challenges related to large language models, ethical considerations, and cross-lingual deployment. The findings reveal that while transformer-based architectures have substantially advanced analytical accuracy, critical gaps persist regarding interpretability, domain adaptation, and ethical deployment of sentiment models in high-stakes business environments. This review concludes by proposing targeted future research directions that address both technical limitations and the broader societal implications of automated sentiment analysis in business decision-making.

Key Words: sentiment analysis, natural language processing, text mining, opinion mining, business intelligence

JEL Classification: C55, M31, D83

1. Introduction

The capacity to systematically measure, classify, and interpret human sentiment from textual data has transformed how businesses understand their customers, competitors, and markets. Sentiment analysis — broadly understood as the computational identification of opinions, attitudes, and emotional states expressed in natural language — represents one of the most commercially significant applications of natural language processing (NLP) to emerge in the past two decades. As organizations increasingly rely on digital platforms, social media, earnings calls, customer reviews, and news feeds as primary sources of market intelligence, the methodological sophistication of sentiment analysis has become a critical determinant of competitive advantage.

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

The business relevance of sentiment analysis is underscored by its application across a wide range of domains. In financial markets, analysts deploy sentiment models to predict stock price movements and assess firm-level risk exposures. In marketing, sentiment analysis supports consumer segmentation, brand monitoring, and the personalization of customer experiences. In human resources and organizational management, text-based sentiment tools are applied to employee feedback, corporate communications, and stakeholder engagement data. The breadth of these applications reflects a broader trend toward data-driven decision-making enabled by artificial intelligence (AI), a trend that has prompted scholars from disciplines as diverse as information systems, finance, and marketing to engage seriously with the methodological foundations and limitations of sentiment analysis.

Despite this growing scholarly and practitioner interest, the literature on sentiment analysis in business research remains fragmented. Studies have tended to develop within disciplinary silos, with computer scientists advancing algorithmic performance on benchmark datasets while business researchers adapt existing tools to domain-specific problems without always critically evaluating their methodological assumptions. The result is a literature in which technical innovation often outpaces rigorous empirical validation, and in which important questions about model interpretability, ethical deployment, and cross-domain generalizability remain insufficiently addressed.

This literature review aims to address this gap by synthesizing current knowledge on sentiment analysis methods as applied in business research. Drawing on twenty scholarly publications encompassing technical, applied, and critical perspectives, the review seeks to provide a coherent account of the state of the field, identify recurring themes and unresolved debates, and articulate a research agenda for future work. The review proceeds as follows. Section 2 establishes the theoretical background, defining key concepts and situating sentiment analysis within broader frameworks of AI and NLP. Section 3 presents a thematically organized synthesis of the literature, organized around three principal clusters. Section 4 identifies research gaps and proposes future directions. Section 5 offers conclusions and notes the limitations of the review.

2. Theoretical Background

This chapter establishes the foundational definitions of sentiment analysis—also known as opinion mining—and situates these techniques within the broader frameworks of Artificial Intelligence (AI) and Natural Language Processing (NLP). It explores the methodological progression from early lexicon-based dictionaries to advanced deep learning architectures, such as Recurrent Neural Networks (RNNs) and transformer-based models. Furthermore, this section examines the critical roles of transfer learning and topic modeling as complementary frameworks that enhance the ability of organizations to derive actionable intelligence from unstructured textual data.

2.1 Defining Sentiment Analysis and Its Business Relevance

Sentiment analysis, also referred to as opinion mining, encompasses a family of computational techniques designed to identify and quantify subjective information from text. At its core, the task involves assigning a polarity — positive, negative, or neutral — to a unit of text, which may range from a single word or phrase to a full document. More advanced formulations extend this basic framework to include aspect-level sentiment analysis, emotion classification, and stance detection, each of which offers richer and more granular insights into the opinions expressed in a corpus.

The business case for sentiment analysis is grounded in the recognition that unstructured textual data constitutes a vast and largely underutilized source of market intelligence. Dwivedi et al. (2019) note that AI technologies, including those underlying sentiment analysis, offer transformative potential for the augmentation of human analytical tasks across a wide range of industrial, intellectual, and social applications. This transformative potential is particularly evident in marketing, where the ability to process large-scale consumer feedback in real time enables a shift from periodic survey-based research to continuous, adaptive market intelligence. Huang and Rust (2020) formalize this insight within a strategic framework for AI in marketing, distinguishing between mechanical AI for automating repetitive tasks, thinking AI for data-driven decision-making, and feeling AI for analyzing human emotions and interactions — a typology that maps directly onto the core functions of sentiment analysis systems.

2.2 Foundational Approaches: From Lexicons to Deep Learning

Early approaches to sentiment analysis relied heavily on lexicon-based methods, in which the sentiment of a text was inferred by aggregating the polarity scores assigned to individual words in a predefined sentiment dictionary. These approaches offered the advantages of transparency and domain independence but were limited by their inability to capture contextual meaning, sarcasm, negation, and domain-specific language. The development of knowledge-based resources such as SenticNet, which Cambria et al. (2018) describe as a system for discovering conceptual primitives for sentiment analysis through context embeddings, represented a significant advance in this tradition by coupling sub-symbolic machine learning with symbolic commonsense reasoning.

The subsequent rise of deep learning substantially changed the landscape of sentiment analysis. Architectures based on recurrent neural networks (RNNs) and convolutional neural networks (CNNs) demonstrated superior ability to capture sequential and local dependencies in text. Shen et al. (2018) proposed the Directional Self-Attention Network (DiSAN), a lightweight architecture that replaced traditional RNN and CNN components with a directional, multi-dimensional attention mechanism, achieving state-of-the-art performance on multiple NLP benchmarks while enabling significantly faster training. This work exemplified a broader shift in the field toward attention-based architectures that culminated in the development of transformer models.

2.3 Transfer Learning and Sentence Embeddings

A pivotal development in recent NLP research has been the widespread adoption of transfer learning, whereby models pre-trained on large general-purpose corpora are fine-tuned on domain-specific tasks. Hosna et al. (2022) provide a comprehensive introduction to transfer learning, highlighting its ability to bridge the gap between data-scarce domain-specific applications and the large data requirements of deep learning models. In the context of sentiment analysis, transfer learning has enabled practitioners to achieve high performance on business-specific tasks — such as financial sentiment classification or product review analysis — without requiring extensive labeled training data.

Gao et al. (2021) advanced the quality of sentence-level representations through SimCSE, a contrastive learning framework that uses standard dropout as minimal data augmentation to produce more uniform and expressive sentence embeddings. The quality of sentence embeddings is directly consequential for sentiment analysis tasks, as the capacity of a model to encode fine-grained semantic distinctions determines its ability to differentiate between subtly different sentiments. The SimCSE approach demonstrated that substantial

improvements in embedding quality could be achieved through simple and computationally efficient modifications to training objectives, making it directly relevant to applied business NLP systems operating under resource constraints.

2.4 Topic Modeling as a Complementary Framework

Sentiment analysis rarely operates in isolation within business research. It is frequently combined with topic modeling — a class of unsupervised methods that identify latent thematic structures in large text corpora — to provide a more complete picture of what is being said and how it is being evaluated. Hannigan et al. (2019) describe topic modeling as a method for rendering new theory from textual data in management research, emphasizing its capacity to reveal phenomenon-based constructs that are difficult to identify through deductive coding alone. Egger and Yu (2022) extend this discussion by comparing the performance of four prominent topic modeling approaches — LDA, NMF, Top2Vec, and BERTopic — on social media data, finding that transformer-based approaches such as BERTopic offer superior coherence and interpretability for the short, noisy texts characteristic of online business discourse. The integration of topic modeling with sentiment analysis — often operationalized as aspect-based sentiment analysis — represents one of the most methodologically sophisticated approaches available to business researchers.

3. Literature Review

This chapter presents a synthesized review of scholarly research organized into three distinct thematic clusters. The first cluster details the methodological evolution from rule-based and classical machine learning to modern neural approaches, highlighting the impact of attention mechanisms and sentence embeddings. The second cluster focuses on domain-specific applications, demonstrating how sentiment analysis is utilized in financial markets, marketing analytics, and organizational management. Finally, the third cluster addresses emerging challenges, specifically the integration of Large Language Models (LLMs) and the associated ethical and practical concerns regarding bias, transparency, and data privacy.

3.1 Thematic Cluster One: Methodological Evolution from Rule-Based to Neural Approaches

The history of sentiment analysis in business research is one of progressive methodological sophistication, driven by the limitations of earlier approaches and enabled by advances in computational resources and algorithmic innovation. The earliest systematic applications of sentiment analysis in business relied on dictionary-based methods, which assigned sentiment scores to texts by matching words against hand-crafted lexicons. While such methods remain in use due to their interpretability and low computational cost, their inability to capture context, domain-specific language, and complex linguistic phenomena such as irony and negation significantly limited their utility in real-world business applications.

The introduction of machine learning methods marked the first major transition in the field. Supervised classifiers trained on labeled datasets — including Naive Bayes, Support Vector Machines, and logistic regression models — demonstrated substantially improved performance over pure lexicon-based approaches by learning statistical associations between textual features and sentiment labels. Sohangir et al. (2018) investigated the application of deep learning models for financial sentiment analysis, comparing several architectures including deep belief networks and long short-term memory (LSTM) networks on stock market-related text data. Their findings demonstrated that deep learning models consistently

outperformed traditional machine learning baselines on financial sentiment classification tasks, particularly when applied to large datasets where the representational power of neural architectures could be fully exploited.

The second major transition came with the widespread adoption of attention mechanisms and transformer-based pre-trained language models. The self-attention mechanism, explored in the context of NLP by Shen et al. (2018), allowed models to weigh the relevance of different parts of an input sequence when computing representations, enabling them to capture long-range dependencies that had previously been challenging for RNN-based models. This architectural innovation laid the groundwork for the transformer models — most notably BERT and its successors — that now dominate applied NLP research. Gao et al. (2021) built on these foundations with SimCSE, demonstrating that contrastive learning objectives could substantially improve the semantic quality of sentence embeddings without increasing model complexity, a finding with direct implications for downstream sentiment classification tasks in business applications.

Transfer learning has emerged as a particularly critical enabler for business applications of sentiment analysis, where labeled domain-specific training data is often scarce. Hosna et al. (2022) document the theoretical basis for transfer learning and its practical advantages in real-world applications, emphasizing that pre-trained representations encode rich linguistic knowledge that can be rapidly adapted to specialized domains. In financial text analysis, for example, models pre-trained on general corpora can be fine-tuned on relatively small collections of labeled earnings call transcripts or analyst reports to achieve performance levels that would be unattainable through training from scratch. This capability has substantially lowered the barrier to entry for businesses seeking to deploy sentiment analysis at scale.

3.2 Thematic Cluster Two: Domain-Specific Applications in Business Contexts

A substantial body of literature has moved beyond methodological development to examine the deployment of sentiment analysis in specific business domains. The most extensively studied application areas include financial markets, consumer behavior and marketing analytics, and public health and crisis communication — the last of which has direct organizational implications for business continuity and reputation management.

In financial research, sentiment analysis has been applied to a diverse range of textual sources, including earnings call transcripts, financial news, social media posts, and analyst reports. Sautner et al. (2023) provide a compelling demonstration of the potential of machine learning-augmented text analysis in a financial context, developing a methodology to identify and quantify firm-level climate change exposures from earnings call transcripts. Their approach adapted a keyword discovery algorithm grounded in machine learning principles to construct exposure measures for more than 10,000 firms across 34 countries, demonstrating the scalability and empirical validity of text-based analytics in financial research. Crucially, their measures predicted real-world firm-level outcomes including green patenting activity and job creation in renewable energy sectors, confirming that carefully constructed sentiment and exposure indicators derived from textual data can serve as meaningful inputs to financial and strategic analysis.

The application of sentiment analysis to marketing and consumer behavior research has been shaped significantly by the strategic AI framework proposed by Huang and Rust (2020). Their conceptualization of "feeling AI" — which encompasses the analysis of human emotions and interactions — provides a theoretical rationale for deploying advanced sentiment models in marketing contexts, from brand sentiment monitoring and customer experience

management to advertising effectiveness evaluation. This framework positions sentiment analysis not merely as a data processing tool but as a component of strategic marketing intelligence that enables firms to achieve more precise segmentation, targeting, and positioning.

The COVID-19 pandemic provided an unusually rich natural experiment for the application of sentiment analysis to real-time public and organizational discourse. Xue et al. (2020) demonstrated the utility of machine learning-based sentiment analysis applied to Twitter data during the pandemic, identifying dominant topics and emotional trajectories across millions of posts. Their findings revealed systematic patterns in public fear, information-seeking behavior, and reactions to government policy — patterns with direct relevance to organizational risk communication and crisis management. While their study focused primarily on public health outcomes, the methodology is directly applicable to business contexts such as reputational risk monitoring, supply chain disruption communication, and employee sentiment surveillance during organizational crises.

In the domain of organizational and management research, topic modeling — closely related to and often combined with sentiment analysis — has gained increasing traction as a method for inductively analyzing large corpora of organizational documents, interview transcripts, and online discourse. Hannigan et al. (2019) provided a foundational methodological account of topic modeling in management research, arguing that the technique enables researchers to render new theoretical constructs from textual data in a manner that complements deductive hypothesis-testing approaches. Their work situates text-based analytical methods, including sentiment analysis, within the broader methodological repertoire of management scholarship, providing theoretical legitimacy for their use in organizational studies. Egger and Yu (2022) extended this literature through comparative evaluation of topic modeling algorithms, finding that modern transformer-based methods such as BERTopic outperform classical approaches when applied to the short, heterogeneous texts typical of social media business discourse.

3.3 Thematic Cluster Three: Large Language Models, Ethics, and Emerging Challenges

The most recent and rapidly evolving cluster in the sentiment analysis literature concerns the rise of large language models (LLMs) and the ethical, practical, and epistemological challenges their widespread adoption poses for business research. LLMs — generative AI systems trained on vast corpora using self-supervised objectives — have demonstrated remarkable zero-shot and few-shot capabilities on a wide range of NLP tasks, including sentiment classification, opinion summarization, and emotional tone analysis.

Ziems et al. (2023) conducted a comprehensive evaluation of thirteen LLMs on twenty-five computational social science tasks, including several that overlap directly with business-relevant sentiment analysis applications such as persuasion detection and political ideology classification. Their findings demonstrated that while LLMs achieved competitive zero-shot performance on many tasks, their reliability varied substantially across tasks and domains, and they were prone to systematic biases that could distort downstream analytical conclusions. This finding has important implications for business researchers who may be tempted to deploy LLMs as off-the-shelf sentiment analysis tools without critical evaluation of their outputs. Hassani and Silva (2023) adopted a more optimistic perspective on the transformative potential of conversational AI models such as ChatGPT in data science, highlighting their potential to democratize access to advanced text analytics. However, they also acknowledged significant

limitations related to hallucination, inconsistency, and data privacy that must be addressed before LLMs can be responsibly integrated into business intelligence pipelines.

Ethical dimensions of AI deployment in sentiment analysis contexts have received growing scholarly attention. Weidinger et al. (2022) developed a comprehensive taxonomy of the ethical and social risks posed by language models, identifying twenty-one distinct risks across six categories including discrimination and hate speech, misinformation, malicious use, and environmental harms. Several of these risks are directly relevant to business applications of sentiment analysis: the potential for models to produce systematically biased sentiment assessments that disadvantage certain demographic groups, the risk of deliberate manipulation of sentiment analysis systems through adversarial inputs, and the privacy implications of processing large-scale personal communications data. Huang et al. (2022) similarly surveyed AI ethics frameworks and identified privacy leakage, algorithmic discrimination, and accountability deficits as central concerns for AI systems deployed in commercial contexts, calling for the development of governance frameworks that ensure responsible innovation.

The intersection of LLMs and argument mining represents another emerging frontier with significant business relevance. Lawrence and Reed (2019) surveyed argument mining — the automatic identification and extraction of argumentative structure from natural language — and identified its potential applications in financial market prediction and public relations, noting that understanding not merely the sentiment but the reasoning behind expressed positions provides substantially richer analytical value. Cambria et al. (2018) made a complementary argument from the perspective of commonsense AI, contending that progress in natural language understanding requires coupling statistical pattern recognition with structured commonsense knowledge, a perspective that anticipates the hybrid neuro-symbolic architectures now attracting research attention. Together, these perspectives suggest that the future of business-oriented sentiment analysis lies not in maximally large generative models alone, but in architectures that combine statistical power with interpretable, structured representations of meaning. Zhang and Chen (2020) reinforced this point in the context of explainable recommendation systems, arguing that the ability to provide transparent explanations for model outputs is essential for user trust and system adoption — a principle that applies with equal force to sentiment-informed business decision systems.

The relationship between anomaly detection and sentiment analysis also merits attention, particularly in financial applications where identifying unusual patterns in expressed sentiment may signal emerging market risks or opportunities. Ruff et al. (2021) provided a systematic review of deep and shallow anomaly detection methods, noting the growing importance of deep learning-based approaches for detecting unusual patterns in complex, high-dimensional data including text. The application of anomaly detection principles to sentiment time series — identifying unusual shifts in expressed sentiment around corporate events, regulatory announcements, or geopolitical shocks — represents a promising methodological direction that bridges the technical literature on anomaly detection with the applied literature on financial text analysis.

The role of sentiment analysis within the broader AI ecosystem in business has also been addressed from a strategic and macro-level perspective. Dwivedi et al. (2019) discussed the multidisciplinary challenges and opportunities associated with AI adoption across industries, emphasizing that technical capability must be matched with organizational readiness, regulatory compliance, and ethical governance to realize sustainable value. Kamalov et al. (2023) extended this analysis to educational and institutional contexts, noting that the rapid

diffusion of AI-driven language technologies is reshaping professional practices across sectors in ways that demand critical engagement from researchers and practitioners alike. These broader perspectives remind researchers that sentiment analysis methods do not exist in isolation but are embedded in complex sociotechnical systems whose design and governance shape their ultimate impact.

4. Research Gaps and Future Directions

Despite the significant advances documented in this review, several important gaps in the existing literature on sentiment analysis methods in business research warrant systematic attention in future work.

The first and arguably most pressing gap concerns the limited interpretability and explainability of state-of-the-art sentiment analysis models in business-critical applications. As transformer-based architectures have displaced simpler rule-based and classical machine learning approaches, the trade-off between predictive accuracy and model transparency has become increasingly consequential. Zhang and Chen (2020) argue persuasively that explainability is essential for user trust and responsible system design in AI-powered recommendation contexts, and the same argument applies directly to sentiment analysis systems informing investment decisions, strategic communications, or human resources practices. Future research should develop and evaluate explanation frameworks specifically tailored to business-oriented sentiment analysis, moving beyond generic feature attribution methods to provide decision-relevant justifications that can be meaningfully interpreted by domain experts without deep NLP expertise.

The second gap concerns domain adaptation and the persistent challenge of deploying sentiment models across heterogeneous business contexts. The current literature reveals a tendency toward developing and evaluating models within narrow domain boundaries — financial text, product reviews, social media posts — without systematic investigation of how performance degrades when models are applied to adjacent domains or novel text types. Hosna et al. (2022) establish the theoretical foundations of transfer learning as a mitigation strategy, but empirical evidence on the boundary conditions of effective transfer in business sentiment analysis remains sparse. Future research should systematically map the domain-specificity of leading sentiment models, develop benchmarks that capture cross-domain variability in business language, and propose training strategies — including domain-adaptive pre-training and few-shot fine-tuning — that improve model robustness across diverse organizational contexts.

The third gap involves the ethical governance of sentiment analysis in business research and practice. Weidinger et al. (2022) and Huang et al. (2022) provide foundational taxonomies of AI risk and ethics, but their frameworks have not been comprehensively operationalized in the specific context of business sentiment analysis. There is a notable absence of empirically grounded research examining how sentiment analysis systems may amplify existing social biases — for example, in employee performance evaluations or customer satisfaction assessments — and what governance mechanisms can mitigate these effects. Future research should develop domain-specific ethical audit frameworks for business sentiment analysis systems, drawing on both technical bias detection methods and organizational ethics scholarship.

A fourth, less frequently acknowledged gap concerns the integration of multilingual and cross-cultural sentiment analysis in globally operating businesses. The majority of

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

methodological advances documented in this review are grounded in English-language data, yet multinational corporations routinely encounter sentiment-bearing text in dozens of languages and across widely varying cultural contexts. The implications of this linguistic bias for the validity of global sentiment analytics represent an important practical and methodological challenge that has received insufficient attention in the business research literature.

5. Conclusion

This review has synthesized twenty scholarly publications to provide a systematic account of sentiment analysis methods in business research, organized around three thematic clusters: methodological evolution from rule-based to neural approaches, domain-specific applications in finance, marketing, and organizational management, and emerging challenges associated with large language models and ethical deployment. Several overarching conclusions emerge from this synthesis.

First, the field has undergone a profound methodological transformation over the past decade, with deep learning and transformer-based architectures displacing classical approaches to become the dominant paradigm. Second, domain-specific applications have demonstrated the substantial practical value of sentiment analysis in financial prediction, consumer intelligence, and crisis communication, though the quality and comparability of methods varies considerably across studies. Third, the rise of LLMs introduces both new capabilities — particularly zero-shot and few-shot analytical flexibility — and new risks, including systematic bias, opacity, and ethical exposure that require careful governance.

This review makes several contributions to the literature. It offers one of the first integrative syntheses specifically focused on sentiment analysis as a business research method, connecting technical and applied literatures that have often developed in parallel. It identifies concrete research gaps and proposes a structured agenda for future work. It situates sentiment analysis within the broader context of AI adoption in business, connecting methodological choices to strategic, organizational, and ethical considerations.

The review is subject to several limitations. The source base, while deliberately selected to span methodological and applied perspectives, is not exhaustive. The emphasis on English-language publications may underrepresent important contributions from non-Anglophone research traditions. Future reviews should incorporate a more systematic search protocol and extend coverage to practitioner-oriented and non-English-language publications to provide a more complete picture of the global state of the field.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Akundi, A., Euresti, D., Luna, S., Ankobiah, W., & Lopes, A. (2022). State of Industry 5.0—Analysis and identification of current research trends. Applied System Innovation, 5(1), 27. https://doi.org/10.3390/asi5010027

Cambria, E., Poria, S., Hazarika, D., & Kwok, K. (2018). SenticNet 5: Discovering conceptual primitives for sentiment analysis by means of context embeddings. Proceedings of the AAAI Conference on Artificial Intelligence, 32(1). https://doi.org/10.1609/aaai.v32i1.11559

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., & Coombs, C. (2019). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Egger, R., & Yu, J. (2022). A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify Twitter posts. Frontiers in Sociology, 7, 886498. https://doi.org/10.3389/fsoc.2022.886498

Gao, T., Yao, X., & Chen, D. (2021). SimCSE: Simple contrastive learning of sentence embeddings. Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, 6894–6910. https://doi.org/10.18653/v1/2021.emnlp-main.552

Hannigan, T. R., Haans, R. F. J., Vakili, K., Tchalian, H., & Glaser, V. (2019). Topic modeling in management research: Rendering new theory from textual data. Academy of Management Annals, 13(2), 586–632. https://doi.org/10.5465/annals.2017.0099

Hassani, H., & Silva, E. S. (2023). The role of ChatGPT in data science: How AI-assisted conversational interfaces are revolutionizing the field. Big Data and Cognitive Computing, 7(2), 62. https://doi.org/10.3390/bdcc7020062

Hosna, A., Merry, E., Gyalmo, J., Alom, Z., & Aung, Z. (2022). Transfer learning: A friendly introduction. Journal of Big Data, 9(1), 102. https://doi.org/10.1186/s40537-022-00652-w

Huang, C., Zhang, Z., Mao, B., & Yao, X. (2022). An overview of artificial intelligence ethics. IEEE Transactions on Artificial Intelligence, 4(4), 799–819. https://doi.org/10.1109/tai.2022.3194503

Huang, M.-H., & Rust, R. T. (2020). A strategic framework for artificial intelligence in marketing. Journal of the Academy of Marketing Science, 49(1), 30–50. https://doi.org/10.1007/s11747-020-00749-9

Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. Sustainability, 15(16), 12451. https://doi.org/10.3390/su151612451

Lawrence, J., & Reed, C. (2019). Argument mining: A survey. Computational Linguistics, 45(4), 765–818. https://doi.org/10.1162/coli_a_00364

Ruff, L., Kauffmann, J. R., Vandermeulen, R. A., Montavon, G., & Samek, W. (2021). A unifying review of deep and shallow anomaly detection. Proceedings of the IEEE, 109(5), 756–795. https://doi.org/10.1109/jproc.2021.3052449

Sautner, Z., van Lent, L., Vilkov, G., & Zhang, R. (2023). Firm-level climate change exposure. The Journal of Finance, 78(3), 1449–1498. https://doi.org/10.1111/jofi.13219

Shen, T., Zhou, T., Long, G., Jiang, J., & Pan, S. (2018). DiSAN: Directional self-attention network for RNN/CNN-free language understanding. Proceedings of the AAAI Conference on Artificial Intelligence, 32(1). https://doi.org/10.1609/aaai.v32i1.11941

Sohangir, S., Wang, D., Pomeranets, A., & Khoshgoftaar, T. M. (2018). Big data: Deep learning for financial sentiment analysis. Journal of Big Data, 5(1), 3. https://doi.org/10.1186/s40537-017-0111-6

Weidinger, L., Uesato, J., Rauh, M., Griffin, C., & Huang, P.-S. (2022). Taxonomy of risks posed by language models. Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, 214–229. https://doi.org/10.1145/3531146.3533088

Xue, J., Chen, J., Hu, R., Chen, C., & Zheng, C. (2020). Twitter discussions and emotions about the COVID-19 pandemic: Machine learning approach. Journal of Medical Internet Research, 22(11), e20550. https://doi.org/10.2196/20550

Zhang, Y., & Chen, X. (2020). Explainable recommendation: A survey and new perspectives. Foundations and Trends® in Information Retrieval, 14(1), 1–101. https://doi.org/10.1561/1500000066

Ziems, C., Held, W. A., Shaikh, O. A., Chen, J., & Zhang, Z. (2023). Can large language models transform computational social science? Computational Linguistics, 50(1), 237–291. https://doi.org/10.1162/coli_a_00502

Managerial Attitudes Toward Artificial Intelligence: Literature Review of Emerging Challenges, Opportunities, and Organizational Implications

Pallérová Diana[1] and Romanová Anita[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: diana.pallerova@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: anita.romanova@euba.sk

Abstract

The rapid proliferation of artificial intelligence across industries has placed managerial attitudes at the center of debates concerning technology adoption, workforce transformation, and organizational change. Despite growing scholarly interest in AI's organizational implications, the literature remains fragmented, with limited synthesis of how managers perceive, evaluate, and respond to AI-driven change. This literature review examines 14 peer-reviewed studies to consolidate current understanding of managerial attitudes toward AI across multiple organizational contexts. Drawing on frameworks from technology adoption, human resource management, and digital transformation research, the review identifies three thematic clusters: organizational readiness and strategic orientation toward AI adoption, managerial perceptions of AI's impact on workforce skills and job redesign, and ethical and governance considerations shaping managerial decision-making. Key findings suggest that managerial attitudes are shaped by a complex interplay of technological capability, organizational culture, ethical awareness, and competitive pressure. Significant research gaps persist regarding cross-cultural variations in managerial attitudes and longitudinal dynamics of AI-related decision-making. The review contributes a structured synthesis of the field and proposes a future research agenda with direct implications for practitioners and policymakers.

Key Words: managerial attitudes, artificial intelligence, organisational behaviour, technology leadership, AI acceptance

JEL Classification: M10, O33, M54

1. Introduction

The advent of artificial intelligence (AI) as a transformative organizational force has fundamentally altered the landscape of managerial decision-making, strategic planning, and human resource governance. From automating routine cognitive tasks to enabling sophisticated data-driven insights, AI technologies are reshaping the very fabric of organizational life at an unprecedented pace (Dwivedi et al., 2019). Yet while the technical capabilities of AI systems continue to advance, the human dimension of AI adoption—particularly the attitudes, perceptions, and strategic orientations of managers—remains insufficiently understood and theorized.

Managerial attitudes occupy a pivotal role in the technology adoption process. Managers function as gatekeepers who evaluate emerging technologies, allocate organizational

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

resources, and champion or resist change initiatives. Their attitudes toward AI are not formed in a vacuum; they are shaped by perceived utility, fear of disruption, ethical concerns, competitive pressures, and organizational culture. As Jöhnk et al. (2020) demonstrate, organizational readiness for AI is fundamentally a function of leadership commitment and managerial disposition, underscoring the centrality of attitudes in determining whether AI adoption succeeds or stalls. Similarly, Neumann et al. (2022) find that even in public sector organizations with strong institutional mandates, the individual orientations of organizational leaders substantially mediate AI adoption trajectories.

Despite this recognized importance, the scholarly literature on managerial attitudes toward AI remains dispersed across disciplines, including information systems, management science, organizational behavior, and public administration. Systematic efforts to consolidate and synthesize this evidence base are scarce. Existing reviews tend to focus on technical dimensions of AI adoption, sector-specific applications, or broad organizational outcomes, often neglecting the managerial agency and attitudinal dynamics that drive implementation. This constitutes a meaningful research gap, both for theory development and for practical guidance to organizations navigating AI integration.

The objective of this literature review is therefore to systematically examine and synthesize the available evidence concerning managerial attitudes toward AI, drawing on a corpus of 14 peer-reviewed studies spanning organizational behavior, digital transformation, marketing strategy, public administration, and human resource management. By identifying thematic clusters within the literature and surfacing areas of agreement, tension, and ambiguity, this review aims to provide a coherent analytical foundation for future empirical and theoretical inquiry.

The paper is organized as follows. Section 2 establishes the theoretical background by defining key constructs and reviewing foundational frameworks relevant to the study of managerial AI attitudes. Section 3 outlines the methodology of the review, including the literature search strategy, selection criteria, and thematic synthesis process. Section 4 presents the thematic synthesis of the literature, organized around three core clusters. Section 5 identifies research gaps and proposes directions for future inquiry. Section 6 concludes with a summary of contributions and an acknowledgment of the review's limitations.

2. Theoretical Background

This section establishes the conceptual and theoretical foundations for examining managerial attitudes toward artificial intelligence in organizational settings. It first clarifies the meaning of managerial attitudes toward AI and then outlines the key theoretical frameworks that help explain how such attitudes are formed, shaped, and translated into organizational decisions.

2.1 Defining Managerial Attitudes Toward AI

Managerial attitudes toward AI can be conceptualized as the evaluative orientations—encompassing cognitive, affective, and behavioral dimensions—that managers hold regarding the deployment, governance, and consequences of artificial intelligence within organizational settings. This tripartite conception draws on classical attitude theory in social psychology, which holds that attitudes comprise beliefs about an object, emotional responses to it, and behavioral intentions stemming from those responses. In the organizational context, this means that a manager's attitude toward AI is shaped simultaneously by beliefs about AI's technical

capabilities, emotional reactions to AI-induced change, and intentions to adopt, resist, or modify AI-driven processes.

As comprehensively characterized by Dwivedi et al. (2019), AI encompasses a broad range of technologies, including machine learning, natural language processing, robotics, and computer vision, which together offer transformative potential to augment or replace human tasks across industrial, intellectual, and social domains. Because AI is inherently multidisciplinary, its organizational manifestations are equally diverse, ranging from back-office automation to front-line customer interaction systems, each with distinct implications for managerial agency and oversight.

2.2 Foundational Theoretical Frameworks

Several theoretical frameworks inform the study of managerial attitudes toward AI. The Technology-Organization-Environment (TOE) framework has been widely applied to understand how technological, organizational, and environmental factors interact to shape adoption decisions. Neumann et al. (2022) employ this framework in their comparative case study of Swiss public organizations, finding that the relative weight of technological and organizational factors shifts depending on the organization's stage in the adoption lifecycle. This dynamic perspective is important because it suggests that managerial attitudes are not static but evolve as experience with AI accumulates.

Task-technology fit theory offers a complementary lens, focusing on the congruence between the capabilities of a technology and the demands of the tasks to which it is applied. This is particularly relevant in understanding why managers may embrace AI in some functional domains while resisting it in others. Huang and Rust (2020) develop a strategic framework that distinguishes between mechanical AI (for automating repetitive functions), thinking AI (for data-driven decision support), and feeling AI (for emotion recognition and interaction analysis), each mapping onto distinct managerial tasks and generating distinct attitudinal responses. Managers who perceive AI as enhancing their decision-making capacity may hold more favorable attitudes than those who perceive AI as encroaching upon domains requiring human judgment and empathy.

The literature on digital transformation provides additional theoretical grounding. Brunetti et al. (2020) propose that digital transformation generates systemic challenges for organizations that require coordinated strategic responses across multiple stakeholder levels, including senior management, operational staff, and external partners. Dąbrowska et al. (2022) extend this analysis to a multi-level research agenda, arguing that the impact of digital technologies cannot be fully understood at the organizational level alone but must be examined across individual, organizational, ecosystem, and societal dimensions. These multi-level perspectives are directly relevant to managerial attitudes because they situate individual managerial responses within broader structural and cultural contexts.

Labor economics also contributes to the theoretical landscape. Acemoğlu and Restrepo (2019) present a task-based framework in which automation displaces labor in tasks previously performed by humans while also creating new tasks that reinstate labor demand. This framework has important implications for how managers conceptualize AI's impact on their workforce, as the net employment effects of AI depend critically on the balance between displacement and reinstatement—a balance that managers themselves help to determine through their investment and deployment decisions.

Finally, ethical theory is increasingly recognized as foundational to understanding managerial AI attitudes. Rodgers et al. (2022) develop a Throughput model that describes how

individuals navigate ethical decision-making in AI-augmented human resource management contexts, highlighting the tensions between algorithmic efficiency and normative principles of fairness, transparency, and human dignity. These ethical dimensions are not peripheral concerns but central to how managers evaluate the legitimacy and acceptability of AI systems, particularly in high-stakes domains such as hiring, performance management, and resource allocation.

3. Methodology

This study applies literature review approach to identify, select, and synthesize scholarly work on managerial attitudes toward artificial intelligence in organizational settings. The review was designed as a structured qualitative synthesis rather than a meta-analysis, reflecting the conceptual diversity and methodological heterogeneity of the existing literature.

The literature search was conducted through keyword-based searches in the academic databases Web of Science, Scopus, and Google Scholar, combining terms such as artificial intelligence, managerial attitudes, AI adoption, technology leadership, organizational change, and AI acceptance. To strengthen the coverage of relevant studies, backward reference checking was also used to identify additional articles cited in key publications.

The review included peer-reviewed studies published in English between 2019 and 2025, reflecting the recent acceleration of scholarly and practical interest in AI adoption and its organizational implications. Studies were selected if they examined managers', leaders', or organizational decision-makers' perceptions of AI, with particular attention to adoption, implementation, workforce implications, and ethical or governance concerns. Studies were excluded if they focused primarily on technical aspects of AI without a managerial perspective, or if they examined only consumer or employee attitudes without addressing managerial decision-making. Following screening for relevance, conceptual fit, duplication, and publication period, a final sample of 14 studies published between 2019 and 2025 was retained for analysis.

The selected studies were analyzed through thematic synthesis. Recurring concepts, findings, and theoretical arguments were coded and compared across the sample in order to identify dominant patterns in the literature. This process resulted in three thematic dimensions (clusters): organizational readiness and strategic orientation toward AI adoption, managerial perceptions of AI's impact on workforce skills and job redesign, and ethical governance and accountability in AI-related decision-making.

4. Literature Review

This section presents the findings of the reviewed literature through synthesis of the selected studies. The analysis is organized into three major thematic clusters that capture the main dimensions through which managerial attitudes toward AI have been examined in the existing research.

4.1 Thematic Cluster I.: Organizational Readiness and Strategic Orientation Toward AI Adoption

A foundational theme in the literature concerns the conditions under which organizations and their managers develop genuine readiness to adopt AI, moving beyond exploratory interest to committed implementation. Jöhnk et al. (2020), in a qualitative interview study with organizational leaders across multiple industries, identify a set of interdependent readiness factors that determine whether AI adoption achieves its intended outcomes. These factors

encompass technological infrastructure, data governance capabilities, human capital, and, crucially, managerial commitment and strategic clarity. Their findings underscore that organizations frequently underestimate the complexity of AI adoption, approaching it as a technical problem rather than as a strategic and cultural transformation that requires sustained managerial engagement.

In the public sector, Madan and Ashok (2022) conduct a systematic literature review of AI adoption in public administration, finding that institutional factors, regulatory environments, and accountability norms substantially shape managerial attitudes and adoption trajectories. Public sector managers, they argue, face distinctive tensions between efficiency imperatives, which favor AI adoption, and normative commitments to fairness, transparency, and democratic accountability, which may generate managerial hesitancy or resistance. Neumann et al. (2022) corroborate these findings in their comparative case study, demonstrating that even organizations with high technological readiness may stall in adoption due to managerial uncertainty about governance and accountability. This suggests that favorable attitudes toward AI in the public sector are contingent not only on perceived technical utility but also on the availability of credible governance frameworks.

The entrepreneurial context presents a contrasting picture. Chalmers et al. (2020) examine AI's implications for new venture creation, arguing that entrepreneurial managers may hold more positive attitudes toward AI because they perceive it as a lever for scaling activities that would otherwise require disproportionate human capital investment. However, they also caution that entrepreneurial ventures face distinctive liabilities stemming from over-reliance on AI-driven processes, particularly when AI systems encounter novel situations outside their training distributions. This introduces an important qualification: managerial enthusiasm for AI adoption is not necessarily correlated with successful AI deployment outcomes.

4.2 Thematic Cluster II.: AI's Impact on Workforce Skills, Job Redesign, and Managerial Responsibility

A second major thematic cluster concerns how managers perceive and respond to AI's transformative effects on the nature of work, workforce skills, and job design. This theme sits at the intersection of organizational behavior, human resource management, and labor economics, drawing on diverse theoretical traditions to illuminate a shared empirical concern.

Morandini et al. (2023) provide a comprehensive review of AI's impact on professional skills, documenting how AI-driven automation is simultaneously reducing the demand for certain cognitive competencies while creating new requirements for digital literacy, critical thinking, and human-AI collaboration skills. They argue that organizations face an upskilling and reskilling imperative that is both urgent and strategically complex, requiring managers to take active responsibility for workforce development rather than treating it as a downstream consequence of technology investment. The managerial attitude embedded in this argument is one of proactive stewardship: managers who view AI adoption as inseparable from workforce development tend to achieve more sustainable adoption outcomes than those who treat labor adjustment as an afterthought.

This perspective is reinforced by Acemoğlu and Restrepo's (2019) task-based framework, which demonstrates that automation's net effects on employment and wage structures depend on whether organizations invest in the creation of new tasks that reinstate labor demand. The implication for managerial attitudes is significant: managers who understand AI as a tool for

augmenting human capabilities, rather than simply replacing human labor, are more likely to make investment decisions that balance efficiency gains with workforce sustainability. Conversely, managers who hold predominantly displacement-oriented attitudes toward AI risk creating internal resistance and productivity losses that offset technological gains.

Wang et al. (2020) examine how information and communication technology use affects work design and individual well-being, providing a framework applicable to AI-enabled work environments. Their review of empirical studies reveals that technology use has heterogeneous effects on job demands, autonomy, skill variety, and social interaction, with managerial choices about technology deployment playing a central role in determining whether these effects are positive or negative. Managers who design AI-augmented work systems with attention to human factors, including meaningful task variety, adequate autonomy, and social support, generate higher levels of employee well-being and engagement than those who optimize exclusively for task efficiency.

The emotional dimensions of AI-driven workplace change also merit consideration. Diener et al. (2019) review the role of positive emotions in organizational settings, arguing that emotional climate substantially influences how employees respond to change initiatives. While their review does not focus specifically on AI, their findings have direct implications for managers navigating AI-induced uncertainty: managerial attitudes that convey confidence, transparency, and genuine concern for employee well-being are likely to generate more constructive emotional climates than those characterized by ambiguity, defensiveness, or exclusive focus on efficiency gains. Managers who acknowledge the emotional dimensions of AI adoption, including anxieties about job security and skill obsolescence, may be better positioned to build the organizational trust necessary for successful implementation.

Furthermore, Gerlich (2025) raises a distinctive concern about the effects of AI tool proliferation on human cognitive functioning, specifically the phenomenon of cognitive offloading, which refers to the tendency to delegate cognitive tasks to AI systems rather than exercising independent judgment. This finding carries significant implications for managerial attitudes: managers who uncritically embrace AI as a cognitive prosthesis risk degrading precisely the capabilities of critical thinking, ethical reasoning, and strategic judgment that are most essential for responsible organizational leadership. The implication is that positive managerial attitudes toward AI must be tempered by a reflexive awareness of the potential long-term cognitive consequences of deep AI dependence.

4.3 Thematic Cluster III.: Ethical Governance, Accountability, and the Normative Dimensions of AI Leadership

The third thematic cluster addresses the ethical and governance dimensions of AI adoption as they are experienced and navigated by managers. This cluster has gained increasing prominence in the literature as AI systems have been deployed in consequential decision-making contexts involving hiring, performance evaluation, credit allocation, and public service delivery, generating intense scrutiny of the normative assumptions embedded in algorithmic systems.

Rodgers et al. (2022) offer the most theoretically developed account of ethical decision making in AI-augmented organizational contexts. Their Throughput model describes how individuals' ethical frameworks, including utilitarian, deontological, and virtue-based orientations, interact with algorithmic processes to produce organizational outcomes. They find that the introduction of AI into human resource management processes does not eliminate ethical discretion but rather transforms the forms and contexts in which it is exercised. Managers retain significant ethical agency in decisions about which AI systems to adopt, how to configure their objectives, how to interpret their outputs, and how to adjudicate conflicts between algorithmic recommendations and human judgment. Managerial attitudes that treat AI as an ethical black box, delegating normative responsibility to technical systems, are therefore not only theoretically untenable but practically dangerous.

Madan and Ashok (2022) develop a related argument in the context of public administration, identifying what they term "AI tensions", conflicts between the efficiency imperatives that motivate AI adoption and the normative commitments to fairness, transparency, privacy, and human rights that define the public sector's accountability obligations. Their review suggests that public sector managers who adopt AI without adequate attention to these tensions risk generating legitimacy crises that ultimately undermine the value of AI investments. The attitudes most conducive to sustainable AI adoption in public settings are those that treat ethical governance not as a compliance burden but as a fundamental dimension of responsible leadership.

The digital transformation literature provides additional context for understanding how governance attitudes toward AI are shaped by broader organizational and institutional dynamics. Brunetti et al. (2020) find that successful digital transformation in multi-stakeholder regional innovation systems requires coordinated governance strategies that align the interests of companies, public administrators, and educational institutions. Managerial attitudes that are narrowly focused on organizational self-interest—prioritizing competitive advantage over shared governance norms—are less likely to produce the collaborative arrangements necessary for responsible AI ecosystem development.

Dąbrowska et al. (2022) extend this analysis, arguing that digital transformation must be understood as a socioeconomic change with implications across individual, organizational, ecosystem, and societal levels. Their multi-level perspective highlights an important asymmetry in how the costs and benefits of AI adoption are distributed: while efficiency gains tend to accrue to organizations, the disruptive effects of AI—including skill displacement, privacy erosion, and the concentration of algorithmic power—are often borne by workers, consumers, and communities. This asymmetry has direct implications for managerial attitudes, suggesting that managers who adopt exclusively firm-centric perspectives on AI adoption may systematically underweight the externalities their decisions generate, with significant long-term consequences for organizational legitimacy and social license.

Finally, Wang et al. (2020) contribute to this cluster by documenting how technology use affects not only individual productivity but also the social and relational dimensions of work, including trust, collaboration, and psychological safety. Managers who design AI-enabled work

environments with attention to these relational dimensions are better positioned to maintain the organizational cohesion necessary for adaptive governance of AI systems over time.

5. Research Gaps and Future Directions

Despite the substantial body of evidence reviewed above, several important gaps in the current literature warrant systematic attention in future research.

Cross-Cultural and Contextual Variation in Managerial AI Attitudes. The existing literature is heavily weighted toward Western organizational contexts, with particular concentration on North American and European settings. While Jöhnk et al. (2020) and Neumann et al. (2022) provide valuable insights from German-speaking and Swiss contexts respectively, comparative studies examining how managerial attitudes toward AI vary across cultural, institutional, and regulatory environments are conspicuously absent. Cultural dimensions such as uncertainty avoidance, power distance, and long-term orientation are theoretically likely to shape how managers evaluate AI's risks and benefits, yet these relationships remain empirically underexplored. Future research should employ cross-national comparative designs to examine whether the attitudinal patterns identified in Western contexts generalize to other settings, including emerging economies in Asia, Latin America, and sub-Saharan Africa where AI adoption is accelerating but governance infrastructure remains underdeveloped.

Longitudinal Dynamics of Managerial AI Attitudes. The preponderance of existing studies employs cross-sectional or snapshot designs, capturing managerial attitudes at a single point in time. This design limitation obscures the dynamic processes through which attitudes form, evolve, and are revised in response to AI adoption experiences. Morandini et al. (2023) document how AI's impact on workforce skills unfolds over time, but longitudinal research tracking managerial attitudes through the full arc of AI adoption—from initial awareness through implementation, consolidation, and review—is virtually absent from the literature. Longitudinal panel studies and organizational ethnographies would provide more nuanced accounts of attitudinal change and its organizational consequences.

The Micro-Foundations of Managerial Ethical Reasoning in AI Contexts. While Rodgers et al. (2022) make important strides in theorizing ethical decision-making in AI-augmented HRM contexts, the micro-level cognitive and emotional processes through which individual managers construct and navigate ethical judgments about AI remain poorly understood. Gerlich (2025) raises concerns about cognitive offloading that have direct implications for the quality of managerial ethical reasoning, yet the relationship between AI dependence and ethical judgment quality has not been empirically investigated. Future research should integrate insights from behavioral ethics, cognitive psychology, and AI studies to examine how extended exposure to AI decision-support systems affects managers' capacity for independent normative reasoning—a question with profound implications for organizational governance and accountability.

Integration of AI Attitudes with Practical Managerial Training. A broader implication of the identified gaps concerns the translation of research findings into practical interventions. As both Brunetti et al. (2020) and Dąbrowska et al. (2022) suggest, multi-stakeholder approaches to digital transformation governance that include educational institutions alongside businesses and public administrators may be necessary to cultivate the managerial AI literacy required for responsible adoption. Future research should therefore examine the effectiveness of targeted

management education and training programs in shaping more reflective, ethically grounded, and strategically sophisticated managerial attitudes toward AI.

6. Conclusion

This literature review has examined the state of scholarly knowledge concerning managerial attitudes toward artificial intelligence, synthesizing evidence from 14 peer-reviewed studies across organizational behavior, human resource management, public administration, entrepreneurship, marketing, and digital transformation research. The review makes several contributions to the literature.

First, it provides an integrated theoretical framework for understanding managerial AI attitudes, drawing on technology adoption theory, task-based frameworks from labor economics, digital transformation scholarship, and ethical decision-making theory to situate managerial orientations within their broader organizational and institutional contexts. Second, it identifies three thematic clusters—organizational readiness and strategic orientation, workforce skill impacts and job redesign, and ethical governance and normative leadership that together capture the principal dimensions along which managerial AI attitudes have been studied. Third, it surfaces important research gaps and proposes a targeted future research agenda addressing cross-cultural variation, longitudinal dynamics, and the micro-foundations of managerial ethical reasoning.

The review's principal findings can be summarized as follows: managerial attitudes toward AI are multidimensional and context-dependent, shaped by perceived technical utility, ethical awareness, workforce implications, and governance frameworks. Positive attitudes toward AI are not uniformly associated with successful adoption outcomes; rather, the quality and sophistication of managerial engagement with AI—including attention to ethical tensions, workforce development, and multi-level governance—appears to be the critical determinant of sustainable organizational value creation from AI investments.

Several limitations of this review should be acknowledged. The source corpus, while diverse in disciplinary scope, is limited to 14 studies, which constrains the comprehensiveness of thematic synthesis. The review also reflects the existing literature's bias toward Western organizational contexts, potentially limiting the generalizability of its conclusions. Future systematic reviews incorporating broader source corpora and more diverse geographic coverage would strengthen the evidence base for the conclusions advanced here. Despite these limitations, the review offers a structured and theoretically grounded synthesis that can serve as a foundation for both future academic inquiry and informed managerial practice in the AI age.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Acemoğlu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. Journal of Economic Perspectives, 33(2), 3–30. https://doi.org/10.1257/jep.33.2.3

Brunetti, F., Matt, D. T., Bonfanti, A., De Longhi, A., & Pedrini, 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

Chalmers, D., MacKenzie, N., & Carter, S. (2020). Artificial intelligence and entrepreneurship: Implications for venture creation in the Fourth Industrial Revolution. Entrepreneurship Theory and Practice, 45(5), 1028–1053. https://doi.org/10.1177/1042258720934581

Dąbrowska, J., Almpanopoulou, A., Brem, A., Chesbrough, H., & Järvi, K. (2022). Digital transformation, for better or worse: A critical multi-level research agenda. R&D Management, 52(5), 930–954. https://doi.org/10.1111/radm.12531

Diener, E., Thapa, S., & Tay, L. (2019). Positive emotions at work. Annual Review of Organizational Psychology and Organizational Behavior, 7, 451–477. https://doi.org/10.1146/annurev-orgpsych-012119-044908

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., & Coombs, C. (2019). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. Societies, 15(1), 6. https://doi.org/10.3390/soc15010006

Huang, M.-H., & Rust, R. T. (2020). A strategic framework for artificial intelligence in marketing. Journal of the Academy of Marketing Science, 49(1), 30–50. https://doi.org/10.1007/s11747-020-00749-9

Jöhnk, J., Weißert, M., & Wyrtki, K. (2020). Ready or not, AI comes—An interview study of organizational AI readiness factors. Business & Information Systems Engineering, 63(1), 5–20. https://doi.org/10.1007/s12599-020-00676-7

Madan, R., & Ashok, M. (2022). AI adoption and diffusion in public administration: A systematic literature review and future research agenda. Government Information Quarterly, 39(4), 101774. https://doi.org/10.1016/j.giq.2022.101774

Morandini, S., Fraboni, F., De Angelis, M., Puzzo, G., & Giusino, D. (2023). The impact of artificial intelligence on workers' skills: Upskilling and reskilling in organisations. Informing Science: The International Journal of an Emerging Transdiscipline, 26, 39–68. https://doi.org/10.28945/5078

Neumann, O., Guirguis, K., & Steiner, R. (2022). Exploring artificial intelligence adoption in public organizations: A comparative case study. Public Management Review, 25(5), 970–1000. https://doi.org/10.1080/14719037.2022.2048685

Rodgers, W., Murray, J. M., Stefanidis, A., Degbey, W. Y., & Tarba, S. Y. (2022). An artificial intelligence algorithmic approach to ethical decision-making in human resource management processes. Human Resource Management Review, 33(1), 100925. https://doi.org/10.1016/j.hrmr.2022.100925

Wang, B., Liu, Y., & Parker, S. K. (2020). How does the use of information communication technology affect individuals? A work design perspective. Academy of Management Annals, 14(2), 695–725. https://doi.org/10.5465/annals.2018.0127

Emotional Dimensions of Technology Resistance: A Thematic Literature Review

Černý Matej[1] and Baziuk Andriana[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: matej.cerny@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: andriana.baziuk@euba.sk

Abstract

Technology resistance represents a pervasive yet underexplored phenomenon in the broader discourse on digital transformation and innovation adoption. While substantial scholarly attention has been directed toward understanding the factors that facilitate technology acceptance, comparatively little systematic inquiry has addressed the emotional and psychosocial mechanisms underpinning resistance. This review synthesizes evidence from 22 peer-reviewed sources to examine the emotional dimensions of technology resistance, drawing on foundational theoretical frameworks including the Technology Acceptance Model, the Unified Theory of Acceptance and Use of Technology, technology readiness theory, and innovation resistance theory. Three thematic clusters are identified: the affective antecedents of resistance, including anxiety, discomfort, and distrust; the role of perceived threat and identity disruption in amplifying negative emotional responses; and the moderating influence of contextual and organizational factors on emotional resistance. The review reveals that emotional barriers are not peripheral concerns but constitute core determinants of whether individuals adopt or reject emerging technologies. Practical implications for designers, policymakers, and organizational leaders are discussed, and a set of targeted future research directions is proposed to advance theoretical and empirical understanding of this underexplored domain.

Key Words: technology resistance, emotional dimensions, adoption barriers, behavioural intention, user acceptance

JEL Classification: O33, D91, M15

1. Introduction

The accelerating proliferation of digital technologies across virtually every sector of contemporary life has generated a rich body of research on the antecedents of technology adoption. Organizations invest substantially in automation, artificial intelligence, and novel digital platforms, operating under the assumption that demonstrable utility will translate into user acceptance. This assumption is especially fragile in AI-intensive contexts, where emerging evidence suggests that the diffusion of systems such as ChatGPT is accompanied not only by expectations of efficiency gains, but also by concerns about worker substitution, changing skill requirements, and the erosion of human connection in service and knowledge work (Carvalho & Ivanov, 2023). Yet this assumption is frequently confounded by the persistence of technology resistance—a phenomenon characterized not merely by the rational calculation of costs and benefits, but by deeply embedded emotional responses that shape individuals' orientations toward innovation. Understanding these emotional dimensions is not an academic abstraction;

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

it carries direct implications for the success of digital transformation initiatives, the design of human-centered technologies, and the management of organizational change.

Despite the voluminous literature on technology acceptance, the emotional substrate of resistance remains comparatively undertheorized. Dominant frameworks such as the Technology Acceptance Model and the Unified Theory of Acceptance and Use of Technology have historically privileged cognitive constructs—perceived usefulness and perceived ease of use—while treating affective states as secondary or moderating variables rather than as primary explanatory mechanisms (Lee et al., 2025). This intellectual imbalance has left scholars and practitioners poorly equipped to account for the persistence of resistance even when technologies are demonstrably useful, intuitive, and widely available. Emotions including anxiety, distrust, discomfort, and the fear of displacement operate at a level that cognitive evaluations alone cannot fully capture.

The present review addresses this gap by synthesizing the existing literature on the emotional dimensions of technology resistance across multiple domains, including healthcare, logistics, food technology, marketing, and the workplace. The objective is threefold: first, to map the conceptual landscape of emotional barriers to technology adoption; second, to identify the principal thematic clusters that organize current understanding; and third, to delineate the research gaps that must be addressed to advance both theory and practice. By integrating findings from studies that span individual, organizational, and societal levels of analysis, this review offers a comprehensive account of why emotional resistance persists and what conditions intensify or attenuate it.

The remainder of the article is structured as follows. The theoretical background section defines the key constructs and reviews the foundational frameworks most relevant to emotional technology resistance. The literature review then outlines the methodological approach and synthesizes prior research across key thematic areas, including affective antecedents, perceived threat and identity disruption, and contextual and organizational influences. The subsequent section identifies research gaps and proposes future directions, followed by a conclusion that summarizes the review's contributions and acknowledges its limitations.

2. Theoretical Background

Understanding technology resistance requires both a clear definition of the concept and an overview of the theoretical frameworks that inform its analysis. In particular, existing models of technology adoption and resistance provide important insights into the role of emotional and cognitive factors in shaping user responses to new technologies.

2.1 Defining Technology Resistance

Technology resistance is broadly understood as the refusal or reluctance of individuals or collectives to adopt, use, or accept a new technology, even when that technology is available and potentially beneficial. This construct is conceptually distinct from simple non-adoption, which may result from lack of access or awareness; resistance implies an active or dispositional opposition rooted in cognitive, emotional, or social factors. Lyu et al. (2024) operationalize resistance in the context of autonomous delivery vehicles as a latent construct encompassing both functional and psychological barriers, thereby drawing attention to the multidimensionality of the phenomenon. This multidimensional understanding is essential because it foregrounds the possibility that individuals may simultaneously recognize a

technology's utilitarian value while harboring affective reservations that override rational endorsement.

2.2 Foundational Theoretical Frameworks

The Technology Acceptance Model and the Unified Theory of Acceptance and Use of Technology have provided the dominant scaffolding for empirical research on technology adoption and rejection. Lee et al. (2025) systematically reviewed both frameworks in the healthcare context and concluded that, while they successfully predict behavioral intention, they inadequately capture the psychosocial and emotional determinants of resistance, including affective trust, anxiety, and social norms related to identity. This limitation is particularly consequential in high-stakes domains where emotional barriers are especially pronounced.

Technology readiness theory offers a complementary perspective by conceptualizing individuals' propensity to embrace technology as a multidimensional psychological state comprising innovativeness, optimism, discomfort, and insecurity (Blut & Wang, 2019). Crucially, the discomfort and insecurity dimensions are intrinsically affective, capturing feelings of being overwhelmed, threatened, or controlled by technology. Blut and Wang's (2019) meta-analysis of 163 articles and over 69,000 individuals confirmed that these negative dimensions independently and significantly suppress technology usage, even when positive dimensions such as optimism are present. This finding establishes technology readiness not as a uniformly positive disposition but as a dynamic interplay between facilitative enthusiasm and inhibitory emotional states.

Innovation resistance theory, as applied by Lyu et al. (2024), posits that innovations face resistance when they demand significant behavioral changes or when they are perceived as incompatible with existing values, habits, or self-concepts. The theory is particularly illuminating for understanding how emotional dissonance arises when a technology challenges established routines or professional identities. In organizational settings, Wang et al. (2020) demonstrated that information and communication technology use fundamentally reshapes work design, altering task characteristics in ways that can either enhance autonomy and meaning or generate role ambiguity and affective strain—a duality that renders the emotional consequences of technology integration inherently context-dependent.

2.3 Emotions as Primary Constructs

While cognitive appraisals of usefulness and ease of use remain important, a growing body of evidence supports the primacy of emotion in shaping technology-related attitudes. The concept of techno-anxiety—diffuse apprehension about interacting with technology—has been linked to avoidance behavior across a range of empirical contexts (Blut & Wang, 2019; Lee et al., 2025). Davenport et al. (2019) noted that artificial intelligence's capacity to autonomously perform tasks previously considered the exclusive province of human expertise generates anxiety not merely about job loss but about the erosion of professional identity and the competence-based emotions that sustain it. Morandini et al. (2023) similarly found that the fear of skills obsolescence is a potent emotional driver of resistance to AI integration in organizational settings, motivating avoidance behaviors that may ultimately be counterproductive for both workers and organizations.

The concept of disgust has received particular attention in the context of novel food technologies. Bryant and Barnett (2020) documented that consumer responses to cultured meat are overwhelmingly shaped by visceral emotional reactions—including disgust,

unnaturalness, and what researchers have termed the "yuck factor"—rather than by deliberative cost-benefit calculations. This work illustrates how primary emotions, rooted in evolutionary psychology, can constitute nearly insurmountable barriers to the adoption of demonstrably viable technologies, irrespective of rational arguments in their favor. Collectively, these theoretical perspectives establish that any comprehensive account of technology resistance must place emotional experience at the center of analysis rather than treat it as an epiphenomenon of cognitive evaluation.

3. Methodology

This study employs a thematic literature review design to explore the emotional dimensions of technology resistance. The aim of paper is to systematically identify, analyze, and synthesize existing academic knowledge on emotional and psychosocial factors influencing resistance to technology.

The literature search was conducted using major academic databases, including Web of Science, Scopus, and Google Scholar. The search focused on peer-reviewed articles published between 2017 and 2025, ensuring a contemporary perspective on emerging technologies such as Generative AI, the Metaverse, and autonomous systems. Key search terms included: technology resistance, technology acceptance, emotions, anxiety, trust, technology readiness, and innovation resistance.

From the initial pool of identified studies, a total of 22 peer-reviewed articles were selected based on their relevance. The inclusion criteria required that the selected studies focus on non-functional barriers to technology adoption. To ensure a comprehensive perspective, the sample included research from diverse sectors, such as healthcare, education, marketing, and human resource management. The review was limited to peer-reviewed journal articles published in English.

The analysis employed a thematic synthesis approach consisting of three stages: initial coding of emotional constructs (e.g., anxiety, fear of replacement, discomfort), followed by their grouping into broader psychosocial patterns such as identity disruption or contextual factors. These patterns were then consolidated into three core themes: affective antecedents, perceived threat and identity, and contextual moderators.

4. Literature Review

4.1 Affective Antecedents of Technology Resistance: Anxiety, Discomfort, and Distrust

The most extensively documented emotional antecedents of technology resistance cluster around anxiety, generalized discomfort with novel systems, and distrust of the technologies or institutions deploying them. These three constructs, while analytically distinguishable, are empirically interrelated and frequently co-occur in populations exhibiting strong resistance to adoption.

Anxiety in technology contexts is typically operationalized as a state of apprehension, tension, or worry associated with the prospect of using a technology, its outcomes, or the social consequences of its use. Blut and Wang (2019) confirmed through meta-analytic synthesis that the insecurity dimension of technology readiness—which captures precisely this form of anxiety—is a robust negative predictor of technology usage intentions across diverse samples and contexts. Notably, the effect of insecurity on usage was found to be mediated by attitudes and subjective norms, indicating that anxiety does not operate in isolation but shapes the

broader evaluative and social context through which individuals form behavioral intentions. Lee et al. (2025) extended this picture to healthcare settings, documenting that patients and clinicians alike exhibit heightened anxiety when confronted with AI-driven diagnostic systems, remote monitoring platforms, and electronic health record systems, even when these technologies are endorsed by trusted professional bodies. This anxiety is not merely about the technology itself but is entangled with concerns about data privacy, algorithmic transparency, and the perceived dehumanization of care relationships.

Discomfort, conceived as a sense of loss of control or feeling overwhelmed by technology, operates through somewhat different mechanisms. Lyu et al. (2024) found that discomfort with the autonomous and opaque decision-making of delivery robots was a significant predictor of both passive non-adoption and active resistance, the latter manifesting as negative word-of-mouth and opposition to infrastructure development. Importantly, their structural equation model revealed that discomfort was partially mediated by perceived risk, suggesting that the emotional experience of being out of control translates into heightened cognitive risk appraisals that further depress adoption intentions. This mediation pathway implies that interventions targeting discomfort—for example, through transparency mechanisms or graduated exposure protocols—may have downstream benefits for risk perception as well.

Distrust represents a particularly consequential affective-cognitive barrier because it operates at the level of the institution or system rather than the individual technology encounter. Mariani et al. (2021) observed that consumer distrust of AI-powered marketing systems is frequently rooted not in negative personal experience but in media narratives and socially circulating beliefs about surveillance, manipulation, and data exploitation. This observation aligns with Bryant and Barnett's (2020) finding that distrust of regulatory bodies and food technology corporations was among the strongest predictors of resistance to cultured meat adoption, even among consumers who expressed abstract support for sustainable protein alternatives. The temporal dynamics of distrust are also significant: once established, it tends to be resistant to disconfirmation through positive experience, creating a path-dependent structure in which early emotional responses powerfully constrain subsequent adoption trajectories.

Taken together, these findings suggest that emotional resistance often begins with immediate affective responses to technology itself; the next thematic cluster shows how such resistance deepens when technologies are perceived as threats to identity, competence, or established ways of life.

4.2 Perceived Threat, Identity Disruption, and Emotional Resistance

A second thematic cluster in the literature addresses the emotional responses generated when technology is perceived as threatening to individuals' professional identities, sense of competence, or established ways of life. This dimension of resistance is particularly pronounced in workplace contexts but extends to consumer domains as well.

The integration of artificial intelligence and automation into professional settings has generated considerable anxiety about job displacement and skills obsolescence. Davenport et al. (2019) argued that AI's growing capacity to perform cognitive and analytical tasks previously associated with high-status knowledge workers creates a distinctive form of identity threat that extends beyond instrumental concerns about employment security. When a technology is perceived as capable of replicating or surpassing human expertise, it implicitly challenges the competence-based identities through which many professionals derive meaning and self-esteem. This emotional dynamic is documented by Morandini et al. (2023), who found that

workers' resistance to upskilling and reskilling initiatives was not primarily driven by rational assessments of the initiatives' utility but by the emotional threat implicit in acknowledging that their existing skills are becoming obsolete. The requirement to reskill carries a stigmatizing implication—that one's current knowledge is inadequate—which activates shame and defensiveness, motivating resistance to the very programs designed to facilitate adaptation.

Singh et al. (2019) documented an analogous pattern in sales professions, where the introduction of AI-powered customer relationship management and predictive analytics platforms generated resistance rooted in the threat these tools posed to relational expertise as the defining competency of sales professionals. Salespeople who had built careers on interpersonal skills and tacit relationship knowledge experienced AI tools not merely as productivity aids but as challenges to the legitimacy of their professional contributions. This identity-protective resistance was manifested in selective engagement with the technology—using it for routine tasks while avoiding functionalities that might expose them to unfavorable comparisons with algorithmic recommendations.

Identity-related resistance may become even more salient in immersive digital environments, where technologies do not merely support work or consumption but mediate the presentation of self. In metaverse-oriented systems, avatars and extended-reality environments are closely tied to social meaning, self-representation, and the continuity between online and offline identity, suggesting that emotional resistance may be intensified when users perceive such technologies as destabilizing how they present themselves or are recognized by others (Park & Kim, 2022).

The disruption of established routines and habits constitutes a related but distinct source of emotionally resonant resistance. Kaur et al. (2021) identified that habitual attachment to existing food ordering behaviors created emotional inertia that resisted the adoption of food delivery applications, even among consumers who acknowledged the applications' functional advantages. This inertia was not experienced as deliberate choice but as an affective pull toward the familiar—an emotional valence attached to established practices that competing technologies must overcome rather than simply satisfy. Longo et al. (2020) similarly noted in their examination of Industry 5.0 that workers' resistance to human-machine collaboration was frequently rooted in the disruption of established patterns of craft knowledge and embodied skill, the loss of which was experienced as a form of deskilling-induced grief.

The emotional dimensions of resistance also manifest in collective and cultural contexts. Serdyukov (2017) noted in the domain of educational innovation that resistance among educators was frequently animated by feelings of professional threat—the perception that innovative pedagogical technologies implied a critique of existing teaching practices and, by extension, of the professional identities built upon them. This emotional resistance to educational technology adoption has proven remarkably durable across decades of reform efforts, suggesting that identity-based emotional responses constitute structural rather than contingent barriers to innovation diffusion.

While this cluster highlights the internal emotional consequences of perceived threat and disruption, the following section shows that the intensity of resistance is also shaped by external conditions, particularly organizational, social, and cultural context.

4.3 Contextual and Organizational Moderators of Emotional Resistance

A third thematic cluster addresses the conditions under which emotional resistance is intensified or attenuated, directing attention to contextual and organizational factors that modulate the affective experience of technology encounters. This literature reveals that

emotional responses to technology are not fixed individual traits but are highly sensitive to the social, organizational, and institutional environments in which adoption occurs. More broadly, this view is consistent with systems-oriented perspectives emphasizing that individual responses are shaped by feedbacks between personal experience, institutional arrangements, and collective social dynamics rather than by isolated cognition alone (Di Baldassarre et al., 2019).

Organizational support and leadership behavior have been identified as significant moderators of technology-related anxiety and resistance. Wang et al. (2020) demonstrated that the work design implications of ICT use are strongly conditioned by organizational context: technologies that enhance autonomy and task variety in supportive organizational climates may generate anxiety and role ambiguity in climates characterized by high surveillance and performance monitoring. This finding suggests that the emotional valence of technology is not intrinsic to the technology itself but is partly constructed through organizational practices and relational dynamics. Klein and Todesco (2021) observed that SMEs' digital transformation during the COVID-19 pandemic was significantly impeded by employees' emotional resistance—rooted in anxiety about competence and fear of error—when organizational support structures were absent or inadequate, whereas organizations that invested in psychological safety and peer learning networks experienced markedly lower resistance levels.

The role of trust in mediating the relationship between organizational context and emotional resistance has been elaborated in the AI adoption literature. Garibay et al. (2023) argued that algorithmic transparency and explainability are not merely technical design requirements but fundamentally emotional ones: individuals who cannot understand how an AI system arrives at its decisions experience a loss of epistemic control that generates anxiety and resistance, regardless of the system's objective accuracy. This argument implies that emotional resistance to AI is partly a function of design choices—specifically, the degree to which systems are legible to non-expert users—and that seemingly technical decisions about interpretability carry profound affective consequences for adoption.

Social norms and peer influence constitute additional contextual moderators of emotional resistance. Barrot et al. (2021) found that Filipino students' emotional responses to online learning platforms were significantly shaped by peer experience and social comparison: students who perceived their peers as struggling with the same platforms experienced solidarity and reduced shame, whereas those who felt isolated in their difficulties developed heightened anxiety and avoidance behavior. Budhwar et al. (2023) similarly noted that employees' emotional responses to generative AI tools such as ChatGPT were modulated by the social climate within their organizations—whether AI use was normalized and openly discussed, or treated with ambivalence and stigma. Normalization reduced anxiety by reframing AI engagement as a collective competency-building activity rather than an individual admission of inadequacy.

Finally, domain-specific cultural and value-based factors shape the intensity of emotional resistance. Bryant and Barnett (2020) documented that disgust responses to cultured meat were significantly stronger in cultural contexts where natural food consumption is deeply embedded in identity and tradition, pointing to the importance of cultural context in calibrating the emotional threshold that novel technologies must overcome. Yung and Khoo-Lattimore (2017) noted in the tourism context that initial emotional discomfort with virtual and augmented reality experiences was substantially attenuated by guided exposure and social framing, suggesting that contextual interventions can effectively recalibrate affective responses that might otherwise constitute durable resistance.

Overall, the three thematic clusters indicate that emotional resistance is shaped not only by immediate affective reactions, but also by perceived identity threats and the broader contexts in which technological adoption takes place.

5. Research Gaps and Future Directions

Despite the advances documented in this review, several significant gaps in the current literature warrant systematic attention from future researchers.

First, the temporal dynamics of emotional resistance remain insufficiently theorized and empirically underdeveloped. The overwhelming majority of studies on technology resistance employ cross-sectional designs that capture emotional responses at a single point in time, typically at or shortly after initial technology encounter. This methodological limitation precludes understanding of how emotional responses evolve over time—whether initial anxiety attenuates with familiarity, whether distrust deepens or dissipates following repeated use, and whether identity-based resistance is amenable to resolution through sustained engagement or organizational intervention. Longitudinal research designs that track the emotional trajectories of technology resistance across extended adoption cycles are therefore a critical priority. Such designs would not only advance theoretical understanding but would provide practitioners with more nuanced guidance on the timing and nature of interventions (Blut & Wang, 2019; Lyu et al., 2024).

Second, there is a notable paucity of research examining emotional technology resistance at the collective or organizational level as a distinct phenomenon irreducible to the aggregation of individual responses. Organizational cultures, teams, and communities develop shared emotional orientations toward technologies that may differ substantially from the sum of individual attitudes. This line of inquiry could also benefit from stronger engagement with organizational scholarship showing that persistent collective responses are often reproduced through broader institutionalized practices, status hierarchies, and taken-for-granted assumptions rather than through individual attitudes alone (Amis, Mair, & Munir, 2020). The mechanisms through which collective emotional resistance is formed, maintained, and potentially resolved—including the role of opinion leaders, institutional narratives, and shared identity threats—remain largely unexplored (Morandini et al., 2023; Serdyukov, 2017). Research incorporating ethnographic and organizational case study methodologies alongside established survey instruments would advance understanding of these collective dynamics.

Third, the design implications of emotional resistance research remain underdeveloped. While several studies note that transparency, explainability, and graduated exposure can attenuate emotional barriers, there is as yet no systematic framework for translating these insights into actionable design principles. Longo et al. (2020) and Garibay et al. (2023) have begun to articulate human-centered design principles relevant to emotional acceptance, but these efforts remain largely conceptual rather than empirically validated. Future research should develop and test design intervention protocols specifically targeting identified emotional barriers, measuring outcomes not only in terms of adoption rates but in terms of sustained positive affect, trust, and sense of competence associated with technology use. Such research would bridge the gap between basic science on emotional resistance and applied human-computer interaction design, with significant implications for practitioners across healthcare, logistics, education, and consumer technology contexts. Future work should also pay greater attention to preventive and psychosocial perspectives when technologies are introduced into sensitive domains such as healthcare and mental health support, where

emotional resistance may interact with vulnerability, trust, and perceived stigma in especially consequential ways (Fusar-Poli et al., 2021).

6. Conclusion

This literature review has synthesized evidence from 22 peer-reviewed sources to advance understanding of the emotional dimensions of technology resistance—a domain that has been systematically underrepresented in the dominant technology acceptance literature. By organizing the existing evidence around three thematic clusters, the review has demonstrated that affective antecedents including anxiety, discomfort, and distrust; processes of identity threat and emotional disruption; and contextual and organizational moderators collectively constitute a rich and structurally coherent account of why individuals and groups resist technologies that may otherwise be demonstrably beneficial.

The central contribution of this review is to show more precisely that technology resistance is often driven by three interrelated emotional mechanisms: anxiety and discomfort in response to unfamiliar or opaque systems, identity-based threat linked to perceived loss of competence or professional relevance, and contextual amplification through weak organizational support, low transparency, or unsupportive social environments. For theory, this means that dominant acceptance models should not treat affect as secondary to cognition, especially in high-stakes, AI-intensive, or identity-relevant settings. For practice, the review suggests that successful adoption strategies must go beyond usability and utility by addressing trust, perceived control, psychological safety, and the social framing of technology use. Several limitations of this review must be acknowledged. The source base, while methodologically and thematically diverse, does not include studies from all global regions or cultural contexts, limiting the generalizability of conclusions. The review is further constrained by the available literature's predominant reliance on self-report measures of emotional states, which may not fully capture the depth or complexity of affective experience. Additionally, the rapid pace of technological change means that some findings may be context-specific to the particular technologies examined and may not generalize to emerging systems. In practical terms, the reviewed literature indicates that resistance is more likely to be reduced when technologies are introduced with explainability, guided exposure, peer normalization, and organizational support rather than with purely efficiency-oriented implementation logic. Despite these limitations, this review provides a theoretically grounded and empirically informed foundation for future inquiry into the emotional dimensions of technology resistance.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Barrot, J. S., Llenares, I. I., & del Rosario, L. S. (2021). Students' online learning challenges during the pandemic and how they cope with them: The case of the Philippines. Education and Information Technologies, 26(6), 7321–7338. https://doi.org/10.1007/s10639-021-10589-x

Blut, M., & Wang, C. (2019). Technology readiness: A meta-analysis of conceptualizations of the construct and its impact on technology usage. Journal of the Academy of Marketing Science, 48(4), 649–669. https://doi.org/10.1007/s11747-019-00680-8

Bryant, C., & Barnett, J. (2020). Consumer acceptance of cultured meat: An updated review (2018–2020). Applied Sciences, 10(15), 5201. https://doi.org/10.3390/app10155201

Budhwar, P., Chowdhury, S., Wood, G., Aguinis, H., & Bamber, G. J. (2023). Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT. Human Resource Management Journal, 33(3), 606–659. https://doi.org/10.1111/1748-8583.12524

Carvalho, I., & Ivanov, S. (2023). ChatGPT for tourism: Applications, benefits and risks. Tourism Review, 78(4), 1119–1134. https://doi.org/10.1108/tr-02-2023-0088

Davenport, T. H., Guha, A., Grewal, D., & Breßgott, T. (2019). How artificial intelligence will change the future of marketing. Journal of the Academy of Marketing Science, 48(1), 24–42. https://doi.org/10.1007/s11747-019-00696-0

Di Baldassarre, G., Sivapalan, M., Rusca, M., Cudennec, C., & Garcia, M. (2019). Sociohydrology: Scientific challenges in addressing the Sustainable Development Goals. Water Resources Research, 55(8), 6327–6355. https://doi.org/10.1029/2018wr023901

Fusar-Poli, P., Correll, C. U., Arango, C., Berk, M., & Patel, V. (2021). Preventive psychiatry: A blueprint for improving the mental health of young people. World Psychiatry, 20(2), 200–221. https://doi.org/10.1002/wps.20869

Garibay, Ö. Ö., Winslow, B., Andolina, S., Antona, M., & Bodala, I. P. (2023). Six human-centered artificial intelligence grand challenges. International Journal of Human-Computer Interaction, 39(3), 391–437. https://doi.org/10.1080/10447318.2022.2153320

Kaur, P., Dhir, A., Talwar, S., & Ghuman, K. (2021). The value proposition of food delivery apps from the perspective of theory of consumption value. International Journal of Contemporary Hospitality Management, 33(4), 1129–1159. https://doi.org/10.1108/ijchm-05-2020-0477

Klein, V. B., & Todesco, J. L. (2021). COVID-19 crisis and SMEs responses: The role of digital transformation. Knowledge and Process Management, 28(2), 117–133. https://doi.org/10.1002/kpm.1660

Lee, A. T., Ramasamy, R., & Subbarao, A. (2025). Understanding psychosocial barriers to healthcare technology adoption: A review of TAM Technology Acceptance Model and Unified Theory of Acceptance and Use of Technology and UTAUT frameworks. Healthcare, 13(3), 250. https://doi.org/10.3390/healthcare13030250

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. https://doi.org/10.3390/app10124182

Lyu, T., Huang, K., & Chen, H. (2024). Exploring the impact of technology readiness and innovation resistance on user adoption of autonomous delivery vehicles. International Journal of Human-Computer Interaction, 41(2), 1–15. https://doi.org/10.1080/10447318.2024.2400387

Mariani, M. M., Perez-Vega, R., & Wirtz, J. (2021). AI in marketing, consumer research and psychology: A systematic literature review and research agenda. Psychology and Marketing, 39(4), 755–778. https://doi.org/10.1002/mar.21619

Morandini, S., Fraboni, F., De Angelis, M., Puzzo, G., Giusino, D., & Pietrantoni, L. (2023). The impact of artificial intelligence on workers' skills: Upskilling and reskilling in organisations.

Informing Science: The International Journal of an Emerging Transdiscipline, 26, 39–68. https://doi.org/10.28945/5078

Park, S., & Kim, Y.-G. (2022). A metaverse: Taxonomy, components, applications, and open challenges. IEEE Access, 10, 4209–4251. https://doi.org/10.1109/access.2021.3140175

Serdyukov, P. (2017). Innovation in education: What works, what doesn't, and what to do about it? Journal of Research in Innovative Teaching & Learning, 10(1), 4–33. https://doi.org/10.1108/jrit-10-2016-0007

Singh, J., Flaherty, K., Sohi, R. S., Deeter-Schmelz, D., Habel, J., Le Meunier-FitzHugh, K., Malshe, A., Mullins, R., & Onyemah, V. (2019). Sales profession and professionals in the age of digitization and artificial intelligence technologies: Concepts, priorities, and questions. Journal of Personal Selling and Sales Management, 39(1), 2–22. https://doi.org/10.1080/08853134.2018.1557525

Wang, B., Liu, Y., & Parker, S. K. (2020). How does the use of information communication technology affect individuals? A work design perspective. Academy of Management Annals, 14(2), 695–725. https://doi.org/10.5465/annals.2018.0127

Yung, R., & Khoo-Lattimore, C. (2017). New realities: A systematic literature review on virtual reality and augmented reality in tourism research. Current Issues in Tourism, 22(17), 2056–2081. https://doi.org/10.1080/13683500.2017.1417359

Amis, J., Mair, J., & Munir, K. A. (2020). The organizational reproduction of inequality. Academy of Management Annals, 2020, Vol. 14, No. 1, 195–230. https://doi.org/10.5465/annals.2017.0033

Open Innovation and Knowledge Dissemination in the Digital Era: A Thematic Literature Review

Banda Federico [1] and Bolek Vladimír [2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: federico.banda@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: vladimir.bolek@euba.sk

Abstract

The digital transformation of knowledge production and dissemination has fundamentally reshaped the paradigms through which innovation is conceived, shared, and implemented across industries and institutions. This literature review examines the intersection of open innovation principles and digital knowledge dissemination, synthesizing evidence from seventeen peer-reviewed sources spanning education, technology, and knowledge management. The objective is to critically map the theoretical underpinnings and empirical developments in this domain, with particular attention to how digital platforms, pedagogical innovations, and cross-sectoral collaborations mediate knowledge co-creation. Key findings indicate that open innovation digital platforms significantly enhance knowledge accessibility and collaborative problem-solving, while digital transformation in higher education emerges as a critical vector through which knowledge dissemination practices are being renegotiated. The review further identifies a persistent tension between the rapid adoption of digital tools and the structural, institutional, and equity-related barriers that constrain their effectiveness. These findings carry substantial implications for policymakers, educators, and organizational leaders seeking to leverage open innovation ecosystems for sustainable knowledge dissemination in an increasingly digitized world.

Key Words: open innovation, knowledge dissemination, digital collaboration, knowledge sharing, innovation ecosystems

JEL Classification: O31, O33, D83

1. Introduction

The proliferation of digital technologies over the past two decades has fundamentally altered the landscape of knowledge production, sharing, and utilization. In parallel, the concept of open innovation—broadly understood as the deliberate use of inflows and outflows of knowledge to accelerate internal innovation and expand the markets for external use of innovation—has gained considerable traction across academic, industrial, and public sector contexts. The convergence of these two phenomena represents one of the most consequential developments in contemporary knowledge management, yet the literature remains fragmented across disciplinary boundaries, offering limited synthesis of how digital infrastructure and open innovation principles interact to shape knowledge dissemination at scale.

The significance of this topic has been amplified by a series of global disruptions, most notably the COVID-19 pandemic, which compelled institutions worldwide to rapidly reconfigure their knowledge-sharing infrastructures and adopt digital modalities with

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

unprecedented urgency. Higher education institutions, research organizations, and industrial enterprises alike were forced to confront longstanding assumptions about the primacy of physical co-location in knowledge transfer, accelerating experimentation with digital platforms, open data repositories, and hybrid collaborative models. Despite this acceleration, a coherent theoretical and empirical account of the mechanisms through which open innovation and digital knowledge dissemination interact remains elusive.

A critical gap exists at the nexus of open innovation theory and digital pedagogy literature. While substantial scholarly attention has been devoted separately to open innovation frameworks and to digital transformation in education and research, comparative and integrative analyses that draw upon both streams remain underrepresented in the proceedings literature. Furthermore, existing reviews tend to privilege technologically advanced contexts, leaving questions about equity, scalability, and cross-cultural transferability inadequately addressed.

This literature review addresses these gaps by synthesizing evidence from seventeen peer-reviewed sources to accomplish three primary objectives. First, it establishes the theoretical foundations that link open innovation to digital knowledge dissemination. Second, it identifies and critically examines thematic clusters within the literature, encompassing digital platforms and knowledge co-creation, digital transformation in institutional knowledge practices, and the role of emerging technologies in enabling open innovation. Third, it delineates unresolved tensions and proposes specific directions for future inquiry. The paper proceeds as follows: Section 2 provides the theoretical background, Section 3 presents the thematic literature synthesis, Section 4 identifies research gaps and future directions, and Section 5 offers conclusions.

2. Theoretical Background

2.1 Defining Open Innovation in the Digital Context

Open innovation, as a conceptual framework, departs from the closed model in which firms and institutions rely exclusively on internal knowledge resources. In its digital manifestation, the framework encompasses the use of networked platforms, community-driven repositories, and collaborative digital environments to orchestrate the flow of knowledge across organizational and disciplinary boundaries. Abbate, Codini, and Aquilani (2019) provide a foundational account of how Open Innovation Digital Platforms (OIDPs) function as enabling infrastructures for knowledge co-creation, arguing that such platforms support all phases of interactive, coupled open innovation processes by providing structured tools for ideation, co-development, and knowledge transfer. This insight is particularly germane in the current context, where digital mediation has become the default modality for inter-organizational collaboration.

The theoretical lineage of open innovation draws on resource-based views of the firm, knowledge management theory, and network theory, all of which find renewed relevance in digital settings. Knowledge, understood as a dynamic resource that appreciates through sharing rather than depreciating through use, is central to this framework. In digital environments, the marginal cost of knowledge reproduction and dissemination approaches zero, fundamentally altering the economics of knowledge sharing and creating structural incentives for openness.

2.2 Digital Transformation and Institutional Knowledge Practices

Digital transformation is not merely a technological phenomenon but a sociotechnical process that reconfigures the organizational, cultural, and epistemic structures through which knowledge is produced and disseminated. Bond, Marín, Dolch, Bedenlier, and Zawacki-Richter (2018) document this reconfiguration in the context of German higher education, demonstrating that digitalization reshapes both the tools and the underlying pedagogical assumptions that govern knowledge transfer. Their findings highlight that the effectiveness of digital transformation depends less on the availability of technology per se than on the institutional capacity to integrate digital tools within coherent pedagogical and organizational frameworks.

Carayannis and Morawska (2022) extend this analysis to the macro-societal level, arguing that the emergence of Society 5.0 and Industry 5.0 paradigms demands a reconceptualization of the university's role as a node in broader open innovation ecosystems. In this vision, universities are not merely repositories of codified knowledge but active participants in knowledge co-creation networks that span industrial, civic, and governmental actors. This reconceptualization has significant implications for how knowledge dissemination is theorized and operationalized.

2.3 Foundational Frameworks: From Transmission to Co-Creation

A critical theoretical evolution in this domain concerns the shift from transmission-based to co-creation-based models of knowledge dissemination. Traditional models conceptualized knowledge as a static commodity transmitted from expert to novice through hierarchical channels. Contemporary frameworks, informed by constructivist learning theory and systems thinking, emphasize the iterative, dialogic, and collaborative dimensions of knowledge production. Kukulska-Hulme and colleagues (2020) identify several emerging pedagogical innovations, including collective learning and machine learning-enhanced personalization, that embody this co-creative orientation and suggest new pathways for disseminating knowledge in open, networked environments.

The design thinking framework, as analyzed by Panke (2019), provides a complementary theoretical lens, foregrounding the iterative, human-centered, and collaborative dimensions of problem-solving and knowledge generation. Design thinking's emphasis on prototyping, iteration, and cross-disciplinary collaboration resonates strongly with open innovation principles and offers a process model applicable across educational and organizational settings. Similarly, the Education 4.0 paradigm, as elaborated by Almeida and Simões (2019), positions learning as an active, experiential process enabled by Industry 4.0 tools, including gamification, simulation, and connected devices, all of which can be understood as instruments of open, participatory knowledge dissemination.

2.4 Community-Curated Knowledge Infrastructures

A further theoretical dimension concerns the governance of open knowledge infrastructures. Williams and colleagues (2018) offer an instructive case study through the Neotoma Paleoecology Database, a community-curated, openly accessible data resource that exemplifies the principles of distributed scientific governance and collaborative knowledge stewardship. The Neotoma model demonstrates how open data practices, scalable governance structures, and interdisciplinary community-building can sustain high-quality knowledge resources over time. Analogously, the MetaboLights database (Yürekten et al., 2023) illustrates how open data repositories in the life sciences have matured into internationally recognized infrastructures for knowledge sharing, demonstrating that openness and rigor are not merely compatible but mutually reinforcing.

These theoretical anchors—open innovation platforms, digital transformation, co-creation frameworks, and community-curated infrastructures—collectively constitute the conceptual architecture within which the following thematic synthesis is situated.

3. Methodology

3.1 Review approach

This study adopts a systematic literature review approach in order to synthesize existing scholarly knowledge on the relationship between open innovation and knowledge dissemination in the digital era. Given that the topic spans several overlapping fields, including open innovation, digital transformation, knowledge management, digital pedagogy, and collaborative digital infrastructures, a systematic approach was selected to provide a transparent and academically grounded procedure for identifying, selecting, and interpreting relevant literature. At the same time, due to the conceptual and interdisciplinary character of the topic, the review is based on qualitative thematic synthesis rather than on statistical meta-analysis.

3.2 Search strategy and selection of sources

The literature selection focused on scholarly publications addressing the intersection of open innovation, digital knowledge sharing, knowledge co-creation, and digitally mediated dissemination practices. Particular attention was paid to studies examining digital platforms, institutional digital transformation, open knowledge infrastructures, and the sociotechnical conditions of knowledge exchange. The final body of literature consisted of seventeen sources selected on the basis of their thematic relevance, conceptual contribution, and analytical usefulness for the objectives of the review. In selecting the sources, preference was given to recent academic publications reflecting current developments in digital transformation, especially in the period associated with the expansion of platform-based collaboration, hybrid knowledge environments, and post-pandemic digitalization.

3.3 Inclusion logic and analytical procedure

The review included studies that provided theoretical, empirical, or review-based insights into the role of digital environments in enabling, reshaping, or constraining knowledge dissemination within open and collaborative innovation contexts. Sources focused exclusively on technical infrastructure without broader relevance to knowledge processes, organizational learning, or innovation dynamics were excluded. After the literature had been selected, all studies were read in detail and comparatively analyzed with attention to their conceptual focus, sectoral context, and principal findings. The synthesis proceeded through iterative thematic coding, which made it possible to identify recurrent patterns and organize the literature into broader analytical clusters. This process resulted in three dominant thematic areas: digital platforms as mechanisms of knowledge co-creation and open innovation, digital transformation of knowledge dissemination in higher education, and the equity, governance, and sociotechnical barriers affecting digitally mediated knowledge ecosystems.

The purpose of the review is not to provide exhaustive bibliometric coverage of all available studies, but rather to develop a structured and critically informed synthesis of relevant literature across disciplinary contexts. Accordingly, the methodology emphasizes thematic coherence, interpretive comparison, and analytical relevance. While this approach necessarily involves a degree of authorial judgment, the use of explicit selection logic and systematic thematic analysis helps ensure consistency and academic rigor in the interpretation of the reviewed literature.

4. Literature Review

4.1 Theme One: Digital Platforms as Engines of Knowledge Co-Creation and Open Innovation

The most directly relevant strand of literature for this review concerns the role of digital platforms in facilitating open innovation and knowledge co-creation. Abbate et al. (2019) occupy a central position in this thematic cluster, demonstrating through a detailed single case study of the Open Innovation Platform Regione Lombardia that digital platforms can successfully support the full cycle of coupled open innovation processes, including the generation, development, and diffusion of collaborative knowledge. The platform in question enabled multi-stakeholder participation, structured knowledge sharing, and iterative co-design, suggesting that the affordances of digital infrastructure can align closely with the process requirements of open innovation.

This finding resonates with broader observations about the Internet of Things (IoT) as a knowledge-enabling technology in domains beyond traditional knowledge industries. Ayaz and colleagues (2019) demonstrate, in the context of smart agriculture, how IoT-based sensor networks can transform data-poor environments into richly informative knowledge systems, effectively creating a new form of open, distributed knowledge infrastructure. While the agricultural context may appear distant from conventional open innovation discourse, the underlying dynamic—whereby networked sensing and data sharing create new possibilities for distributed, evidence-based decision-making—is directly applicable to knowledge dissemination more broadly. The capacity of IoT to make "fields talk," as the authors metaphorically describe it, illustrates a generalizable principle: that digital connectivity transforms previously siloed knowledge into shareable, actionable resources.

Open data repositories represent another significant instantiation of this principle. Williams et al. (2018) document how the Neotoma Paleoecology Database has succeeded in lowering the costs of knowledge management, broadening access to specialized scientific data, and sustaining high-quality curation through distributed community governance. The critical design insight embedded in Neotoma's architecture is that openness must be accompanied by structured governance to prevent quality degradation—a finding that has important implications for the design of open innovation platforms across sectors. Yürekten et al. (2023) reinforce this conclusion in the metabolomics domain, observing that MetaboLights' recognition as a recommended repository by leading journals and European research infrastructure has created positive feedback loops between openness and scientific credibility.

Research on renewable energy systems by Breyer and colleagues (2022) further illustrates how open, collaborative knowledge dissemination within a research community can accelerate the maturation of an entire scientific field. The rapid growth of 100% renewable energy systems research since the mid-2000s is attributable, in significant part, to the cumulative, open exchange of findings, methods, and data across an expanding network of research groups globally. This case underscores that open innovation in knowledge-intensive fields is not merely a platform-design problem but a community-building and norm-setting challenge that requires sustained institutional attention.

Despite this broadly positive picture, tensions and contradictions within this thematic cluster merit acknowledgment. The Abbate et al. (2019) analysis, while richly detailed, is based on a single case study, limiting the generalizability of its conclusions. Furthermore, the platform model it describes is embedded in a specific regional governance context, raising questions

about whether similar structures can be replicated in less institutionally dense or resource-rich settings.

4.2 Theme Two: Digital Transformation of Knowledge Dissemination in Higher Education

Higher education institutions constitute one of the most consequential arenas in which digital transformation and knowledge dissemination intersect. The COVID-19 pandemic served as a dramatic stress test for higher education knowledge systems worldwide, compelling institutions to operationalize digital knowledge delivery at unprecedented scale and speed. Crawford and colleagues (2020) provide a sweeping comparative analysis of intra-period responses across twenty countries, documenting a wide range of institutional reactions from minimal adaptation to comprehensive online curriculum redevelopment. The diversity of these responses reflects not only the heterogeneity of institutional readiness but also the absence of established frameworks for rapid digital knowledge dissemination in crisis conditions.

Bozkurt and colleagues (2020) extend this analysis across thirty-one countries, drawing on the experiences of over sixty percent of the world's population to argue that the pandemic-induced interruption of education has surfaced both the possibilities and the profound inequities embedded in digital knowledge dissemination systems. Their findings reveal that while technology enabled a form of continuity for many learners, the shift to fully online modalities disadvantaged students with limited digital access, creating a troubling divergence between the emancipatory rhetoric of digital openness and the exclusionary realities of digital divide dynamics.

Bond et al. (2018) offer a pre-pandemic baseline for understanding digital transformation in higher education, demonstrating that even within a comparatively well-resourced national system such as Germany, significant heterogeneity exists in students' and teachers' perceptions and uses of digital media. Their data indicate that while digital tools are widely adopted at the surface level, deeper transformation of pedagogical practice—involving genuine shifts in epistemological assumptions and knowledge-sharing norms—remains partial and contested. This finding suggests that digital transformation discourse risks conflating tool adoption with genuine knowledge practice transformation, a distinction of considerable theoretical and practical importance.

The emergence of blended learning as a hybrid knowledge dissemination modality occupies an important position in this thematic cluster. Anthony and colleagues (2020) conduct a theoretical and systematic review of blended learning adoption in higher education, identifying a complex web of factors at the student, lecturer, and administrative levels that shape implementation outcomes. Their findings highlight the importance of alignment between technological affordances and pedagogical intentions, suggesting that effective knowledge dissemination in blended environments requires not only digital infrastructure but also deliberate instructional design and institutional support structures.

Almeida and Simões (2019) contribute a complementary perspective through their analysis of Education 4.0 in Portuguese higher education, documenting how serious games and Industry 4.0 tools can transform passive knowledge reception into active, experiential co-creation. The twenty-five case studies they examine reveal a consistent pattern: that when digital tools are embedded within co-creative pedagogical frameworks rather than deployed as substitutes for conventional instruction, they generate significantly richer knowledge outcomes. This finding is congruent with open innovation principles, insofar as it emphasizes the generative potential of participatory knowledge processes over hierarchical transmission.

Artificial intelligence's emerging role in digital knowledge dissemination warrants specific attention. Almusaed and colleagues (2023) examine the application of AI in hybrid education settings, finding that AI-driven personalization tools can substantially enhance student engagement and adaptive knowledge delivery. However, they also note significant concerns regarding the ethical governance of AI-mediated knowledge systems and the risk that algorithmic personalization may inadvertently reinforce rather than transcend existing knowledge inequities. This tension—between AI's potential to democratize knowledge access and its capacity to entrench structural disparities—represents one of the most consequential unresolved debates in this thematic cluster.

4.3 Theme Three: Equity, Barriers, and the Sociotechnical Dimensions of Open Knowledge Dissemination

A third and critically important thematic cluster concerns the structural and sociotechnical barriers to open, digital knowledge dissemination, particularly in resource-constrained or structurally disadvantaged contexts. Okoye and colleagues (2022) provide a systematic analysis of digital technology adoption in Latin American higher education, documenting a constellation of barriers including inadequate infrastructure, insufficient teacher digital competency, limited institutional support, and entrenched cultural norms that privilege conventional knowledge delivery modalities. Their findings underscore that the mere availability of digital platforms does not translate automatically into open, equitable knowledge dissemination; rather, contextual factors mediate the relationship between technological affordance and knowledge practice in ways that existing theoretical frameworks have insufficiently theorized.

Kukulska-Hulme et al. (2020) identify the concept of "learning in the wild"—knowledge acquisition in informal, technology-enabled contexts beyond formal institutional boundaries—as an emerging practice that both extends and complicates conventional notions of knowledge dissemination. This innovation suggests that the boundaries between formal and informal, institutional and community-based knowledge ecosystems are becoming increasingly porous in digital environments, opening new possibilities for open innovation but also generating governance challenges regarding quality assurance and credentialing.

The circular economy literature provides a cross-disciplinary lens on the challenge of knowledge dissemination to diverse stakeholder groups. Van Langen and colleagues (2021) demonstrate, through a multi-stakeholder perceptions study, that knowledge about complex sociotechnical transitions—such as circular economy adoption—is unevenly distributed and differentially interpreted across researcher, economist, and administrator communities. This finding has direct implications for open innovation and knowledge dissemination more broadly, suggesting that effective knowledge dissemination requires not only the technical infrastructure of open platforms but also active attention to the communicative, interpretive, and cultural dimensions of knowledge reception.

Gilbert and colleagues (2020) contribute a rapid-implementation case study from the healthcare sector, documenting how virtual clinics were rapidly deployed in response to COVID-19. While the healthcare context is specific, the underlying dynamics—rapid digitalization under crisis conditions, quality improvement through iterative feedback, and distributed knowledge production among clinical teams—mirror patterns observable across other knowledge-intensive sectors navigating digital transformation. The case illustrates that open, digitally-mediated knowledge systems can achieve rapid maturation when contextual

pressures are sufficiently intense, but also highlights the importance of structured evaluation frameworks for sustaining quality in accelerated innovation cycles.

Taken together, these three thematic clusters reveal a literature characterized by productive convergence on the enabling potential of digital platforms and open knowledge practices, alongside significant unresolved debate regarding equity, governance, and the sociotechnical conditions under which openness translates into genuine knowledge democratization.

5. Research Gaps and Future Directions

Despite the richness of the literature reviewed, several significant gaps impede a comprehensive understanding of how open innovation and digital knowledge dissemination interact in practice. Identifying these gaps is essential not only for advancing theoretical development but also for informing the design of more effective and equitable knowledge systems.

A first and particularly salient gap concerns the governance of open innovation digital platforms at scale. Abbate et al. (2019) and Williams et al. (2018) each provide detailed accounts of specific platform governance models, but the field lacks comparative studies that systematically evaluate the conditions under which different governance architectures—centralized versus distributed, institutional versus community-led—produce superior outcomes in terms of knowledge quality, accessibility, and innovation impact. Future research should develop comparative governance frameworks applicable across sectors and institutional contexts, drawing on both qualitative case studies and quantitative outcome assessments to build cumulative evidence.

A second gap concerns the equity dimensions of digital knowledge dissemination within open innovation ecosystems. The literature reviewed here, particularly Bozkurt et al. (2020), Okoye et al. (2022), and Bond et al. (2018), consistently documents disparities in digital access, competency, and institutional support that undermine the democratizing potential of open platforms. However, existing studies predominantly diagnose rather than theorize these disparities, and intervention-oriented research capable of generating actionable guidance for reducing digital inequities within open innovation contexts remains sparse. Future research should prioritize longitudinal, mixed-methods studies that examine how targeted policy interventions—such as digital literacy programs, infrastructure investment, and inclusive platform design—can reduce structural barriers to participation in open knowledge ecosystems. This work carries direct practical implications for international development agencies, national education ministries, and platform designers.

A third gap concerns the role of artificial intelligence in mediating open innovation and knowledge dissemination. Almusaed et al. (2023) and Kukulska-Hulme et al. (2020) both gesture toward the transformative potential of AI-enabled knowledge systems, yet the ethical governance of AI-mediated knowledge processes—particularly concerning issues of algorithmic bias, data sovereignty, and knowledge attribution—remains theoretically and empirically underdeveloped. As AI systems become increasingly integrated into open innovation platforms and digital knowledge repositories, future research must develop robust ethical frameworks and empirical evaluation tools capable of assessing whether AI-mediated knowledge dissemination reinforces or disrupts existing power asymmetries in knowledge ecosystems.

A fourth, cross-cutting gap concerns the cross-disciplinary integration of open innovation research. The sources reviewed here span agriculture (Ayaz et al., 2019), healthcare (Gilbert et

al., 2020), energy research (Breyer et al., 2022), environmental science (Williams et al., 2018), and education (Crawford et al., 2020; Bond et al., 2018), yet each disciplinary strand largely develops in relative isolation. Future review and synthesis work should systematically bridge these disciplinary literatures to identify transferable principles and context-specific contingencies that can inform a genuinely transdisciplinary theory of open innovation and digital knowledge dissemination.

6. Conclusion

This literature review has examined the intersection of open innovation and digital knowledge dissemination across a range of disciplinary contexts, synthesizing evidence from seventeen peer-reviewed sources to map the theoretical landscape and identify key thematic clusters, unresolved debates, and directions for future inquiry.

The review's principal contribution is to demonstrate that the relationship between open innovation and digital knowledge dissemination is mediated by three interconnected dynamics. First, digital platforms—when designed with appropriate governance structures and aligned with co-creative epistemological frameworks—can substantially enhance the accessibility, quality, and collaborative potential of knowledge resources, as evidenced by both platform-based studies and open data repository research. Second, digital transformation in higher education represents both an opportunity and a site of contestation, where the promise of open, accessible knowledge is frequently constrained by structural inequities, pedagogical conservatism, and institutional inertia. Third, the equity and governance dimensions of digital knowledge dissemination constitute the most consequential and undertheorized aspects of this domain, demanding urgent theoretical and empirical attention.

The review reaffirms that openness, in both its technological and epistemological senses, is a necessary but insufficient condition for effective knowledge dissemination. Governance, competency, equity, and cultural context are equally determinative of whether open innovation platforms fulfill their transformative potential. Future research must grapple seriously with these structural dimensions if the field is to move beyond techno-optimism toward a more nuanced and actionable understanding of how digital knowledge systems can serve the broader public good.

This review is subject to several limitations. The reliance on seventeen sources, while sufficient for a focused synthesis, necessarily constrains the comprehensiveness of thematic coverage. Additionally, the sources span multiple disciplinary domains, and the integration across these domains involves interpretive choices that introduce an element of analytical subjectivity. Notwithstanding these limitations, the review provides a coherent and evidence-grounded contribution to understanding open innovation and knowledge dissemination in the digital era.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Abbate, T., Codini, A., & Aquilani, B. (2019). Knowledge co-creation in Open Innovation Digital Platforms: Processes, tools and services. Journal of Business & Industrial Marketing, 34(7), 1434–1447. https://doi.org/10.1108/JBIM-09-2018-0276

Almeida, F., & Simões, J. (2019). The role of serious games, gamification and Industry 4.0 tools in the Education 4.0 paradigm. Contemporary Educational Technology, 10(2), 120–136. https://doi.org/10.30935/cet.554469

Almusaed, A., Almssad, A., Yitmen, İ., & Homod, R. Z. (2023). Enhancing student engagement: Harnessing "AIED"'s power in hybrid education—A review analysis. Education Sciences, 13(7), 632. https://doi.org/10.3390/educsci13070632

Anthony, B., Kamaludin, A., Romli, A., Raffei, A. F. M., & Danakor, N. (2020). Blended learning adoption and implementation in higher education: A theoretical and systematic review. Technology, Knowledge and Learning, 27(2), 551–578. https://doi.org/10.1007/s10758-020-09477-z

Ayaz, M., Uddin, M. A., Sharif, Z., Mansour, A., & Aggoune, E.-H. M. (2019). Internet-of-Things (IoT)-based smart agriculture: Toward making the fields talk. IEEE Access, 7, 129551–129583. https://doi.org/10.1109/access.2019.2932609

Bond, M., Marín, V. I., Dolch, C., Bedenlier, S., & Zawacki-Richter, O. (2018). Digital transformation in German higher education: Student and teacher perceptions and usage of digital media. International Journal of Educational Technology in Higher Education, 15(1), 48. https://doi.org/10.1186/s41239-018-0130-1

Bozkurt, A., Jung, I., Xiao, J., Vladimirschi, V., & Schuwer, R. (2020). A global outlook to the interruption of education due to COVID-19 pandemic: Navigating in a time of uncertainty and crisis. Asian Journal of Distance Education, 15(1), 1–126. https://doi.org/10.5281/zenodo.3878572

Breyer, C., Khalili, S., Bogdanov, D., Ram, M., & Aghahosseini, A. (2022). On the history and future of 100% renewable energy systems research. IEEE Access, 10, 78176–78218. https://doi.org/10.1109/access.2022.3193402

Carayannis, E. G., & Morawska, J. (2022). The futures of Europe: Society 5.0 and Industry 5.0 as driving forces of future universities. Journal of the Knowledge Economy, 13(4), 3445–3471. https://doi.org/10.1007/s13132-021-00854-2

Crawford, J., Butler-Henderson, K., Rudolph, J., Malkawi, B. H., Glowatz, M., Burton, R., Magni, P. A., & Lam, S. (2020). COVID-19: 20 countries' higher education intra-period digital pedagogy responses. Journal of Applied Learning & Teaching, 3(1), 1–20. https://doi.org/10.37074/jalt.2020.3.1.7

Gilbert, A., Billany, J. C. T., Adam, R., Martin, L., & Tobin, R. (2020). Rapid implementation of virtual clinics due to COVID-19: Report and early evaluation of a quality improvement initiative. BMJ Open Quality, 9(2), e000985. https://doi.org/10.1136/bmjoq-2020-000985

Kukulska-Hulme, A., Beirne, E., Conole, G., Costello, E., Coughlan, T., Ferguson, R., FitzGerald, E., Gaved, M., Herodotou, C., Holmes, W., Mac Lochlainn, C., Nic Giolla Mhichíl, M., Rienties, B., Sargent, J., Scanlon, E., Sharples, M., & Whitelock, D. (2020). Innovating pedagogy 2020: Open University Innovation Report 8. The Open University. https://doi.org/10.13140/rg.2.2.26971.72480

Okoye, K., Haruna, H., Arrona-Palacios, A., Quintero, H. N., & Sanchez, A. L. (2022). Impact of digital technologies upon teaching and learning in higher education in Latin America: An

outlook on the reach, barriers, and bottlenecks. Education and Information Technologies, 28(2), 2291–2360. https://doi.org/10.1007/s10639-022-11214-1

Panke, S. (2019). Design thinking in education: Perspectives, opportunities and challenges. Open Education Studies, 1(1), 281–306. https://doi.org/10.1515/edu-2019-0022

Van Langen, S. K., Vassillo, C., Ghisellini, P., Restaino, D., & Ulgiati, S. (2021). Promoting circular economy transition: A study about perceptions and awareness by different stakeholder groups. Journal of Cleaner Production, 316, 128166. https://doi.org/10.1016/j.jclepro.2021.128166

Williams, J. W., Grimm, E. C., Blois, J. L., Charles, D. F., & Davis, E. (2018). The Neotoma Paleoecology Database, a multiproxy, international, community-curated data resource. Quaternary Research, 89(1), 156–177. https://doi.org/10.1017/qua.2017.105

Yürekten, Ö., Payne, T., Tejera, N., Amaladoss, F. X., & Steinbeck, C. (2023). MetaboLights: Open data repository for metabolomics. Nucleic Acids Research, 52(D1), D640–D646. https://doi.org/10.1093/nar/gkad1045

Capacity Building for Artificial Intelligence Adoption in Organizations: A Systematic Literature Review

Hojdik Vladimír [1] and Zuzana Hajduová[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: vladimir.hojdik@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: zuzana.hajduova@euba.sk

Abstract

The accelerating integration of artificial intelligence (AI) into organizational processes presents both transformative opportunities and formidable challenges. Despite growing scholarly interest in digital transformation, a coherent understanding of the conditions under which organizations successfully build the capacity required to adopt AI remains elusive. This literature review synthesizes existing research on organizational AI readiness, leadership, strategy, and workforce development to construct a consolidated framework for understanding capacity building in the context of AI adoption. Drawing on fifteen peer-reviewed sources spanning industrial, educational, and entrepreneurial contexts, the review identifies three principal thematic clusters: organizational readiness and cultural prerequisites, strategic leadership and governance frameworks, and human capital development. Key findings indicate that successful AI adoption demands simultaneous investment in technological infrastructure, leadership agility, ethical governance, and workforce upskilling. The review further highlights significant gaps in the literature concerning longitudinal assessments of AI adoption capacity and sector-specific implementation models. The implications of these findings extend to policymakers, organizational managers, and academic researchers seeking to operationalize AI capacity building as a strategic imperative for sustainable organizational transformation.

JEL Classification: O33, M53, J24

Keywords: artificial intelligence adoption, capacity building, organizational readiness, digital transformation, leadership, human capital

1. Introduction

The proliferation of artificial intelligence technologies across industrial, educational, entrepreneurial, and governmental sectors has fundamentally altered the landscape of organizational strategy and management. No longer confined to the domain of computer science, AI has emerged as a systemic force reshaping decision-making processes, operational workflows, competitive dynamics, and the very nature of human work (Chalmers et al., 2021; Peres et al., 2020). Yet the widespread recognition of AI's transformative potential has not been matched by an equally broad organizational capacity to harness it effectively. Organizations across sectors continue to struggle with implementing AI in ways that are scalable, ethical, and aligned with broader strategic goals.

Capacity building, understood as the deliberate process of developing the skills, structures, leadership dispositions, and institutional arrangements necessary to adopt and sustain a new technological paradigm, has become one of the most pressing concerns for scholars and

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

practitioners alike. The literature on digital transformation offers useful reference points, revealing that organizations which master agility, cultivate visionary leadership, and align technological investment with coherent strategy are more likely to realize the benefits of digital innovation (AlNuaimi et al., 2022; Brunetti et al., 2020). However, AI adoption represents a qualitatively distinct challenge that goes beyond earlier waves of digital transformation. The opacity of AI algorithms, the ethical dimensions of automated decision-making, the need for data literacy at all organizational levels, and the disruption posed to existing job roles collectively demand a more nuanced and comprehensive approach to capacity building.

Despite a growing body of work addressing isolated dimensions of AI adoption—including AI readiness assessment (Jöhnk et al., 2021), AI policy frameworks (Roberts et al., 2021), and AI's implications for entrepreneurship (Chalmers et al., 2021)—the literature lacks an integrative synthesis that examines capacity building across these dimensions simultaneously. Existing reviews tend to be either narrowly sector-specific or excessively broad in their treatment of digital transformation, failing to draw the connections necessary to inform actionable organizational strategy. This gap constitutes both the motivation and the justification for the present review.

The purpose of this literature review is to consolidate current scholarly knowledge on how organizations build the capacity required for meaningful AI adoption. More specifically, the paper seeks to explain the main concepts of AI capacity building, describe the dominant thematic patterns in the literature, and identify gaps that should inform future research and organizational practice.

This review proceeds in four steps. It first outlines the conceptual framing of organizational AI capacity and clarifies the analytical perspective adopted in the paper. It then synthesizes the literature across three interrelated dimensions: organizational readiness and cultural prerequisites, strategic leadership and governance frameworks, and human capital development. The following section identifies unresolved issues and promising directions for future research, while the final section offers concluding remarks and reflects on the broader implications of the review.

2. Review Design and Conceptual Framing

This study adopts a structured literature review approach aimed at synthesizing existing research on capacity building for artificial intelligence (AI) adoption in organizations. Relevant publications were identified through searches in major academic databases, including Web of Science, Scopus, and Google Scholar, using keyword combinations such as "AI adoption," "capacity building," "organizational readiness," and "digital transformation." The selection process focused on peer-reviewed studies with clear relevance to organizational, strategic, and human dimensions of AI adoption, while excluding purely technical or non-comparative contributions. The selected literature was analyzed using a thematic synthesis approach, involving the identification of key concepts, grouping studies into thematic clusters, and interpreting relationships and research gaps. Although the review does not strictly follow a formal PRISMA protocol and is limited to English-language sources, it ensures a transparent and analytically consistent synthesis of current knowledge in the field.

2.1 Defining Capacity Building in the Context of AI

Capacity building, as a theoretical construct, originates in organizational development and public administration scholarship, where it refers to the processes by which individuals, organizations, and societies enhance their ability to perform core functions, solve problems,

and achieve objectives sustainably (Martínez-Peláez et al., 2023). In the context of AI adoption, capacity building encompasses a broader set of dimensions than technical skill acquisition alone. It includes the development of data infrastructure, the cultivation of AI-literate leadership, the reconfiguration of organizational culture, and the establishment of governance mechanisms capable of managing the ethical and regulatory complexities inherent to AI deployment.

A foundational distinction must be drawn between AI adoption and AI integration. Adoption, in the narrower sense, refers to the decision to implement AI tools within organizational processes, whereas integration implies a deeper and more transformative embedding of AI capabilities into the strategic and operational fabric of the organization (Jöhnk et al., 2021). Capacity building, as conceptualized in this review, is concerned primarily with the preconditions and processes that enable not merely the adoption but the sustained and strategic integration of AI.

2.2 AI Readiness as a Theoretical Construct

A central theoretical construct in this field is AI readiness, defined by Jöhnk et al. (2021) as the extent to which an organization's assets, capabilities, and commitments are sufficiently developed to support a given AI adoption purpose. This construct acknowledges that readiness is not a binary condition but a multidimensional and context-dependent spectrum. Organizations may be technically ready possessing the necessary data infrastructure and computational resources while remaining organizationally unready due to inadequate leadership, resistant cultures, or insufficient human capital development.

The concept of AI readiness intersects with broader theoretical frameworks in organizational management, including dynamic capabilities theory, which holds that organizations create competitive advantage not through static resources but through the capacity to reconfigure and adapt resources in response to environmental change (Sturgeon, 2019, Mikalef,2019). Within this framework, AI capacity building can be understood as the deliberate cultivation of dynamic capabilities that enable organizations to sense, seize, and reconfigure in response to the AI-driven transformation of their competitive environments.

2.3 Digital Transformation Theories as Contextual Frameworks

AI adoption does not occur in isolation; it is embedded within the broader trajectory of digital transformation. Abad-Segura et al. (2020) document the exponential growth of scholarly interest in digital transformation across sectors, noting its intimate connection with sustainability and institutional change. Digital transformation theory posits that technological adoption fundamentally restructures organizational routines, value chains, and stakeholder relationships - a claim that applies with particular force to AI, given its capacity for autonomous decision-making and continuous learning.

AlNuaimi et al. (2022), drawing on new institutional theory, demonstrate that digital transformational leadership and organizational agility jointly and interactively influence the success of digital transformation initiatives. Their model, which positions digital strategy as a moderating variable, provides a theoretically grounded architecture for understanding how leadership behavior and organizational flexibility translate into effective technology adoption. This framework is especially relevant to AI adoption, where the strategic alignment of human and technological resources is critical.

2.4 The Society 5.0 and Industry 5.0 Frameworks

An emerging theoretical horizon relevant to capacity building for AI adoption is offered by the Society 5.0 and Industry 5.0 paradigms. Carayannis and Morawska (2022) argue that these frameworks transcend the productivity-focused logic of Industry 4.0 by centering human well-being, social responsibility, and sustainability at the core of technological innovation. In organizational terms, this implies that capacity building for AI adoption must not be narrowly conceived as a performance optimization exercise but must incorporate ethical, social, and ecological considerations. This human-centric perspective aligns with the ethical governance dimensions of AI policy, as analyzed in comparative international research (Roberts et al., 2021), and introduces a normative dimension that enriches purely instrumental accounts of AI capacity building.

2.5 Human Capital Theory and AI Adoption

Human capital theory, which conceptualizes workforce skills and knowledge as productive assets that can be deliberately developed through training and education, provides a further theoretical foundation for the present review. Kamalov et al. (2023) argue that the integration of AI into educational and organizational environments requires a fundamental reconsideration of competency frameworks, as AI not only automates routine tasks but demands new forms of creative, critical, and collaborative intelligence from human workers. In this sense, capacity building is inseparable from human capital development, and organizations that invest systematically in AI literacy and related competencies are better positioned to leverage AI for sustained competitive and social value.

3. Core Dimensions of Organizational AI Capacity

3.1 Organizational Readiness and Cultural Prerequisites for AI Adoption

A recurring and foundational theme in the literature concerns the internal organizational conditions that must be present before effective AI adoption can occur. Jöhnk et al. (2021) provide the most systematic treatment of this theme, identifying through in-depth interviews with organizational practitioners a set of AI readiness factors spanning technological assets, data governance, process alignment, and cultural openness. Their findings reveal that organizations frequently underestimate the cultural dimensions of AI readiness, tending to focus on technical infrastructure while neglecting the attitudinal and behavioral changes required at the workforce level. Crucially, Jöhnk et al. (2021) observe that an informed and proactive readiness assessment significantly increases the probability of successful AI adoption, suggesting that self-diagnostic organizational capacity is itself a component of the readiness it seeks to measure.

This concern for cultural readiness is echoed across the digital transformation literature more broadly. Brunetti et al. (2020), in their multi-stakeholder study of digital transformation challenges in the Tyrol-Veneto macroregion, find that organizational resistance to change—manifested as risk aversion, siloed departmental thinking, and inadequate cross-functional collaboration—constitutes one of the most formidable barriers to successful digital adoption. While their study is not AI-specific, the mechanisms they identify are highly applicable to AI adoption contexts, where the opacity of algorithmic systems and the perceived threat of automation can amplify cultural resistance considerably.

Martínez-Peláez et al. (2023) extend this analysis to the context of small and medium enterprises (SMEs), arguing that for organizations with constrained resources, readiness for

digital transformation—including AI adoption—depends critically on the mobilization of stakeholder networks and the cultivation of key organizational capabilities. Their systematic review underscores the importance of relational capital and ecosystem embeddedness as preconditions for sustainable digital adoption, suggesting that capacity building cannot be conceived as a purely internal organizational project but must encompass the management of external partnerships and institutional relationships.

The construction sector provides a further illustration of how organizational culture and technical infrastructure must co-evolve for AI-adjacent technologies to deliver value. Sacks et al. (2020), examining digital twin adoption in the construction industry, demonstrate that the realization of data-centric management systems requires not only investment in sensing and modeling technology but also deep changes in project governance, data ownership norms, and cross-organizational information-sharing practices. Their findings highlight the systemic nature of capacity building challenges, wherein isolated technical investments are unlikely to produce transformative outcomes in the absence of corresponding organizational and cultural change.

3.2 Strategic Leadership, Governance, and AI Policy Frameworks

A second major thematic cluster in the literature addresses the roles of leadership, strategic orientation, and governance in enabling or constraining AI adoption. AlNuaimi et al. (2022) offer a theoretically grounded and empirically tested model demonstrating that digital transformational leadership—characterized by visionary orientation, intellectual stimulation, and change facilitation—positively influences both organizational agility and the overall success of digital transformation. Their finding that organizational agility mediates the relationship between leadership and transformation outcomes is particularly significant for AI adoption, as it suggests that leaders cannot directly determine AI outcomes but must first build the adaptive organizational capacity that makes effective AI integration possible.

Strategic alignment emerges as a critical dimension of this leadership challenge. Sturgeon (2019) argues that in the context of digital economy upgrading, organizations and policymakers must develop coherent strategic frameworks that connect technology investment to broader competitive and developmental goals. Organizations that adopt AI in an ad hoc or reactive manner, without embedding AI initiatives within a larger strategic logic, are likely to generate fragmented and suboptimal outcomes. This argument resonates with the findings of Brunetti et al. (2020), who identify the absence of clear digital strategy as a central challenge for organizations seeking to navigate digital transformation, and with the recommendations of AlNuaimi et al. (2022), who position digital strategy as a moderating variable that amplifies the positive effects of leadership and agility on transformation success.

The governance of AI adoption—encompassing not only internal organizational governance but also regulatory frameworks and policy environments—represents a further dimension of this thematic cluster. Roberts et al. (2021) provide a detailed analysis of China's national AI strategy, examining how policy instruments, ethical guidelines, and regulatory architectures shape the conditions under which organizations develop AI capabilities. Their analysis reveals a governance model that emphasizes state-led coordination, substantial public investment in AI research and infrastructure, and the proactive articulation of ethical norms as instruments of both domestic legitimacy and international competitive positioning. This macro-level governance perspective complements the organizational-level leadership models discussed above, illustrating how national policy environments constitute an important enabling or constraining context for organizational AI capacity building.

Carayannis and Morawska (2022) contribute to this theme by arguing that the Society 5.0 and Industry 5.0 frameworks represent a normative governance vision in which the development of AI and other advanced technologies is deliberately oriented toward human flourishing and social sustainability. In organizational terms, this implies that governance frameworks for AI adoption must incorporate mechanisms for stakeholder participation, impact assessment, and ethical accountability. George and Wooden (2023), examining strategic AI adoption in higher education institutions, similarly emphasize the importance of governance structures that balance technological ambition with institutional accountability, arguing that the development of "smart university" frameworks requires coordinated governance involving academic leadership, administrative management, and external stakeholders.

A notable tension within this thematic cluster concerns the pace and directionality of AI governance. While some authors emphasize the importance of enabling regulatory environments that reduce friction for AI adoption (Sturgeon, 2019; Roberts et al., 2021), others caution that insufficiently regulated AI adoption generates significant ethical risks, particularly in relation to algorithmic bias, data privacy, and the displacement of vulnerable workers (Chalmers et al., 2021; Kamalov et al., 2023). This debate reflects a broader and unresolved tension in the digital transformation literature between the imperatives of competitive dynamism and the requirements of social responsibility—a tension that capacity building frameworks must explicitly address rather than paper over.

3.3 Human Capital Development, Education, and AI Literacy

The third and arguably most practically consequential thematic cluster concerns the development of human capital as a foundational component of AI adoption capacity. The literature is broadly convergent on the proposition that the successful integration of AI into organizational contexts requires not only technical expertise in AI development and deployment but also widespread AI literacy among non-specialist organizational members—including managers, frontline workers, and policymakers—who must interact with, oversee, and adapt to AI-driven systems.

Kamalov et al. (2023) examine this challenge in the context of educational institutions, arguing that AI's penetration of academic environments demands a comprehensive reconfiguration of pedagogical approaches, curriculum design, and assessment frameworks. Their analysis of the ChatGPT phenomenon serves as a heuristic for a broader argument: that AI is not merely a tool to be instrumentally deployed but a force that reorganizes the conditions of knowledge production, skill formation, and institutional authority. For organizations, the implication is that investment in AI literacy must be continuous, adaptive, and embedded in organizational learning processes rather than treated as a discrete, one-time training intervention.

George and Wooden (2023) extend this argument to the strategic management of higher education institutions, proposing that the development of AI capacity within universities requires deliberate investment in faculty development, administrative upskilling, and the creation of institutional knowledge management systems capable of supporting AI-augmented academic and administrative functions. Their framework positions human capital development not as a peripheral concern but as a central strategic priority for organizations seeking to leverage AI effectively.

The entrepreneurial context illuminated by Chalmers et al. (2021) adds further texture to this theme, demonstrating that AI's impact on organizational human capital is simultaneously augmentative and disruptive. On the augmentative side, AI tools can enhance entrepreneurial

cognition, accelerate opportunity identification, and reduce the resource constraints that limit new venture creation. On the disruptive side, AI's capacity to automate tasks previously performed by human workers creates new organizational design challenges, including the need to reconfigure roles, develop hybrid human-AI workflows, and manage the psychological and motivational dimensions of working alongside autonomous systems. Chalmers et al. (2021) argue that these challenges are particularly acute for less-resourced and more traditional entrepreneurial ventures, for whom the transition costs of AI adoption may outweigh the benefits in the short term.

The COVID-19 pandemic, which catalyzed a dramatic acceleration of digital and AI adoption across organizations and educational institutions worldwide, provides an important contextual backdrop for understanding the urgency and the unevenness of AI capacity building challenges. Bozkurt et al. (2020) document the global educational response to pandemic-induced disruption, revealing both the remarkable adaptive capacities that organizations can mobilize under conditions of crisis and the deep inequalities in digital infrastructure, skills, and readiness that determine the distributive consequences of rapid technological adoption. Their findings underscore the point that capacity building for AI adoption is not a level playing field: organizations and systems with prior investments in digital infrastructure, leadership capability, and human capital are substantially better positioned to benefit from AI integration, while those lacking these foundations face compounding disadvantages.

Peres et al. (2020) complement this perspective from the industrial AI angle, noting that the implementation of AI in Industry 4.0 manufacturing contexts requires not only technical investment in cyber-physical systems and predictive analytics but also the development of hybrid competency profiles that combine domain expertise with data science capabilities. Their systematic review identifies workforce development as one of the primary challenges facing organizations seeking to adopt industrial AI, arguing that the scarcity of personnel who combine deep engineering knowledge with AI fluency represents a critical bottleneck to the realization of AI's industrial potential. This finding has broad applicability beyond manufacturing, pointing to a general pattern in which the interdisciplinary human capital required for effective AI adoption is in short supply and must be deliberately cultivated through sustained educational and organizational investment.

4. Open Questions in AI Capacity-Building Research

Despite the substantial and growing body of literature reviewed in the preceding sections, significant gaps remain in the scholarly understanding of capacity building for AI adoption in organizations. Main principal areas of underdevelopment are identified below, each accompanied by specific directions for future research.

The absence of longitudinal research designs represents perhaps the most significant methodological gap in the current literature. The vast majority of existing studies—including the interview-based investigation of Jöhnk et al. (2021) and the systematic reviews of Peres et al. (2020) and Martínez-Peláez et al. (2023)—adopt cross-sectional or retrospective research designs that are ill-suited to capturing the temporal dynamics of capacity building. AI readiness is not a static attribute but a developmental trajectory, shaped by learning, failure, iteration, and environmental feedback. Future research should employ longitudinal methodologies-including panel surveys, organizational ethnographies, and longitudinal case studies-to track how organizations build, lose, and reconfigure AI capacity over time. Such research would generate insights of far greater practical utility for managers and policymakers seeking to design effective capacity building interventions.

Sector-specific capacity building models represent a second underexplored domain. While the existing literature encompasses educational institutions (George & Wooden, 2023; Kamalov et al., 2023), manufacturing environments (Peres et al., 2020), construction firms (Sacks et al., 2020), and entrepreneurial ventures (Chalmers et al., 2021), these sectoral analyses remain largely isolated from one another, with limited cross-sectoral synthesis or comparison. The institutional, regulatory, competitive, and cultural conditions that shape AI capacity building differ substantially across sectors, and a capacity building framework optimized for a large industrial manufacturer is unlikely to translate directly to a small educational institution or a new entrepreneurial venture. Future research should develop and empirically validate sector-specific capacity building models while also identifying cross-cutting principles that apply across organizational types. Comparative research designs that systematically examine AI adoption across sectors would make an especially valuable contribution.

The ethical and social dimensions of AI capacity building constitute a third area warranting substantially greater scholarly attention. While authors such as Roberts et al. (2021) and Carayannis and Morawska (2022) gesture toward the importance of ethical governance in AI adoption, the literature as a whole treats ethical dimensions as peripheral concerns rather than as central components of organizational capacity. Future research should investigate how organizations develop the moral reasoning capabilities, participatory governance structures, and accountability mechanisms required to adopt AI in socially responsible ways. This research agenda should attend particularly to the distributional consequences of AI adoption across workforce segments differentiated by gender, age, educational attainment, and socioeconomic status - a dimension that Bozkurt et al. (2020) touch on in the educational context but that remains systematically underexplored in organizational management scholarship. The practical implication of this gap is that current capacity building frameworks risk generating AI adoption processes that amplify existing organizational inequalities rather than challenging them.

5. Conclusion

Taken together, the reviewed literature indicates that successful AI adoption depends not only on access to advanced technologies, but on the broader organizational capacity to absorb, govern, and scale them. Across the reviewed studies, following interdependent dimensions emerge as especially important: organizational readiness and cultural prerequisites, strategic leadership and governance frameworks, and human capital development. Considered jointly, these dimensions suggest that AI adoption is best understood as a long-term organizational capability-building process rather than as a narrowly technical implementation task.

Across these themes, a consistent finding emerges: effective AI adoption is not primarily a technical problem but a complex organizational, strategic, and human challenge that requires sustained, multidimensional investment. Organizations that invest narrowly in AI tools and infrastructure while neglecting leadership development, cultural change management, ethical governance, and workforce upskilling are likely to realize only a fraction of AI's potential value. Conversely, organizations that approach AI capacity building as a systemic and strategic priority—aligning technological investment with human capital development, leadership cultivation, and governance innovation—are substantially better positioned to derive durable competitive and social benefit from AI integration.

The review is subject to several limitations that should inform the interpretation of its findings. First, the relatively small corpus of sources constrains the comprehensiveness of the thematic synthesis and may introduce selection bias. Second, the sources span multiple sectors and national contexts, making direct comparisons difficult. Third, the preponderance of

conceptual and qualitative studies in the reviewed literature limits the strength of causal inferences that can be drawn about the relationships between specific capacity building interventions and AI adoption outcomes.

Notwithstanding these limitations, the review makes a meaningful contribution by consolidating a fragmented literature, identifying areas of scholarly consensus and debate, and delineating a research agenda that can guide future empirical inquiry and practical intervention in this rapidly evolving domain.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Abad-Segura, E., González-Zamar, M.-D., Infante-Moro, J. C., & Ruipérez García, G. (2020). Sustainable management of digital transformation in higher education: Global research trends. Sustainability, 12(5), 2107. https://doi.org/10.3390/su12052107

AlNuaimi, B. K., Singh, S. K., Ren, S., Budhwar, P., & Vorobyev, D. (2022). Mastering digital transformation: The nexus between leadership, agility, and digital strategy. Journal of Business Research, 145, 636–648. https://doi.org/10.1016/j.jbusres.2022.03.038

Bozkurt, A., Jung, I., Xiao, J., Vladimirschi, V., Schuwer, R., & Egorov, G. (2020). A global outlook to the interruption of education due to COVID-19 pandemic: Navigating in a time of uncertainty and crisis. Asian Journal of Distance Education, 15(1), 1–126. https://doi.org/10.5281/zenodo.3878572

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), 697–722. https://doi.org/10.1108/tqm-12-2019-0309

Carayannis, E. G., & Morawska, J. (2022). The futures of Europe: Society 5.0 and Industry 5.0 as driving forces of future universities. Journal of the Knowledge Economy, 13(4), 3445–3471. https://doi.org/10.1007/s13132-021-00854-2

Chalmers, D., MacKenzie, N., & Carter, S. (2021). Artificial intelligence and entrepreneurship: Implications for venture creation in the Fourth Industrial Revolution. Entrepreneurship Theory and Practice, 45(5), 1028–1053. https://doi.org/10.1177/1042258720934581

George, B., & Wooden, O. S. (2023). Managing the strategic transformation of higher education through artificial intelligence. Administrative Sciences, 13(9), 196. https://doi.org/10.3390/admsci13090196

Jöhnk, J., Weißert, M., & Wyrtki, K. (2021). Ready or not, AI comes—An interview study of organizational AI readiness factors. Business & Information Systems Engineering, 63(1), 5–20. https://doi.org/10.1007/s12599-020-00676-7

Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. Sustainability, 15(16), 12451. https://doi.org/10.3390/su151612451

Martínez-Peláez, R., Ochoa-Brust, A., Rivera Manrique, S. I., & Félix, V. G. (2023). Role of digital transformation for achieving sustainability: Mediated role of stakeholders, key

capabilities, and technology. Sustainability, 15(14), 11221. https://doi.org/10.3390/su151411221

Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. A. (2019). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational performance. Information & Management, 56(7), 103167. https://doi.org/10.1016/j.im.2019.103167

Peres, R. S., Jia, X., Lee, J., Sun, K., Colombo, A. W., & Barata, J. (2020). Industrial artificial intelligence in Industry 4.0: Systematic review, challenges and outlook. IEEE Access, 8, 220121–220139. https://doi.org/10.1109/access.2020.3042874

Roberts, H., Cowls, J., Morley, J., Taddeo, M., Wang, V., & Floridi, L. (2021). The Chinese approach to artificial intelligence: An analysis of policy, ethics, and regulation. AI & Society, 36(1), 59–77. https://doi.org/10.1007/s00146-020-00992-2

Sacks, R., Brilakis, I., Pikas, E., Xie, H., & Girolami, M. (2020). Construction with digital twin information systems. Data-Centric Engineering, 1, e14. https://doi.org/10.1017/dce.2020.16

Sturgeon, T. J. (2019). Upgrading strategies for the digital economy. Global Strategy Journal, 11(1), 34–57. https://doi.org/10.1002/gsj.1364

Measuring Effectiveness of Knowledge Transfer Interventions: Literature Review

Pallérová Diana[1] and Stovíček Filip[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: diana.pallerova@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: filip.stovicek@euba.sk

Abstract

Knowledge transfer interventions constitute a critical mechanism through which organisations, institutions, and health systems translate evidence into practice. Despite their growing prominence across sectors including education, public health, and workplace training, systematic methods for measuring their effectiveness remain inconsistently applied and poorly standardised. This review examines the existing literature on knowledge transfer interventions and the frameworks employed to evaluate their impact, drawing on twenty primary sources spanning health sciences, education, and organisational behaviour. The analysis identifies three principal thematic clusters: evaluation frameworks and models for training transfer, dissemination strategies and their measurable outcomes, and contextual determinants that moderate intervention effectiveness. Key findings indicate that while sophisticated conceptual models such as the Integrated Training Transfer and Effectiveness Model have been proposed, practical application of robust evaluation methodologies lags considerably behind theoretical development. Social support structures, stakeholder engagement, and multi-channel dissemination emerge as central enablers of effective knowledge transfer. The review concludes by identifying critical gaps in longitudinal evaluation design, cross-sectoral comparability, and digital-context measurement, and proposes targeted future research directions to advance the field.

Key Words: knowledge transfer effectiveness, training evaluation, Kirkpatrick model, organisational learning measurement, ROI of training

JEL Classification: O33, M53, D83

1. Introduction

The translation of knowledge into practice represents one of the most persistent and consequential challenges confronting modern organisations and public institutions. Whether in clinical health settings where evidence-based guidelines must be adopted by frontline practitioners, in workplaces where training programmes seek to cultivate durable competencies, or in educational environments undergoing rapid digital transformation, the fundamental question remains constant: how effectively has knowledge been transferred, and how can that effectiveness be reliably measured? This question is not merely academic. The gap between the generation of knowledge and its application in real-world settings carries substantial costs, measured in suboptimal patient outcomes, underperforming workforces, and persistent social inequities (Brownson et al., 2017).

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

Despite decades of research on knowledge management and dissemination, the field lacks consensus on what constitutes an adequate measure of transfer effectiveness. Evaluation frameworks have proliferated, yet they tend to be domain-specific, seldom validated across contexts, and insufficiently attentive to the dynamic, multilevel processes through which knowledge moves from its source to its intended application. The problem is further compounded by the acceleration of knowledge production in the digital era, which demands that transfer mechanisms be not only effective but also timely, scalable, and adaptive (Carayannis & Morawska, 2022).

The onset of the COVID-19 pandemic dramatically exposed the fragility of existing knowledge transfer infrastructures. Educational institutions were compelled to migrate pedagogical processes online within days, producing a natural experiment that revealed both the potential and the profound limitations of digital knowledge transfer (Bozkurt et al., 2020; Sandars et al., 2020). Similarly, health systems scrambled to implement rapidly evolving clinical guidance, underscoring the urgent need for robust dissemination and evaluation mechanisms (Teede et al., 2018a).

Against this backdrop, the present literature review pursues two interconnected objectives. First, it seeks to synthesise the current state of knowledge regarding frameworks and methodologies used to measure the effectiveness of knowledge transfer interventions. Second, it aims to identify the principal research gaps that constrain the development of a more unified, evidence-based approach to evaluation in this domain. By drawing on sources from public health, organisational behaviour, clinical guideline implementation, and educational technology, the review adopts a deliberately cross-disciplinary perspective, reflecting the reality that knowledge transfer challenges transcend any single field.

The remainder of the article is structured as follows. Section 2 establishes the theoretical background by defining core constructs and reviewing foundational models. Section 3 presents the thematic synthesis of the literature across three identified clusters. Section 4 delineates the research gaps and proposes future directions, and Section 5 offers concluding observations including the limitations of this review.

2. Theoretical Background

Understanding the effectiveness of knowledge transfer interventions requires a clear theoretical foundation, as the concept itself spans multiple disciplines and operates across different levels of analysis. Knowledge transfer is not a single, uniform process but rather a complex interaction involving the production, communication, interpretation, and application of knowledge within specific organisational and social contexts. Consequently, evaluating its effectiveness necessitates engagement with theories from organisational learning, training evaluation, dissemination science, and educational psychology. These theoretical perspectives provide the conceptual tools needed to define what constitutes successful transfer and to identify the mechanisms through which knowledge influences behaviour and outcomes.

This section establishes the theoretical grounding for the study by first defining the concept of knowledge transfer and its effectiveness, and then reviewing the principal theoretical models that have informed evaluation approaches in the literature. Together, these foundations

support the analytical framework used in the subsequent methodological and empirical sections of the paper.

2.1 Defining Knowledge Transfer and Its Effectiveness

Knowledge transfer is broadly understood as the process by which knowledge, skills, attitudes, or practices are communicated from a source to a recipient with the intent that the recipient will assimilate, apply, and potentially further disseminate that knowledge (Brownson et al., 2017). This definition encompasses a wide range of activities, from clinical guideline dissemination and workplace training to formal educational instruction and community-level health interventions. Critically, knowledge transfer is not synonymous with knowledge transmission; the latter implies a one-directional flow, whereas transfer implies successful uptake and application at the receiving end.

Effectiveness, in the context of knowledge transfer, is therefore not reducible to exposure or reach. A practitioner who receives a clinical guideline but does not modify their clinical behaviour has been exposed to knowledge but has not been the subject of effective transfer (Teede et al., 2018b). Effectiveness must consequently be operationalised at multiple levels: at the level of awareness and comprehension, at the level of behavioural change, and at the level of measurable outcomes in the practice setting. This multilevel conceptualisation is foundational to virtually all contemporary evaluation frameworks and is consistent with classical learning taxonomies that distinguish between knowledge acquisition, skill application, and transfer of learning to new contexts.

2.2 Foundational Theoretical Models

The dominant theoretical tradition in training effectiveness evaluation traces its origins to Kirkpatrick's four-level model, which distinguishes between learner reactions, learning, behaviour change, and results. While Kirkpatrick's framework has been widely adopted, its limitations particularly its linearity and its failure to account for contextual moderators—have prompted the development of more integrative approaches. Nielsen and Shepherd (2022) propose the Integrated Training Transfer and Effectiveness Model (ITTEM), which synthesises key elements of training transfer and training effectiveness literatures into a dynamic, multilevel framework. The ITTEM explicitly acknowledges that training outcomes are shaped not only by the content and design of an intervention but also by individual characteristics, social support structures, and organisational climate. This represents a significant theoretical advance, insofar as it positions effectiveness measurement as an inherently contextual and relational exercise rather than a simple input-output calculation.

Complementary theoretical grounding is provided by dissemination science, which offers a distinct but related perspective on knowledge transfer (Brownson et al., 2017). Dissemination science foregrounds the role of message characteristics, source credibility, audience attributes, and communication channels in determining whether knowledge reaches and influences its intended recipients. The theoretical frameworks drawn from this tradition, including diffusion of innovations theory and social marketing principles, emphasise that effective knowledge transfer is an active, purposive process requiring strategic design rather than passive distribution.

In the educational domain, transformative learning theory provides an additional lens through which to understand knowledge transfer effectiveness. This theory, invoked in studies of online learning satisfaction during the COVID-19 pandemic, holds that effective learning involves not merely the acquisition of new information but a fundamental restructuring of the

learner's frames of reference (Almusharraf & Khahro, 2020). From this perspective, measuring effectiveness requires attending to whether learners' underlying assumptions and practices have been durably altered.

Design thinking, as a framework for educational and organisational problem-solving, has also emerged as a relevant conceptual apparatus for understanding how knowledge transfer interventions are designed and evaluated (Panke, 2019). Design thinking emphasises iterative prototyping, user-centred inquiry, and empirical testing, all of which are directly applicable to the development of effective and measurable knowledge transfer mechanisms.

Finally, the concept of knowledge economy institutions elaborated by Carayannis and Morawska (2022) provides a macro-level theoretical frame, situating individual knowledge transfer interventions within broader systemic transformations driven by Society 5.0 and Industry 5.0 imperatives. This perspective highlights the importance of aligning measurement frameworks with the evolving technological and social contexts in which knowledge transfer occurs, particularly as artificial intelligence and digital platforms become dominant channels of knowledge dissemination (Ahmad et al., 2022).

3. Methodology

This study employs literature review approach to examine how the effectiveness of knowledge transfer interventions is conceptualised and measured across organisational, educational, and health-related contexts. Given the interdisciplinary breadth of the topic, the review was designed as a qualitative and interpretive synthesis aimed at identifying dominant approaches, recurring themes, and key research gaps in the literature.

The literature search was conducted in the academic databases Web of Science, Scopus, and Google Scholar using keyword combinations such as knowledge transfer effectiveness, training evaluation, training transfer, dissemination, Kirkpatrick model, organisational learning measurement, and ROI of training. To broaden the scope of relevant sources, backward reference checking was also applied to selected key studies.

The review process followed a three-stage procedure: identification and selection of relevant literature, analytical classification of studies, and thematic synthesis of findings. This approach enabled the inclusion of heterogeneous sources from different disciplinary traditions while maintaining a coherent analytical focus on the evaluation of knowledge transfer interventions.

The review included peer-reviewed studies published in English between 2017 and 2025. Studies were considered relevant if they addressed frameworks or models for evaluating knowledge transfer and training effectiveness, dissemination strategies and their measurable outcomes, or contextual factors influencing transfer effectiveness. Studies focusing exclusively on technical knowledge management systems without an evaluative perspective, or on learning outcomes without clear relevance to transfer into practice, were excluded. On this basis, a final set of 20 primary sources published between 2017 and 2025 was selected for analysis.

The selected studies were analysed using thematic synthesis. Recurring concepts, methodological approaches, and empirical findings were compared in order to identify dominant patterns in the literature. This process resulted in three thematic clusters: evaluation frameworks and models for training transfer and effectiveness, dissemination strategies and their measurable outcomes, and contextual determinants moderating intervention effectiveness.

4. Literature Review

This section synthesises the main insights derived from the reviewed literature by organising the discussion around three overarching thematic areas. These themes represent the key ways in which previous research has approached the effectiveness of knowledge transfer interventions, with particular attention to evaluation models, dissemination outcomes, and the contextual conditions influencing their success.

4.1 Evaluation Frameworks for Training Transfer and Effectiveness

A central preoccupation of the knowledge transfer literature is the development of conceptual models capable of capturing the complexity of the transfer process and guiding its measurement. Nielsen and Shepherd (2022) make a foundational contribution in this respect by presenting the ITTEM, designed specifically to evaluate whether mental health and wellbeing training interventions are effectively transferred to the workplace and result in meaningful changes in employees' emotions, cognitions, and behaviours. What distinguishes the ITTEM from earlier models is its dynamic character; it does not treat transfer as a binary event but as an ongoing process mediated by a constellation of individual, social, and organisational variables. The model integrates dominant training transfer models with effectiveness evaluation frameworks, thereby enabling researchers to assess not only whether learning occurred but whether it was applied and sustained over time.

Mehner et al. (2023) provide important empirical texture to these theoretical developments through a mixed-method study involving 541 employees who had undergone workplace training within the preceding two years. Their findings reveal that both training transfer and knowledge sharing exert significant and independent effects on individual organisational performance outcomes. Notably, they identify social support—understood as collegial encouragement, supervisory feedback, and organisational recognition—as a critical moderating variable that amplifies the likelihood of both transfer and sharing. This finding has direct implications for effectiveness measurement, as it suggests that evaluating a training intervention solely on the basis of individual learning gains would substantially underestimate its true impact when social support structures are functioning well, and overestimate it when they are absent.

The practical implications of these frameworks are substantial. If effectiveness measurement is confined to immediate post-intervention assessments of knowledge acquisition, as is common in many organisational training evaluations, the most consequential outcomes—behavioural change and performance improvement—will remain unmeasured. This methodological limitation is widely acknowledged in the literature, yet it persists in practice, partly because longitudinal evaluation designs are resource-intensive and partly because organisations frequently lack the measurement infrastructure required to track long-term outcomes (Nielsen & Shepherd, 2022; Mehner et al., 2023).

There is, moreover, an emerging consensus that effectiveness frameworks must be sensitive to the nature of the knowledge being transferred. Complex, tacit knowledge—such as clinical judgment or interpersonal skill—requires different transfer mechanisms and different evaluation methodologies than explicit, codified knowledge. The evaluation frameworks currently available tend to be better suited to the latter than the former, representing a significant gap in the field.

4.2 Dissemination Strategies and Their Measurable Outcomes

A second major thematic cluster concerns the strategies employed to disseminate knowledge and the challenge of measuring their outcomes. Brownson et al. (2017) provide a comprehensive and influential treatment of this theme in the public health context, arguing that the gap between knowledge discovery and its application in practice settings is attributable in substantial part to ineffective dissemination. Their analysis draws on lessons from communication science, agricultural extension, social marketing, and political science to argue that effective dissemination requires careful attention to the characteristics of the message, the source, the audience, and the channel. Crucially, they distinguish between dissemination—the active, strategic communication of evidence to a target audience—and the passive expectation that knowledge will diffuse naturally once published.

The clinical guideline literature offers compelling case studies of structured, multifaceted dissemination programmes and the challenges of evaluating their impact. Teede et al. (2018a, 2018b) describe the development and dissemination of the international evidence-based guideline for polycystic ovary syndrome (PCOS) management, which incorporated a comprehensive multifaceted dissemination and translation programme supported by an integrated evaluation component. This model, subsequently updated in the 2023 edition (Teede et al., 2023), illustrates the aspirational standard for guideline dissemination: clear advice grounded in the best available evidence, delivered through multiple channels, with explicit mechanisms for monitoring uptake and impact. The involvement of expert multidisciplinary groups and attention to consumer preferences in the guideline development process reflects a participatory approach that is increasingly recognised as a prerequisite for effective knowledge transfer in health contexts.

Saul et al. (2018) provide a further example from the public health domain, describing the DREAMS core package of interventions targeting HIV prevention among adolescent girls and young women in sub-Saharan Africa. The DREAMS programme exemplifies a multi-component knowledge transfer initiative in which the effectiveness of individual intervention elements must be disentangled from the effects of the package as a whole. This poses a distinctive measurement challenge: conventional evaluation designs are ill-equipped to attribute outcomes to specific knowledge transfer components within a complex, multi-layered programme.

The measurement of dissemination outcomes has been further complicated by the proliferation of digital communication channels. Brownson et al. (2017) discuss the availability of digital tools for tracking dissemination, including web analytics, social media metrics, and citation counts, but note that these tools measure reach rather than uptake or impact. A document that is widely downloaded is not necessarily one that changes practice. This distinction between reach, uptake, and impact constitutes one of the most fundamental and unresolved methodological challenges in the field. The rapid shift to virtual care delivery prompted by the COVID-19 pandemic, documented for example in the context of virtual clinics (Gilbert et al., 2020), has added further urgency to this challenge, as digital dissemination has become the primary mode of knowledge transfer across numerous sectors simultaneously.

4.3 Contextual Determinants Moderating Knowledge Transfer Effectiveness

The third thematic cluster addresses the contextual factors that moderate whether knowledge transfer interventions achieve their intended effects. This literature converges on the conclusion that effectiveness is not an intrinsic property of an intervention but is profoundly shaped by the environment in which it is deployed. Several categories of contextual

determinants have been identified, including organisational culture, stakeholder engagement, technological infrastructure, and the nature of learner characteristics.

The COVID-19 pandemic has provided an involuntary stress test for knowledge transfer systems globally, generating a substantial body of evidence on the contextual determinants of effectiveness in crisis conditions. Bozkurt et al. (2020) document the global response to the interruption of education due to the pandemic across 31 countries, revealing dramatic variation in the capacity of educational institutions to maintain effective knowledge transfer during the transition to remote learning. Their analysis highlights the role of pre-existing technological infrastructure, institutional agility, and learner digital literacy as critical contextual moderators. Institutions with established educational technology cultures navigated the transition more effectively, while those lacking such infrastructure experienced significant disruption to knowledge transfer continuity.

Almusharraf and Khahro (2020) contribute evidence from the higher education context in Saudi Arabia, demonstrating that student satisfaction with online learning experiences—itself a proxy measure for perceived knowledge transfer effectiveness—was significantly influenced by platform usability, instructor responsiveness, and the degree to which online environments replicated the interactive dimensions of face-to-face learning. Korkmaz and Toraman (2020) extend this analysis by examining educator perspectives on online learning practices during the pandemic, finding that educators identified substantial barriers to effective knowledge transfer in the online environment, including inadequate preparation time, limited familiarity with digital pedagogical approaches, and the loss of non-verbal communicative cues that facilitate real-time assessment of learner comprehension.

Sandars et al. (2020) offer practical guidance on rapid migration to online learning in health professions education, acknowledging that contextual readiness varies dramatically and that effectiveness measurement must account for these baseline differences. Their twelve-tip framework implicitly operationalises effectiveness in terms of learner engagement, instructional coherence, and the preservation of assessment validity—dimensions that are directly relevant to any evaluation of knowledge transfer in digital contexts.

At the organisational level, van Langen et al. (2021) demonstrate the importance of stakeholder perceptions and awareness in shaping the uptake of new knowledge frameworks, drawing on evidence from circular economy transition. Their study illustrates that even when high-quality knowledge is available and actively disseminated, its transfer into practice is contingent on stakeholder receptivity, which itself varies systematically across groups defined by professional role, institutional affiliation, and prior knowledge exposure. This finding resonates with the broader knowledge transfer literature, which has consistently identified stakeholder engagement as a central determinant of intervention effectiveness (Brownson et al., 2017; Nielsen & Shepherd, 2022).

The role of artificial intelligence in shaping the context of knowledge transfer is an emerging theme of considerable importance. Ahmad et al. (2022) document the growing penetration of AI applications in educational settings, arguing that AI tools are transforming both the delivery and the administration of knowledge transfer activities. The implications for effectiveness measurement are significant: AI-enabled platforms generate rich behavioural data that could, in principle, support more granular and timely assessment of transfer outcomes than traditional evaluation methods allow. However, the trustworthiness and deployability of AI-based knowledge transfer and evaluation tools remain areas of active concern, as reflected in the FUTURE-AI framework developed by Lekadir et al. (2025) for trustworthy AI deployment in

healthcare contexts. The alignment of AI capabilities with ethical and methodological standards for evaluation represents a frontier challenge for the field.

5. Research Gaps and Future Directions

Notwithstanding the substantial body of literature reviewed above, several significant gaps constrain the development of a coherent, evidence-based approach to measuring knowledge transfer effectiveness.

The most fundamental gap concerns longitudinal evaluation design. The vast majority of studies examining knowledge transfer effectiveness rely on cross-sectional or short-term pre-post designs that are poorly suited to capturing the durable behavioural and organisational changes that represent the most meaningful outcomes of transfer (Nielsen & Shepherd, 2022). Training that produces measurable learning gains at immediate assessment may fail to influence practice weeks or months later, while interventions that appear modestly effective in the short term may produce cumulative impacts that only become visible with sustained observation. Future research should prioritise longitudinal study designs, incorporating repeated measurement across multiple time points and employing statistical methods capable of modelling the dynamic trajectories of knowledge application over time.

A second gap concerns the cross-sectoral comparability of evaluation frameworks. The literature reviewed in this article spans public health, clinical practice, workplace training, and formal education, yet the evaluation models developed within each domain remain largely insular. The ITTEM proposed by Nielsen and Shepherd (2022) was developed specifically for mental health training in workplace settings; the dissemination science framework articulated by Brownson et al. (2017) was primarily elaborated in the context of public health; the guideline implementation models described by Teede et al. (2018a, 2023) are oriented toward clinical practice. There is a need for overarching frameworks capable of accommodating the distinctive features of knowledge transfer across these varied contexts while enabling meaningful comparison. Future research should engage in systematic cross-domain synthesis and validation of evaluation frameworks, potentially drawing on implementation science as an integrative discipline.

A third gap concerns measurement in digital and AI-mediated knowledge transfer contexts. The rapid expansion of digital platforms as primary channels for knowledge transfer has not been matched by the development of appropriate evaluation methodologies. Existing frameworks were largely designed for face-to-face or hybrid contexts and require significant adaptation to capture the distinctive dynamics of online learning environments, including asynchronous engagement, automated feedback, and algorithm-driven content personalisation (Ahmad et al., 2022; Lekadir et al., 2025). Future research should develop and validate measurement instruments specifically designed for digital knowledge transfer contexts, leveraging the behavioural trace data generated by digital platforms while addressing the associated ethical and methodological challenges.

Finally, the role of social and relational factors in moderating transfer effectiveness merits more systematic investigation. Mehner et al. (2023) and Nielsen and Shepherd (2022) both highlight the importance of social support, yet the mechanisms through which social relationships facilitate or impede knowledge transfer remain insufficiently theorised and empirically documented. Future research should employ social network analysis and qualitative methods to illuminate these mechanisms, with direct implications for the design of interventions that build relational infrastructure alongside content-based knowledge transfer.

6. Conclusion

This literature review has examined the current state of research on measuring the effectiveness of knowledge transfer interventions, drawing on sources from clinical guideline implementation, public health dissemination, workplace training, and educational technology. The review has identified three principal thematic clusters: the development and application of evaluation frameworks for training transfer and effectiveness, the design and measurement of dissemination strategies, and the contextual determinants that moderate intervention outcomes.

Across these themes, the literature converges on several key conclusions. Effective knowledge transfer is a multilevel, dynamic process that cannot be adequately captured by single-point, outcome-only evaluation designs. Social support structures, stakeholder engagement, and channel selection are critical moderating variables whose influence must be incorporated into evaluation frameworks. The disruption wrought by the COVID-19 pandemic has both exposed the vulnerabilities of existing knowledge transfer systems and generated novel evidence on the contextual conditions that support or undermine effective transfer in digital environments. The emergence of AI-enabled knowledge transfer tools presents both significant opportunities for richer, more granular evaluation and important ethical and methodological challenges that the field has yet to fully address.

This review is not without limitations. The 20 sources consulted, while spanning multiple disciplines, do not constitute an exhaustive mapping of the literature, and the absence of a formal systematic search protocol means that relevant works may have been omitted. Several of the included sources address knowledge transfer only tangentially, and the cross-domain synthesis attempted here necessarily glosses over domain-specific nuances. Nevertheless, the review offers a valuable cross-sectoral synthesis that identifies gaps and directions with broad applicability. Future scholarship should pursue the longitudinal, cross-sectoral, and digitally-focused research directions identified herein, with the ultimate aim of developing robust, transferable frameworks for the evaluation of knowledge transfer interventions across the full range of contexts in which they occur.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Ahmad, S. F., Alam, M. M., Rahmat, M. K., & Shujaat, M. (2022). Academic and administrative role of artificial intelligence in education. Sustainability, 14(3), 1101. https://doi.org/10.3390/su14031101

Almusharraf, N., & Khahro, S. H. (2020). Students' satisfaction with online learning experiences during the COVID-19 pandemic. International Journal of Emerging Technologies in Learning, 15(21), 246–267. https://doi.org/10.3991/ijet.v15i21.15647

Bozkurt, A., Jung, I., Xiao, J., Vladimirschi, V., Schuwer, R., Egorov, G., Lambert, S. R., Al-Freih, M., Pete, J., Olcott, D., Jr., Rodes, V., Aranciaga, I., Bali, M., Alvarez, A. V., Roberts, J., Pazurek, A., Raffaghelli, J. E., Panagiotou, N., de Coëtlogon, P., … Paskevicius, M. (2020). A global outlook to the interruption of education due to COVID-19 pandemic: Navigating in a time of uncertainty

and crisis. Asian Journal of Distance Education, 15(1), 1–126. https://doi.org/10.5281/zenodo.3878572

Brownson, R. C., Eyler, A. A., Harris, J. K., Moore, J. B., & Tabak, R. G. (2018). Getting the word out: New approaches for disseminating public health science. Journal of Public Health Management and Practice, 24(2), 102–111. https://doi.org/10.1097/PHH.0000000000000673

Carayannis, E. G., & Morawska-Jancelewicz, J. (2022). The futures of Europe: Society 5.0 and Industry 5.0 as driving forces of future universities. Journal of the Knowledge Economy, 13(4), 3445–3471. https://doi.org/10.1007/s13132-021-00854-2

Gilbert, A. W., Billany, J. C. T., Adam, R., Martin, L., Tobin, R., Bagdai, S., Galvin, N., Farr, I., Allain, A., Davies, L., & Bateson, J. (2020). Rapid implementation of virtual clinics due to COVID-19: Report and early evaluation of a quality improvement initiative. BMJ Open Quality, 9(2), e000985. https://doi.org/10.1136/bmjoq-2020-000985

Korkmaz, G., & Toraman, Ç. (2020). Are we ready for the post-COVID-19 educational practice? An investigation into what educators think as to online learning. International Journal of Technology in Education and Science, 4(4), 293–309. https://doi.org/10.46328/ijtes.v4i4.110

Lekadir, K., Frangi, A. F., Porras, A. R., Glocker, B., Cintas, C., & FUTURE-AI Consortium. (2025). FUTURE-AI: International consensus guideline for trustworthy and deployable artificial intelligence in healthcare. BMJ, 388, e081554. https://doi.org/10.1136/bmj-2024-081554

Mehner, L., Rothenbusch, S., & Kauffeld, S. (2024). How to maximize the impact of workplace training: A mixed-method analysis of social support, training transfer and knowledge sharing. European Journal of Work and Organizational Psychology. Advance online publication. https://doi.org/10.1080/1359432X.2024.2319082

Nielsen, K., & Shepherd, R. (2022). Understanding the outcomes of training to improve employee mental health: A novel framework for training transfer and effectiveness evaluation. Work & Stress, 36(3), 208–228. https://doi.org/10.1080/02678373.2022.2028318

Panke, S. (2019). Design thinking in education: Perspectives, opportunities and challenges. Open Education Studies, 1(1), 281–306. https://doi.org/10.1515/edu-2019-0022

Sandars, J., Correia, R., Dankbaar, M., de Jong, P., Goh, P. S., Hege, I., Masters, K., Oh, S. Y., Patel, R., Premkumar, K., Webb, A., & Pusic, M. (2020). Twelve tips for rapidly migrating to online learning during the COVID-19 pandemic. MedEdPublish, 9, 82. https://doi.org/10.15694/mep.2020.000082.1

Saul, J., Bachman, G., Allen, S., Toiv, N., Cooney, C., & Beamon, T. A. (2018). The DREAMS core package of interventions: A comprehensive approach to preventing HIV among adolescent girls and young women. PLOS ONE, 13(12), e0208167. https://doi.org/10.1371/journal.pone.0208167

Teede, H. J., Misso, M. L., Costello, M. F., Dokras, A., Laven, J., Moran, L., Piltonen, T., & Norman, R. J. (2018a). Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. Clinical Endocrinology, 89(3), 251–268. https://doi.org/10.1111/cen.13795

Teede, H. J., Misso, M. L., Costello, M. F., Dokras, A., Laven, J., Moran, L., Piltonen, T., & Norman, R. J. (2018b). Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. Human Reproduction, 33(9), 1602–1618. https://doi.org/10.1093/humrep/dey256

Teede, H. J., Tay, C. T., Laven, J. J. E., Dokras, A., Moran, L., Piltonen, T., Norman, R. J., & International PCOS Network. (2023). Recommendations from the 2023 international evidence-based guideline for the assessment and management of polycystic ovary syndrome. The Journal of Clinical Endocrinology & Metabolism, 108(10), 2447–2469. https://doi.org/10.1210/clinem/dgad463

Van Langen, S. K., Vassillo, C., Ghisellini, P., Restaino, D., Passaro, R., & Ulgiati, S. (2021). Promoting circular economy transition: A study about perceptions and awareness by different stakeholder groups. Journal of Cleaner Production, 316, 128166. https://doi.org/10.1016/j.jclepro.2021.128166

Evaluation of AI Applications in Logistics from the Perspective of ESG Criteria: A Scoping Review

Otmanns Laura[1] and Hanák Róbert[2]

[1]Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: laura.oltmanns@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: robert.hanak@euba.sk

Abstract

The use of artificial intelligence (AI) in the logistics industry is growing rapidly, driven by several factors. One of key factors is raising expectations for efficiency and increasing regulatory requirements in sustainability. While existing research has primarily focused on technological developments and operational efficiency, the impact of AI on sustainability aspects has not yet been sufficiently explored and scientific effort is fragmented. The aim of this study is therefore to evaluate selected AI applications in logistics from the perspective of ESG (Environmental, Social, Governance) criteria, with a particular focus on AI-supported route optimization and forecasting in warehouse management. The analysis is based on a structured review of scientific publications from 2020 to 2025. . The results show that environmental impacts are addressed most comprehensively in the literature, with particular emphasis on the potential to reduce emissions, energy consumption, and resource use through optimized transport and warehouse processes. Social impacts, such as changes in working conditions or skill requirements, are, however, only considered marginally. In the area of governance, the focus is primarily on challenges related to data quality, transparency of algorithmic decisions, and regulatory requirements. Overall, the study shows that AI offers significant potential for supporting sustainable logistics. At the same time, a lack of holistic ESG assessments becomes apparent, highlighting the need for further research and integrative approaches for practical implementation.

Key Words: ESG, Logistics, AI

JEL Classification: Q, R4, C61, L86, D83

1. Introduction

The logistics sector, which represents an important part of the global economy, faces a wide range of challenges and conflicting objectives. Logistics companies are required to continuously optimize their processes in favor of quality while keeping their internal cost structures as low as possible. Artificial intelligence (AI), as one of the most important technological developments of recent years, promises cost savings, efficiency gains, and a reduction in employee workload for the logistics sector (Fatorachian, 2024; Mojumder & Nuruzzaman, 2025). This also includes, to some extent, the fulfillment of regulatory requirements to which the industry is subject. Global climate targets and the resulting local regulations further increase the pressure on the logistics sector. Therefore, sustainability, understood as the interplay of ESG criteria, will remain at the top of the agenda in the logistics industry, although local regulations currently need to be complied with to varying degrees

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

(Sattar et al., 2025).Research activities on the use of AI in logistics are increasing; however, from a scientific perspective, it is still insufficiently documented how the use of AI in logistics information systems affects sustainability. Existing studies are technology-oriented and primarily explain efficiency gains from a process perspective.

The aim of this article is therefore to analyze two selected AI application areas in logistics regarding their environmental, social, and governance-related impacts. The respective effects are presented in a structured manner along the ESG dimensions to make existing research gaps visible and to highlight the need for further research.

2. Theoretical Background

As in many other industries, the use of AI in the logistics sector is understood as part of the broader digitalization process. The literature also shows consensus that digitalization supports the achievement of ESG goals but requires balanced management. Digital transformation therefore requires strategic planning and operational discipline to effectively integrate ESG requirements.

2.1 The Use of AI in Logistics Information Systems

The central technological infrastructure components of modern logistics companies are their information systems (IS). They enable internal transparency, synchronization, and automation, as well as standardized communication between suppliers, logistics sites, and customers. All logistics processes can be fully mapped within an IS system: from order entry to storage and picking, through to shipping and transportation. The typical types of IS systems are:

- WMS (Warehouse Management Systems): Systems supporting all warehouse processes, including inventory management, order processing, reporting, and documentation, as well as electronic communication with partners.
- TMS (Transport Management Systems): These systems provide digital support throughout the entire goods transportation process. They assist in the optimization, execution, and monitoring of transport activities.
- ERP systems (Enterprise Resource Planning): These systems integrate multiple business areas (including logistics) into a higher-level platform. This enables cross-process data integration and resource planning.
- SCM systems (Supply Chain Management): These systems enable the planning, monitoring, and control of the entire supply chain—from raw material procurement to delivery—across company boundaries (Dyczkowska et al., 2025; Kihél, 2022).

AI can be integrated into all the systems mentioned above and is used in logistics in various forms. At the same time the AI paradigms most frequently described in the literature are:

- Machine Learning (ML): Supervised and unsupervised learning methods capable of identifying patterns in data and generating predictions based on them. It processes

smaller datasets and uses a limited number of variables. In logistics, it is used for demand forecasting or delivery time estimation (Ran, 2021).

- Deep Learning (DL): A subset of ML that can process large datasets, learn autonomously in a more complex manner, and can therefore achieve higher accuracy. It is frequently used in logistics to analyze sensor data or deviations in transport and warehouse processes (Sattar et al., 2025).
- Reinforcement Learning (RL): An adaptive learning approach in which optimal solutions are learned through reward or penalty signals. It is particularly suitable for real-time environments and is therefore used in logistics for route planning or fleet management (Durlik et al., 2024).
- Predictive Analytics (PA): A data-driven analytical method that generates forecasts for the future based on historical and current data. It is used in logistics, for example, to predict material flows or maintenance requirements (Paul et al., 2024).
- Natural Language Processing (NLP): A subfield of AI that uses algorithms and statistical methods or neural models to understand and interpret human language and generate meaningful outputs. It is used, for example, in corporate chatbots, text analysis, or automated communication (Toorajipour et al., 2021).

The prerequisites for integrating AI into an existing IS system depend on the technical requirements of the specific AI application. However, in general, it can be stated that a solid database must already be available within the relevant process, as the output of AI is only as good as the data it is provided with. A reduction in data quality of 5% can already lead to a reduction in model performance of 15–20%. High data quality strongly depends on temporal consistency. For example, communication between different companies may involve different time zones or delayed timestamps for goods receipt, which can impair AI performance and lead to inaccurate results. In addition, it can always be assumed in practice that heterogeneity exists between companies. If the data of different companies vary in format, content, or accuracy, the work of AI becomes even more complex. The logistics industry also exhibits a particular characteristic: the terminology used varies significantly and is defined differently depending on the company. For correct AI interpretation, such attributes must be standardized or additionally explained to the AI (Zhu et al., 2025).

2.2 The use of AI in route optimization

After outlining how AI can be integrated into logistics information systems, a closer examination of specific applications in different areas of logistics is provided. Route optimization is one of the most frequently described AI applications in logistics literature. These systems use real-time data such as traffic volume, capacities, delivery priorities, or weather conditions to minimize transport routes, travel times, and resource utilization. The objective of their use is to increase delivery reliability and customer satisfaction (Ogorodnyk et al., 2025; Necula & Rieder, 2025). In addition, the use of AI enables demand forecasting based on historical and real-time data, as well as the dynamic adjustment of routes in real time. A holistic application is achieved when fleet and energy management, as well as IoT sensor technologies, are integrated into the decision-making processes. AI is not limited to optimizing individual transport segments, such as last-mile or cold chain logistics, but can also comprehensively

control and optimize complex transport chains such as intermodal transport (Durlik et al., 2024).

Regarding the first ESG criterion, namely Environmental (E), the use of AI in route optimization has been shown to reduce pollution and resource consumption. For example, one study reports emission and fuel reductions of 15–30% when AI is implemented as part of a holistic digital infrastructure, for instance when AI support is applied jointly in route optimization and fleet management (Sarioguz, 2025). In a study conducted in an urban logistics environment, the introduction of AI-supported route optimization resulted in an average reduction of 13.9% in fuel consumption and 17.3% in energy consumption (Mohsen & Mohsen, 2025). Another study from an urban context reports a reduction in delivery times of 15–22%. Particularly significant efficiency gains were achieved through the combined use of electric vehicles and AI-based route optimization and energy management, with up to 40% emission reduction, 10–25% improvements in energy efficiency, and 15–20% shorter delivery times (Ferreira & Esperança, 2024). The environmental benefits reported in these studies depend on the overall context, such as the use of electric vehicles or combustion engines, or whether the study focuses on an urban or more rural environment. Existing digital and infrastructural systems may also negatively affect the overall outcome. At the same time, it can be argued that maximum resource savings can primarily be achieved through holistic, environmentally oriented strategies (Sarioguz, 2025; Felder et al., 2023; Ferreira & Esperança, 2024). However, AI-supported route optimization does not only provide environmental benefits but also entails challenges that cannot be solved by software alone. Regional infrastructure conditions such as road usage, lack of parking space, loading structures, or insufficient consolidation points remain outside the solution scope (Ferreira & Esperança, 2024; Popoola et al., 2025). Likewise, the sustainable design of the entire supply chain, such as intermodal transport, depends on policy and infrastructure conditions (Sarioguz, 2025).

For the social ESG dimension (S), no explicit advantages, or disadvantages of AI-supported route optimization in logistics were identified in the analyzed literature. The focus of existing research lies clearly on the environmental dimension (E), which is examined in much greater detail. It is conceivable that the reduction in delivery time has a positive effect on drivers' workload or stress levels; however, there are no direct studies on this. Changes in work that were previously performed manually and are now automated using AI are, if addressed at all, only mentioned marginally in most articles. It can therefore only be assumed that the shift from manual to automated work represents a positive change for the affected employees. Similarly, it can be assumed that the previously mentioned reduction in emissions may lead to improved air quality and increased satisfaction among affected populations (Oršić et al., 2022). The increase in customer satisfaction due to higher delivery punctuality is one of the main reasons why companies invest in AI-supported route optimization. Therefore, improved customer satisfaction can be mentioned at this point. However, it must be considered that the optimal route calculated by AI does not always correspond to customer preferences, which can lead to decision-making challenges in dispatching (Sarioguz, 2025; Leogrande, 2024; Liu, Hou & Cai, 2025; Veluru, 2023; Oršić et al., 2022).

In the analyzed studies, the governance dimension (G) repeatedly highlights that AI-based route optimization can only be considered sustainable and compliant if transparency in AI decision-making processes, a sufficient degree of standardization, and clear institutional and regulatory frameworks are ensured. In particular, the traceability of algorithmic decisions is identified as a key prerequisite for building trust among companies, employees, and other stakeholders (Mohsen, 2024; Popoola et al., 2025). Furthermore, it is noted that AI-supported route optimization can also be used to analyze and monitor driving behavior. By identifying

risky driving patterns, there is potential to reduce accident risks and thus improve both traffic safety and operational risk management (Popoola et al., 2025). At the same time, data protection and data security are identified as key challenges and risks in the application of AI in route optimization. The handling of sensitive personal and corporate data requires appropriate technical and organizational measures to prevent misuse, data breaches, or legal violations (Durlik et al., 2024; Paul et al., 2024).

2.3 AI forecasts in warehouse management

In warehouse management, AI can provide support by generating forecasts of future goods, material, and process flows based on extensive data sets. These predictions typically relate to demand development, picking volumes, resource requirements, as well as inbound and outbound quantities. In addition, process-relevant events such as delays or operational measures, for example necessary reorders, can be anticipated at an early stage. The objective is to improve operational planning and supporting decision-making processes to meet customer requirements more effectively and reliably (Sodiya et al., 2024; Yan et al., 2022).

Environmental effects (E) are addressed in several studies, with the main benefits lying in the optimization of warehouse processes. For example, AI can be used to dynamically reorganize operations by triggering automated processes once predefined thresholds are reached. Based on historical data, AI can determine an optimal safety stock level, thereby preventing supply shortages and overstocking. Manual controls can also be optimized, as inventory anomalies can be detected more quickly with AI and reviewed in a more targeted manner (Khmara, 2025). This leads to savings in resources such as storage space and energy consumption, walking distances within warehouses and empty runs, demand for fleet or production materials, as well as warehouse waste (Sodiya et al., 2024; Yan et al., 2022; Paul et al., 2024; Amosu et al., 2024). In most cases, the potential savings are discussed but not quantitatively substantiated. Amosu et al. report in their study that an international retail company was able to reduce its excess inventory by 30% after implementing an AI-driven model, resulting in significant cost savings. They also describe the case of an online retailer that used AI-based forecasts during intensive campaign periods, strategically increased its inventory levels, and achieved a 20% increase in revenue. Another retailer reduced its excess stock by 15% within a few months using a predictive AI model, while simultaneously decreasing the frequency of stockouts by 10% (Amosu et al., 2024). Other studies report cost reductions of 20–30% in the examined application cases (Shankaran, 2023). Successful integration of AI-based forecasting requires high data quality and the complex integration of various data sources (e.g., WMS, historical purchasing data, customer data, etc.). This, along with the need for continuous data maintenance, creates personnel and time burdens, particularly for small enterprises. In addition, the acquisition costs of AI systems are high.

Within the social ESG dimension (S), the need for additional qualification measures to address existing skill gaps is emphasized. The required training is often time-consuming and costly, while the shortage of qualified personnel further increases organizational effort. Furthermore, managing complex AI systems without adequate training can pose safety risks in terms of proper and error-free application (Sodiya et al., 2024; Khmara, 2025). AI-supported warehouse optimization may also increase data protection and privacy risks in certain cases, which can negatively affect employees or even customers. This is particularly relevant when personal performance, movement, or sensor data are processed on a large scale and when technical and/or organizational safeguards are insufficient (Sodiya et al., 2024; Amosu et al., 2024). In addition, it is critically noted that algorithmic bias in AI-based forecasting models is possible. This may negatively affect certain customer groups, as they could be systematically

disadvantaged (Amosu et al., 2024; Sodiya et al., 2024; Nicoletti et al., 2024). If the actual demand for sensitive goods such as pharmaceuticals is underestimated by AI due to incomplete or biased data, this may lead to supply shortages. At the same time, Amosu et al. report that, in addition to internal quantitative improvements, an increase in customer satisfaction was also achieved in the studies they reviewed.

For the governance criterion (G), the high value of forecasting performance is, the system output itself—should be emphasized. This enables more transparent and data-driven business decisions. Incorrect decisions and the resulting corporate risks, caused by misinterpreted or misanalyzed information, can be reduced. Furthermore, collaboration between departments, such as warehouse and sales, can benefit from increased data transparency and the traceability of decisions (Amosu et al., 2024; Yan et al., 2022; Khmara, 2025). At the same time, certain AI models are only partially interpretable due to their high level of complexity. This represents a governance risk, as the responsible human decision-maker must first understand and interpret the results to build trust in the system and make informed decisions. As with all IT systems that enallage volumes of data, the use of AI-based forecasting systems entails an increased risk of cyberattacks, which may lead to the loss of sensitive customer or company data (Amosu et al., 2024).

3 Research design

The data collection for this publication follows a scoping review approach based on Arksey and O'Malley (2005) and the refined methodological guidelines of Levac et al. (2010). The aim of a scoping review is to explore and map a research field, systematically organize existing findings, and identify existing research gaps. Unlike traditional systematic reviews, the focus is not on comprehensively capturing all available studies, but rather on the breadth of content and conceptual classification of a research area. The scoping review process was structured around five key methodological steps:

Definition of the research question

The objective of the data collection was to analyze the current state of the scientific literature on the impacts of artificial intelligence (AI) applications in logistics from the perspective of ESG criteria (Environmental, Social, Governance). The focus was on identifying which environmental, social, and governance-related effects could be identified through the use of AI in key logistics application areas and which research gaps exist in this context.

Identification of relevant studies

The literature review was conducted in the scientific databases ScienceDirect, Scopus, Semantic Scholar, and Google Scholar. The time frame covered the years 2020 to 2025 to account for the dynamic development of AI technologies and their current applications in logistics. The search strategy was iterative and based on the combination of key terms using Boolean operators. In an initial exploratory phase, search terms such as "artificial intelligence logistics," "AI in logistics," and "machine learning logistics" were used to identify relevant fields of application. Based on this initial analysis, AI-supported route optimization and AI-based forecasting in warehouse management were identified as key application areas. Building on this, the search was expanded in a second phase to specifically target ESG-related issues, for example through combinations such as "AI and logistics and sustainability", "AI route optimization and ESG", "AI warehouse forecasting and sustainability" and "machine learning supply chain environmental impact".

Selection of suitable studies

The selection of studies was based on defined inclusion and exclusion criteria. The inclusion criteria were:

- Peer-reviewed scientific publications (journal articles, conference papers)
- Publication year: 2020–2025
- Thematic focus: AI applications in logistics (particularly route optimization and warehouse management)
- Explicit reference to at least one ESG dimension (Environmental, Social, Governance)
- Availability in English or German.

The exclusion criteria were:

- Purely technical studies without a recognizable sustainability or ESG reference
- Literature without a peer-review process
- Studies outside the defined time
- Publications without full-text access

The initial search yielded a large number of potentially relevant publications, which were then systematically screened based on title, abstract, and content relevance. Following this multi-stage selection process, a total of 31 scientific articles were selected for in-depth analysis. The final selection was based on thematic relevance, coverage of the ESG dimensions, and the clarity and comprehensibility of the reported results.

Analysis and Structuring of Data

The selected studies were analyzed for content and structured according to the three ESG dimensions. In doing so, the reported environmental, social, and governance-related impacts of AI applications were systematically recorded and evaluated comparatively. Data extraction was performed using a structured framework that included the following information:

- AI application area (route optimization, warehouse management, other)
- Addressed ESG dimensions
- Reported impacts (positive, negative, ambivalent)
- Key findings and research gaps

The aim in this step was to identify recurring patterns, consistent results, and divergent findings. Additionally, the studies were categorized according to their thematic focuses to enable a structured presentation of the application areas and their respective sustainability impacts.

4. Research Results

Overall, the reviewed literature clearly shows that the application of AI in the logistics industry holds significant potential for improving efficiency and conserving resources. At the same time, there are marked differences in the depth and systematic approach with which the individual ESG dimensions are addressed in the literature. The following provides an overview of the key findings across the ESG dimensions.

Environmental (E): Emission Reduction with Context Dependency

The environmental dimension represents the most comprehensively studied ESG area. Most of our analyzed studies address environmental aspects. When using route optimization with the help of AI, direct, measurable environmental effects are documented in the literature:

- 15–30% reduction in emissions and fuel consumption with holistic integration (Sarioguz, 2025)
- 13.9% fuel savings and 17.3% energy savings in urban environments (Mohsen & Mohsen, 2025)
- Up to 40% reduction in emissions through the combined use of electric vehicles and AI (Ferreira & Esperança, 2024)
- 15–22% reduction in delivery times through dynamic adjustments (Ferreira & Esperança, 2024).

The analysis shows strong context dependency: vehicle technology, the logistics environment, digital maturity, and infrastructure significantly influence the extent of the effects. The greatest resource savings result from systemic integration into comprehensive sustainability strategies (Sarioguz, 2025; Ferreira & Esperança, 2024). When using AI forecasts in warehouse management, environmental benefits are primarily achieved indirectly through process optimization:

- up to 30% reduction in excess inventory (Amosu et al., 2024)
- 15% reduction in excess inventory while simultaneously reducing delivery bottlenecks by 10% (Amosu et al., 2024)
- 20–30% cost reductions through AI-supported forecasting models (Amosu et al., 2024; Shankaran, 2023).

Compared to route optimization, there are significantly fewer quantitative studies available on warehouse management.

Social (S): Significant Research Gap

Compared to the environmental dimension, the social dimension of AI use in logistics has been studied much less systematically. Only about half of our analyzed studies address social aspects, and these do so predominantly only marginally and without a robust empirical foundation. In the context of route optimization, potential positive effects are primarily inferred indirectly. These include a possible reduction in stress through more efficient route planning, increased delivery reliability and customer satisfaction, as well as improved air quality due to lower emissions. Furthermore, a shift in job profiles in the wake of increasing automation is discussed. However, these effects remain largely hypothetical due to a lack of empirical validation (Oršić et al., 2022). For the use of AI in warehouse management, individual positive social effects are also identified, particularly improved product availability and the resulting higher customer satisfaction (Amosu et al., 2024). Compared to route optimization, however, the evidence base here is even more fragmented.

The potential benefits are offset by several recurring challenges that apply to both route optimization and AI-based forecasting systems in warehouse management. Key problem areas include:

- Skill requirements: The use of AI-based systems requires new competencies and time- and cost-intensive training measures. The existing shortage of skilled workers further complicates implementation (Sodiya et al., 2024; Khmara, 2025)
- Data protection and surveillance: AI systems process extensive personal performance and movement data from employees and customers. Risks arise from inadequate technical or organizational safeguards (Sodiya et al., 2024; Amosu et al., 2024)
- Algorithmic biases and conflicting objectives: AI models may exhibit systematic forecasting errors or biases that disadvantage certain customer groups.

Furthermore, conflicts of interest arise between algorithmic optimization and individual user preferences, for example in the form of undesirable delivery time windows or route decisions (Amosu et al., 2024; Sodiya et al., 2024; Nicoletti et al., 2024; Sarioguz, 2025; Leogrande, 2024; Liu et al., 2025; Veluru, 2023).

In the field of AI applications in logistics, there is a significant research gap regarding social impacts. There is a lack of robust empirical evidence on actual effects on employment, working conditions, skill development, as well as on issues of fairness and data protection across the entire implementation and usage cycle.

Governance (G): The Tension Between Control and Transparency

The governance dimension is addressed in roughly half of the studies, though with less depth than the environmental dimension. In both the use of AI for route optimization and warehouse management, the key areas lie particularly in:

- Improving decision quality: AI systems enable the analysis of complex interrelationships and thus help reduce the number of wrong decisions (Amosu et al., 2024; Yan et al., 2022)
- Increasing data transparency: Centralized data management can improve collaboration between business units (Khmara, 2025)
- Improving risk management: AI can be used to analyze driving behavior to identify risky patterns and increase traffic safety (Popoola et al., 2025; Mohsen, 2024).

However, these potentials are offset by key governance risks and requirements:

- Algorithmic opacity: With highly complex models, it is often unclear why a system makes a particular recommendation (black-box problem), which raises issues of trust and liability (Amosu et al., 2024)
- Data protection and IT security: AI systems process large amounts of sensitive data, increasing the risks of cyberattacks, data loss, or data breaches. Several studies emphasize the need for appropriate technical and organizational safeguards (Durlik et al., 2024; Paul et al., 2024; Amosu et al., 2024).

Despite the identified opportunities and risks, there is currently a lack of robust empirical evidence regarding the actual effectiveness of governance mechanisms. Research gaps exist regarding the practical implementation of Explainable AI (XAI), specific governance practices in companies, and the clarification of responsibilities and liability issues in operational use.

The results show that research on AI in logistics has so far been primarily technology- and environment-oriented, while social and governance dimensions of a holistic ESG perspective have not been sufficiently considered.

5. Research Gaps and Future Directions

The most fundamental gap concerns longitudinal evaluation design. The vast majority of studies examining knowledge transfer effectiveness rely on cross-sectional or short-term pre-post designs that are poorly suited to capturing the durable behavioural and organisational changes that represent the most meaningful outcomes of transfer (Nielsen & Shepherd, 2022). Training that produces measurable learning gains at immediate assessment may fail to influence practice weeks or months later, while interventions that appear modestly effective in the short term may produce cumulative impacts that only become visible with sustained observation. Future research should prioritise longitudinal study designs, incorporating

repeated measurement across multiple time points and employing statistical methods capable of modelling the dynamic trajectories of knowledge application over time.

A second gap concerns the cross-sectoral comparability of evaluation frameworks. The literature reviewed in this article spans public health, clinical practice, workplace training, and formal education, yet the evaluation models developed within each domain remain largely insular. The ITTEM proposed by Nielsen and Shepherd (2022) was developed specifically for mental health training in workplace settings; the dissemination science framework articulated by Brownson et al. (2017) was primarily elaborated in the context of public health; the guideline implementation models described by Teede et al. (2018a, 2023) are oriented toward clinical practice. There is a need for overarching frameworks capable of accommodating the distinctive features of knowledge transfer across these varied contexts while enabling meaningful comparison. Future research should engage in systematic cross-domain synthesis and validation of evaluation frameworks, potentially drawing on implementation science as an integrative discipline.

A third gap concerns measurement in digital and AI-mediated knowledge transfer contexts. The rapid expansion of digital platforms as primary channels for knowledge transfer has not been matched by the development of appropriate evaluation methodologies. Existing frameworks were largely designed for face-to-face or hybrid contexts and require significant adaptation to capture the distinctive dynamics of online learning environments, including asynchronous engagement, automated feedback, and algorithm-driven content personalisation (Ahmad et al., 2022; Lekadir et al., 2025). Future research should develop and validate measurement instruments specifically designed for digital knowledge transfer contexts, leveraging the behavioural trace data generated by digital platforms while addressing the associated ethical and methodological challenges.

Finally, the role of social and relational factors in moderating transfer effectiveness merits more systematic investigation. Mehner et al. (2023) and Nielsen and Shepherd (2022) both highlight the importance of social support, yet the mechanisms through which social relationships facilitate or impede knowledge transfer remain insufficiently theorised and empirically documented. Future research should employ social network analysis and qualitative methods to illuminate these mechanisms, with direct implications for the design of interventions that build relational infrastructure alongside content-based knowledge transfer.

6. Conclusion

Based on the results and the subsequent discussion, several recommendations can be derived for academia, companies, and policymakers. These aim to better leverage the potential of AI applications in logistics in line with ESG criteria and to reduce existing uncertainties.

For scientific research, there is a clear need for a more holistic consideration and evaluation of ESG dimensions in relation to the use of AI. Future studies should not assess the environmental effects of AI applications solely from the perspective of efficiency gains but should also consider and compare the energy and resource consumption of AI systems themselves. The rebound effect also represents an unexplored area in the context of AI implementation, although it could provide additional insights into the net environmental impact. Furthermore, there remains a need for research in the social dimension. Studies are still lacking that specifically examine the effects of AI on working conditions in the logistics sector. It also remains unclear whether AI can serve as a potential solution to mitigate the shortage of skilled labor. Governance aspects such as data quality, the explainability of AI-based decisions,

and the actual implementation effort should also be quantified more extensively to bridge the existing gap between technical requirements and practical feasibility.

For companies, a deliberate and strategically embedded use of AI is recommended. AI should not be introduced solely as a technological solution but should be integrated into a broader sustainability and digitalization strategy from the outset. Prior to implementation, the availability and quality of data should be carefully assessed. Companies should also be prepared to implement organizational and process-related adjustments, for example regarding data provision or transparent decision-making processes. The need for qualification measures, as well as the associated costs and time requirements, should not be underestimated. For small and medium-sized enterprises, it is advisable to rely on scalable solutions, partnerships, or external service providers to conserve financial and human resources and reduce competitive disadvantages.

For regulatory bodies and other policymakers, a focus on practical and harmonized solutions can be recommended. For Europe, the EU AI Act represents a key step toward regulating AI, but at the same time increases documentation and organizational requirements for companies. At the international level, stronger alignment of regulatory frameworks would be desirable to facilitate the development of standardized IT infrastructures for companies.

In summary, it can be argued that a forward-looking, deliberate, and sustainable approach to the use of AI is more advantageous in the long term—both at the corporate and political level—than a purely reactive approach. Experience shows that preventive investments in appropriate structures, governance measures, and regulatory frameworks are less costly and more effective than the subsequent correction of mis developments or the management of resulting damages.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); 100% share.

References

Amosu, O. R., Kumar, P., Ogunsuji, Y. M., Oni, S., & Faworaja, O. (2024). AI-driven demand forecasting: Enhancing inventory management and customer satisfaction. World Journal of Advanced Research and Reviews, 23(2), 708–719. https://doi.org/10.30574/wjarr.2024.23.2.2394

Andiyappillai, N. (2019). Data analytics in warehouse management systems (WMS) implementations: A case study. International Journal of Computer Applications, 181(47)Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. International Journal of Social Research Methodology, 8(1), 19–32. https://doi.org/10.1080/1364557032000119616

Durlik, I., Miller, T., Kostecka, E., Łobodzińska, A., & Kostecki, T. (2024). Harnessing AI for sustainable shipping and green ports: Challenges and opportunities. Applied Sciences, 14(15), 5994. https://doi.org/10.3390/app14155994

Dyczkowska, J. A., Olkiewicz, M., & Królikowski, T. (2025). The use of artificial intelligence in logistics information systems. Procedia Computer Science, 270, 5198–5206. https://doi.org/10.1016/j.procs.2025.09.647

European Commission. (n.d.). Corporate sustainability and responsibility. Available at: https://single-market-economy.ec.europa.eu/industry/sustainability/corporate-sustainability-and-responsibility_en, [accessed 24.11.2025].

European Parliament and Council. (n.d.). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Available at: https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=OJ:L_202401689, [accessed 10.12.2025].

Fatorachian, H. (2024). Leveraging artificial intelligence for optimizing logistics performance: A comprehensive review. Global Journal of Business & Social Science Review, 12(3), 146–160. https://doi.org/10.35609/gjbssr.2024.12.3(5

Ferreira, J. C., & Esperança, M. (2025). Enhancing Sustainable Last-Mile Delivery: The Impact of Electric Vehicles and AI Optimization on Urban Logistics. World Electric Vehicle Journal, 16(5), 242. https://doi.org/10.3390/wevj16050242

Khmara, M. P. (2025). The impact of artificial intelligence application on the optimization of logistics and warehouse management. Journal of Strategic Economic Research, 3(26), 97–109. https://doi.org/10.30857/2786-5398.2025.3.8

Kihél, Y. E. (2022). Digital transition methodology of a warehouse in the concept of sustainable development with an industrial case study. Sustainability, 14(23), 15282. https://doi.org/10.3390/su142315282

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. Implementation Science, 5, Article 69. https://doi.org/10.1186/1748-5908-5-69

Liu, Y., Hou, J., & Cai, C. (2025). Research on the optimization of cold chain logistics distribution routes considering time-dependent networks and simultaneous pick-up and delivery from the perspective of sustainability. PLOS ONE, 20(9), Article e0330535. https://doi.org/10.1371/journal.pone.0330535

Luccioni, A. S., Strubell, E., & Crawford, K. (2025). From efficiency gains to rebound effects: The problem of Jevons' paradox in AI's polarized environmental debate. Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT '25), 1–13. https://doi.org/10.1145/3715275.3732007

Mohsen, B. M., & Mohsen, M. (2025). AI-driven energy optimization in urban logistics: Implications for smart SCM in Dubai. Sustainability, 17(18), Article 8301. https://doi.org/10.3390/su17188301

Mojumder, M. U., & Nuruzzaman, M. (2025). AI-driven optimization of warehouse layout and material handling: A quantitative study on efficiency and space utilization. Review of Applied Science and Technology, 4(2), 233–273.

Necula, S.-C., & Rieder, E. (2025). Generative and adaptive AI for sustainable supply chain design. Journal of Theoretical and Applied Electronic Commerce Research, 20(3), Article 240. https://doi.org/10.3390/jtaer20030240

Nicoletti, B., & Appolloni, A. (2024). Green logistics 5.0: A review of sustainability-oriented innovation with foundation models in logistics. European Journal of Innovation Management, 27(9), 542–561. https://doi.org/10.1108/EJIM-07-2024-0787

Ogorodnyk, O., Stendal, J. A., Leirmo, T., & Harik, E. H. C. (2025). Explainable AI for delivery route optimization using reinforcement learning. IFAC-PapersOnLine, 59(10), 2136–2140. https://doi.org/10.1016/j.ifacol.2025.09.359

Oršič, J., Jereb, B., & Obrecht, M. (2022). Sustainable operations of last mile logistics based on machine learning processes. Processes, 10(12), Article 2524. https://doi.org/10.3390/pr10122524

Okpeke Paul, P., Aderoju, A. V., Shitu, K., Ononiwu, M. I., Igwe, A. N., Ofodile, O. C., & Ewim, C. P.-M. (2024). Predictive analytics and AI in sustainable logistics: A review of applications and impact on SMEs. Magna Scientia Advanced Research and Reviews, 12(1), 231–251. Retrieved March 27, 2026, from https://msarr.magnascientiapub.com/content/predictive-analytics-and-ai-sustainable-logistics-review-applications-and-impact-smes

Popoola, R., Aisosa, O. F., Abiodun, L. S., Adeolu, O., Mbamalu, P. O., Rachael, A. A., & Adimchi, C. (2025). Artificial intelligence-driven supply chain optimization for sustainable last-mile delivery in smart cities: An electric vehicle routing approach. Asian Journal of Advanced Research and Reports, 19(10), 16–31. https://doi.org/10.9734/ajarr/2025/v19i1011615

Ran, J. (2021). Intelligent method of supply chain circulation industry structure based on machine learning. Mobile Information Systems, 2021, 1–10. https://doi.org/10.1155/2021/4116497

Sarioguz, O. (2025). Artificial intelligence for sustainable logistics: Reducing carbon emissions and fuel consumption through route optimization. International Journal of Science and Research Archive, 15(3), 1527–1537. https://doi.org/10.30574/ijsra.2025.15.3.1933

Sattar, M. U., Dattana, V., Hasan, R., Mahmood, S., Khan, H. W., & Hussain, S. (2025). Enhancing supply chain management: A comparative study of machine learning techniques with cost–accuracy and ESG-based evaluation for forecasting and risk mitigation. Sustainability, 17(13), Article 5772. https://doi.org/10.3390/su17135772.

Shankaran, S. (2023). Revolutionizing warehouse management: How AI drives operational efficiency. International Journal of All Research Education and Scientific Methods (IJARESM), 11(11). https://www.researchgate.net/publication/389516065_Revolutionizing_Warehouse_Management_How_AI_Drives_Operational_Efficiency

Sodiya, E. O., Umoga, U. J., Amoo, O. O., & Atadoga, A. (2024). AI-driven warehouse automation: A comprehensive review of systems. GSC Advanced Research and Reviews, 18(2), 272–282. https://doi.org/10.30574/gscarr.2024.18.2.0063

Toorajipour, R., Sohrabpour, V., Nazarpour, A., Oghazi, P., & Fischl, M. (2021). Artificial intelligence in supply chain management: A systematic literature review. Journal of Business Research, 122, 502–517. https://doi.org/10.1016/j.jbusres.2020.09.009

Veluru, C. S. (2023). A comprehensive study on optimizing delivery routes through generative AI using real-time traffic and environmental data. Journal of Scientific and Engineering Research, 10(10), 168–175.

Yan, Y., Chow, A. H. F., Ho, C. P., Kuo, Y. H., Wu, Q., & Ying, C. (2022). Reinforcement learning for logistics and supply chain management: Methodologies, state of the art, and future opportunities. Transportation Research Part E: Logistics and Transportation Review, 162, Article 102712. https://doi.org/10.1016/j.tre.2022.102712

Zhu, C., Xin, J., & Trinh, T. K. (2025). Data quality challenges and governance frameworks for AI implementation in supply chain management. Pinnacle Academic Press Proceedings Series, 2, 1–14. https://pinnaclepubs.com/index.php/PAPPS/article/view/101/102

Artificial Intelligence in Rural Businesses: A Structured Literature Review

Veronika Volfová[1] and Nora Grisáková[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: veronika.volfova@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: nora.grisakova@euba.sk

Abstract

This article presents a structured literature review of research on the use of artificial intelligence in rural businesses. Drawing on publications identified in the Web of Science database, the review aims to map the current state of knowledge and to identify the main thematic areas addressed in the existing literature. A total of 16 articles were included in the final review after relevance-based screening.

The analysis identified three main thematic clusters: the technological transformation of agriculture associated with Agriculture 4.0, the role of AI in broader rural business development, and the application of AI-based methods to analyse entrepreneurial success factors in rural contexts. The findings indicate that the current literature is concentrated primarily on the agricultural sector, while non-agricultural rural enterprises and the broader socio-economic impacts of AI remain less explored. At the same time, the reviewed studies suggest that AI may help rural businesses overcome geographical barriers, address labour shortages, improve efficiency, and strengthen competitiveness.

The article contributes by providing a structured overview of a still emerging field of research and by identifying gaps that may inform future studies on AI adoption and its impacts in rural business environments

Key Words: artificial intelligence, innovation, rural businesses

JEL Classification: O33, R11, Q12

1. Introduction

The emergence of artificial intelligence has created new opportunities for sustainable and inclusive development not only in urban environments but also in rural areas. AI creates new possibilities for the development of the green and circular economy, agriculture, and tourism, while also contributing to improvements in production, service quality, and business processes in rural settings. It can also help reduce disadvantages related to remoteness and limited infrastructure (European Parliament, 2023). In particular, the ability of AI to analyse large volumes of data and optimise processes remotely may be especially important for regions that have long struggled with limited access to services, expertise, and markets.

Although the use of artificial intelligence in business has received growing scholarly attention, research focusing specifically on rural businesses remains relatively limited. This article therefore aims to map the current state of knowledge on the use of AI in rural businesses, identify the main thematic areas addressed in the literature, and highlight research gaps that may inform future research.

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

The article is based on a structured literature review conducted using the Web of Science database, focusing on publications addressing both artificial intelligence and rural business contexts. An initial set of 33 scientific articles was identified using the keywords "AI" and "rural businesses." In the next step, the articles were screened for their thematic relevance to the topic of the study. Only publications explicitly addressing the use of AI in relation to rural enterprises, rural entrepreneurship, or business activities in rural areas were retained for further analysis. As a result, 16 articles were included in the final review.

Given the exploratory nature of the topic and the relatively limited number of directly relevant publications, the review does not follow a formal systematic review protocol. Instead, it applies a structured and thematically oriented approach to identify key research areas, recurring topics, and gaps in the existing literature. The selected studies were subsequently analysed and grouped according to their thematic proximity, which made it possible to identify the main research clusters within the field.

2. Theoretical Background

In order to interpret the selected literature coherently, it is first necessary to define the main concepts that frame this review. Since the analysed studies operate across different disciplinary contexts, clarification of the concepts of rural enterprises and rural areas, innovation, and artificial intelligence provides the conceptual basis for understanding how AI is discussed in relation to business activity in rural settings.

Definition of rural enterprises and rural areas

Rural enterprises are regarded as business entities that make a significant contribution to the national economy (Kamar and Ali, 2023). A characteristic feature of rural enterprises is not only their geographical location but also their connection to the local environment, including the use of local resources and an irreplaceable contribution to the local community (Rachwał, 2022). As noted by Beckmann et al. (2023), rural enterprises also play an important role in the sustainable development of rural areas, support innovation, and contribute to preventing rural depopulation by creating new jobs. According to the OECD definition, rural areas are considered "territories where the population density is lower than 150 inhabitants/km^2." At the same time, the OECD divides rural areas into the local level (LAU2) and the regional level (NUTS3). At the local level, rural space is defined as an area that does not meet the definition of urban areas, meaning its population density is lower than 150 inhabitants/km^2. At the regional level, rural space is defined according to the share of the population living in rural cells. This level is further divided into, first, "predominantly rural regions, where more than 50% of the regional population lives in rural municipalities." Second, "significantly rural regions, where 15–50% of the regional population lives in rural municipalities." And "predominantly urban regions, where less than 15% of the regional population lives in rural municipalities" (Brezzi et al., 2011).

Definition of innovation

Innovation provides businesses with opportunities for faster growth and greater competitiveness. It is therefore important to examine the factors that shape innovative activities in rural enterprises (Gamito, Madureira, and Santos, 2021). In general, rural enterprises are not perceived as highly innovative. They are characterized by limited availability of resources and labor, poorer infrastructure, and also a limited customer base (Steinerowska-Streb et al., 2024). Gamito et al. (2021) defined categories of resources that generally shape the dynamics of rural

innovation. These are natural resources, cultural amenities, traditions, the rural population, and the territory. According to the OECD report (2025), innovations in rural areas have a positive impact on society in adapting to global changes. Innovations in rural areas are not built on general research and development but are specific, influenced, and driven by people and creativity in a given place. Thanks to digitalization, automation, and AI, new business models are emerging through which distance and resource unavailability can be overcome.

Definition of artificial intelligence

According to the European Parliament (2023), artificial intelligence may be defined as "the ability of machines to mimic human capabilities, such as reasoning, learning, planning, or creativity." AI enables technical systems to respond more efficiently, solve problems, and achieve defined goals. At the same time, AI-based systems are capable of operating independently and adapting their behaviour on the basis of previous actions. The implementation of artificial intelligence brings a wide range of benefits. In business, it offers opportunities for the development of new products and services across various sectors, including agriculture. It can improve machine maintenance, increase production and quality, save energy, or improve product sustainability. At the same time, the implementation of AI in enterprises also brings risks, such as job displacement, which may affect local communities (European Parliament, 2025).

These concepts provide the basis for the structured review presented in the following section. Together, these concepts make it possible to distinguish between studies focused on technological innovation in agriculture, broader forms of rural business development, and research addressing the determinants of entrepreneurial success in rural contexts. The conceptual clarification is therefore important not only for defining the scope of the topic, but also for interpreting the main thematic clusters identified in the literature.

3. Literature Review

The selected studies indicate that the existing literature on artificial intelligence in rural businesses is concentrated around three main thematic areas. These areas include the technological transformation of agriculture, the use of AI in broader rural business development, and the application of AI-based methods to analyse entrepreneurial success factors in rural contexts. The following review is therefore organised according to these thematic clusters.

3.1 Artificial intelligence in agriculture and Agriculture 4.0

In the context of artificial intelligence in rural areas, the most frequently discussed topic is the use of AI in agriculture. Takemura et al. (2021) examined an educational program focused on linking agriculture with modern IT tools. This program aims to address labor shortages caused by an aging population and rural decline due to depopulation. The proposed response to these issues is the introduction of the "super-smart society" concept based on AI and IoT. Furthermore, in the context of implementing Agriculture 4.0 and its importance for the development and growth of agricultural production in developing countries, da Silveira et al. (2023) investigated the barriers affecting the adoption of artificial intelligence, IoT, machine learning, and other modern technologies.

Pastor et al. (2025) explored the use of Model-Based Systems Engineering (MBSE), combined with generative artificial intelligence and natural language processing, to streamline the management and supervision of agricultural projects. They concluded that these

technologies help better manage knowledge and automate control while reducing both technological and business risks. Bešić et al. (2024) focused on the analysis of Society 5.0 and its influence on agricultural business and innovation, primarily through the implementation of artificial intelligence, IoT, robotics, and other technologies. In addition to these more specific contributions, Mulongo (2025) provided an overview of current progress in the application of AI in agriculture, together with the associated benefits, barriers, and consequences. Similarly, Pimenow et al. (2025) published an extensive review examining the impact of artificial intelligence in local contexts, including its positive influence on the digital transformation of businesses. On the other hand, Kohli (2020) focused on the pitfalls arising from the use of AI and the digitalization of enterprises. The author highlights the negative impacts of digitalisation and e-commerce, particularly for traditional and small businesses. In the context of rural enterprises, he draws attention to the situation of the agricultural workforce, which is losing jobs due to technological progress. The impact on agriculture is also addressed by Shaikh et al. (2022), who focused on the benefits and challenges associated with using not only artificial intelligence but also robotics, IoT, and drones. Their article summarizes global state-of-the-art agricultural systems and future trends, including the challenges they entail.

3.2 Artificial intelligence in broader rural business development

Beyond the agricultural sector, the literature also points to a broader understanding of the role of AI in rural entrepreneurship and rural development. As noted by Barrera-Perales and Burgos (2023), it is necessary to distinguish between innovations concerning artificial intelligence in agriculture and purely rural innovations. Therefore, they proposed a conceptual framework for grasping this issue. According to the authors, innovations associated with artificial intelligence in agriculture focus on technology, agrobusiness, and the competitiveness of agricultural enterprises, whereas rural innovations are centered on social change and the development of the internal community.

Several studies address issues related to AI in rural entrepreneurship beyond agriculture. Lehner et al. (2024) proposed a smart grid concept using technologies such as AI that could support the distribution network in Ethiopia. Since residents in some parts, especially rural areas, lack access to electricity, this concept aims to support business development in these regions. Roy (2025) examined the impact of AI in relation to social media management and its effect on rural tourism businesses. The findings show that AI has a positive effect on the personalization of marketing content, automation of customer interactions, and content strategy, while also helping to overcome geographical barriers and facilitating access to further markets. The issue of AI adoption in rural enterprises was also addressed by Dowell et al. (2026), who focused on small and medium-sized enterprises in the UK. Through extensive panel research and probit regression, they found that networking is a key factor for the implementation of AI in rural firms. At the same time, they recommend better communication of the benefits of AI adoption and improved support for the development of business networks.

3.3 AI-based analysis of entrepreneurial success factors in rural contexts

Artificial intelligence affects rural enterprises not only as an implemented technology but also as a tool for analysing the factors that shape entrepreneurial activity in rural areas. Several studies have used AI and machine learning methods to identify these determinants. For example, Celbis (2021) used five different machine learning techniques to analyze factors affecting start-up entrepreneurs in rural areas, thereby identifying areas that have a fundamental impact on the success or failure of rural entrepreneurs. In the results, he highlighted the main factors as “capital constraints, age, factors related to trust and over-trust,

awareness of current trends, the use of various media tools, a competitive character, institutional factors, and education." A similar contribution was provided by Wu et al. (2022), who used artificial intelligence to identify the main factors influencing the entrepreneurial success of rural workers. They emphasize that examining these factors is important, as supporting entrepreneurship in these areas is a key element for poverty alleviation and structural transformation.

This analytical perspective is also reflected in studies focusing on the adoption of AI and the allocation of resources in rural business environments. Polas et al. (2026) examined the factors influencing the willingness of owners of family agricultural businesses in rural Bangladesh to adopt AI for sustainable business. They found that the willingness to adopt AI depends primarily on the perceived value of the technology's benefits to their business and the innovativeness of the farmers themselves. Rolando et al. (2025) used data-driven AI to analyze funded projects. The aim was to determine how financial resources are distributed and whether these investments effectively support the competitiveness of local agri-food chains and small and medium-sized enterprises.

Taken together, the reviewed studies reveal a clear thematic structure in the existing literature. This structure provides the basis for the synthesis presented in the following section.

4. Research Areas and Future Directions

Building on the thematic structure identified in the reviewed literature, several broader patterns emerge. The strongest concentration of existing research is in the field of agriculture. This suggests that AI in rural contexts is still most commonly associated with technological transformation in farming and agri-business. By contrast, the use of AI in non-agricultural rural business activities remains less represented in the current literature, even though this area appears to offer considerable potential for business development, market expansion, and the reduction of geographical disadvantages.

Another important finding is that the existing literature pays increasing attention to the determinants of AI adoption and entrepreneurial success in rural settings. At the same time, however, the review indicates that several important dimensions remain underexplored. These include the long-term socio-economic impacts of AI on rural labour markets and communities, the digital divide between technologically advanced and traditional enterprises, and the specific barriers that hinder AI implementation in smaller rural firms outside the agricultural sector.

Future research should therefore focus more strongly on non-agricultural rural enterprises, comparative analyses across different types of rural businesses, and the broader social consequences of AI adoption in rural areas. Greater attention should also be paid to empirical studies examining how rural enterprises integrate AI into everyday business practices and how this affects their resilience, competitiveness, and long-term development.

5. Conclusion

Based on the structured literature review of 16 articles from the Web of Science database, it can be concluded that research on artificial intelligence in rural businesses remains relatively limited, although interest in this topic has been increasing. The literature reviewed in this article points to an uneven distribution of attention across thematic areas, with the strongest concentration of research remaining in the agricultural sector.

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

The findings show that AI is primarily associated with addressing labour shortages, improving efficiency, and supporting technological innovation in agriculture. At the same time, the review suggests that AI also holds significant potential for non-agricultural rural enterprises, particularly in overcoming geographical barriers, improving access to markets, and strengthening business competitiveness. However, the adoption of these technologies is shaped by several barriers, including financial constraints, insufficient infrastructure, and differences in the innovativeness and capabilities of entrepreneurs.

The main contribution of this article lies in providing a structured overview of the current state of knowledge and in identifying gaps in the existing literature. In particular, the long-term socio-economic impacts of AI on rural labour markets, communities, and smaller non-agricultural enterprises remain insufficiently explored. These areas represent important directions for future research.

Funding

The article is an (partial) output of research project VEGA 1/0085/25, titled: 'Synergistic framework for designing modular models in corporate strategic decision-making using machine and deep learning'; 100% share.

References

Barrera-Perales, O. T., & Burgos, A. L. (2023). Conceptual distinction between agricultural innovation and rural innovation: Implications for scientific research and public policy. Innovation and Development, 13(3), 471–492. https://doi.org/10.1080/2157930X.2022.2049534.

Beckmann, M., Garkisch, M., & Zeyen, A. (2023). Together we are strong? A systematic literature review on how SMEs use relation-based collaboration to operate in rural areas. Journal of Small Business & Entrepreneurship, 35(4), 515–549. https://doi.org/10.1080/08276331.2021.1874605.

Bešić, C., Ćoćkalo, D., Bakator, M., Stanisavljev, S., & Bogetić, S. (2024). Society 5.0 and its impact on agricultural business and innovation: A new paradigm for rural development. Ekonomika Poljoprivrede / Economics of Agriculture, 71(3), 803–819. https://doi.org/10.59267/ekoPolj2403803B.

Brezzi, M., Dijkstra, L., & Ruiz, V. (2011). OECD extended regional typology: The economic performance of remote rural regions (OECD Regional Development Working Papers No. 2011/06). OECD Publishing. https://doi.org/10.1787/5kg6z83tw7f4-en.

Celbiş, M. G. (2021). A machine learning approach to rural entrepreneurship. Papers in Regional Science, 100(4), 1079–1104. https://doi.org/10.1111/pirs.12595.

da Silveira, F., da Silva, S. L. C., Machado, F. M., Barbedo, J. G. A., & Amaral, F. G. (2023). Farmers' perception of the barriers that hinder the implementation of agriculture 4.0. Agricultural Systems, 208, Article 103656. https://doi.org/10.1016/j.agsy.2023.103656.

Dowell, D., Bowen, R., & Morris, W. (2026). Artificial Intelligence (AI) in rural business: The drivers and effects on AI adoption in rural SMEs. Journal of Rural Studies, 123, Article 104104. https://doi.org/10.1016/j.jrurstud.2026.104104.

European Parliament. (2025, February 20). Artificial intelligence: risks and opportunities. https://www.europarl.europa.eu/topics/cs/article/20200918STO87404/umela-inteligence-jake-jsou-vyhody-a-nevyhody

European Parliament. (2023, June 20). What is artificial intelligence and how is it used? https://www.europarl.europa.eu/topics/en/article/20200827STO85804/what-is-artificial-intelligence-and-how-is-it-used

Gamito, T. M., Madureira, L., & Santos, J. M. L. (2021). Unveiling and typifying rural resources underpinned by innovation dynamics in rural areas. Regional Science Policy & Practice, 13(3), 457–478. https://doi.org/10.1111/rsp3.12228.

Kamar, M. A. A., & Ali, M. A. F. (2023). The impacts of the COVID-19 pandemic crisis on rural entrepreneurs. Malaysian Journal of Society and Space, 19(1), 109–118. https://doi.org/10.17576/geo-2023-1901-08.

Kohli, G. (2020). Pitfalls of digitization with respect to traditional businesses and employment. JIMS8M: The Journal of Indian Management & Strategy, 25(1), 28–33. https://doi.org/10.5958/0973-9343.2020.00004.6.

Lehner, M., Gemeda, D. B., Simani, B., Kebede, N., Guye, A., & Stork, W. (2024). Developing a smart distribution grid: Case study in Ethiopia. In 2024 3rd International Conference on Power Systems and Electrical Technology (PSET) (pp. 540–545). IEEE. https://doi.org/10.1109/PSET62496.2024.10808857.

Mulongo, N. Y. (2025). Artificial intelligence for smart agriculture industry. In R. Paul (Ed.), 2025 IEEE World AI IoT Congress (AIIoT) (pp. 898–901). IEEE. https://doi.org/10.1109/AIIoT65859.2025.11105360.

Pastor, R., Rodriguez, P. G., Lecuona, A., & Cortés, J. P. (2025). Smart agri-region and value engineering. Systems, 13(6), Article 430. https://doi.org/10.3390/systems13060430.

Pimenow, S., Pimenowa, O., Prus, P., & Niklas, A. (2025). The impact of artificial intelligence on the sustainability of regional ecosystems: Current challenges and future prospects. Sustainability, 17(11), Article 4795. https://doi.org/10.3390/su17114795.

Polas, M. R. H., Jahanshahi, A. A., Islam, M. E., Falahat, M., Raju, V., & Upoma, K. S. H. (2026). Toward sustainable agrientrepreneurship: Behavioral drivers of AI adoption among rural agrientrepreneurs in emerging economies. Journal of Enterprising Communities: People and Places in the Global Economy. Advance online publication. https://doi.org/10.1108/JEC-04-2025-0122.

Rachwał, T., Kulawiak, A., & Suliborski, A. (2022). Research on rural entrepreneurship in terms of the literature: Definition problems and selected research issues. Quaestiones Geographicae, 41(2), 7–19. https://doi.org/10.2478/quageo-2022-0020.

Rolando, D., Barreca, A., Malavasi, G., & Rebaudengo, M. (2025). Data-driven AI approach to address territorial strategies: Why investing in agri-food sector to enhance the Valsesia inner area. In O. Gervasi, B. Murgante, C. Garau, Y. Karaca, M. N. F. Lago, F. Scorza, & A. C. Braga (Eds.), Computational science and its applications – ICCSA 2025 workshops (Vol. 15889, pp. 177–198). Springer. https://doi.org/10.1007/978-3-031-97603-2_12.

Roy, N. C. (2025). AI-driven social media and economic empowerment in rural tourism entrepreneurship: An empirical analysis. Journal of Entrepreneurship and Public Policy. https://doi.org/10.1108/JEPP-04-2025-0111.

Shaikh, T. A., Mir, W. A., Rasool, T., & Sofi, S. (2022). Machine learning for smart agriculture and precision farming: Towards making the fields talk. Archives of Computational Methods in Engineering, 29(7), 4557–4597. https://doi.org/10.1007/s11831-022-09761-4.

Steinerowska-Streb, I., Peterková, J., & Steiner, A. (2024). Innovation and rural context: An exploratory case study of a small rural enterprise from the Czech Republic. The International Journal of Entrepreneurship and Innovation, 25(2), 116–127. https://doi.org/10.1177/14657503241241415.

Takemura, Y., Kamei, K., Sanada, A., & Ishii, K. (2021). Smart Agriculture IoT Education Course in enPiT-everi (Education Network for Practical Information Technologies - Evolving and Empowering Regional Industries). In M. Sugisaka, Y. Jia, T. Ito, & J. J. Lee (Eds.), Proceedings of the 2021 International Conference on Artificial Life and Robotics (ICAROB 2021) (pp. 401–404). ALife Robot Corp. Ltd.

Wu, S., Chang, Y., & Liao, X. (2022). Analysis of the influence of back home to start undertaking and rural revitalization based on artificial intelligence. Mobile Information Systems, 2022, Article 4616959. https://doi.org/10.1155/2022/4616959.

Artificial Intelligence in Human Resource Management: A Structured Literature Review of Recruitment, Fairness, Talent Management, and Employee Experience

Roman Tranžík[1] and Miroslava Jozekova[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: roman.tranzik@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: miroslava.jozekova@euba.sk

Abstract

This literature review examines the use of artificial intelligence in human resource management, with a specific focus on recruitment and selection, fairness and governance, talent management, retention, and employee-related outcomes. The growing integration of AI into HR processes has transformed the way organizations attract, evaluate, develop, and retain employees, while simultaneously raising important concerns regarding transparency, accountability, and the role of human oversight. The review synthesizes recent scholarly literature and identifies three major thematic domains: the application of AI in recruitment and selection, the ethical and governance implications of AI-supported HR decision-making, and the strategic use of AI in talent management, retention, and employee experience. The findings indicate that AI can improve efficiency, support data-driven decision-making, and enhance the analytical capacity of HR functions. At the same time, the literature consistently highlights risks associated with algorithmic bias, procedural fairness, lack of explainability, and the possible negative effects of AI on employee trust and workplace well-being. The review further identifies several research gaps, especially the dominance of recruitment-focused studies, the lack of longitudinal research, and the insufficient integration of technical and organizational perspectives. The article concludes by outlining future research directions that emphasize broader HR functional coverage, stronger governance frameworks, and a more human-centered understanding of AI adoption in HRM.

Key Words: artificial intelligence, human resource management, recruitment, talent management, fairness, employee experience

JEL Classification: M12, M15, O33

1. Introduction

In recent years, artificial intelligence has become one of the most influential technologies shaping organizational operations. Its expanding application no longer affects only technical, production, or marketing processes; it is also penetrating the field of human resource management, where it is transforming the ways in which employees are attracted, evaluated, developed, and retained (Dima et al., 2024; Fenwick et al., 2024). Human resource management, which has long been grounded in a combination of administrative tasks, professional judgment, and interpersonal interaction, is thus gradually evolving into a domain in which data analytics, automation, and algorithmically supported decision-making are becoming increasingly important (Madanchian, 2024).

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

This transformation is significant not only from the perspective of technological progress but also from that of strategic organizational management. AI in HRM is associated with expectations of greater efficiency, faster information processing, better candidate matching, more accurate talent identification, and more robust prediction of employee behavior, for example in relation to turnover or engagement (Marín Díaz et al., 2023; Rožman et al., 2022). At the same time, however, AI raises questions that directly affect the legitimacy of personnel decision-making. The literature increasingly points to the risks of algorithmic bias, ambiguous accountability, limited transparency, and the weakening of human oversight in decisions with substantial consequences for individuals (Bujold et al., 2024; Kazim et al., 2021).

These issues are especially pronounced in recruitment and selection, where AI is currently used most extensively. Research shows that AI can accelerate candidate screening, reduce administrative burdens, and support greater consistency in decision-making, while at the same time shaping applicants' perceptions of fairness, trust, and organizational attractiveness (Almeida et al., 2025; Ligeiro et al., 2024). As the use of AI in HR processes expands, however, attention is also moving beyond recruitment. Recent literature addresses AI-supported talent management, retention, employee well-being, and the broader effects of AI on employees' work experiences, thereby extending the understanding of AI in HRM from an operational tool to a broader sociotechnical and organizational phenomenon (Faqihi & Miah, 2023; Soulami et al., 2024).

Despite growing scholarly interest in this issue, the literature remains fragmented to some extent. Some studies focus on the adoption of AI in specific HR functions, especially recruitment; others concentrate primarily on technical or ethical questions; and still others examine the consequences of AI for employees and organizations. What is still lacking, however, is a sufficiently integrated perspective that connects these lines of inquiry and makes it possible to understand more systematically how AI is used across the main domains of HRM and what key benefits, risks, and research gaps emerge from this use (Bujold et al., 2024; Dima et al., 2024; Madanchian, 2024).

2. Theoretical Background

2.1 Artificial Intelligence in Human Resource Management

Artificial intelligence in human resource management refers to a set of technological solutions that support data analysis, process automation, and decision-making in the personnel domain. The current literature emphasizes that AI is no longer confined to recruitment and selection but now extends to onboarding, retention, talent management, and broader HR decision-making. At the theoretical level, AI in HRM therefore cannot be understood merely as a tool of operational efficiency; rather, it should be seen as part of a broader digital transformation of personnel systems that is reshaping the work of HR departments and their strategic role within organizations (Dima et al., 2024; Fenwick et al., 2024; Madanchian, 2024).

This shift is accompanied by a broader redefinition of the HR function itself. Whereas traditional personnel management was based primarily on administrative and support tasks, more recent approaches show that AI is moving HR toward a data-driven and strategically integrated form of people management. Fenwick et al. (2024) therefore describe a growing convergence between humans and machines in the workplace, while Tinguely et al. (2023) argue that the expanding adoption of AI is fundamentally reshaping HRM practice and requires new thinking about how HR systems should be designed in environments characterized by algorithmic support.

2.2 Main Areas of AI Use in HRM

From a theoretical perspective, it is essential to recognize that AI in HRM does not operate within a single isolated function but across multiple personnel processes. Madanchian (2024) shows that AI tools are currently most relevant in recruitment, onboarding, and retention, and that their importance lies in their ability to support faster, more data-precise, and more consistent decision-making. Dima et al. (2024) likewise emphasize that the effects of AI in HR cannot be reduced to automation alone; they also involve a transformation of HR activities themselves and of the roles of the actors participating in them. Theoretically, therefore, AI in HRM represents not only a technological instrument but also a driver of personnel process redesign.

This perspective is also supported by studies focused on specific HR decisions. In recruitment and selection, AI is associated with résumé screening, predictive candidate matching, and automated support for selection processes, whereas in retention and talent management it is used to identify turnover risk, support employee development, and better target decisions related to human capital (Ligeiro et al., 2024; Marín Díaz et al., 2023).

2.3 Human–AI Collaboration in HR Decision-Making

One of the central theoretical questions concerns the relationship between human judgment and algorithmic support. Recent literature no longer confines itself to the dilemma of whether AI will replace humans; instead, it increasingly emphasizes the need to design HR systems as forms of collaboration between people and technology. In this regard, Fenwick et al. (2024) discuss a growing convergence between humans and machines in the workplace, while Tinguely et al. (2023) explicitly argue that HR systems in the age of AI must be designed with regard to the nature of tasks, the degree of decision risk, and the social acceptability of such solutions. Theoretically, this means that AI in HRM should not be understood as an autonomous decision-making mechanism but as part of a broader sociotechnical system.

This understanding is especially important because HR decisions concern domains with substantial consequences for individuals, such as hiring, career progression, or retention. For this reason, the literature underscores the need to preserve human oversight, the possibility of interpreting outputs, and clear responsibility for final decisions. Explainable AI in HR, for example in attrition prediction, is particularly important in this respect because it makes it possible not only to generate predictions but also to better understand the factors behind them and subsequently connect analytical outputs to practical HR interventions (Marín Díaz et al., 2023).

2.4 Responsible AI in HRM: Fairness, Transparency, and Auditability

A second key theoretical line is the concept of responsible AI in HRM. In their review of the empirical literature, Bujold et al. (2024) show that responsible AI in personnel management is closely linked to fairness, trust, privacy protection, transparency, explainability, and the preservation of the human role in decision-making. From a theoretical standpoint, these principles cannot be regarded as merely supplementary; in HRM they constitute fundamental conditions of AI's legitimacy because personnel decisions directly affect individuals' opportunities, careers, and positions in the workplace.

This framework is closely connected to the literature on diversity, equity, and inclusion. Naoum et al. (2026) point out that AI in HRM has a dual effect with respect to DEI: it may streamline decision-making and reduce certain forms of subjectivity, but it may also reproduce or intensify historical inequalities embedded in data and models. The scholarly literature

therefore emphasizes the importance of auditing algorithmic systems. In this context, Kazim et al. (2021) stress the need to systematize auditing in algorithmic recruitment, since without scrutiny of models, data, and decision criteria, neither fairness nor trustworthiness in algorithmically supported HR processes can be ensured.

2.5 AI Adoption in HRM and the Responses of Candidates and Employees

The theoretical foundations of AI research in HRM also include the perspective of technology acceptance among users. In recruitment, classical technology acceptance models remain relevant for AI tools. Almeida et al. (2025) use the Technology Acceptance Model to explain recruiters' acceptance of AI, while Tanantong and Wongras (2024), drawing on UTAUT, highlight the importance of factors such as perceived value, autonomy, social influence, facilitating conditions, and trust in AI. Theoretically, this means that the adoption of AI in HRM is not merely a matter of technical availability but also of how users perceive its benefits, their control over the process, and the trustworthiness of the system.

The perspectives of candidates and employees are equally important. Ligeiro et al. (2024) show that, in AI-supported recruitment processes, trust in the process, perceived novelty, and candidates' motivation to continue in the selection procedure play a significant role. At the employee level, Soulami et al. (2024) emphasize that the adoption of AI in the workplace opens questions concerning the future of work and employee well-being, while Xu et al. (2023) empirically confirm a relationship between perceiving AI as an opportunity and workplace well-being. AI in HRM therefore cannot be examined solely through the lens of organizational efficiency; its psychological and social consequences for individuals are equally important.

2.6 Synthesis of the Theoretical Foundations

Based on the foregoing, it can be concluded that the theoretical foundations for examining AI in HRM are multidimensional. They include a functional perspective on the use of AI across HR processes, a sociotechnical perspective on collaboration between humans and technology, an ethical perspective centered on responsible AI, and a behavioral perspective on AI acceptance among recruiters, candidates, and employees. These perspectives complement one another and provide an appropriate basis for the subsequent literature synthesis focused on recruitment, talent management, employee experience, and the broader consequences of implementing AI in human resource management (Bujold et al., 2024; Fenwick et al., 2024; Madanchian, 2024; Soulami et al., 2024).

3. Literature Review

3.1 Artificial Intelligence in Recruitment and Employee Selection

The most strongly represented area in research on AI in HRM remains recruitment and selection. The recent literature agrees that this stage of the employee lifecycle is where AI is most visibly applied, especially in résumé screening, preliminary candidate assessment, applicant–position matching, and the organization of selection processes. Madanchian's (2024) review article shows that hiring, onboarding, and retention constitute the three most common areas in which AI is reshaping HR decision-making, with recruitment serving as the entry point to the broader digitization of the HR function. Similarly, Dima et al. (2024), in their review, show that AI is changing the very character of HR activities by combining task automation with data optimization and decision support in personnel processes.

At the same time, the empirical literature shows that the effectiveness of AI in recruitment cannot be assessed solely through the technical dimension of the process. Ligeiro et al. (2024) found that AI-enabled recruitment and selection processes are positively associated with organizational attractiveness, applicants' intrinsic motivation, perceived novelty, and trust in the process, with higher trust and more favorable perceptions increasing candidates' willingness to apply, participate, and complete the selection procedure. This line of research is important because it shows that the success of AI in recruitment depends not only on the accuracy of algorithms but also on how the technology is perceived by candidates as participants in the decision-making process.

The user perspective is equally salient on the side of recruiters. Almeida et al. (2025) apply TAM to AI acceptance in recruitment and show that perceived usefulness and ease of use shape recruiters' attitudes toward AI tools, which subsequently influence their intention to use them. Their findings further suggest that statistically significant relationships exist among perceived usefulness, perceived ease of use, attitude, and intention to use, with the model explaining a substantial proportion of the variance in the intention to adopt AI tools. Such findings move the research beyond the simple claim that AI is "useful" toward a more precise understanding of the mechanisms that lead to its actual acceptance in HR practice.

3.2 Fairness, Transparency, and Governance of AI in HR Decision-Making

A second prominent thematic line in the literature concerns fairness, explainability, and governance in AI-enabled HRM. This line builds on the broader concept of responsible AI and starts from the premise that personnel decisions belong among high-stakes decisions because they directly affect individuals' professional opportunities. In their review of the empirical literature, Bujold et al. (2024) show that research on responsible AI in HRM focuses primarily on bias and discrimination, trust and perceived fairness, privacy, transparency and explainability, as well as preserving the role of humans in the decision-making process. Their findings suggest that ethical principles in HR cannot be viewed as secondary add-ons to technology but as conditions of its legitimacy and acceptance.

Questions of fairness are particularly well developed in the literature on algorithmic candidate screening. In scenario-based research, Cai et al. (2024) show that applicants attribute higher procedural and distributive fairness to a human evaluator than to an AI system. They also find that perceptions of fairness are moderated by outcome favorability and the perceived level of AI expertise: when the outcome is positive or when AI is presented as a specialized expert, perceived fairness increases, but it still does not reach the level attributed to a human decision-maker. This result is important because it demonstrates that the problem of fairness in AI recruitment is not merely technical, but also perceptual and relational.

The importance of governance is further confirmed by more recent stakeholder-oriented studies. Malin et al. (2025) show that, from the perspective of workers' representatives, the most prominent concerns regarding AI in HRM relate to transparency, accountability, objectivity, and human oversight, with these concerns being especially strong in personnel selection, retention, HR analyses, and development and training. The authors also identify six core requirements for the adoption of AI in HRM, some organizational in character—namely control, responsibilities, and human oversight—and others technical, including lawful AI, transparency, and data security. These findings move the literature from abstract ethical principles toward more concrete organizational and technical conditions of implementation.

An equally important dimension of this discussion is diversity, equity, and inclusion. In their systematic review, Naoum et al. (2026) emphasize the dual impact of AI on DEI: when tools are

well designed, AI may increase consistency and contribute to fairer decisions, but it may also intensify existing inequalities if it is based on biased data or deployed without adequate organizational and ethical oversight. This conclusion is of central importance for HRM because it suggests that fairness cannot be ensured solely through the technical accuracy of a model; it also requires participatory design, human oversight, and ethical governance mechanisms.

3.3 AI in Talent Management, Retention, and Employee Experience

A third significant line of research concerns the use of AI beyond recruitment itself, particularly in talent management, retention, and broader employee outcomes. It is in this area that the literature suggests a shift from narrowly process-oriented uses of AI to a more strategic use of HR data. Madanchian (2024) explicitly argues that the literature on AI in HR can no longer remain confined to hiring but must also include onboarding and retention, because only these domains reveal whether AI contributes to longer-term value creation within organizations. This shift is important because it moves the discussion from the efficiency of individual tasks to the question of how AI reshapes broader decision-making about human capital.

In talent management, AI is most commonly associated with talent identification, employee development, the personalization of learning, and support for career-related decisions. Faqihi and Miah (2023) frame AI-driven talent management systems as an area of high potential, while at the same time highlighting the need for a stronger theoretical foundation and the risks that arise when AI is adopted without adequate risk governance. Their argument is complemented by the study of Rožman et al. (2022), which shows that AI-supported talent management, training and development, leadership, and organizational culture are positively associated with employee engagement and enterprise performance. This stream of literature therefore suggests that AI in talent management can function not only as a tool for more efficient decision-making but also as a means of strengthening employee engagement and performance.

A marked increase in research can also be observed in the area of retention and turnover prediction. Marín Díaz et al. (2023) show that explainable AI may be particularly valuable here because it offers not only attrition prediction itself but also an interpretation of the factors that drive it. In the HRM context, this aspect is crucial: without explainability, models may be analytically interesting, but they are much more difficult to use in practice because HR managers need to know which areas to target through interventions, development efforts, or changes in working conditions. The literature thus increasingly suggests that the value of AI in retention lies not only in prediction but in bridging data analysis and managerial action.

Alongside performance and retention indicators, broader effects of AI on employees are also being studied with increasing intensity. In their bibliometric and systematic review, Soulami et al. (2024) show that research is increasingly oriented toward employee well-being, workplace stress, and the future position of employees in AI-enabled workplaces. At the same time, Rožman et al. (2022) emphasize that AI can reduce workloads in routine tasks and thereby support engagement, productivity, and better work coordination, although this positive line of inquiry stands in tension with literature that highlights uncertainty, control, and fears of replacement. In this area, therefore, the literature does not offer a one-sidedly optimistic picture but rather a dual perspective in which the potential benefits of AI intersect with risks to employees' psychological well-being and job security.

3.4 Synthesis of the Literature Review

The synthesis of the analyzed literature reveals three relatively stable findings. First, the core of research on AI in HRM remains concentrated in recruitment and selection, where the benefits of automation, as well as the limits associated with trust and acceptance, are most clearly visible. Second, a growing portion of the literature shifts attention to fairness, explainability, and governance, thereby emphasizing that AI in HR is not only a technical issue but also an ethical and organizational one. Third, research is gradually expanding into talent management, retention, and employee outcomes, where AI is increasingly understood as a tool of strategic human capital management rather than merely a means of operational efficiency.

4. Research Gaps and Future Directions

4.1 The Dominance of Recruitment and the Limited Coverage of Other HR Functions

The first prominent gap in the literature is the uneven distribution of research attention across HR functions. Current review studies show that the largest volume of research is concentrated in recruitment and selection, whereas onboarding, learning and development, performance management, talent management, and retention remain far less systematically developed. Although Madanchian (2024) explicitly shifts the discussion "from recruitment to retention," the very need for such a review indicates that the literature has long been unbalanced. Likewise, Dima et al. (2024) show that AI is changing a wide range of HR activities, yet empirical research still does not cover these domains evenly. Future research should therefore focus more strongly on how AI is used in individual HR functions beyond recruitment and on the specific mechanisms, benefits, and risks that emerge in different parts of the employee lifecycle.

This gap also has an important theoretical implication. As long as the core of research remains focused mainly on hiring, there is a risk that AI in HRM will be understood too narrowly as a tool for candidate selection rather than as a broader sociotechnical phenomenon that reshapes decision-making about human capital within organizations. Future research should therefore seek to develop functionally differentiated models that take into account the differences among recruitment, talent management, development, performance evaluation, and retention, rather than subsuming all these areas under the general label of "AI in HRM." This is partly suggested in the integrative framework of Dima et al. (2024), but the literature still lacks a sufficient number of comparative studies across individual HR functions.

4.2 The Lack of Longitudinal and Post-Adoption Studies

A second important gap lies in the predominance of cross-sectional and largely exploratory research designs. Review articles show that current research often captures AI in HRM at the point of introduction, in terms of perceived benefits, risks, or user attitudes, but pays far less attention to what happens after implementation over time. Soulami et al. (2024) demonstrate growing interest in employee well-being and the broader effects of AI on employees, yet their systematic analysis also indicates that the field is still emerging and that research remains primarily descriptive rather than longitudinal. Similarly, the stakeholder-oriented study by Malin et al. (2025) provides valuable insight into requirements and concerns surrounding AI adoption, but it reflects mainly the stage of acceptance and rule setting rather than long-term observation of organizational consequences.

For future research, this means a shift from asking whether AI has been adopted to asking how AI is actually used, how its use changes over time, and what secondary consequences it produces for HR processes, employees, and organizational outcomes. Longitudinal studies

should examine whether initial adoption leads to stabilized usage, underutilization of the technology, or, conversely, to the expansion of AI into additional HR activities. It will also be important to investigate how employee and managerial trust changes over time, how perceptions of fairness evolve, and whether originally declared benefits translate into measurable outcomes in performance, retention, or workplace well-being. Based on the available literature, this post-adoption dynamic appears to be one of the least developed areas of current research.

4.3 Weak Integration of the Technical and Social Perspectives

A third gap is the persistent separation between the technical and social lines of research. Bujold et al. (2023/2024) explicitly note that earlier reviews often failed to adequately account for the multidomain character of AI in HRM and for the combination of the technical aspect—how AI is used—with the social aspect—what principles of responsible AI are applied in this use. Their review is important precisely because it attempts to connect these two dimensions, but in doing so it also indirectly shows that much of the previous literature had examined them separately. A similar problem can be seen in the audit literature, where Kazim et al. (2021) systematize risks, bias, transparency, and audit opportunities in algorithmic recruitment, but such approaches remain only weakly connected with organizational behavior, employee reactions, and HR decision-making practice.

Future research should therefore deliberately connect explainability, auditability, bias mitigation, and governance with the way AI actually enters HR decision-making practice and with how it is perceived by candidates, employees, managers, and other stakeholders. Particularly promising is the integration of explainable AI approaches with research on trust, procedural justice, and perceived legitimacy. Without such bridging, part of the research will remain overly technical and part overly behavioral, thereby weakening the field's ability to develop more robust theoretical models of AI in HRM. More comprehensive models should examine not only whether a system is accurate or fair, but also whether it is organizationally usable, socially accepted, and practically controllable.

4.4 The Need for a Broader Stakeholder Perspective and Stronger Contextualization

A fourth gap is the insufficient embedding of research in a broader stakeholder and organizational context. Malin et al. (2025) show that workers' representatives view AI in HRM critically across all HR phases, especially in personnel selection, and identify six key requirements: control, human oversight, responsibilities, transparency and explainability, lawful AI, and data security. This finding suggests that the adoption of AI in HRM is not merely a matter of technological readiness or recruiters' attitudes, but also a question of labor legitimacy, the organization of accountability, and acceptance by broader workplace constituencies. Even so, stakeholder-specific research remains the exception rather than the rule.

Closely related to this is the need for stronger contextualization of research. Naoum et al. (2026) examine AI-augmented HRM through a DEI lens and emphasize that the impact of AI on fairness and inclusion may be either positive or negative depending on how systems are designed, what data they use, and the organizational setting in which they operate. This suggests that findings cannot automatically be generalized across industries, countries, and organizational types. Future research should therefore pay greater attention to contextual comparisons—for example, between large firms and SMEs, between more and less regulated

sectors, or between different cultural and labor-law environments. Without such contextual grounding, a substantial part of the recommendations will remain overly general.

4.5 Future Research Directions

On the basis of the identified gaps, several directions for future research can be proposed. First, research should move beyond its dominant focus on recruitment toward a broader examination of AI across business functions within HRM, especially in onboarding, learning and development, performance management, and retention. Second, longitudinal studies are needed to capture post-adoption dynamics and to show whether AI becomes a stable component of HR decision-making or remains only a partially used technology. Third, future research should more systematically connect technical themes such as explainability, auditing, and bias mitigation with social themes such as trust, fairness, autonomy, and employee well-being. Fourth, there is a clear need to strengthen stakeholder-oriented research that takes seriously the perspectives not only of recruiters and managers but also of employees, candidates, workers' representatives, and other actors in the organizational environment.

Overall, the further development of research on AI in HRM will require a move away from narrowly technological or narrowly process-oriented perspectives. Future literature will be most useful when it is able to integrate functional, ethical, organizational, and behavioral dimensions into a single analytical framework. Such a direction could better explain not only where AI is used in HR, but also under what conditions it generates sustainable value for organizations while remaining fair, understandable, and acceptable to the people it affects.

5. Conclusion

This literature review examined the use of artificial intelligence in human resource management, with specific attention to recruitment, fairness and governance, talent management, retention, and broader employee-related effects. Based on the analyzed literature, it can be concluded that AI in HRM is gradually moving from the realm of isolated technological tools into that of a broader sociotechnical and strategic phenomenon. The literature shows that AI can improve the efficiency of multiple HR processes, support data-driven decision-making, and expand the analytical capacities of HR departments, while at the same time raising new questions related to fairness, transparency, trust, and human oversight (Dima et al., 2024; Fenwick et al., 2024; Madanchian, 2024).

The first key finding is that recruitment and selection remain the dominant domain of both research and practical AI application. It is here that the benefits of automation, acceleration, and standardization of processes are most visible, but also the limitations associated with perceived fairness, candidate trust, and technology acceptance among recruiters (Almeida et al., 2025; Ligeiro et al., 2024). The literature therefore suggests that the success of AI in recruitment depends not only on the technical accuracy of systems but also on their social acceptability and on the extent to which AI outputs are understandable and legitimate to those involved in the decision-making process.

The second major finding is the growing importance of the ethical and governance perspective. Research on responsible AI in HRM shows that fairness, explainability, privacy, auditability, and human oversight are among the central conditions of legitimacy for AI systems in personnel management. Of particular importance is the fact that AI may have a dual effect with respect to diversity, equity, and inclusion: when appropriately designed, it may contribute to more consistent decision-making, but with inadequate data and insufficient oversight it may

reproduce or intensify existing inequalities (Bujold et al., 2024; Kazim et al., 2021; Naoum et al., 2026). It follows that the technical functionality of AI alone is insufficient and must be complemented by clearly defined rules of accountability, transparency, and control.

A third important finding is that the literature is gradually extending beyond recruitment, particularly into talent management, retention, and employee outcomes. The research suggests that AI can support talent identification, personalized development, turnover prediction, and more precisely targeted HR interventions, with explainable AI being especially important in this context because it connects analytical models with practical managerial decision-making (Faqihi & Miah, 2023; Marín Díaz et al., 2023; Rožman et al., 2022). At the same time, however, the literature warns that the effects of AI on employees are not unambiguously positive, since increased efficiency and potentially higher engagement coexist with themes of uncertainty, stress, control, and fears of replacement (Soulami et al., 2024).

The review also identified several research gaps. The most prominent is the excessive concentration of research in recruitment, while other HR functions remain less systematically examined. Another important weakness is the lack of longitudinal studies and the limited integration of technical and social perspectives, which constrains the literature's ability to explain the longer-term organizational consequences of using AI in HRM. Future research should therefore focus on broader coverage of individual HR functions, post-adoption dynamics, stakeholder-oriented inquiry, and a stronger connection between explainability, fairness, governance, and actual HR practice (Malin et al., 2025; Soulami et al., 2024).

This study also has certain limitations. It is based on a selected set of more recent scholarly sources and does not constitute a fully exhaustive mapping of the entire literature on AI in HRM. Moreover, the thematic structuring of the article necessarily involves a degree of interpretive selection, since individual studies often overlap across multiple domains, such as recruitment, ethics, and employee outcomes. Despite these limitations, however, the article offers a coherent view of the main thematic lines of research and provides a foundation for further investigation of AI in human resource management as a field in which technological innovation, organizational decision-making, and human consequences intersect.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-ImpactSK); VEGA 1/109/24.

References

Aizenberg, E., Dennis, M. J., & van den Hoven, J. (2025). Examining the assumptions of AI hiring assessments and their impact on job seekers' autonomy over self-representation. AI & Society, 40, 919–927. https://doi.org/10.1007/s00146-023-01783-1

Albaroudi, E., Mansouri, T., & Alameer, A. (2024). A comprehensive review of AI techniques for addressing algorithmic bias in job hiring. AI, 5(1), 383–404. https://doi.org/10.3390/ai5010019

Almeida, F., Junça Silva, A., Lopes, S. L., & Braz, I. (2025). Understanding recruiters' acceptance of artificial intelligence: Insights from the Technology Acceptance Model. Applied Sciences, 15(2), 746. https://doi.org/10.3390/app15020746

Bujold, A., Roberge-Maltais, I., Parent-Rocheleau, X., Boasen, J., Sénécal, S., & Léger, P.-M. (2024). Responsible artificial intelligence in human resources management: A review of the empirical literature. AI and Ethics, 4, 1185–1200. https://doi.org/10.1007/s43681-023-00325-1

Cai, F., Zhang, J., & Zhang, L. (2024). The impact of artificial intelligence replacing humans in making human resource management decisions on fairness: A case of resume screening. Sustainability, 16(9), 3840. https://doi.org/10.3390/su16093840

Căvescu, A. M., & Popescu, N. (2025). Predictive analytics in human resources management: Evaluating AIHR's role in talent retention. AppliedMath, 5(3), 99. https://doi.org/10.3390/appliedmath5030099

Dalain, A. F., & Yamin, M. A. (2025). Examining the influence of AI-supporting HR practices towards recruitment efficiency with the moderating effect of anthropomorphism. Sustainability, 17(6), 2658. https://doi.org/10.3390/su17062658

Dima, J., Gilbert, M.-H., Dextras-Gauthier, J., & Giraud, L. (2024). The effects of artificial intelligence on human resource activities and the roles of the human resource triad: Opportunities and challenges. Frontiers in Psychology, 15, 1360401. https://doi.org/10.3389/fpsyg.2024.1360401

Faqihi, A., & Miah, S. J. (2023). Artificial intelligence-driven talent management system: Exploring the risks and options for constructing a theoretical foundation. Journal of Risk and Financial Management, 16(1), 31. https://doi.org/10.3390/jrfm16010031

Fenwick, A., Molnar, G., & Frangos, P. (2024). Revisiting the role of HR in the age of AI: Bringing humans and machines closer together in the workplace. Frontiers in Artificial Intelligence, 6, 1272823. https://doi.org/10.3389/frai.2023.1272823

García-Madurga, M.-Á., Gil-Lacruz, A.-I., Saz-Gil, I., & Gil-Lacruz, M. (2024). The role of artificial intelligence in improving workplace well-being: A systematic review. Businesses, 4(3), 389–410. https://doi.org/10.3390/businesses4030024

Jiang, L., Qin, X., Yam, K. C., Dong, X., Liao, W., & Chen, C. (2023). Who should be first? How and when AI-human order influences procedural justice in a multistage decision-making process. PLOS ONE, 18(7), e0284840. https://doi.org/10.1371/journal.pone.0284840

Kassir, S., Baker, L., Dolphin, J., & Polli, F. (2023). AI for hiring in context: A perspective on overcoming the unique challenges of employment research to mitigate disparate impact. AI and Ethics, 3, 845–868. https://doi.org/10.1007/s43681-022-00208-x

Kazim, E., Koshiyama, A. S., Hilliard, A., & Polle, R. (2021). Systematizing audit in algorithmic recruitment. Journal of Intelligence, 9(3), 46. https://doi.org/10.3390/jintelligence9030046

Koman, G., Boršoš, P., & Kubina, M. (2024). The possibilities of using artificial intelligence as a key technology in the current employee recruitment process. Administrative Sciences, 14(7), 157. https://doi.org/10.3390/admsci14070157

Ligeiro, N., Dias, I., & Moreira, A. (2024). Recruitment and selection process using artificial intelligence: How do candidates react? Administrative Sciences, 14(7), 155. https://doi.org/10.3390/admsci14070155

Madanchian, M. (2024). From recruitment to retention: AI tools for human resource decision-making. Applied Sciences, 14(24), 11750. https://doi.org/10.3390/app142411750

Malin, C., Fleiß, J., & Thalmann, S. (2025). Stakeholder-specific adoption of AI in HRM: Workers' representatives' perspective on concerns, requirements, and measures. Frontiers in Artificial Intelligence, 8, 1561322. https://doi.org/10.3389/frai.2025.1561322

Marín Díaz, G., Galán Hernández, J. J., & Galdón Salvador, J. L. (2023). Analyzing employee attrition using explainable AI for strategic HR decision-making. Mathematics, 11(22), 4677. https://doi.org/10.3390/math11224677

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. Advance online publication. https://doi.org/10.1007/s11301-025-00580-y

Necula, S.-C., Fotache, D., & Rieder, E. (2024). Assessing the impact of artificial intelligence tools on employee productivity: Insights from a comprehensive survey analysis. Electronics, 13(18), 3758. https://doi.org/10.3390/electronics13183758

Rožman, M., Oreški, D., & Tominc, P. (2022). Integrating artificial intelligence into a talent management model to increase the work engagement and performance of enterprises. Frontiers in Psychology, 13, 1014434. https://doi.org/10.3389/fpsyg.2022.1014434

Soulami, M., Benchekroun, S., & Galiulina, A. (2024). Exploring how AI adoption in the workplace affects employees: A bibliometric and systematic review. Frontiers in Artificial Intelligence, 7, 1473872. https://doi.org/10.3389/frai.2024.1473872

Tanantong, T., & Wongras, P. (2024). A UTAUT-based framework for analyzing users' intention to adopt artificial intelligence in human resource recruitment: A case study of Thailand. Systems, 12(1), 28. https://doi.org/10.3390/systems12010028

Tinguely, P. N., Lee, J., & He, V. F. (2023). Designing human resource management systems in the age of AI. Journal of Organization Design, 12, 263–269. https://doi.org/10.1007/s41469-023-00153-x

Xu, G., Xue, M., & Zhao, J. (2023). The relationship of artificial intelligence opportunity perception and employee workplace well-being: A moderated mediation model. International

Journal of Environmental Research and Public Health, 20(3), 1974. https://doi.org/10.3390/ijerph20031974

Artificial Intelligence in Customer Relationship Management: A Literature Review of Chatbots, Predictive Analytics, and Customer Loyalty Outcomes

Adam Čardáš[1] and Zuzana Hajduová[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: adam.cardas@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: zuzana.hajduova@euba.sk

Abstract

The integration of artificial intelligence into customer relationship management (CRM) is changing how companies acquire customers, keep them, and build long-term value. Although adoption in practice is growing quickly and the number of empirical studies is increasing, research on the combined impact of AI-driven CRM tools remains scattered. This applies in particular to conversational agents (chatbots), predictive analytics and customer lifetime value (CLV) modelling, which are often examined separately across the fields of information systems, marketing and strategic management. This literature review brings together 17 peer-reviewed studies published between 2019 and 2024. Its aim is to summarise the current state of knowledge, highlight the main theoretical approaches and point out key research gaps. The analysis focuses on three areas: first, the role of AI chatbots in shaping customer satisfaction and loyalty; second, the use of predictive analytics and CLV models to support proactive customer retention; and third, the organisational and ethical challenges linked to the implementation of AI in CRM. The findings suggest that factors such as chatbot service quality, perceived efficiency and trust play an important role in how AI influences customer loyalty. At the same time, predictive CLV models show noticeably higher accuracy compared to traditional RFM-based approaches. However, several challenges remain unresolved, including concerns about data privacy, limited transparency of algorithms, the complexity of system integration, and the lack of long-term performance evidence. Addressing these issues represents an important direction for future research, particularly in the context of Central and Eastern European markets.

Key Words: artificial intelligence, customer relationship management, CRM, chatbot, customer loyalty, customer lifetime value, predictive analytics, literature review

JEL Classification: D12, M31, O33

1. Introduction

Customer relationship management has changed significantly since it first emerged as a strategic discipline in the 1990s. What was originally a tool for structuring customer data and supporting sales processes has developed into a data-driven capability that increasingly relies on artificial intelligence. Today, companies use AI to better understand customer needs, personalise communication and estimate future customer value. Technologies such as machine learning, natural language processing and deep learning are now a standard part of many CRM systems across industries. They support a wide range of applications, from conversational customer service tools to real-time churn prediction and more advanced loyalty programme management enhanced by AI (Ozay et al., 2024).

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

The business case for AI in CRM is increasingly clear. Empirical studies show that companies using predictive CLV models can achieve higher revenues, mainly due to more precise customer segmentation and better allocation of retention investments (Egorenkov, 2024). At the same time, chatbots based on large language models are helping to shorten response times and improve first-contact resolution. When applied to less complex interactions, they can reach customer satisfaction levels comparable to those delivered by human agents (Hsu & Lin, 2023). Despite this progress in practice, academic research is not keeping pace. Although bibliometric reviews confirm the growing interest in AI within CRM, they also highlight a lack of comprehensive studies that would connect key areas such as chatbots, predictive analytics and customer loyalty into a unified perspective (Ozay et al., 2024; Saura et al., 2021).

This literature review aims to address the identified gap by bringing together findings from 17 peer-reviewed studies and offering a clearly organised view of how AI-driven CRM tools influence customer loyalty. It focuses on three main objectives. First, it identifies the theoretical frameworks most commonly used at the intersection of AI, CRM and customer loyalty. Second, it summarises empirical insights across three key thematic areas. Third, it outlines a research agenda that responds to the most important gaps in current knowledge. The topic is particularly relevant for Slovakia and the wider Central and Eastern European region, where the adoption of AI in CRM is growing rapidly, but local empirical evidence is still relatively limited.

The paper is structured as follows. Section 2 outlines the theoretical background and explains the review methodology. Section 3 presents the findings through a thematic synthesis built around three analytical areas. Section 4 highlights the main research gaps and suggests directions for future research. Section 5 concludes the paper.

2. Theoretical Background and Methodology

This review uses a literature review approach, following the framework proposed by Tranfield et al. (2003) and later applied to AI in marketing by Mariani et al. (2022). The literature was collected by searching Web of Science, Scopus, and Google Scholar with terms combining "artificial intelligence" or "machine learning" with "customer relationship management" or "CRM", as well as "customer loyalty", "chatbot", "customer lifetime value" and "predictive analytics". Studies were included if they were peer-reviewed, published in English between 2019 and 2024, and offered a meaningful theoretical or empirical contribution to understanding how AI-driven CRM tools affect customer loyalty or retention. Grey literature, practitioner reports, and papers focused solely on technical model performance without implications for marketing or management were excluded. The final set of 17 studies was organised thematically through iterative reading and coding.

The analysis is organised around three thematic clusters identified through iterative reading: first, AI-powered chatbots and their role in driving customer satisfaction and loyalty; second, predictive analytics and CLV modelling as tools for managing customer retention; and third, the organisational and ethical challenges involved in integrating AI into CRM. These clusters align with the main streams of AI-CRM research highlighted in recent bibliometric studies (Ozay et al., 2024; Saura et al., 2021) and reflect the key pathways through which AI affects customer loyalty outcomes.

Theoretical Frameworks

The reviewed literature draws on a heterogeneous set of theoretical frameworks reflecting the interdisciplinary nature of AI-CRM research. The Resource-Based View (RBV) of the firm, which positions unique and inimitable capabilities as sources of sustained competitive advantage, provides the dominant strategic lens for understanding why AI-CRM integration generates performance differentiation. Goswami and Rainy (2024) explicitly situate their synthesis within the RBV, arguing that AI-integrated CRM systems constitute dynamic capabilities that enable firms to continuously reconfigure their customer engagement strategies in response to shifting behavioural signals. This framing links AI-CRM investment to the broader literature on digital transformation and IT-enabled competitive advantage.

At the customer level, the Technology Acceptance Model (TAM) and its extensions remain the main framework for understanding why consumers adopt or resist AI-powered CRM tools. For example, Hsu and Lin (2023) use TAM concepts, especially perceived usefulness and perceived ease of use, to explain satisfaction with chatbots and the intention to continue using them. Their findings show that both practical and relational assessments of chatbot quality influence outcomes related to loyalty. The S-O-R (Stimulus-Organism-Response) framework has also been applied, viewing AI-driven service interactions as environmental stimuli that shape internal cognitive and emotional states, which then affect loyalty behaviour.

Trust appears across theoretical frameworks as the key factor mediating the relationship between AI-powered CRM and customer loyalty. Unlike traditional CRM interactions, AI-driven touchpoints create uncertainty about system competence, transparency of intent, and the accuracy of algorithmic recommendations. Several studies draw on trust theory from information systems and consumer psychology to show how perceived chatbot reliability, response accuracy, and empathic ability influence the cognitive and emotional foundations of customer loyalty (Hsu & Lin, 2023; Zhand et al., 2023). The idea of Explainable AI (XAI), which refers to the extent to which AI systems can justify their recommendations in ways humans can understand, is also emerging as a relevant theoretical extension, especially in high-stakes CRM settings such as financial services and healthcare (Zhang et al., 2023).

Customer Lifetime Value theory constitutes a third foundational framework, bridging marketing strategy and analytical modelling. Traditional CLV frameworks conceptualise the customer relationship as a stream of future cash flows discounted to present value, with retention probability as the key variable (Kumar & Reinartz, 2018). AI-powered CLV models extend this framework by replacing static retention probability estimates with dynamic, behaviourally-informed predictions generated by machine learning algorithms that incorporate hundreds of transactional, demographic, and engagement signals (Egorenkov, 2024).

3. Thematic Synthesis of the Literature

AI-Powered Chatbots and Customer Loyalty

The chatbot strand constitutes the most extensively researched area at the intersection of AI-CRM and customer loyalty. Jenneboer et al. (2022), in their systematic literature review of 27 studies, identify four primary mechanisms through which chatbots influence customer loyalty: service convenience and availability, personalisation quality, perceived empathy, and resolution effectiveness. Their synthesis confirms that customers exhibit stronger loyalty intentions toward brands deploying chatbots when the chatbot is perceived as responsive,

contextually accurate, and capable of handling the complexity of the customer's query without human escalation. Critically, the authors find that chatbot-driven loyalty is contingent on service quality: poorly designed chatbots that fail to resolve queries or communicate in natural language can actively damage brand relationships.

Hsu and Lin (2023), studying chatbot satisfaction and loyalty in a retail setting with 412 participants, show that chatbot service quality affects customer loyalty through two pathways. First, it has a direct effect through cognitive trust in the system's competence, and second, an indirect effect through affective satisfaction with the service encounter. Their results indicate that the quality of AI-mediated interactions goes beyond simply solving problems and also includes relational aspects such as perceived warmth, naturalness of communication, and a sense of being acknowledged. This supports Pitardi and Marriott's (2021) finding that anthropomorphic cues strongly influence trust in voice-based AI assistants, and extends the principle to text-based chatbots.

A clear pattern in the literature is that task complexity moderates the relationship between chatbots and customer loyalty. For routine, low-complexity tasks such as order tracking, FAQ responses, or basic account management, customers often see AI as equal or superior to human agents and report higher loyalty when AI services are available continuously and respond instantly. In contrast, for complex, emotionally charged, or ambiguous interactions, customers tend to prefer human agents and may respond negatively if limited to AI-only channels (Hsu & Lin, 2023; Nicolescu & Tudorache, 2022). This suggests that chatbot strategies in AI-CRM should focus on selective automation rather than replacing humans entirely.

Predictive Analytics, CLV Modelling, and Retention Management

The second thematic cluster concerns the use of AI-powered predictive analytics, particularly customer lifetime value (CLV) modelling and churn prediction, as tools for proactive customer retention management. Traditional CLV frameworks, which rely on Recency-Frequency-Monetary (RFM) analysis or simple probabilistic models, have clear limitations. They depend heavily on historical purchase data and cannot incorporate real-time behavioural signals that might indicate changing customer needs or intentions. AI-based CLV models address these limitations by using advanced machine learning algorithms, including gradient boosting, neural networks, and survival analysis. These models can integrate multiple data sources, such as transactional history, behavioural patterns, demographic information, and engagement metrics, to generate dynamic, customer-level predictions of future value. By doing so, firms can more accurately identify high-value customers, anticipate churn risks, and tailor retention strategies accordingly, creating measurable improvements in revenue and long-term customer loyalty (Egorenkov, 2024; Goswami & Rainy, 2024).

The performance advantages of AI-based CLV models over traditional approaches are well documented in the reviewed literature. Egorenkov (2024) reports that firms adopting AI-powered predictive CLV models achieve accuracy improvements of 20–35% over RFM-based benchmarks, with churn prediction models incorporating behavioural signals reaching 85–92% accuracy under controlled conditions. A case study of a major retail organisation demonstrates that AI-driven CLV segmentation enabled the reallocation of retention investment toward high-predicted-value customers, resulting in measurable improvements in purchase frequency and revenue per customer. In financial services, CLV models integrating transaction history, account activity, and customer service interaction data have been shown to identify at-risk high-value customers at sufficient lead time to enable targeted retention intervention (Egorenkov, 2024).

The link between predictive analytics and customer loyalty works through two main mechanisms. The first is proactive churn prevention. By detecting early behavioural signals of disengagement, such as a drop in purchase frequency, lower interaction levels, or unresolved service issues, AI-CRM systems can trigger personalised retention actions before loyalty is significantly affected. The second mechanism is CLV-based personalisation. By segmenting customers according to predicted lifetime value, companies can tailor loyalty programme design, communication frequency, and offer structures to different value tiers, making retention efforts more targeted and efficient (Ozay et al., 2024; Kumar & Reinartz, 2018).

The B2B context introduces additional complexity. Uddin et al. (2023) demonstrate that CLV prediction in B2B Software-as-a-Service settings requires hierarchical modelling approaches that account for the heterogeneity of organisational customer populations, the multi-product nature of B2B relationships, and temporal data constraints arising from longer sales cycles. These methodological requirements differentiate B2B AI-CLV research from the predominantly B2C focus of the broader literature, identifying a substantive gap in the current evidence base.

Organisational and Ethical Challenges of AI-CRM Integration

The third thematic cluster addresses the organisational and ethical dimensions of AI-CRM adoption, which the reviewed literature consistently identifies as the primary constraints on value realisation. Ozay et al. (2024), in their bibliometric review of 810 scholarly articles spanning two decades, document four major challenge categories: data quality and integration complexity, organisational capability gaps, ethical and regulatory concerns, and the interpretability of AI-generated recommendations.

Data quality and integration complexity represent the most frequently cited operational barrier. AI-powered CRM systems require unified, high-quality customer data repositories that integrate transactional, behavioural, and interaction data across channels - a condition that many organisations, particularly SMEs, lack due to legacy system fragmentation and siloed data governance structures. Saura et al. (2021) identify data architecture readiness as the single strongest predictor of AI-CRM adoption success, noting that organisations with mature data governance capabilities consistently outperform those with fragmented data environments in realising AI-driven loyalty improvements.

Ethical issues, especially concerning data privacy, algorithmic fairness, and transparency in AI-driven customer decisions, are becoming an important focus in the AI-CRM literature. Using AI to predict customer value and allocate retention resources selectively raises questions about unequal service quality and the risk that models trained on biased historical data may reproduce discriminatory treatment of certain customer segments (Egorenkov, 2024). The introduction of the EU AI Act (2024) and the General Data Protection Regulation (GDPR) creates a regulatory framework that limits several high-impact AI-CRM applications, including predictive scoring for credit-linked loyalty programmes, behavioural profiling for dynamic pricing, and AI-mediated personalisation involving sensitive customer data. Companies operating in Slovakia and the wider Central and Eastern European region must carefully comply with these regulations to ensure responsible AI-CRM implementation.

Explainability appears as a recurring and unresolved challenge in AI-CRM. Zhang et al. (2023) show that customers' willingness to engage with AI-powered systems is strongly affected by their perceived understanding of how AI recommendations are generated. This finding underlines the business relevance of research on Explainable AI (XAI) for applications that directly interact with customers. When customers see AI suggestions as opaque or

potentially manipulative, their trust, and therefore their loyalty, is undermined, even if the recommendation is objectively sound. These results have clear implications for the design of AI-CRM interfaces and communication strategies, suggesting that mechanisms promoting transparency may be just as important commercially as the accuracy or performance of the algorithms.

4. Research Gaps and Future Directions

A first key gap in the literature is the lack of longitudinal evidence on the impact of AI-CRM on customer loyalty. Most studies reviewed rely on cross-sectional surveys or short-term case studies that measure loyalty-related outcomes at a single point in time. However, the relationship between AI-powered CRM and customer loyalty is inherently dynamic. Trust in AI systems develops gradually through repeated interactions, churn prevention benefits accumulate over multiple retention cycles, and gains in customer lifetime value become visible only over longer-term relationships. Research using longitudinal designs, such as panel studies that track loyalty metrics over several periods or natural experiments linked to AI-CRM implementation, is therefore urgently needed to understand the sustainability of AI-driven loyalty outcomes and to differentiate lasting loyalty effects from short-term satisfaction.

Second, the integration of chatbot and predictive analytics research streams remains underdeveloped. The existing literature tends to treat AI-powered chatbots and predictive CLV modelling as separate topics examined by distinct research communities. In practice, however, these technologies are increasingly deployed as complementary components of integrated AI-CRM architectures: chatbot interaction data informs CLV models, and CLV predictions trigger personalised chatbot communication strategies. Research examining the combined and potentially synergistic effects of these AI-CRM components on loyalty outcomes would substantially advance both theoretical understanding and managerial guidance.

Third, the Central and Eastern European (CEE) context represents a significant gap in the literature. Most existing studies are based on samples from North America, Western Europe, or East Asia, with little attention given to AI-CRM adoption in the CEE region. The region has distinctive characteristics, including rapidly advancing digital transformation, specific GDPR implementation conditions, generally lower organisational AI maturity, and culturally shaped patterns of consumer trust. As a result, findings from other regions cannot be assumed to apply directly. Research focused on AI-CRM adoption barriers, customer trust in AI-mediated services, and AI-driven CLV outcomes in Slovakia and the wider CEE context would provide valuable contributions to regional scholarship and the broader AI-ImpactSK research agenda.

Fourth, the ethical aspects of AI-CRM, especially the fairness and explainability of predictive CLV models and churn scoring systems, require much more theoretical and empirical investigation. While current research recognises these issues, empirical examination is limited. As regulatory requirements under the EU AI Act and GDPR grow stricter, organisations need practical, evidence-based guidance on designing AI-CRM systems that are both effective and ethically responsible. Research exploring the intersection of Explainable AI, responsible AI practices, and CRM performance would provide valuable insights for both academia and policy-making.

Fifth, the context of small and medium-sized enterprises (SMEs) is systematically underrepresented in AI-CRM research. Most existing studies focus on large enterprises that have mature data infrastructures and dedicated analytical teams. SMEs, which represent the vast majority of businesses and employment in Slovakia and across the EU, operate under fundamentally different conditions. They face tighter resource constraints, have limited data

governance capabilities, and depend more on cloud-based CRM platforms with embedded AI functionality. Research that investigates how SMEs can effectively adopt and leverage AI-CRM tools to enhance customer loyalty, and identifies which organisational capabilities most strongly predict successful adoption in these contexts, would address a gap with direct relevance for both managerial practice and policy development.

5. Conclusion

This literature review has synthesised 17 peer-reviewed studies to provide an analytically organised account of the current state of knowledge on AI-powered CRM and its implications for customer loyalty. Three principal conclusions emerge from this synthesis.

First, AI-powered chatbots and predictive analytics both produce measurable outcomes relevant to customer loyalty, but they operate through different mechanisms. Chatbots affect loyalty mainly by shaping perceptions of service quality, building trust, and driving satisfaction, with their effectiveness strongly influenced by the complexity of the task and the match between AI capabilities and customer interaction needs. In contrast, predictive analytics and CLV modelling affect loyalty through proactive churn prevention and segmentation-based personalisation. These approaches generate loyalty outcomes over longer time horizons and require ongoing investment in organisational capabilities to fully realise their potential.

Second, trust emerges as the pivotal construct mediating the relationship between AI-CRM deployment and customer loyalty across both thematic clusters. The design of transparency mechanisms, explainability features, and consent architectures in AI-CRM systems is therefore not merely a regulatory compliance requirement but a commercially strategic decision that directly influences the loyalty outcomes firms seek to generate.

Third, despite the clear commercial importance of AI-CRM, the academic literature remains fragmented across disciplinary communities. Research continues to concentrate on large firms and Western market contexts, and there is a notable lack of longitudinal evidence. Closing these gaps is essential, particularly through studies that integrate the insights from both chatbot and predictive analytics streams, investigate AI-CRM adoption and customer loyalty in Central and Eastern European contexts, and capture the long-term dynamics of AI-driven loyalty outcomes. This represents the most pressing priority for advancing understanding in the field.

For marketing practitioners, these findings highlight that realising the value of AI-CRM requires investment not only in technology, but also in the organisational capabilities that allow these technologies to build and sustain customer loyalty. Such capabilities include data governance, analytical expertise, cross-functional integration, and ethical AI frameworks, all of which are essential for generating trust-based and long-lasting customer relationships.

Funding

This paper is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I05-03-V02-00003/2025/VA (AI-impactSK); Vega 1/0109/24.

References

Egorenkov, D. (2024). AI-powered predictive customer lifetime value: Maximizing long-term profits. International Journal of Scientific Research and Management, 12(9), EC-2024-1430–EC-2024-1437. https://ijsrm.net/index.php/ijsrm/article/view/5672

Eliashberg, J., Elberse, A., & Leenders, M. A. A. M. (2006). The motion picture industry: Critical issues in practice, current research, and new research directions. Marketing Science, 25(6), 638–661. https://doi.org/10.1287/mksc.1050.0177

European Commission. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689

Goswami, D., & Rainy, T. (2024). Mechanisms by which AI-enabled CRM systems influence customer retention and overall business performance: A systematic literature review of empirical findings. SSRN Working Paper. https://doi.org/10.2139/ssrn.5317471

Hsu, C. L., & Lin, J. C. C. (2023). Understanding the user satisfaction and loyalty of customer service chatbots. Journal of Retailing and Consumer Services, 71, 103211. https://doi.org/10.1016/j.jretconser.2022.103211

Jenneboer, L., Herrando, C., & Constantinides, E. (2022). The impact of chatbots on customer loyalty: A systematic literature review. Journal of Theoretical and Applied Electronic Commerce Research, 17(1), 212–229. https://doi.org/10.3390/jtaer17010011

Kumar, V., & Reinartz, W. (2018). Customer relationship management: Concept, strategy, and tools (3rd ed.). Springer. https://doi.org/10.1007/978-3-662-55381-7

Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. Journal of Marketing, 80(6), 69–96. https://doi.org/10.1509/jm.15.0420

Mariani, M. M., Perez-Vega, R., & Wirtz, J. (2022). AI in marketing, consumer research and psychology: A systematic literature review and research agenda. Psychology & Marketing, 39(4), 755–776. https://doi.org/10.1002/mar.21619

Nicolescu, L., & Tudorache, M. T. (2022). Human-computer interaction in customer service: The experience with AI chatbots — A systematic literature review. Electronics, 11(10), 1579. https://doi.org/10.3390/electronics11101579

Ozay, D., Simsek, S., & Tabak, A. (2024). Artificial intelligence (AI)-based customer relationship management (CRM): A comprehensive bibliometric and systematic literature review with outlook on future research. Enterprise Information Systems, 18(7), 2351869. https://doi.org/10.1080/17517575.2024.2351869

Palmatier, R. W., Kumar, V., & Harmeling, C. M. (Eds.). (2018). Customer engagement marketing. Palgrave Macmillan.

Pitardi, V., & Marriott, H. R. (2021). Alexa, she's not human but… Unveiling the drivers of consumers' trust in voice-based artificial intelligence. Psychology & Marketing, 38(4), 626–642. https://doi.org/10.1002/mar.21457

Saura, J. R., Ribeiro-Soriano, D., & Palacios-Marqués, D. (2021). Setting B2B digital marketing in artificial intelligence-based CRMs: A review and directions for future research. Industrial Marketing Management, 98, 161–178. https://doi.org/10.1016/j.indmarman.2021.08.006

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

Uddin, M. N., Haque, M., & Islam, S. (2023). A novel approach to predicting customer lifetime value in B2B SaaS companies. Journal of Marketing Analytics, 11, 412–429. https://doi.org/10.1057/s41270-023-00234-6

Zhang, Y., Hua, L., Jiao, Y., Zhang, J., & Saini, R. (2023). More than watching: An empirical and experimental examination on the impacts of live streaming user-generated video consumption. Information & Management, 60(3), 103771. https://doi.org/10.1016/j.im.2023.103771

Organizational Dynamics and Temporal Patterns in AI Adoption: An Analytical Synthesis of Leadership, Departmental Influence, and Implementation Sequencing

Mohammadreza Tatlari[1] and Nora Grisáková[2]

[1] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: mohammandreza.tatlari@euba.sk

[2] Bratislava University of Economics and Business, Faculty of Business Management, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia (EU); e-mail: nora.grisakova@euba.sk

Abstract

This paper examines organizational AI adoption through leadership, departmental dynamics, and temporal sequencing. It synthesizes existing knowledge on how artificial intelligence is embedded in organizations, focusing on governance arrangements, cross-functional variation, organizational adaptation, and implementation pacing. Methodologically, the paper is an exploratory analytical study based on secondary sources, combining scientific literature, survey-based analytical reports, and selected organizational examples.

The synthesis indicates that AI adoption is uneven across organizational functions and depends not only on technological readiness, but also on leadership involvement, workflow redesign, governance mechanisms, and the ability to align experimentation with strategic oversight. The reviewed evidence also shows that implementation typically unfolds through iterative stages of piloting, adjustment, and scaling rather than as a single-step deployment. Temporal sequencing therefore emerges as an important analytical dimension of AI adoption, closely linked to organizational learning and adaptation.

The paper contributes to the current discussion by offering an analytically structured perspective that connects organizational dynamics with implementation timing and the governance-experimentation tension. Rather than presenting original primary empirical findings, it develops a synthesis of existing knowledge that supports a more context-sensitive understanding of AI adoption in organizational settings.

Key Words: Artificial intelligence adoption; Organizational dynamics; Leadership; Departmental influence; Temporal sequencing; AI governance; Organizational adaptation; Innovation management

JEL Classification: O33, O31, M15

1. Introduction

Organizations across sectors are increasingly integrating artificial intelligence (AI) into their operations, decision-making processes, and service delivery models. As a result, AI adoption has become not only a technological issue, but also an organizational one, involving changes in governance arrangements, managerial responsibilities, workflow design, and cross-functional coordination. While prior research has examined technological, environmental, and strategic determinants of AI adoption, less attention has been devoted to the internal

Financované Európskou úniou NextGenerationEU | PLÁN OBNOVY | AI IMPACT SK

organizational dynamics through which adoption unfolds over time (Neumann et al., 2024; Radhakrishnan & Chattopadhyay, 2020).

Existing literature suggests that AI adoption is shaped not only by technological readiness, but also by leadership involvement, interdepartmental differences, and the organization's capacity to align experimentation with strategic oversight. These issues become particularly important when AI implementation extends beyond isolated pilot projects and begins to affect multiple organizational functions, established routines, and governance structures (Romeo & Lacko, 2025; McKinsey & Company, 2025).

This paper approaches AI adoption from the perspective of organizational dynamics and temporal sequencing. It examines how governance structures, departmental variation, and implementation pacing interact in shaping adoption trajectories, while also drawing attention to interpretive flexibility as the capacity of organizations to adapt AI tools and implementation practices to specific institutional and operational contexts (Gunasekaran & Ngai, 2012; Wu et al., 2019).

The aim of the paper is to synthesize and interpret existing knowledge on organizational AI adoption with particular attention to three interconnected dimensions: leadership and governance, functional differences across departments, and the timing and sequencing of implementation. Rather than presenting original primary empirical findings, the paper develops an analytically structured synthesis based on scientific literature, survey-based analytical reports, and selected organizational examples.

The paper contributes to the current discussion in three ways. First, it highlights how leadership and governance arrangements influence the organizational embedding of AI. Second, it emphasizes that AI adoption unfolds unevenly across organizational functions and over time, making sequencing, piloting, and adaptation important analytical dimensions. Third, it interprets AI adoption as a process of organizational adjustment in which experimentation, governance, and institutional learning must be balanced.

The remainder of the paper is structured as follows. The next section outlines the research design and methodological approach. This is followed by a section synthesizing recurring patterns identified across the reviewed literature, analytical reports, and selected organizational materials. The final sections discuss the implications of these patterns and conclude by summarizing the main insights of the paper.

2. Research Design and Methods

This paper is an exploratory analytical study based on secondary sources. It synthesizes and interprets existing knowledge on organizational AI adoption, focusing on leadership structures, departmental dynamics, and the temporal sequencing of implementation. Rather than presenting original primary empirical research, the paper draws on scientific studies, survey-based analytical reports, and selected organizational materials to identify recurring themes, tensions, and implementation patterns across organizational contexts.

The methodological approach combines literature-based synthesis, qualitative content analysis, and the comparative interpretation of selected organizational examples. Its purpose is not statistical generalization, but an analytically grounded account of how AI adoption is shaped by governance arrangements, functional differences across departments, and implementation pacing. In this sense, the paper sits between theory-oriented synthesis and application-oriented interpretation.

The source base used for the purposes of this paper consisted of three groups of materials. The first group included scientific literature on AI adoption, technology adoption, organizational change, and governance. In this category, particular attention was paid to works such as Radhakrishnan and Chattopadhyay (2020), Neumann et al. (2024), Romeo and Lacko (2025), Gunasekaran and Ngai (2012), and Wu et al. (2019). These studies were used primarily to frame the conceptual discussion and to identify key determinants, barriers, and organizational implications of AI adoption. The second group consisted of recent analytical and industry reports addressing the organizational use of AI and generative AI. For the purposes of this paper, this category included, in particular, McKinsey's The State of AI and Capgemini Research Institute's Harnessing the Value of Generative AI, as these reports provide explicit methodological notes and empirically grounded observations on the adoption of AI across business functions, governance arrangements, and workflow redesign. The third group consisted of publicly accessible organizational materials used for the illustrative application-oriented dimension of the paper. These included company reports, public strategy documents, official communications, and analytical texts related to selected firms discussed in the article, such as JPMorgan, Walmart, Pfizer, Siemens, and Delta.

The analytical procedure followed three interconnected steps. First, the selected scientific literature and analytical reports were reviewed to identify the main thematic dimensions relevant to the paper: leadership and governance, differences in AI adoption across business functions, organizational adaptation mechanisms, and the role of sequencing, piloting, and scaling. Second, selected organizational examples were compared to illustrate how these dimensions appear in specific organizational and sectoral settings. Third, insights from the literature and reports were interpreted together with the selected organizational materials to identify recurring patterns and analytically relevant points of convergence.

The interpretive framework combined deductive and inductive logic. Initial thematic categories were derived from the literature on technology adoption, organizational transformation, and AI governance. These categories were then refined through the analysis of the selected reports and organizational materials. The analysis focused on recurring motifs and interpretive patterns rather than on testing causal relationships in a statistical sense.

Where the paper refers to percentages or other quantified statements, these should be understood as descriptive indicators derived from the reviewed secondary sources rather than as outputs of original data collection conducted by the authors. Such figures are used to illustrate broader tendencies and arguments reported in the existing literature and analytical reports. For this reason, the rhetoric of the article remains interpretive and cautious, emphasizing tendencies, recurring themes, and literature-supported observations rather than definitive causal claims.

This methodological approach has several limitations. As the paper is based on secondary evidence, its analytical strength depends on the methodological quality, transparency, and comparability of the original studies and reports on which it draws. In addition, the source base is heterogeneous, since scientific articles, consultancy reports, and organizational materials differ in purpose, design, and depth. The findings presented in the paper should therefore be understood as an analytically structured synthesis of existing knowledge supported by illustrative examples from organizational practice, rather than as the results of an original primary empirical study.

3. Synthesized Findings

Leadership and Departmental Dynamics

The reviewed literature and analytical reports suggest that leadership involvement plays an important role in shaping organizational AI adoption. In particular, recent survey-based evidence indicates that CEO oversight of AI governance is associated with stronger reported bottom-line effects of generative AI deployment, while 28% of respondents in organizations using AI reported that responsibility for AI governance rests with the CEO (McKinsey & Company, 2025). At the same time, the evidence suggests that AI adoption is not evenly distributed across organizational functions. Existing reports consistently indicate that AI use is most visible in IT and marketing and sales, with service operations also emerging as an important area of deployment (McKinsey & Company, 2025; Capgemini Research Institute, 2024). This pattern points to the uneven functional diffusion of AI, where adoption tends to concentrate in functions with clearer opportunities for automation, process optimization, or customer-facing value creation.

The reviewed sources also indicate that leadership support and governance arrangements matter not only for formal oversight but also for the pace and coordination of implementation across departments. Rather than suggesting a single universal governance model, the available evidence points to a tendency toward selective centralization of certain governance mechanisms combined with differentiated implementation across business functions (McKinsey & Company, 2025). This creates organizational tension between the need for control, risk management, and alignment on the one hand, and the practical need for experimentation and adaptation within individual departments on the other.

Interpretive Flexibility and Adaptation

The reviewed literature and analytical reports suggest that AI adoption is not limited to the technical deployment of new tools, but also involves substantial organizational adaptation. One recurring theme across the examined sources is interpretive flexibility, that is, the capacity of organizations to adjust AI systems and related workflows to specific operational contexts. Recent survey-based evidence indicates that some organizations have fundamentally redesigned at least part of their workflows in response to generative AI, suggesting that implementation often requires more than the simple insertion of a new technology into existing routines (McKinsey & Company, 2025). Table 1 summarizes selected organizational indicators that illustrate this broader process of adaptation.

Table 1. Selected organizational indicators of AI adoption and adaptation

Indicator	Value	Interpretation
Organizations in which the CEO oversees AI governance	28%	Indicates strong top-management involvement in AI governance
Organizations reporting fundamental redesign of at least some workflows due to generative AI	21%	Suggests that AI adoption often requires process-level adaptation
Organizations hiring AI compliance specialists	13%	Reflects growing emphasis on governance and regulatory alignment
Organizations hiring AI ethics specialists	6%	Indicates emerging but still limited formalization of AI ethics roles

Respondents reporting that up to 5% of employees were reskilled due to AI in the past year	44%	Suggests that reskilling is present, but often still limited in scope
Organizations most often using generative AI for text outputs	63%	Shows that text-based applications remain the most common use case
Organizations using generative AI for image outputs	36%	Indicates broader experimentation with non-textual outputs
Organizations using generative AI for code generation	27%	Highlights the importance of AI use in technical and development functions

Source: Authors' elaboration based on McKinsey & Company (2025).

Across the reviewed materials, organizational adaptation appears in several forms. These include reskilling and upskilling initiatives, adjustments to internal processes, the introduction of governance- or compliance-related roles, and experimentation with different forms of AI-supported output and task allocation (McKinsey & Company, 2025; Neumann et al., 2024). Rather than representing uniform responses, these adaptations vary across sectors and organizational functions, depending on the regulatory environment, the complexity of work processes, and the degree of managerial support. The examined sources therefore suggest that successful adoption depends not only on the availability of AI tools, but also on the organization's capacity to redesign work practices and align them with broader institutional and strategic requirements.

The literature and selected organizational examples also indicate that hybrid forms of human-AI collaboration remain important, particularly in contexts where accuracy, accountability, or regulatory compliance are critical (Neumann et al., 2024; Romeo & Lacko, 2025). In such settings, AI is often introduced in ways that supplement rather than fully replace existing work arrangements. This suggests that organizational adaptation should be understood not as a one-time adjustment, but as an ongoing process of aligning technological possibilities with human oversight, operational realities, and institutional constraints.

Temporal Sequencing Effects

The reviewed literature and analytical reports suggest that the timing and sequencing of AI implementation represent an important dimension of organizational adoption. Rather than unfolding as a single-step transformation, AI adoption typically progresses through stages of experimentation, piloting, adjustment, and broader organizational scaling (McKinsey & Company, 2025; Romeo & Lacko, 2025). In this sense, implementation pacing should be understood not merely as an operational detail, but as a strategic factor that influences whether AI initiatives remain isolated experiments or become embedded in organizational routines.

The examined sources also indicate that the pace of adoption differs across sectors and organizational settings. In practice, the sequencing of implementation tends to reflect differences in regulatory requirements, process complexity, data readiness, and managerial coordination. As a result, organizations do not move from initial experimentation to scaled deployment at the same speed, and the path from pilot initiatives to wider organizational use often depends on the ability to align technological opportunities with workflow redesign, governance structures, and cross-functional cooperation. Recent survey-based evidence further suggests that organizations are increasingly recognizing the importance of redesigning workflows as part of AI deployment, which reinforces the view that sequencing and adaptation are closely interconnected rather than analytically separate processes (McKinsey & Company, 2025).

From this perspective, phased implementation may be interpreted as an organizational mechanism for reducing uncertainty and supporting learning during adoption. The reviewed materials suggest that AI implementation is more sustainable when organizations combine experimentation with iterative adjustment, rather than treating deployment as a purely technical rollout (Accenture, 2022; Romeo & Lacko, 2025). This is particularly relevant in settings where implementation affects multiple departments or where operational reliability, compliance, and human oversight remain central concerns.

Governance-Experimentation Tension

The reviewed literature and analytical reports suggest that one of the central challenges of organizational AI adoption lies in balancing formal governance with the need for experimentation. Survey-based evidence indicates that organizations continue to face a gap between early experimentation and broader institutionalization. For example, Accenture reports that only 12% of firms have reached an advanced level of AI maturity, while 63% are still largely testing the waters. The same report also suggests that only 30% of AI pilot initiatives are subsequently scaled, pointing to the difficulty of transforming experimentation into organization-wide value creation (Accenture, 2022).

At the same time, recent evidence indicates that stronger governance arrangements are increasingly associated with more robust organizational outcomes. McKinsey reports that CEO oversight of AI governance is one of the elements most strongly correlated with higher self-reported bottom-line impact from generative AI, while 28% of respondents in organizations using AI state that responsibility for overseeing AI governance rests with the CEO. This suggests that experimentation alone is insufficient unless it is connected to clear strategic accountability and organizational coordination (McKinsey & Company, 2025).

The tension between speed and control is also visible in recent Deloitte findings. Deloitte reports that many leaders feel pressure to realize value from generative AI quickly while simultaneously managing risks, and 60% of surveyed organizations report that they are effectively balancing rapid implementation with risk management. At the same time, Deloitte identifies trust, workforce implications, and scaling challenges as persistent barriers to broader deployment. Taken together, these findings suggest that the governance-experimentation relationship should not be understood as a trade-off between rigid control and unrestricted innovation, but rather as a continuing process of aligning experimentation, trust, and strategic oversight (Deloitte, 2024).

The literature further suggests that organizational roles related to compliance, ethics, data governance, and implementation coordination are becoming more important as AI adoption matures, even though their specific institutional form varies across organizations (McKinsey & Company, 2025; Romeo & Lacko, 2025). In this sense, effective AI adoption appears to depend less on a single ideal governance model than on the organization's ability to connect strategic oversight with iterative learning and cross-functional adaptation.

4. Discussion

The reviewed literature, analytical reports, and selected organizational examples suggest that AI adoption is shaped not only by technological readiness, but also by internal organizational conditions that influence how implementation unfolds over time. In particular, the findings synthesized in this paper point to the importance of leadership involvement, cross-functional variation, organizational adaptation, and implementation sequencing as mutually connected dimensions of adoption. This perspective supports earlier work emphasizing that AI

adoption should be understood as a broader organizational process rather than a purely technical act of deployment (Neumann et al., 2024; Radhakrishnan & Chattopadhyay, 2020).

One of the clearest patterns in the reviewed sources is that AI adoption is uneven across organizational functions. Implementation is often most visible in IT, marketing and sales, and service operations, where opportunities for automation, optimization, and customer-facing value creation are easier to identify (McKinsey & Company, 2025; Capgemini Research Institute, 2024). AI diffusion within organizations therefore does not follow a uniform path. Instead, it reflects differences in functional priorities, task characteristics, risk exposure, and managerial support. Departmental variation should thus be viewed as a structural feature of organizational AI adoption rather than as a temporary implementation imbalance.

The analysis also highlights the importance of interpretive flexibility and organizational adaptation. AI adoption often requires changes in workflows, skill structures, governance arrangements, and human oversight mechanisms. This supports the view that AI implementation is not simply a matter of introducing tools into existing routines, but often involves redesigning work practices and aligning them with institutional and operational demands (McKinsey & Company, 2025; Romeo & Lacko, 2025). These findings are consistent with broader organizational perspectives that stress the interaction between strategy, structure, and operational practice in shaping innovation outcomes (Gunasekaran & Ngai, 2012; Wu et al., 2019).

A further point emerging from the synthesis is the relevance of temporal sequencing. The reviewed sources indicate that the pace and order of implementation matter for whether AI remains confined to local experimentation or becomes embedded more broadly in organizational routines. Pilot initiatives, phased deployment, iterative adjustment, and workflow redesign appear to be closely interrelated rather than separate dimensions of implementation. This suggests that timing should be regarded as a meaningful analytical dimension of AI adoption, especially in organizations where implementation involves multiple departments, governance constraints, and varying levels of digital maturity (Accenture, 2022; McKinsey & Company, 2025).

The discussion also points to a continuing tension between experimentation and governance. Organizations need structures that address accountability, compliance, and strategic alignment while preserving room for learning, adaptation, and bottom-up experimentation. Governance should therefore not be interpreted solely as a mechanism of control, but also as a coordinating framework that can enable or constrain the sustainable scaling of AI across the organization (Deloitte, 2024; McKinsey & Company, 2025). This tension becomes particularly visible when organizations move beyond isolated pilots and seek wider institutionalization of AI-supported processes.

From a practical perspective, organizations are more likely to benefit from AI when implementation is treated as an organizational transformation process rather than as a narrowly technical initiative. Leadership support, cross-functional coordination, gradual scaling, and the redesign of workflows and capabilities appear to matter more than the mere presence of AI tools. At the same time, the reviewed evidence does not support the idea of a single universal governance model. Successful adoption appears to depend on the organization's ability to combine strategic oversight with iterative learning and context-sensitive adaptation.

The study has several limitations that should be acknowledged in interpreting these conclusions. Because the paper is based on secondary sources, its analytical strength depends on the methodological transparency and comparability of the studies, reports, and

organizational materials included in the review. In addition, the source base is heterogeneous, combining scientific literature, consultancy reports, and selected organizational examples. The discussion should therefore be understood as an analytically structured synthesis of existing knowledge rather than as a basis for strong causal claims or universal generalization. Future research could build on this perspective by comparing sector-specific trajectories of AI adoption, examining the longer-term organizational consequences of phased implementation, and exploring how governance arrangements evolve as AI use matures across functions.

5. Conclusion

This paper has examined organizational AI adoption through the lens of leadership, departmental variation, and temporal sequencing. Drawing on scientific literature, survey-based analytical reports, and selected organizational examples, it argues that AI adoption should be understood not merely as a technological implementation issue, but as a broader organizational process shaped by governance arrangements, functional differences, and implementation pacing.

The synthesis suggests that AI adoption unfolds unevenly across organizational contexts and functions. Leadership involvement, workflow redesign, organizational adaptation, and phased implementation appear to shape whether AI remains limited to isolated experimentation or becomes embedded in organizational routines. Sustainable adoption, in turn, depends on the organization's capacity to align strategic oversight with iterative learning and cross-functional coordination.

The paper contributes to the current discussion by offering an analytically structured perspective that connects organizational dynamics with implementation timing and adaptation processes. Given the secondary-source nature of the paper, the conclusions should be understood as interpretive rather than universally generalizable. Future research could further examine sector-specific adoption trajectories, governance models, and the evolution of AI-related organizational roles as implementation matures.

Funding

The article is a partial output of research project VEGA 1/0085/25, titled: 'Synergistic framework for designing modular models in corporate strategic decision-making using machine and deep learning'; 100% share.

References

Accenture. (2022). *The art of AI maturity: Advancing from practice to performance*. Accenture. Online: https://www.accenture.com/content/dam/system-files/acom/custom-code/ai-maturity/Accenture-Art-of-AI-Maturity-Report-Global-Revised.pdf

Capgemini Research Institute. (2024). *Harnessing the value of generative AI: Top use cases across sectors*. Capgemini Research Institute. Online: https://www.capgemini.com/wp-content/uploads/2024/05/Final-Web-Version-Report-Gen-AI-in-Organization-Refresh.pdf

Deloitte. (2024). *The state of generative AI in the enterprise: Q2 report*. Deloitte. Online: https://www.deloitte.com/ce/en/services/consulting/research/state-of-generative-ai-in-enterprise.html

Gunasekaran, A., & Ngai, E. W. T. (2012). The future of operations management: An outlook and analysis. International Journal of Production Economics, 135(2), 687-701. https://doi.org/10.1016/j.ijpe.2011.11.002

McKinsey & Company. (2025). *The state of AI: How organizations are rewiring to capture value*. McKinsey & Company. Online: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-how-organizations-are-rewiring-to-capture-value

Neumann, O., Guirguis, K., & Steiner, R. (2024). Exploring artificial intelligence adoption in public organizations: A comparative case study. Public Management Review, 26(1), 114-141. https://doi.org/10.1080/14719037.2022.2048685

Radhakrishnan, J., & Chattopadhyay, M. (2020). Determinants and barriers of artificial intelligence adoption: A literature review. In Re-imagining diffusion and adoption of information technology and systems: A continuing conversation (pp. 89-99). Springer. https://doi.org/10.1007/978-3-030-64849-7_9

Romeo, E., & Lacko, J. (2025). Adoption and integration of AI in organizations: A systematic review of challenges and drivers towards future directions of research. Kybernetes. Advance online publication. https://doi.org/10.1108/K-07-2024-2002

Wu, L.-F., Huang, I.-C., Huang, W.-C., & Du, P.-L. (2019). Aligning organizational culture and operations strategy to improve innovation outcomes: An integrated perspective in organizational management. Journal of Organizational Change Management, 32(2), 224-250. https://doi.org/10.1108/JOCM-07-2017-0283