Redefining Boundaries: A Comparative Analysis of Human-AI Collaboration Dynamics in European Healthcare and Financial Services

Abstract

Purpose: This paper examines how human-AI collaboration boundaries are conceptualized and implemented differently across two key European sectors: healthcare and financial services. Using a narrative literature review methodology and the AI-Human Collaboration Theoretical Framework, it analyzes sectoral variations in collaborative dynamics.

Design/methodology/approach: A narrative literature review approach is employed to synthesize existing knowledge across disciplines. The AI-Human Collaboration Theoretical Framework is applied to analyze sector-specific patterns in human-AI interaction based on professional identities, risk models, and expertise traditions. Hypothetical case illustrations demonstrate practical applications.

Findings: The research identifies distinct human-AI collaboration models across sectors: healthcare prioritizes clinical authority; and financial services implements tiered authority models based on risk profiles. Sectoral contexts significantly shape collaborative boundaries through professional traditions, regulatory environments, and knowledge integration patterns.

Research limitations/implications: Limitations include the theoretical approach and use of hypothetical cases, which may not capture the full complexity within sectors. Future research should validate findings through empirical studies, examine boundary evolution longitudinally, and examine cross-sector learning opportunities.

Originality/value: This research contributes to innovation management theory by demonstrating how sector-specific professional identities and expertise traditions shape collaborative boundaries between human and technological agents. It offers a structured comparative framework for analyzing human-AI collaboration and provides actionable insights for designing collaborative systems that improve human capabilities within sectoral contexts.

Keywords: Artificial intelligence, Human-AI collaboration, Innovation management, Healthcare, Financial services.

Introduction

The integration of Artificial Intelligence (AI) into organizational practices has fundamentally transformed contemporary management approaches across European sectors (Appio et al., 2023; Khvatova et al., 2024; Holmström, 2022; Shrestha et al., 2019). As these technologies continue to reshape business strategies, competition dynamics, and organizational structures (Haefner et al., 2023; Johnson et al., 2022), a critical research gap has arisen regarding how different sectors conceptualize and implement human-AI collaboration. While significant attention has been devoted to the technical aspects of AI implementation, less focus has been placed on understanding how sectoral contexts influence the boundaries established between human and AI agents in collaborative settings.

The selection of healthcare and financial services sectors for comparison was determined by three key criteria: their contrasting professional identity structures, differential regulatory frameworks, and divergent approaches to algorithmic decision-making. Healthcare represents a domain where professional autonomy and ethical considerations dominate technology adoption, with sector-specific governance influencing the use of algorithms (Kellogg et al, 2020). In contrast, financial services exemplify data-driven environments with established quantitative traditions, particularly in areas such as consumer credit, where algorithmic decision-making raises concerns about bias, fairness, and normative outcomes (Sargeant, 2023). These distinctions provide optimal conditions for examining how sectoral contexts shape collaborative boundaries.

This research responds directly to calls for deeper study of AI's transformative impact on organizational paradigms and management practices. As highlighted in contemporary discourse, AI solutions are prompting profound changes that necessitate an integral reassessment of acceptance and adoption patterns across different organizational contexts (Khvatova et al., 2024; Appio et al., 2023). The emergence of generative AI further complicates this landscape, requiring critical examination of the distinctions between different classifications of AI and their unique contributions to innovation management (Ooi et al., 2023; Sætra, 2023).

The purpose of this article is to analyze how European healthcare and financial services sectors differ in their conceptualization of human-AI boundaries. Specifically, this research seeks to formulate a sectoral typology that extends current theoretical understanding of human-AI collaboration beyond generic frameworks. This comparative perspective is studied through the research question: How do European healthcare and financial services sectors construct distinct human-AI collaboration processes? The investigation of this question is particularly significant for innovation management theory and practice, as it provides insights into how organizations can design sector-appropriate models for human-AI collaboration that correspond with established professional practices, regulatory environments, and knowledge traditions.

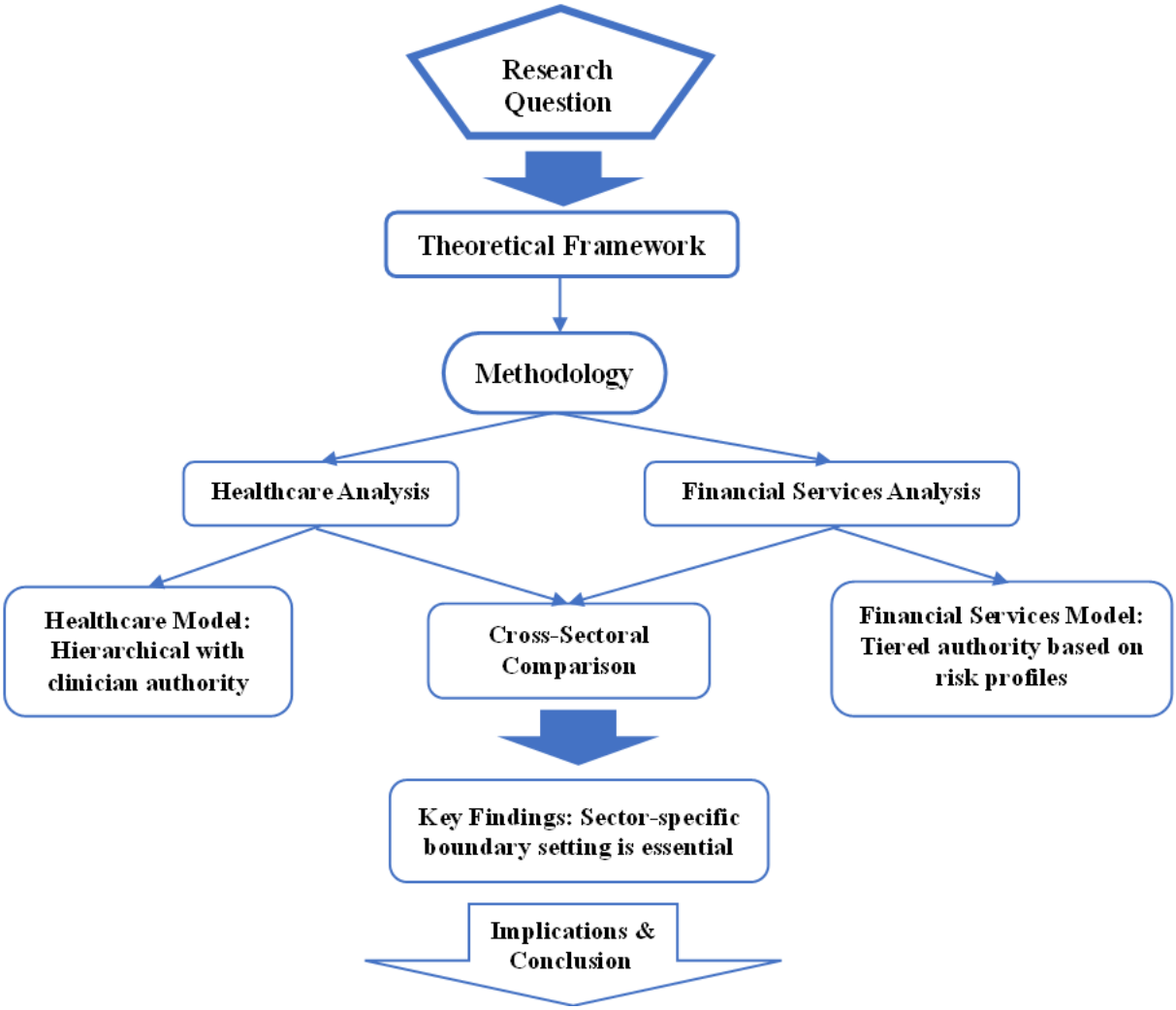
This research contributes to the evolving discourse on AI-enabled transformation in three significant ways. First, it advances theoretical understanding of how sector-specific professional identities and expertise traditions shape collaborative boundaries between human and technological agents (Cannavale et al., 2022; Duan et al., 2019). Second, it offers a structured comparative framework for analyzing human-AI collaboration across different organizational contexts, addressing the need for more sophisticated approaches to AI implementation that acknowledge sectoral variations (Sharma et al., 2024; Chatterjee et al., 2021). Third, it provides actionable insights for innovation managers seeking to design collaborative systems that increase human capabilities within their sectoral context (Garbuio et al., 2021; Agostini et al., 2020).

Through this analysis, a contribution is made to the developing understanding of how organizations across diverse sectors navigate the integration of AI technologies into existing work processes and professional identities. This understanding is relevant for innovation managers seeking to implement AI solutions that complement established expertise and decision-making structures, ultimately increasing human capabilities and professional authority. Furthermore, by examining these processes through a sectoral lens, this research responds to growing recognition that AI adoption strategies must be contextualized instead of generic (Kopalle et al., 2022; Ancillai et al., 2023; Kumar et al., 2024), a perspective that fits with calls for more differentiated approaches to innovation management in AI-enabled environments (Tran & Murphy, 2023; Giuggioli & Pellegrini, 2023).

Theoretical Framework: Human-AI Collaboration Approaches

The following flow chart (Figure 1) visualizes the research structure and key findings of this comparative analysis. It illustrates the logical progression from the central research question through the theoretical framework and methodology to the sector-specific analyses and their synthesis. The chart highlights the distinct collaborative models identified in healthcare and financial services, along with the key differences that come up from cross-sectoral comparison and their implications for theory and practice. This visual representation provides an overview of how sector-specific factors shape human-AI collaboration relationships across European organizational contexts.

Figure 1. Theoretical Framework, Research Structure and Key Findings



Source: Authors

The conceptualization of human-AI collaboration has evolved through multiple theoretical perspectives, encompassing both supportive and critical viewpoints. As identified by Daly et al. (2025), trust emerges as a fundamental dimension in human-AI collaboration, with sectoral variations in how trust is established and maintained. This perspective is complemented by Jiang et al. (2023), who argue that effective collaboration extends beyond context-aware services to encompass flexible role negotiation between human and artificial agents. The user-centered perspective advanced by Jiang et al. (2024) further stresses that collaboration models must be designed around human needs and capabilities instead of technological possibilities.

The theoretical foundation of this analysis is anchored in the AI-Human Collaboration Theoretical Framework, which itself draws upon multiple interdisciplinary perspectives. Central to this framework is Distributed Cognition Theory (Hutchins, 1995), which conceptualizes cognitive processes as extending beyond individual mental spaces to be dynamically distributed across human and technological interfaces. This perspective fundamentally challenges traditional dichotomies between human intelligence and technological systems, positioning collaborative intelligence as an emergent ecosystem of knowledge generation (Holland, 1998; Woolgar, 1990).

However, critical perspectives must be acknowledged. As Razmak et al. (2025) demonstrate, tensions exist between complementing and replacing human labor in AI implementations. Power asymmetries come up when professional expertise is challenged by algorithmic authority, creating resistance patterns that vary by sector. Khosravy et al. (2024) underscore that social values must be integrated into collaborative systems, particularly in Industry 5.0 contexts where human-centricity is important. These critical viewpoints suggest that collaboration is not merely a technical challenge but involves complex negotiations of authority, expertise, and professional identity.

Building upon this foundation, the cognitive amplification layer as conceptualized by Clark (1998) provides a relevant theoretical lens through which the improvement of human cognitive capacities can be understood. These mechanisms are not merely supplementary but represent transformative potential for knowledge production, where AI systems and human intellect engage in a reciprocal, iterative process of mutual enhancement (Engelbart, 1962; Licklider, 1960). The knowledge co-production cycle within this framework represents a practice, non-

linear process of collaborative meaning-making. This cycle encompasses multiple stages: initial human contextual framing, computational analysis, collaborative interpretation, mutual refinement, and integrated knowledge deployment. Such a model stresses emergence and co-evolutionary potential and not predetermined technological determinism (Kauffman, 1995; Pickering, 1995).

From a labor sociology perspective, the integration of AI systems reconfigures traditional workplace hierarchies and professional jurisdictions. Policy frameworks at the European level, particularly regarding algorithmic accountability and worker protection, create additional layers of complexity that shape collaborative boundaries differently across sectors. These interdisciplinary insights reveal that human-AI collaboration cannot be understood solely through technological or organizational lenses but requires attention to broader social and political contexts.

Adaptive learning mechanisms form another critical component of this theoretical framework. Drawing from Bandura's (1986) Social Cognitive Theory, these mechanisms facilitate continuous mutual adaptation between human agents and artificial systems. The learning process is conceptualized not as a unidirectional transfer of information but as a complex, reciprocal interaction where both technological and human components continuously progress and transform (Trist, 1981). Philosophical perspectives on human-technological relationships, as articulated by Simondon (2017), further enrich the theoretical understanding applied in this analysis. The framework challenges traditional ontological boundaries between human and machine, positioning their interaction as a fluid process of mutual becoming that moves beyond instrumentalist views of technology (Floridi, 2014).

Ethical considerations are very relevant in this collaborative framework. Bostrom's (2014) critical examination of AI expansion provides relevant insights into potential risks and strategic considerations that must be incorporated into any analysis of human-AI collaboration. Merchán-Cruz et al. (2025) extend this ethical dimension by proposing trust-by-design principles for collaborative intelligence systems, stressing that ethical considerations must be embedded from inception instead of added retrospectively. This dimension ensures that collaborative interactions

remain fundamentally aligned with human values and societal well-being across different sectoral contexts (Agrawal, 2018; Brynjolfsson & McAfee, 2014).

Methodology

This study employs a narrative literature review approach for integral synthesis across diverse disciplinary perspectives (Juntunen & Lehenkari, 2021), selected for its suitability in analysing complex interdisciplinary topics where technology studies, organizational theory, and sector-specific literature must be integrated (Hart, 2018; Sposato, 2025). The review follows Torraco's (2016) integrative methodology, which facilitates examination of how human-AI collaboration dynamics manifest differently across sectors, identifying patterns and underlying factors that shape these differences (Juntunen & Lehenkari, 2021; Hart, 2018).

The literature selection process was conducted systematically through multiple stages. Initial searches were performed across Web of Science, Scopus, and Google Scholar databases using combinations of keywords including "human-AI collaboration," "healthcare AI," "financial services AI," and "sectoral AI adoption." Inclusion criteria encompassed peer-reviewed articles published between 2018-2025, focusing on European contexts, and addressing collaborative aspects rather than purely technical implementations. Exclusion criteria eliminated purely technical papers without organizational dimensions and studies outside European regulatory contexts. The initial search yielded 487 articles, which were reduced to 126 through title and abstract screening, with final selection of 73 articles based on full-text review and relevance assessment.

Thematic analysis was conducted through iterative coding processes. Initial codes were derived deductively from the theoretical framework, including "professional identity," "authority distribution," and "knowledge integration." Inductive codes arose through careful reading, identifying patterns such as "risk-based differentiation" and "regulatory influence." The coding process involved three rounds of refinement, with inter-coder reliability checks performed on a subset of 20 articles. Themes were synthesized through constant comparison, identifying convergent and divergent patterns across sectors.

The review focused on European organizational contexts, analyzing literature on AI implementation in healthcare and financial services to identify themes related to human-AI boundaries, professional identity, and knowledge integration (Torraco, 2016; Sposato, 2025). Based on this synthesis, sector-specific collaboration models were established and illustrated through hypothetical case studies constructed from patterns identified in the literature (Hart, 2018; Juntunen & Lehenkari, 2021). The construction of hypothetical cases followed a structured process: key patterns from literature were identified, composite scenarios were formulated incorporating multiple real-world elements, and validation was performed through cross-referencing with empirical studies to ensure representativeness. This approach enables structured comparison while acknowledging the context-dependent nature of human-AI collaboration, providing richer understanding of how sectoral contexts shape collaborative boundaries (Torraco, 2016; Sposato, 2025).

Sectoral Analysis: Healthcare

Theoretical Analysis

The healthcare sector in Europe is distinguished by distinctive clinical knowledge traditions that significantly influence approaches to human-AI collaboration. These traditions underscore the integration of scientific evidence, clinical expertise, and patient values in medical decision-making (Secinaro et al., 2021). Within this context, evidence-based practice has become a dominant paradigm that shapes how new technologies, including AI, are integrated into clinical workflows.

However, the integration process is not without tensions. As Calisto et al. (2022) demonstrate, adoption of intelligent agents in medical imaging faces significant barriers related to trust, interpretability, and professional autonomy. These findings challenge assumptions about smooth technology integration, revealing that clinicians often resist AI systems that appear to infringe upon professional judgment. The modeling of adoption patterns reveals complex interactions between technological capabilities, organizational support, and individual clinician characteristics.

Evidence-based practice in European healthcare systems creates a specific framework for evaluating and implementing AI decision support tools. These tools must demonstrate technical performance and clinical validity within established care processes. This approach influences how AI systems are positioned within the clinical decision-making hierarchy, typically as supportive tools instead of autonomous decision-makers (Ali Mohamad et al., 2023).

The integration of AI into healthcare contexts is fundamentally shaped by the sector's distinctive professional identity structures and ethical frameworks. As underlined by Cannavale et al. (2022), healthcare professionals operate within established ethical codes that place primary stress on patient welfare and clinical responsibility. These ethical considerations directly influence how AI systems are positioned within clinical decision-making processes, with clear subordination to human clinical judgment being a defining characteristic of healthcare approaches to human-AI collaboration.

Unintended consequences of AI implementation in healthcare have been documented across multiple contexts. System failures, algorithmic biases, and over-reliance on automated recommendations have led to adverse patient outcomes in some cases. These experiences have reinforced conservative approaches to AI integration, with healthcare organizations implementing multiple validation layers and maintaining strict human oversight protocols. Institutional resistance emerges not merely from professional protectionism but from legitimate concerns about patient safety and accountability.

Professional autonomy and responsibility structures are particularly pronounced in healthcare contexts. Medical professionals operate under ethical and legal frameworks that place ultimate responsibility for patient outcomes on human clinicians. This responsibility structure necessitates clear boundaries in human-AI collaboration, with AI systems explicitly positioned as advisory as opposed to determinative in clinical decisions (Cannavale et al., 2022). The resulting collaborative models reflect what Secinaro et al. (2021) describe as "clinician-led" approaches to AI integration, where technological capabilities increase professional judgment.

The innovation adoption process in healthcare organizations follows distinctive patterns that reflect the sector's professional structures and risk considerations. As noted by Cannavale et al.

(2022), healthcare organizations typically implement highly structured evaluation processes for new technologies, with particular attention to clinical validity and potential impacts on established care processes. This approach influences how AI technologies are assessed and integrated, with accent on gradual implementation that preserves clinical authority while carefully evaluating outcomes.

Recent developments in personalized AI explanations, as examined by Calisto et al. (2025), suggest approaches to human-AI interaction in clinical settings. By adapting communication to clinician expertise levels, these systems acknowledge heterogeneity within the medical profession and recognize that collaborative models must be flexible. This personalization extends beyond technical features to encompass communication styles, uncertainty expression, and recommendation framing.

Patient-centered care values represent another significant influence on how AI is integrated into healthcare contexts. These values underline the importance of personalized, holistic approaches to patient care that respect individual preferences and circumstances. The integration of AI must therefore preserve the human relationship aspects of care while raising technical capabilities for diagnosis and treatment planning. This balance reflects broader tensions in healthcare innovation between technological advancement and the preservation of care quality (Secinaro et al., 2021).

Knowledge management considerations in healthcare AI implementation deserve particular attention. As outlined by Ali Mohamad et al. (2023), effective integration of AI into clinical practice requires sophisticated knowledge management approaches that facilitate the bidirectional flow of information between AI systems and clinical experts. This knowledge integration process must respect established clinical hierarchies while establishing the analytical capabilities available to healthcare professionals. The resulting knowledge ecosystems combine algorithmic analysis with clinical judgment in structured, carefully delineated relationships.

External validation studies, such as those conducted by Abrantes et al. (2023) on breast density classification models, reveal the importance of rigorous testing across diverse clinical populations. These validation processes often uncover performance variations that necessitate model refinement or restricted deployment contexts. Such findings reinforce the sector's cautious

approach to AI integration and underscore the gap between laboratory performance and clinical utility.

Regulatory constraints on AI application in healthcare are particularly stringent compared to other sectors. European regulations regarding medical devices, data protection, and patient safety create a complex compliance landscape that shapes how AI systems can be deployed in clinical settings (López Jiménez et al., 2023). These constraints often result in more conservative approaches to human-AI boundaries, with explicit limitations on AI autonomy in clinical decision-making. As identified by Cannavale et al. (2022), regulatory considerations significantly influence innovation adoption patterns in healthcare organizations, with particular attention to compliance and risk management in technology implementation.

Hypothetical Case Study: Nordic Regional Hospital Network

A public hospital network in the Nordic region illustrates typical human-AI collaboration approaches in the healthcare sector. This hypothetical organization has implemented an AI diagnostic support system to assist radiologists in identifying potential abnormalities in medical imaging. The implementation process was characterized by careful attention to professional roles, regulatory compliance, and integration with existing clinical workflows.

The collaboration model constructed in this context is explicitly clinician-led, with the AI system positioned as a consultative resource within the diagnostic process. Radiologists maintain complete control over the diagnostic workflow, using the AI system as a "second reader" that identifies potential areas of concern for further evaluation. The system provides probability assessments and comparative analyses but does not generate definitive diagnoses.

The implementation faced initial resistance from senior radiologists who perceived the system as challenging their expertise. Trust-building measures were implemented following principles identified by Daly et al. (2025), including transparency workshops, performance benchmarking sessions, and collaborative refinement processes. These interventions transformed skepticism

into cautious acceptance, with clinicians gradually recognizing the system's value in reducing fatigue-related oversights while preserving professional authority.

This approach complies with Secinaro et al.'s (2021) observations regarding the integration of AI in diagnostic contexts, where technological capabilities improve and not replace clinical judgment. The explicit positioning of the AI system as advisory instead of determinative reflects the sector's accent on professional autonomy and responsibility (Cannavale et al., 2022). By preserving the primacy of clinical judgment while raising analytical capabilities, the organization addresses both professional identity considerations and patient safety requirements.

Knowledge integration in this case follows structured pathways that respect clinical authority while increasing diagnostic capabilities. The AI system continuously learns from radiologist decisions, gradually improving its pattern recognition capabilities while preserving the primacy of human judgment in final diagnoses. This bidirectional knowledge flow represents an innovative approach to developing analytical capabilities (Ali Mohamad et al., 2023). The resulting knowledge ecosystem combines algorithmic precision with clinical judgment in a complementary relationship.

Boundary approaches in this case are characterized by a clear delineation of the AI advisory role versus clinical decision-making authority. The system's outputs are presented as suggestions instead of conclusions, with radiologists making independent assessments before reviewing the AI analysis. This approach preserves professional autonomy while benefiting from the pattern recognition capabilities of the AI system. The explicit boundary delineation reflects what Ali Mohamad et al. (2023) describe as "hierarchical" approaches to AI integration in healthcare, where technological capabilities improve professional judgment.

Challenges arose in multi-jurisdictional contexts when the hospital network expanded across national borders. Different regulatory interpretations and professional practice variations required adaptive implementation strategies (López Jiménez et al., 2025). The organization formulated a flexible framework allowing local customization while maintaining core safety protocols, illustrating the complexity of scaling AI solutions in healthcare beyond single regulatory environments.

The implementation process itself reflects the distinctive innovation adoption patterns of healthcare organizations. As observed by Cannavale et al. (2022), healthcare innovations typically undergo rigorous evaluation processes that assess technical performance and impacts on clinical workflows and patient outcomes. The phased implementation approach deployed by the organization, with careful evaluation at each stage, exemplifies these sectoral patterns in innovation adoption.

The outcomes of this collaboration model include augmented diagnostic efficiency while maintaining professional judgment. The organization has reported reductions in reading time and augmented detection rates for certain conditions, particularly in screening contexts. However, the primary authority for diagnostic decisions remains firmly with human clinicians, reflecting the sector's stress on professional responsibility and patient-centered care. This balance between efficiency enhancement and preservation of clinical authority exemplifies what Secinaro et al. (2021) describe as "patient-centered" approaches to AI integration in healthcare.

These collaborative patterns in healthcare reflect broader sectoral tendencies toward cautious, clinician-controlled integration of AI capabilities. The healthcare sector's approach to human-AI collaboration is characterized by explicit preservation of clinical authority, positioning of AI as advisory and not determinative, and careful attention to regulatory compliance and ethical considerations in technology implementation. This approach corresponds with Ali Mohamad et al.'s (2023) observation that successful AI implementation in healthcare requires attention to technical capabilities and to the professional dimensions of clinical practice.

Sectoral Analysis: Financial Services

Theoretical Analysis

The financial services sector in Europe has established data-driven decision traditions that create a distinctive context for human-AI collaboration. These traditions have progressed through the progressive integration of quantitative methods, statistical models, and computational approaches in financial analysis and decision-making (Fares et al., 2023). This history has created a sectoral

culture that is generally receptive to algorithmic approaches while maintaining awareness of their limitations.

Recent research by Xu and Cho (2025) reveals nuanced factors affecting human-AI collaboration performance in financial contexts. Their sustainable service growth perspective identifies that successful collaboration depends not merely on technological sophistication but on alignment with organizational sustainability goals, employee capability development, and stakeholder value creation. This multi-dimensional view challenges purely efficiency-oriented implementations.

The evolution of analytical methods in financial services has significantly influenced how organizations in this sector approach AI implementation. As pointed out by Fares et al. (2023), financial institutions have progressed through multiple generations of quantitative methods, from basic statistical approaches to sophisticated machine learning algorithms. This analytical heritage has created a sectoral context where algorithmic decision support is well-established, providing a foundation for the integration of more advanced AI capabilities. The resulting collaborative relationships reflect what Fares et al. (2023) describe as "graduated autonomy" approaches, where algorithmic authority varies based on context and risk considerations.

Risk management frameworks are particularly influential in shaping AI integration in financial services. These frameworks, crafted through regulatory requirements and industry best practices, establish structured approaches to identifying, assessing, and mitigating various forms of financial risk. AI systems are increasingly incorporated into these frameworks, contributing to risk assessment while operating within established governance structures (Fares et al., 2023). The resulting collaborative models reflect what Hilb (2020) describes as "risk-calibrated" approaches to AI governance, where the level of human oversight is proportional to the potential consequences of algorithmic decisions.

However, conflicting perspectives come up regarding optimal collaboration models. Some institutions have experienced algorithmic trading disasters and credit scoring biases that resulted in significant financial losses and reputational damage. These failures have prompted reassessment of human oversight levels, with some organizations reverting to more conservative

approaches despite efficiency costs. The tension between automation benefits and risk exposure continues to shape sectoral debates about appropriate boundaries.

The competitive practices of the financial services sector significantly influence AI implementation strategies. As identified by Johnson et al. (2022), financial institutions operate in highly competitive environments where technological capabilities can provide significant competitive advantages. This competitive pressure encourages innovation in AI applications while simultaneously raising concerns about algorithmic risk and regulatory compliance. The resulting tension shapes how financial organizations balance efficiency gains with appropriate risk management in their collaborative approaches.

Client trust dynamics represent another significant influence on human-AI boundaries in financial services. Financial institutions must balance the efficiency and analytical power of AI systems with the maintenance of client relationships and trust. This balance often results in the preservation of human touchpoints in client-facing processes, even as back-office functions become increasingly automated (Trittin-Ulbrich et al., 2021). The resulting boundary variations reflect what Fares et al. (2023) describe as "context-sensitive" approaches to AI integration, where human involvement is preserved in relationships-focused functions while allowing greater algorithmic autonomy in analytical processes.

Power asymmetries manifest differently in financial services compared to healthcare. While medical professionals maintain clear hierarchical authority, financial analysts often find their expertise commodified and replaced by algorithmic systems. This displacement creates resistance patterns characterized by skill hoarding, shadow systems, and passive non-compliance. Organizations addressing these approaches through reskilling programs and hybrid role creation report more successful implementations.

Knowledge integration patterns in financial services AI implementation deserve particular attention. As underlined by Kopalle et al. (2022), effective AI integration in financial contexts requires sophisticated approaches to combining algorithmic analysis with human expertise. This knowledge integration process typically involves structured interactions between AI systems and financial analysts, with explicit protocols for reconciling algorithmic recommendations with

human judgment. The resulting knowledge ecosystems combine computational power with contextual understanding in calibrated relationships that vary by function and risk level.

Regulatory compliance and explainability requirements create specific constraints on AI implementation in financial services. European regulations regarding financial services, data protection, and algorithmic decision-making establish requirements for transparency, non-discrimination, and explainability that shape how AI systems can be deployed. These requirements often necessitate human oversight of AI-generated recommendations or decisions, particularly in high-stakes contexts. As observed by Hilb (2020), governance considerations significantly influence how financial organizations design and implement AI systems, with particular attention to accountability and explainability in algorithmic processes.

Recent policy expansion at the European level, particularly the AI Act and revised financial services directives, create additional complexity. These frameworks mandate algorithmic impact assessments, regular audits, and citizen rights to human review of automated decisions. Financial institutions navigate these requirements through layered governance structures that preserve efficiency while ensuring compliance, though implementation costs and operational complexity have increased substantially.

Hybrid decision-making structures have come up as a common approach in financial services, combining algorithmic analysis with human judgment in varying proportions depending on the specific context and risk profile. These structures establish graduated boundaries between human and AI decision authority, with more routine, lower-risk decisions being more fully automated while complex or high-stakes decisions maintain greater human involvement. This graduated approach reflects what Fares et al. (2023) describe as "tiered autonomy" models, where algorithmic authority varies based on contextual factors and risk considerations.

Hypothetical Case Study: Pan-European Banking Group

A multinational banking group operating across multiple European countries provides an illustrative example of human-AI collaboration relationships in the financial services sector. This

hypothetical organization has implemented an AI system for credit decision support, integrating algorithmic analysis into established lending processes while maintaining compliance with diverse national regulations.

The collaboration model established in this context features algorithmic pre-screening with tiered human review based on case complexity and risk profile. The AI system conducts initial assessment of credit applications, analyzing financial data, transaction history, and external information to generate a risk assessment and preliminary recommendation. Applications are then routed through different human review pathways based on complexity, with straightforward cases receiving streamlined review while complex or borderline cases receive more intensive human evaluation.

Initial implementation revealed unexpected challenges in small and medium business (SMB) lending contexts. As identified by Ankrah et al. (2025), the social architecture of SMB relationships required preservation of human judgment for assessing non-quantifiable factors such as entrepreneurial capability, local market knowledge, and relationship history. The organization adapted its model to maintain human-centered approaches for SMB segments while automating consumer lending decisions.

This approach complies with Fares et al.'s (2023) observations regarding the integration of AI in credit decision processes, where algorithmic capabilities raise human judgment in graduated patterns that reflect risk considerations. The tiered review structure exemplifies what Hilb (2020) describes as "risk-calibrated governance" approaches to AI implementation, where the level of human oversight is proportional to the potential consequences of algorithmic decisions. By varying human involvement based on case complexity and risk profile, the organization addresses both efficiency objectives and governance requirements.

Knowledge integration in this case follows structured pathways that combine algorithmic analysis with human expertise in calibrated relationships. The AI system provides standardized assessment based on historical patterns, while human credit officers contribute contextual understanding and ethical judgment to the decision process. This integration of complementary capabilities represents what Fares et al. (2023) describe as "capability-enhancing" approaches to

AI implementation, where technological and human contributions are combined in ways that strengthen overall decision quality.

Boundary approaches in this case establish a tiered decision authority structure based on risk profile and complexity. The AI system has greater autonomy in routine, lower-risk decisions, while human credit officers retain more extensive authority over complex or unusual cases. This tiered approach enables efficiency gains while maintaining appropriate human oversight for situations requiring judgment, contextual understanding, or ethical consideration. The explicit boundary delineation reflects what Hilb (2020) describes as "graduated authority" approaches to AI governance, where technological and human contributions are calibrated based on context-specific factors.

Cross-border operational challenges arose when harmonizing AI systems across jurisdictions with varying regulatory requirements. National interpretations of explainability requirements, data localization rules, and consumer protection standards necessitated flexible architecture. The organization constructed a core-satellite model with standardized algorithmic components adapted through local compliance modules, illustrating the complexity of multi-jurisdictional AI deployment in financial services.

The implementation process itself reflects the distinctive innovation adoption patterns of financial organizations. As noted by Fares et al. (2023), financial institutions typically use structured approaches to technology implementation, with particular attention to risk management and regulatory compliance. The phased implementation approach employed by the organization, with careful validation at each stage, exemplifies these sectoral patterns in innovation adoption.

The outcomes of this collaboration model include increased efficiency in credit processing while maintaining human oversight for ethical considerations and complex cases. The organization has reported reductions in processing time for routine applications while maintaining or improving risk assessment accuracy. The preservation of human judgment for complex cases helps address concerns about algorithmic bias and ensures that unique circumstances can be appropriately considered. This balance between efficiency improvement and appropriate human oversight

exemplifies what Fares et al. (2023) describe as "balanced implementation" approaches to AI integration in financial services.

Performance metrics reveal complex outcomes beyond efficiency gains. Employee satisfaction initially declined due to role uncertainty but recovered following reskilling initiatives and clearer career pathways. Client satisfaction improved for routine transactions but required careful management in complex cases where human expertise remained valued. These mixed outcomes underscore the importance of holistic evaluation beyond technical performance metrics.

These collaborative patterns in financial services reflect broader sectoral tendencies toward tiered, risk-based approaches to human-AI boundaries. The financial services sector's approach to human-AI collaboration is characterized by data-driven decision processes, graduated automation based on risk profiles, and the maintenance of human oversight for regulatory compliance and client trust preservation. This approach conforms to Hilb's (2020) observation that successful AI implementation in financial services requires attention to technical capabilities and to accountability considerations.

Cross-Sectoral Comparison

Ethical Frameworks and Authority Distribution Models

A comparative analysis of boundary-setting mechanisms across healthcare and financial services reveals significant variations in how human-AI collaboration is conceptualized and implemented. In professional identity preservation, healthcare institutions focus on maintaining clinical authority with explicit subordination of AI systems to clinicians' judgment, while financial services adopt graduated approaches varying by function and risk level (Holmström, 2022; Trittin-Ulbrich et al., 2021; Kopalle et al., 2022).

The most pronounced difference lies in ethical frameworks governing AI integration. Healthcare ethics center on the primacy of patient welfare, informed consent, and the principle of "do no harm" (Secinaro et al., 2021; Cannavale et al., 2022). This creates rigid boundaries where AI

systems must demonstrably preserve clinical judgment and patient safety. Conversely, financial services ethics focus on fiduciary responsibility, algorithmic fairness, and transparency in automated decision-making (Hilb, 2020; Fares et al., 2023). These frameworks permit greater AI autonomy provided decisions remain explainable and non-discriminatory.

The implications are stark: healthcare professionals face potential malpractice liability for AI-influenced decisions, creating conservative implementation approaches. Financial professionals operate under regulatory frameworks that actually encourage algorithmic efficiency while mandating audit trails and bias monitoring (Lacmanovic & Skare, 2025). This fundamental difference in accountability structures shapes every aspect of human-AI collaboration design.

Healthcare establishes hierarchical models with clinicians at the apex, using AI in advisory capacities, while financial services implement variable distributions, delegating routine decisions to AI while maintaining human authority for complex cases (Ali Mohamad et al., 2023; Hilb, 2020). These variations reflect fundamental differences in how sectors conceptualize responsibility, with healthcare placing primary responsibility on clinicians and financial services implementing tiered structures varying by function and risk.

Cross-Sectoral Learning Opportunities and Knowledge Transfer

Cross-sectoral learning opportunities arise despite these differences. Healthcare's rigorous validation processes offer models for financial services facing increased scrutiny over algorithmic decisions. Conversely, financial services' experience with graduated automation provides insights for healthcare organizations seeking efficiency gains in administrative processes. However, direct transfer of practices remains limited due to fundamental differences in professional cultures, regulatory frameworks, and risk tolerances.

The findings reveal how professional roles progress differently across sectors in AI-enabled environments, with varying stress on preserving existing expertise versus developing hybrid capabilities (Meissner & Shmatko, 2019; Papa et al., 2020; Gravili et al., 2023; Holmström, 2022). This suggests the need for differentiated approaches to conceptualizing human-AI

collaboration that incorporate contextual factors (Ancillai et al., 2023; Cucari et al., 2023). As organizations navigate emerging technologies like generative AI (Sætra, 2023; Ooi et al., 2023), sectoral contexts will continue to shape collaborative boundaries, underscoring their enduring importance even as technologies progress (Bartoloni et al., 2022; Troisi et al., 2024).

While the identified patterns represent common tendencies within each sector, substantial variation exists within sectors based on organizational size, geographical location, and technological maturity. Teaching hospitals and community clinics may implement different collaborative models despite sharing healthcare sector characteristics. Similarly, investment banks and retail banking operations exhibit distinct approaches to AI integration. These intra-sectoral variations suggest that sector membership provides important but not deterministic influence on collaborative boundaries.

Discussion and Practical Implications

Theoretical Implications

The comparative analysis of human-AI collaboration across European healthcare and financial services sectors reveals how sector-specific cultural factors significantly shape collaborative boundaries. Professional traditions and governance structures influence how organizations conceptualize relationships between human expertise and AI capabilities (Trittin-Ulbrich et al., 2021; Cannavale et al., 2022). In healthcare, the strong professional identity of clinicians drives hierarchical models positioning AI in advisory roles (Ali Mohamad et al., 2023; Secinaro et al., 2021), while in financial services, the evolution of quantitative methods has facilitated more integrated approaches with authority distribution varying by context and risk (Fares et al., 2023; Trittin-Ulbrich et al., 2021).

Regulatory environments further influence boundary definition, with healthcare institutions facing stringent medical device regulations (Cannavale et al., 2022; Secinaro et al., 2021) and financial firms navigating complex data protection and algorithmic transparency requirements (Fares et al., 2023; Hilb, 2020). These findings suggest that effective human-AI collaboration

requires alignment with established professional identities and sectoral norms as opposed to generic implementation approaches (Schiavone et al., 2022; Garbuio et al., 2021).

The observed sectoral variations have profound implications for understanding AI adoption as a complex social negotiation that reflects and potentially transforms existing professional relationships (Chatterjee et al., 2021; Sharma et al., 2024; Tsamados et al., 2021; Frey & Osborne, 2024). Organizations must navigate tensions between technological transformation and preservation of professional knowledge, with healthcare's preservation of clinical authority and financial services' graduated authority distribution representing different responses to this tension (Johnson et al., 2022; Agostini et al., 2020; Lee et al., 2022; Chalmers et al., 2021). Kopalle et al. (2022) and Dittmar (2025) note that cultural factors influence AI adoption, suggesting sectoral cultures shape not just adoption decisions but the specific collaborative structures implemented.

Future empirical validation should address several specific dimensions. Longitudinal studies tracking collaborative boundary evolution as AI capabilities advance would reveal whether current patterns represent stable equilibria or transitional states. Comparative case studies across multiple organizations within each sector could validate the typology while identifying boundary conditions. Quantitative assessment of collaboration outcomes—including efficiency metrics, error rates, and stakeholder satisfaction—would provide evidence for optimal boundary configurations. Mixed-methods approaches combining organizational ethnography with performance analytics could capture both cultural and operational dimensions of human-AI collaboration.

This research contributes to the AI-Human Collaboration Theoretical Framework by demonstrating how abstract principles are instantiated differently in specific sectoral contexts (Hutchins, 1995; Clark, 1998; Engelbart, 1962). It extends understanding of human-AI collaborative boundaries by introducing a sectoral dimension to existing theoretical perspectives that have examined collaborative dynamics primarily through cognitive, technological, or ethical lenses (Floridi, 2014; Bostrom, 2014; Pickering, 1995; Simondon, 2017).

Implementation Frameworks and Policy Recommendations

Healthcare managers should establish clinician-led models following a three-tier approach: (1) AI Advisory Integration: Position AI tools explicitly as decision support rather than decision replacement, with transparent reasoning processes integrated into clinical workflows (Cannavale et al., 2022; Secinaro et al., 2021). (2) Professional Autonomy Preservation: Maintain ultimate clinical authority through mandatory human review protocols, particularly for high-stakes decisions. (3) Continuous Validation: Implement ongoing performance monitoring with clinical outcome tracking to ensure AI recommendations align with patient welfare objectives.

Financial services managers should design tiered authority models through: (1) Risk-Calibrated Automation: Implement graduated AI autonomy based on transaction complexity and potential impact, with routine decisions fully automated while complex cases require human oversight (Fares et al., 2023; Hilb, 2020). (2) Explainability Infrastructure: Ensure all AI decisions include transparent reasoning accessible to both staff and regulators. (3) Dynamic Boundary Management: Establish protocols for adjusting human-AI boundaries based on market conditions and regulatory changes.

Policymakers should develop sector-specific AI governance frameworks recognizing these fundamental differences. Healthcare requires stringent validation protocols with mandatory clinical oversight, while financial services benefit from performance-based regulations emphasizing outcomes over processes (Bartsch et al., 2025). Both sectors need clear liability frameworks addressing AI-influenced decisions and standardized approaches to algorithmic auditing.

Performance Management and Regulatory Compliance

Implementation success depends on sector-appropriate change management. Healthcare organizations must involve clinical leaders throughout development, addressing professional autonomy concerns directly through demonstration of AI's supportive rather than replacement role (Cannavale et al., 2022). Financial institutions should underline efficiency gains while preserving human expertise for relationship management and complex decision-making, using pilot programs to demonstrate value before full deployment (Trittin-Ulbrich et al., 2021).

Organizations must establish sector-specific performance metrics and governance structures to ensure successful human-AI collaboration implementation. Healthcare institutions should implement clinical outcome tracking systems that monitor diagnostic accuracy, patient safety indicators, and clinician satisfaction scores, with mandatory quarterly reviews of AI system performance against established clinical benchmarks (Secinaro et al., 2021; Cannavale et al., 2022).

Financial services organizations should establish real-time monitoring dashboards tracking decision accuracy rates, processing efficiency gains, and regulatory compliance metrics, with automated alerts for performance deviations that trigger immediate human oversight protocols (Fares et al., 2023; Hilb, 2020). Both sectors require independent audit committees with external expertise to validate AI system performance and ensure ongoing alignment with professional standards and regulatory requirements.

Practical implementation demands sector-specific regulatory compliance frameworks with clear escalation procedures and risk mitigation strategies. Healthcare organizations must establish clinical oversight committees with designated AI coordinators responsible for monitoring system performance, managing adverse event reporting, and ensuring compliance with medical device regulations through structured validation protocols (Cannavale et al., 2022; Ali Mohamad et al., 2023).

Financial institutions should implement algorithmic accountability frameworks including bias testing protocols, explainability documentation requirements, and customer complaint resolution procedures, with designated compliance officers responsible for regulatory reporting and stakeholder communication (Hilb, 2020; Fares et al., 2023). Organizations in both sectors must develop incident response protocols specifying when to reduce AI system autonomy, how to communicate system limitations to stakeholders, and procedures for reverting to manual processes during system failures or performance degradation.

Skills Development and Research Implications

Healthcare professionals require training in AI literacy focused on critical evaluation of algorithmic recommendations while preserving clinical reasoning capabilities (Secinaro et al., 2021). Financial services employees need hybrid analytical skills combining algorithmic understanding with contextual judgment, particularly in areas where human insight remains superior to automated analysis (Fares et al., 2023). These sector-specific approaches ensure AI implementations improve established professional practices while maximizing technological benefits within appropriate ethical and regulatory constraints.

Table 1 presents a comparison of human-AI collaboration models across European healthcare and financial services sectors. This comparative analysis systematically examines nine key dimensions that shape collaborative practices, underlining the distinctive approaches designed in each sector. The healthcare sector demonstrates a strong preservation of clinical authority with AI positioned in explicitly advisory roles, while the financial services sector implements more graduated approaches with authority distribution varying by context and risk profile.

For each dimension, the table includes sectoral implications that explain how these differences reflect broader organizational patterns and influence implementation strategies. This structured comparison demonstrates that effective human-AI collaboration requires sector-specific approaches aligned with established professional practices and not universal models. As AI progress, organizations must continue to negotiate appropriate human-AI boundaries within their distinct sectoral contexts.

Table 1: Comparative Analysis of Human-AI Collaboration Models Across European Sectors

Dimension	Healthcare Sector	Financial Services Sector	Sectoral Implications
Authority Distribution	Hierarchical model with clinicians at apex; AI explicitly positioned in advisory capacity	Tiered authority structure based on risk profiles; greater AI autonomy for routine decisions, human authority for complex cases	Reflects fundamental differences in professional responsibility structures
Professional Identity	Strong preservation of clinical authority; explicit subordination of AI to clinician judgment	Graduated approaches varying by function; greater integration in analytical roles	Indicates how sector-specific identity traditions shape technology adoption
Risk Management	Cautious approach with multiple validation layers; explicit limitations on AI autonomy in clinical decisions	Tiered framework calibrating human oversight according to decision risk profile	Demonstrates different sectoral approaches to balancing innovation with risk
Knowledge Integration	Underlines complementary contributions of AI pattern recognition and clinical reasoning; clinician-controlled review processes	More extensive data integration across diverse sources; combines quantitative analysis with contextual judgment in varied proportions	Shows how knowledge traditions influence information processing models
Regulatory Context	Stringent requirements underlining patient safety; medical device regulations shape implementation	Complex requirements related to financial stability, transparency, and non-discrimination	Illustrates regulatory influence on collaboration boundaries

Dimension	Healthcare Sector	Financial Services Sector	Sectoral Implications
Ethical Considerations	Primary stress on patient welfare, informed consent, and non-maleficence	Focus on fairness, transparency, and fiduciary responsibility	Highlights sector-specific values that must be preserved in AI integration
Implementation Approach	Clinician-led development with extensive validation before deployment	Risk-calibrated implementation with tiered oversight structures	Reveals different innovation adoption patterns across sectors
Illustrative Model	AI diagnostic support with explicit clinician review and authority	Credit assessment with tiered review based on case complexity and risk	Exemplifies how abstract principles manifest in practical applications
Skills Requirements	Healthcare professionals need education in AI advisory tool use while preserving clinical reasoning	Financial employees require hybrid analytical skills combining algorithmic understanding with contextual judgment	Indicates implications for workforce development and education

Source: authors

This analysis has several limitations that should be acknowledged. The theoretical approach and use of hypothetical cases may not capture the full complexity within sectors, as real-world implementations likely display greater heterogeneity than the idealized types presented. The illustrative nature of case examples constrains generalizability, as they represent composite patterns instead of specific organizational realities. Empirical validation through detailed case studies of actual implementations is essential before broader application of the proposed typology. The focus on healthcare and financial services necessarily excludes other important

sectors where human-AI collaboration is being implemented, such as education, public administration, and transportation. Additionally, the European context may not reflect collaboration patterns in other global regions with different regulatory environments and professional traditions.

The static nature of the current analysis limits understanding of how these relationships move forward as organizations gain experience with AI systems and as technologies advance. Future research should address this through longitudinal studies examining how boundary conceptions shift over time in response to changing capabilities and evolving professional identities. Empirical validation through comparative case studies of actual AI implementations would provide more robust evidence for the patterns identified and could reveal additional factors influencing collaborative boundaries.

Specific research directions arise from this analysis. Investigation of hybrid organizations spanning multiple sectors could reveal how conflicting collaborative models are reconciled. Examination of failed AI implementations would provide insights into boundary configurations that prove unsustainable. Cross-cultural studies comparing European approaches with those in other regions could identify universal versus culturally specific patterns. Research on emerging AI technologies, particularly large language models and generative AI, could explore whether these capabilities fundamentally alter established sectoral boundaries or reinforce existing patterns.

The relationship between collaborative boundaries and implementation outcomes deserves greater attention. Comparative studies examining how various collaborative models affect performance, workforce satisfaction, and innovation capacity would provide valuable insights for both theory and practice. Research should also investigate how organizational factors such as size, structure, and culture interact with sectoral influences to shape collaborative approaches.

Two additional promising research directions include cross-sector learning opportunities and professional education requirements. While sectoral differences are significant, organizations may benefit from understanding alternative models that could be appropriately adapted to their context. As collaborative systems become more prevalent, educational institutions must adapt to

prepare workers for new roles, with attention to how sector-specific requirements shape these educational needs. Finally, deeper examination of ethical frameworks and their implementation in collaborative systems is needed to provide guidance for responsible innovation across different sectoral contexts.

Conclusion

This article has analyzed how European healthcare and financial services sectors develop distinctive human-AI collaboration patterns based on their particular professional contexts, risk models, and expertise traditions. The findings highlight the importance of sector-specific approaches to boundary management, with healthcare characterized by explicit preservation of clinical authority and advisory AI roles, while financial services exhibit graduated approaches with authority distribution varying by task complexity and risk profile.

These sectoral patterns demonstrate how organizational contexts shape human-AI collaboration beyond technical considerations, as professional traditions and knowledge structures influence how collaborative boundaries are implemented. Rather than converging on a single model, organizations develop approaches reflecting their contextual requirements. This has significant implications for innovation management theory, suggesting that successful AI integration requires attention to social, professional, and regulatory contexts, not just technological capabilities.

The cross-sectoral comparison reveals fundamental differences in how authority is distributed, with healthcare institutions establishing hierarchical models that place clinicians at the apex of decision-making processes, while financial services implement variable authority distributions that delegate more routine decisions to AI systems while maintaining human oversight for complex or high-stakes cases. These variations reflect distinct professional identities, with healthcare professionals operating under ethical frameworks that stress patient welfare and clinical responsibility, and financial services professionals working within risk management traditions that calibrate oversight based on potential consequences.

The research also indicates that knowledge integration patterns differ significantly between sectors, with healthcare underlining the complementary contributions of AI pattern recognition and clinical reasoning through explicitly clinician-controlled review processes, while financial services tend toward more extensive data integration across diverse sources, combining quantitative analysis with contextual judgment in varied proportions based on task requirements. These distinctions suggest that effective AI implementation must fit with established knowledge structures and validation approaches that have progressed within specific sectoral contexts.

Regulatory environments further influence boundary definition, with healthcare institutions facing stringent medical device regulations that limit AI autonomy in clinical decisions, and financial firms navigating complex requirements related to algorithmic transparency, non-discrimination, and explainability. As technology capabilities continue to grow, these regulatory frameworks will likely develop in response, potentially reshaping collaborative boundaries while maintaining sector-specific approaches to risk management and professional responsibility.

The typology developed in this research extends theoretical understanding by demonstrating that human-AI collaboration cannot be conceptualized through universal models but requires sector-specific frameworks that account for professional identities, regulatory contexts, and knowledge traditions. This contribution challenges technology-deterministic views of AI implementation and stresses the enduring influence of sectoral contexts even as technological capabilities advance.

As AI capabilities evolve, their integration into organizational practices will continue to be shaped by the distinctive professional contexts and regulatory environments of different sectors. This ongoing negotiation between technological possibility and contextual requirements will define future human-AI collaboration across European healthcare and financial services. Understanding these sector-specific dynamics will enable practitioners to design collaborative systems that augment human capabilities, preserve professional values, and support organizational success in an increasingly AI-enabled environment.

The findings have important implications for both researchers and practitioners. For researchers, they highlight the need for more tailored theoretical approaches to human-AI collaboration that

incorporate sectoral and organizational contexts as significant variables. For practitioners, they underscore the importance of fitting implementation strategies with established professional structures and knowledge traditions, instead of attempting to impose generic collaboration models that may conflict with sectoral requirements. By designing context-sensitive approaches that recognize and address the distinctive needs of different sectors, organizations can harness the potential of AI technologies while preserving the valuable aspects of human expertise and professional authority that remain essential to effective performance in complex domains.

References

Abrantes, J., Bento e Silva, M. J. N., Meneses, J. P., Oliveira, C., Calisto, F. M. G. F., & Filice, R. W. (2023). External validation of a deep learning model for breast density classification. *European Congress of Radiology 2023*, Vienna, Austria.

Agostini, L., Galati, F., & Gastaldi, L. (2020). The digitalization of the innovation process: Challenges and opportunities from a management perspective. *European Journal of Innovation Management, 23*(1), 1-12.

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines: The simple economics of artificial intelligence*. *Harvard Business Review Press.*

Ali Mohamad, T., Bastone, A., Bernhard, F., & Schiavone, F. (2023). How artificial intelligence impacts the competitive position of healthcare organizations. *Journal of Organizational Change Management, 36*(8), 49-70.

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

Ankrah, E. A., Awori, K., Nyairo, S., Muchai, M., Ochieng, M., Kariuki, M., Hayes, G. R., & O'Neill, J. (2025) Social by Nature: How Socio-tecture Shapes the Work of SMBs and Considerations for Reimagining Collaborative Human-AI Systems. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1-19).

Appio, F. P., La Torre, D., Lazzeri, F., Masri, H., & Schiavone, F. (2023). *Impact of Artificial Intelligence in Business and Society: Opportunities and Challenges*. Routledge.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.

Bartsch, S., Behn, O., Benlian, A., Brownsword, R., Bücker, S., Düwell, M., Formánek, N., Jungtäubl, M., Leyer, M., Richter, A., Schmidt, J. H., & Will-Zocholl, M. (2025). Governance of High-Risk AI Systems in Healthcare and Credit Scoring. *Business & Information Systems Engineering*.

Bartoloni, S., Calo, E., Marinelli, L., Pascucci, F., Dezi, L., Carayannis, E., Revel G. M., & Gregori, G. L. (2022). Towards designing society 5.0 solutions: The new Quintuple Helix-Design Thinking approach to technology. *Technovation, 113*, 102413.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

Calisto, F. M., Abrantes, J. M., Santiago, C., Nunes, N. J., & Nascimento, J. C. (2025). Personalized explanations for clinician-AI interaction in breast imaging diagnosis by adapting communication to expertise levels. *International Journal of Human-Computer Studies*, 197, 103444.

Calisto, F. M., Nunes, N., & Nascimento, J. C. (2022). Modeling adoption of intelligent agents in medical imaging. *International Journal of Human-Computer Studies, 168*, 102922.

Cannavale, C., Esempio Tammaro, A., Leone, D. & Schiavone, F. (2022). Innovation adoption in inter-organizational healthcare networks - the role of artificial intelligence. *European Journal of Innovation Management, 25*(6), 758-774.

Chalmers, D., MacKenzie, N. G., & Carter, S. (2021). Artificial intelligence and entrepreneurship: Implications for venture creation in the fourth industrial revolution. *Entrepreneurship Theory and Practice, 45*(5), 1028-1053.

Chatterjee, S., Rana, N. P., Dwivedi, Y. K., & Baabdullah, A. M. (2021). Understanding AI adoption in manufacturing and production firms using an integrated TAM-TOE model. *Technological Forecasting and Social Change, 170*, 120880.

Clark, A. (1998). *Being there: Putting brain, body, and world together again*. MIT Press.

Cucari, N., Nevi, G., Laviola, F., & Barbagli, L. (2023). Artificial Intelligence and Environmental Social Governance: An Exploratory Landscape of AI Toolkit. Available at *SSRN*.

Daly, S. J., Hearn, G., & Papageorgiou, K. (2025). Sensemaking with AI: The Role of Trust in Human-AI Collaboration Across Health and Creative Industries. *Social Sciences & Humanities Open, 11*, 101346.

Dittmar, E. C. (2025). AI as a catalyst for organizational learning: moving beyond tool implementation to learning transformation. *Development and Learning in Organizations*.

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data-evolution, challenges and research agenda. *International Journal of Information Management, 48*, 63-71.

Engelbart, D. C. (1962). *Augmenting human intellect: A conceptual framework*. Stanford Research Institute.

Fares, O. H., Butt, I., & Lee, S. H. M. (2023). Utilization of artificial intelligence in the banking sector: A systematic literature review. *Journal of Financial Services Marketing, 28*(4), 835-852.

Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford University Press.

Frey, C. B., & Osborne, M. (2024). Generative AI and the Future of Work: A Reappraisal. *Brown Journal of World Affairs, 30*(1), 1-17.

Garbuio, M., & Lin, N. (2021). Innovative idea generation in problem finding: Abductive reasoning, cognitive impediments, and the promise of artificial intelligence. *Journal of Product Innovation Management, 38*(6), 701-725.

Giuggioli, G., & Pellegrini, M. M. (2023). Artificial intelligence as an enabler for entrepreneurs: a systematic literature review and an agenda for future research. *International Journal of Entrepreneurial Behavior & Research, 29*(4), 816-837.

Gravili, G., Hassan, R., Avram, A., & Schiavone, F. (2023). Big data and human resource management: paving the way toward sustainability. *European Journal of Innovation Management, 26*(7), 552-590.

Haefner, N., Parida, V., Gassmann, O., & Wincent, J. (2023). Implementing and scaling artificial intelligence: A review, framework, and research agenda. *Technological Forecasting and Social Change, 197*, 122878.

Hart, C. (2018). *Doing a Literature Review: Releasing the Research Imagination* (2nd ed.). Sage Publications.

Hilb, M. (2020). Toward artificial governance? The role of artificial intelligence in shaping the future of corporate governance. *Journal of Management and Governance, 24*, 851-870.

Holland, J. H. (1998). *Emergence: From chaos to order*. Oxford University Press.

Holmström, J. (2022). From AI to digital transformation: The AI readiness framework. *Business Horizons, 65*(3), 329-339.

Hutchins, E. (1995). *Cognition in the wild*. MIT Press.

Jiang, N., Liu, X., Liu, H., Lim, E. T. K., Tan, C., & Gu, J. (2023). Beyond AI-powered context-aware services: the role of human-AI collaboration. *Industrial Management & Data Systems, 123*(11), 2771–2802.

Jiang, T., Sun, Z., Fu, S. and Lv, Y. (2024). Human-AI interaction research agenda: A user-entered perspective. *Data and Information Management, 8*(4), p. 100078.

Johnson, P. C., Laurell, C., Ots, M., & Sandström, C. (2022). Digital innovation and the effects of artificial intelligence on firms' research and development-Automation or augmentation, exploration or exploitation? *Technological Forecasting and Social Change, 179*, 121636.

Juntunen, M., & Lehenkari, M. (2021). A narrative literature review process for an academic business research thesis. *Studies in Higher Education, 46*(2), 330-342.

Kauffman, S. A. (1995). *At home in the universe: The search for the laws of self-organization and complexity*. Oxford University Press.

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals, 14*(1), 366–410.

Khosravy, M., Gupta, N., Pasquali, A., Dey, N., González Crespo, R., & Witkowski, O. (2024). Human-collaborative artificial intelligence along with social values in Industry 5.0: A survey of the state-of-the-art. *IEEE Transactions on Cognitive and Developmental Systems, 16*(1), 165-176.

Khvatova, T., Appio, F. P., Ray, S., & Schiavone, F. (2024). Exploring the Role of AI in B2B Customer Journey Management: Towards an IPO Model. *IEEE Transactions on Engineering Management, 71*, 13852-13866.

Kim, S. W., Kong, J. H., Lee, S. W., & Lee, S. (2022). Recent advances of artificial intelligence in manufacturing industrial sectors: A review. *International Journal of Precision Engineering and Manufacturing, 23*, 111-129.

Kopalle, P. K., Gangwar, M., Kaplan, A., Ramachandran, D., Reinartz, W., & Rindfleisch, A. (2022). Examining artificial intelligence (AI) technologies in marketing via a global lens: Current trends and future research opportunities. *International Journal of Research in Marketing, 39*(2), 522-540.

Kumar, S., Kumar, V., Chaudhuri, R., Chatterjee, S., & Vrontis, D. (2025). AI capability and environmental sustainability performance: Moderating role of green knowledge management. *Technology in Society*, *81*, 102870.

Lacmanovic, S., & Skare, M. (2025). Artificial intelligence bias auditing–current approaches, challenges and lessons from practice. *Review of Accounting and Finance, 24*(3), 375-400.

Lee, Y. S., Kim, T., Choi, S., & Kim, W. (2022). When does AI pay off? AI-adoption intensity, complementary investments, and R&D strategy. *Technovation, 118*, 102590.

Licklider, J. C. R. (1960). Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics, HFE-1*(1), 4-11.

López Jiménez, D., Dittmar, E.C. & Vargas Portillo, J.P. (2023). The Self-Regulating Jury in the Field of Interactive Advertising. *Law, State and Telecommunications Review, 15*(1), 1-30.

López Jiménez, D., Dittmar, E.C. & Vargas Portillo, J.P. (2025). Implementation of Self-Regulation in Online Gambling Advertising: Organizational Challenges and Benefits. *Public Organization Review.*

Meissner, D., & Shmatko, N. (2019). Integrating professional and academic knowledge: The link between researchers skills and innovation culture. *The Journal of Technology Transfer, 44*, 1273-1289.

Merchán-Cruz, E. A., Gabelaia, I., Savrasovs, M., Hansen, M. F., Soe, S., Rodriguez-Cañizo, R. G., & Aragón-Camarasa, G. (2025). Trust by Design: An Ethical Framework for Collaborative Intelligence Systems in Industry 5.0. *Electronics, 14*(10), 1952.

Ooi, K. B., Tan, G. W. H., Al-Emran, M., Al-Sharafi, M. A., Capatina, A., Chakraborty, A., Dwivedi, Y. K., Huang, T. L., Kar, A. K., Lee, V. H. Loh, X. M., Micu, A., Mikalef, P., Mogaji, E., Pandey, N., Raman, R., Rana, N. P., Sarker, P., Sharma, A., Teng, C. I., Wamba, S. F., & Wong, L. W. (2023). The potential of generative artificial intelligence across disciplines: perspectives and future directions. *Journal of Computer Information Systems, 65*(1), 76-107.

Papa, A., Dezi, L., Gregori, G.L., Mueller, J., & Miglietta, N. (2020). Improving innovation performance through knowledge acquisition: the moderating role of employee retention and human resource management practices. *Journal of Knowledge Management, 24*(3), 589-605.

Pickering, A. (1995). *The mangle of practice: Time, agency, and science*. University of Chicago Press.

Razmak, J., El Refae, G. A., & Farhan, W. (2025). The role of AI applications in production and operations management: complementing or replacing human labour. *International Journal of Economics and Business Research, 29*(13), 1–28.

Sargeant, H. (2023). Algorithmic decision-making in financial services: Economic and normative outcomes in consumer credit. *AI and Ethics, 3*, 1295-1311.

Sætra, H. S. (2023). Generative AI: Here to stay, but for good? *Technology in Society, 75*, 102372.

Schiavone, F., Pietronudo, M. C., Sabetta, A., & Bernhard, F. (2022). Designing AI implications in the venture creation process. *International Journal of Entrepreneurial Behavior & Research, 29*(4), 838-859.

Secinaro, S., Calandra, D., Secinaro, A., Muthurangu, V., & Biancone, P. (2021). The role of artificial intelligence in healthcare: a structured literature review. *BMC Medical Informatics and Decision Making, 21*(125), 1-23.

Sharma, S., Islam, N., Singh, G., & Dhir, A. (2024). Why Do Retail Customers Adopt Artificial Intelligence (AI) Based Autonomous Decision-Making Systems? *IEEE Transactions on Engineering Management, 71*, 1846-1861.

Shrestha, Y. R., Ben-Menahem, S. M., & Von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review, 61*(4), 66-83.

Simondon, G. (2017). *On the mode of existence of technical objects* (Original work published 1958). University of Minnesota Press.

Sposato, M. (2025). Developing Digital Leadership Skills for AI-Driven Research Environments. In *Impacts of Generative AI on the Future of Research and Education* (pp. 235-254). IGI Global.

Torraco, R. J. (2016). Writing integrative literature reviews: Using the past and present to explore the future. *Human Resource Development Review, 15*(4), 404-428.

Tran, H., & Murphy, P. J. (2023). Generative artificial intelligence and entrepreneurial performance. *Journal of Small Business and Enterprise Development, 30*(5), 853-856.

Trist, E. (1981). *The evolution of socio-technical systems*. Occasional Paper No. 2, Ontario Quality of Working Life Centre.

Trittin-Ulbrich, H., Scherer, A. G., Munro, I., & Whelan, G. (2021). Exploring the dark and unexpected sides of digitalization: Toward a critical agenda. *Organization, 28*(1), 8-25.

Troisi, O., Visvizi, A. & Grimaldi, M. (2024). Rethinking innovation through industry and society 5.0 paradigms: a multileveled approach for management and policy-making. *European Journal of Innovation Management, 27*(9), 22-51.

Tsamados, A., Aggarwal, N., Cowls, J., Morley, J., Roberts, H., Taddeo, M., & Floridi, L. (2021). The ethics of algorithms: key problems and solutions. In

Woolgar, S. (1990). Configuring the user: The case of usability trials. *The Sociological Review, 38*(1 suppl), 58-99.

Xu, C., & Cho, S. E. (2025). Factors Affecting Human-AI Collaboration Performances in Financial Sector: Sustainable Service Development Perspective. *Sustainability, 17*(10), 4335.