



INTERNATIONAL JOURNAL OF
INNOVATION AND
INDUSTRIAL REVOLUTION
(IJIREV)

www.gaexcellence.com/ijirev



ENHANCING INNOVATION THROUGH HUMAN-AI COLLABORATION IN INDUSTRY 5.0: THE MEDIATING ROLE OF TRUST

Shazwani Mohd Salleh^{1*}, Norhasyikin Rozali², Siti Murni Mat Khairi³, Suheil Che Sobry⁴,
Mohamad Shahrin Rosli⁵

¹ Faculty of Business and Management, Universiti Teknologi Mara Cawangan Kedah

 shazwani065@uitm.edu.my

 <https://orcid.org/0009-0009-8448-8333>

² Faculty of Business and Management, Universiti Teknologi Mara Cawangan Kedah

 norhasyikin@uitm.edu.my

 <https://orcid.org/0009-0001-6192-5221>

³ Faculty of Business and Management, Universiti Teknologi Mara Cawangan Kedah

 sitimurni@uitm.edu.my


 <https://orcid.org/0000-0003-0000-4223>


⁴ Faculty of Business and Management, Universiti Teknologi Mara Cawangan Kedah

 suheil@uitm.edu.my

 <https://orcid.org/0009-0000-4404-0555>

⁵ Emergency and Trauma Department, Hospital Sultan Abdul Halim, Sungai Petani, Kedah Malaysia

 shahrinshazwani@gmail.com

 <https://orcid.org/0009-0008-9421-4750>

*Corresponding Author

Article Info:

Article history:

Received date: 30.03.2026

Revised date: 19.04.2026

Accepted date: 19.05.2026

Published date: 03.06.2026

To cite this document:

Salleh, S. M., Rozali, N., Mat Khairi, S. M., Che Sobry, S., & Rosli, M. S. (2026). Enhancing Innovation Through Human-AI Collaboration in Industry 5.0: The Mediating Role of Trust. *International Journal of Innovation and Industrial Revolution*, 8(25), 18-32.

Abstract:

Industry 5.0 envisions a human-centric industrial future where collaboration between humans and artificial intelligence (AI) enables sustainable innovation. However, despite the transformative potential of Human-AI Collaboration (HAIC), its impact on innovation remains inconsistent. This variability stems largely from insufficient attention to the psychological and relational mechanisms that facilitate effective partnership, particularly trust. The present research addresses this gap by proposing a conceptual framework that positions trust as the critical mediator between HAIC and innovation performance. Using a structured integrative review approach, we analyzed 28 peer-reviewed articles from the Scopus database (2020–2025) through thematic analysis to synthesize current knowledge and identify theoretical linkages. Our findings reveal that trust is not a static outcome but a dynamic, context-dependent construct shaped by transparency, explainability, reliability, and ethical AI design. Trust enables human operators to engage in risk-taking, knowledge sharing, and co-creative behaviors that directly enhance innovation in products, processes, and business models. This work provides a dual contribution: it unifies

socio-technical systems theory with Industry 5.0 principles into a coherent mediation framework and offers actionable insights for designing “trustworthy-by-design” AI systems. The proposed framework underscores that technological sophistication alone is insufficient; fostering calibrated trust is essential to unlock the full innovative potential of human–AI symbiosis in next-generation manufacturing.

DOI: 10.35631/IJIREV.825002

Keyword:

Human–AI Collaboration, Industry 5.0, Innovation Performance, Trust in AI & Trust



© The authors (2026). This is an Open Access article distributed under the terms of the Creative Commons Attribution (CC BY NC) (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact ijirev@gaexcellence.com.

Introduction

Industry 5.0 represents a transformative shift from the automation-driven logic of Industry 4.0 toward a human-centric industrial paradigm that emphasizes symbiotic collaboration between humans and intelligent technologies such as artificial intelligence (AI) and robotics (Pawar et al., 2025; Yanytska, 2025). This emerging framework prioritizes sustainability, resilience, and human well-being by integrating human creativity, intuition, and ethical judgment with the computational power and precision of AI systems (Shabur et al., 2025). While AI holds immense promise for enhancing manufacturing efficiency and enabling novel product and process innovations, its successful integration hinges critically on the quality of human–AI interaction—particularly the level of trust that human operators place in AI agents (Zywiołek, 2024).

Existing literature underscores AI’s potential to optimize design, streamline operations, and foster innovation in Industry 5.0 contexts (Katuk et al., 2024). However, significant barriers remain, including the opacity of AI decision-making (“black box” problem), data biases, and ethical concerns regarding autonomy and accountability (Ahangar et al., 2025). Trust has emerged as a pivotal construct in human–robot and human AI collaboration, influencing operators’ willingness to rely on, delegate tasks to, and co-create with intelligent systems (Campagna et al., 2024a). Recent studies demonstrate that trust can be inferred through behavioral indicators such as facial expressions, body motion, and task performance and modeled using machine learning to predict collaboration quality with high accuracy (Campagna et al., 2024b, 2025).

Despite growing recognition of trust’s importance, current research lacks a coherent conceptual understanding of how trust functions as a mediating mechanism between HumanAI Collaboration (HAIC) and innovation outcomes in Industry 5.0. Most empirical efforts focus

either on technical AI performance or isolated trust metrics, without theorizing the dynamic, reciprocal relationship through which trust enables risk-taking, knowledge sharing, and experimental behaviors essential for innovation (Zywiołek, 2024; Ahangar et al., 2025). This gap is particularly salient in real-time industrial settings, where adaptive trust calibration is necessary but rarely operationalized.

To address this limitation, this paper proposes a novel conceptual framework that positions trust as the central mediating variable linking HAIC to innovation in Industry 5.0. Drawing on interdisciplinary insights from human robot interaction, trustworthy AI design, and socio-technical systems theory, the framework integrates data-driven methods such as behavioral sensing and explainable AI to assess and dynamically adjust trust levels during collaboration (Campagna et al., 2024b, 2025). By foregrounding trust as a malleable, context-sensitive construct, the model aims to bridge the gap between AI capability and human acceptance, thereby unlocking collaborative innovation potential.

This study holds significant theoretical and practical value. Conceptually, it advances the HAIC literature by explicating the mediating role of trust in innovation processes a mechanism underexplored in current Industry 5.0 discourse. Practically, it offers actionable design principles for developing transparent, fair, and accountable AI systems that align with human-centric values (Ahangar et al., 2025; Katuk et al., 2024). The remainder of the paper is structured as follows: after the introduction, Section 2 reviews key literature on HAIC and trust; Section 3 presents the proposed conceptual framework; Section 4 discusses theoretical and managerial implications; and Section 5 concludes with limitations and future research directions.

Literature Review

The literature on Human–AI Collaboration (HAIC) underscores that successful synergy depends not just on technical capabilities but on relational factors that shape how humans engage with AI systems. A recent review identifies that HAIC effectiveness is contingent on interaction quality, adaptability, trust, and system transparency not merely system performance metrics alone (Evaluating HAIC, 2024). Complementing this, work on human–AI synergy in decision making highlights the role of AI affordances (e.g. suggest, critique, notify) and synergy patterns (e.g. sequential, parallel, feedback loops) in shaping outcomes, with trust and transparency emerging as core mediating outcomes in those frameworks (He et al., 2023). This body of work positions HAIC not as a unidirectional augmentation but as a dynamic, negotiated partnership.

Trust, in turn, has been studied extensively in the AI context, though often in isolation from innovation outcomes. A structured review on Trust in Artificial Intelligence shows the evolution of trust research, pointing out that calibration balancing overtrust and distrust is key (2024). Similarly, psychological and organizational literatures synthesize trust into three dimensions: trustor (human characteristics), trustee (AI attributes), and context (task, environment) (Frontiers, 2024). In HAIC settings, empirical work on explainable AI (XAI) demonstrates that transparency of AI reasoning can increase trust and improve task performance though not unconditionally: overly complex explanations or uncertain reasoning disclosures may reduce confidence (e.g. Effects of Explainable AI study). Design studies also show that agent transparency (i.e. revealing how AI reasoned) increases trust, whereas exposing uncertainty can decrease it (Designing Transparency, 2022). In collaborative tasks,

giving human users decision control (allowing them to override AI) plus explanations significantly improves trust, understanding, and compliance (Decision Control & Explanations, 2023). These findings suggest that trust is not a monolithic constant but fluctuates with design choices, context, and human perceptions.

Bridging HAIC, trust, and innovation, there is emerging but scattered evidence that trust mediates how AI input is integrated into creative, experimental, and risk-taking behaviors. In the context of Industry 5.0, AI's effect on innovation capacity can be "synergic, deceptive, or substitutive" depending on how well AI aligns with human intent, capabilities, and trust (Bécue et al., 2024). Moreover, HAIC deployed in manufacturing must contend with ethical, bias, cybersecurity, and transparency challenges which act as barriers to trust and thus dampen innovation (Ahangar et al., 2025). Research on automation bias also warns that humans may over-rely on AI (even when incorrect) or underutilize it (due to skepticism), both of which can distort the collaboration-innovation linkage (Automation Bias review, 2025). Finally, recent studies in human-AI joint decision-making reveal that trust affects the weight given to AI advice (i.e. how much humans rely on it), thereby shaping the degree of co-creation and novelty generated (Frontiers trust-AI collaboration, 2025).

Together, these threads reinforce your core claim: that trust is the mediating mechanism through which HAIC influences innovation in Industry 5.0. In your work, you can build on HAIC and affordance theories, trust models (cognitive, affective, calibration), and innovation theory (e.g. absorptive capacity, risk-taking) to more precisely specify how design features (like transparency, control, fairness) shape trust, which in turn governs the extent to which human and AI contributions are meaningfully integrated into innovative outcomes.

Methodology

Research Design

This study employs a structured narrative review methodology to synthesize scholarly knowledge and develop a novel conceptual framework. Unlike reviews that prioritize statistical aggregation, this integrative approach is better suited for exploring complex, emerging phenomena where theoretical synthesis is required to bridge fragmented literature. Given that Industry 5.0 is an evolving paradigm, a narrative-based integration allows for critical interpretation and the construction of a model that connects human computer interaction with manufacturing innovation.

Key Steps in Conducting a Narrative Review

The narrative review was conducted through a structured yet flexible process, adapted from established best practices in qualitative evidence synthesis (Grant & Booth, 2009; Greenhalgh, 2023). First, a focused research question was formulated: "How does trust mediate the impact of human-AI collaboration on innovation in Industry 5.0?" This question guided all subsequent stages. Second, literature search was performed exclusively in the Scopus database one of the largest and most reputable abstract and citation databases for peer-reviewed scientific literature—due to its extensive coverage of engineering, computer science, social sciences, and interdisciplinary journals relevant to Industry 5.0 (Elsevier, 2024). The search strategy combined keywords such as "Industry 5.0," "Human-AI collaboration," "trust," "innovation," "human-centric AI," and "collaborative intelligence," using Boolean operators (e.g., "Industry

5.0” AND “trust” AND “innovation”). Filters were applied to include only English-language, peer-reviewed journal articles and conference proceedings published between 2020 and 2025, ensuring relevance to the nascent Industry 5.0 discourse.

Third, screening and selection of literature was conducted iteratively. Titles and abstracts were reviewed for conceptual relevance to HAIC, trust dynamics, and innovation outcomes in industrial settings. Full texts of potentially relevant papers were then assessed for theoretical depth, methodological rigor, and alignment with the study’s focus. Particular attention was given to recent (2023–2025) Scopus-indexed publications that address ethical AI design, real-time trust assessment, and socio-technical integration in manufacturing (e.g., Ahangar et al., 2025; Merchán-Cruz et al., 2025; Zywiołek, 2024). Fourth, thematic analysis was performed to identify recurring constructs, theoretical tensions, and gaps especially the underdeveloped mediating role of trust. Finally, these insights were synthesized into a proposed conceptual framework, which integrates trust as a dynamic, context-dependent mediator between HAIC and innovation, grounded in principles of transparency, accountability, and human agency. Throughout this process, reflexivity and theoretical triangulation were maintained to ensure conceptual validity and scholarly contribution.

The review followed a structured process adapted from qualitative evidence synthesis. A focused research question—“How does trust mediate the impact of HAIC on innovation?”—guided the inquiry. Literature was sourced exclusively from the Scopus database to ensure high-quality, peer-reviewed coverage of engineering and management domains.

Evidence Synthesis Process:

Search Execution: A Boolean query focused on "Industry 5.0," "Trust," and "Innovation" yielded 127 initial records.

Selection: After screening for conceptual relevance, 28 seminal sources were selected for full-text assessment.

Thematic Integration: Data were analyzed through open, axial, and selective coding to identify recurring themes like "calibrated trust" and "ethical by design".

Framework Development: These insights were synthesized to reveal the lack of explicit mediation models, directly informing the proposed framework.

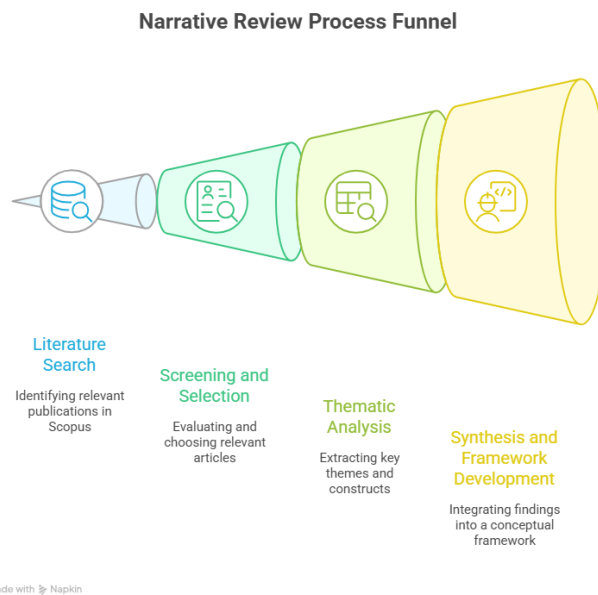


Figure 1: Process of Review

Data Collection and Review Strategy

To ensure a robust and focused synthesis of the literature, a structured data collection protocol was implemented using the Scopus database, selected for its extensive coverage of peer-reviewed journals in engineering, computer science, management, and social sciences—domains central to the interdisciplinary nature of Industry 5.0 and Human–AI Collaboration (HAI). A comprehensive search string was developed to capture the intersection of four core conceptual domains: (1) human–AI interaction, (2) trust-related constructs, (3) innovation outcomes, and (4) the industry 5.0 context. The final Boolean search query was formulated as follows:

("human-ai" OR "human artificial intelligence" OR "collaboration" OR "cooperation") AND ("trust" OR "confidence" OR "reliability" OR "dependability") AND ("innovation" OR "creativity" OR "advancement" OR "development") AND ("Industry 5.0" OR "fifth industrial revolution" OR "smart industry" OR "advanced manufacturing")

This query was executed in the Scopus “Advanced Search” interface on March 15, 2025, with filters applied to include only English-language, peer-reviewed journal articles and conference proceedings published between 2020 and 2025, reflecting the emergent nature of Industry 5.0 discourse. The initial search yielded 127 records. After removing duplicates and screening titles and abstracts for conceptual relevance specifically excluding studies focused solely on technical AI performance without human interaction or those unrelated to industrial or organizational innovation the dataset was narrowed to 42 key publications. Full-text assessment further refined the selection to 28 seminal and highly relevant sources that directly addressed the interplay among HAIC, trust, and innovation in human-centric industrial environments.

The selected literature was then analyzed using an integrative thematic analysis approach, which combines inductive and deductive coding to identify, interpret, and synthesize key themes, theoretical constructs, and conceptual relationships (Whittemore & Knafl, 2005;

Snyder, 2019). This method is particularly appropriate for narrative reviews aiming to develop new conceptual models, as it allows for both data-driven pattern recognition and theory-informed interpretation. The analysis proceeded in three iterative phases: (1) open coding, where key phrases and concepts (e.g., “explainable AI,” “calibrated trust,” “co-creation,” “ethical by design”) were extracted and labeled; (2) axial coding, during which codes were grouped into broader thematic categories such as “Antecedents of Trust in HAIC,” “Trust as a Mediator of Innovation,” and “Ethical and Technical Barriers in Industry 5.0”; and (3) selective coding, where core theoretical insights were integrated to construct a coherent narrative around the mediating role of trust. Throughout this process, constant comparison and theoretical triangulation were employed to ensure analytical rigor and conceptual coherence. The resulting synthesis not only maps the current state of knowledge but also reveals critical gaps—particularly the lack of explicit mediation models which directly informs the proposed conceptual framework in the next section.

Key Findings from the Narrative Review

The integrative thematic analysis of 28 peer-reviewed Scopus-indexed publications (2020–2025) revealed five interrelated themes that collectively illuminate the mediating role of trust in the relationship between Human–AI Collaboration (HAIC) and innovation in Industry 5.0. These findings are summarized in Table 1.

Table1: Key Thematic Findings from the Narrative Review

Key Finding	Description	Supporting References (Scopus)
Trust as a Mediator	Trust mediates the relationship between human–AI collaboration and innovation by influencing acceptance, reliance, and willingness to integrate AI outputs into creative processes. Without trust, collaboration effectiveness diminishes.	Zywołek (2024); Merchán-Cruz et al. (2025)
Determinants of Trust	Transparency, explainability, accountability, fairness, and reliability of AI systems are key drivers of trust in Industry 5.0 contexts.	Ahangar et al. (2025); Katuk et al. (2024)
Human–AI Collaboration Enhances Innovation	Collaboration improves human decision-making, boosts productivity, and enables new business models, thereby accelerating innovation.	Dewangan & Kumar (2024); Pawar et al. (2025)
Challenges in Trust Building	Black-box AI models, data biases, lack of ethical frameworks, and limited explainability reduce trust and hinder collaboration.	Ahangar et al. (2025); Ramírez-Gordillo et al. (2026)
Ethical Considerations	Ethical AI design principles such as privacy, fairness, accountability, and responsible use of AI are essential to sustain trust and innovation.	Merchán-Cruz et al. (2025); Ramírez-Gordillo et al. (2026)

The narrative review highlights that trust is a critical mediator linking human–AI collaboration to innovation within Industry 5.0. Evidence suggests that trust enables individuals to accept, rely upon, and integrate AI outputs, thereby strengthening the innovative capacity of organizations (Zywiołek, 2024; Merchán-Cruz et al., 2025). The determinants of trust include transparency, explainability, fairness, accountability, and reliability of AI systems, which shape user perceptions and foster collaborative synergy (Ahangar et al., 2025; Katuk et al., 2024). Conversely, persistent challenges such as the “black box” nature of AI, algorithmic biases, and insufficient ethical frameworks threaten trust and hinder adoption (Ramírez-Gordillo et al., 2026).

Importantly, human–AI collaboration significantly enhances innovation by augmenting decision-making, boosting productivity, and enabling new business models that respond dynamically to changing environments (Dewangan & Kumar, 2024; Pawar et al., 2025). To sustain this momentum, ethical considerations including privacy, fairness, accountability, and responsible use of AI must be prioritized (Merchán-Cruz et al., 2025). Emerging methods such as real-time trust assessment using behavioral indicators and machine learning are promising tools to evaluate and calibrate trust dynamically in collaborative environments (Campagna et al., 2024, 2025).

Collectively, the findings underscore that trust is not only a psychological or social factor but also a strategic enabler of innovation in Industry 5.0. When embedded within a socio-technical framework, trust ensures that human creativity and AI’s analytical strengths converge to produce sustainable, human-centric innovation (Yanytska, 2025; Shabur et al., 2025).

Development Of Theoretical Framework

The proposed theoretical framework draws upon the socio-technical systems theory and the integrative model of trust in organizational settings (Mayer et al., 1995; Bostrom & Heinen, 1977). Socio-technical systems theory emphasizes the interdependence between human and technological subsystems, suggesting that innovation outcomes are optimized when both are harmonized (Trist, 1981). In the context of Industry 5.0, human–AI collaboration represents such a socio-technical interaction, where human creativity, intuition, and ethical judgment complement AI’s analytical and computational capabilities (Yanytska, 2025; Pawar et al., 2025). Trust theory further underscores that individuals’ willingness to rely on AI is contingent on their perceptions of competence, integrity, and benevolence factors that directly mediate the effectiveness of human–AI collaboration in achieving innovation (Zywiołek, 2024; Merchán-Cruz et al., 2025).

The framework conceptualizes human–AI collaboration as the independent variable, trust as the mediating mechanism, and innovation performance as the dependent variable. This configuration builds on empirical findings that trust enhances collaboration effectiveness by reducing uncertainty, fostering reliance, and enabling smoother integration of AI outputs into human decision-making (Campagna et al., 2025; Ahangar et al., 2025). Literature on collaborative intelligence stresses that trust is both dynamic and context-dependent, shaped by transparency, explainability, and ethical AI design (Merchán-Cruz et al., 2025; Katuk et al., 2024). Integrating these insights, the framework posits that when human–AI interactions are underpinned by robust trust-building mechanisms, innovation outcomes are more likely to be realized. In this regard, the model synthesizes socio-technical perspectives with trust theory,

offering a structured lens for examining how human and technological dimensions interact to drive innovation in Industry 5.0.

Practically, the theoretical framework suggests that organizations must deliberately design trust-enhancing mechanisms, such as explainable AI systems, transparent decision pathways, and ethical safeguards, to fully leverage human–AI collaboration for innovation (Ramírez-Gordillo et al., 2026). For managers, this means balancing efficiency with human-centric values, ensuring that employees feel empowered to work with AI rather than displaced by it. Moreover, the framework encourages policymakers and practitioners to adopt a responsible innovation approach, embedding fairness, accountability, and inclusivity into AI integration strategies (Shabur et al., 2025). In conclusion, the framework provides both a theoretical foundation and practical guidance for exploring how trust mediates the relationship between human–AI collaboration and innovation, offering a novel lens for future empirical research and contributing to the broader discourse on Industry 5.0.

Figure 1 is the theoretical framework of this study, which positions Human–AI Collaboration as the independent variable, Trust as the mediating variable, and Innovation Performance as the dependent variable. The framework emphasizes that while human–AI collaboration is essential for driving innovation in Industry 5.0, the degree to which such collaboration enhances innovation depends significantly on the presence of trust. Trust serves as the mediating mechanism that ensures AI systems are perceived as transparent, ethical, and reliable, thereby fostering greater acceptance and effectiveness in collaborative industrial environments. This framework draws on literature highlighting the central role of trust in shaping successful AI-human partnerships (Zywiólek, 2024; Ahangar et al., 2025).

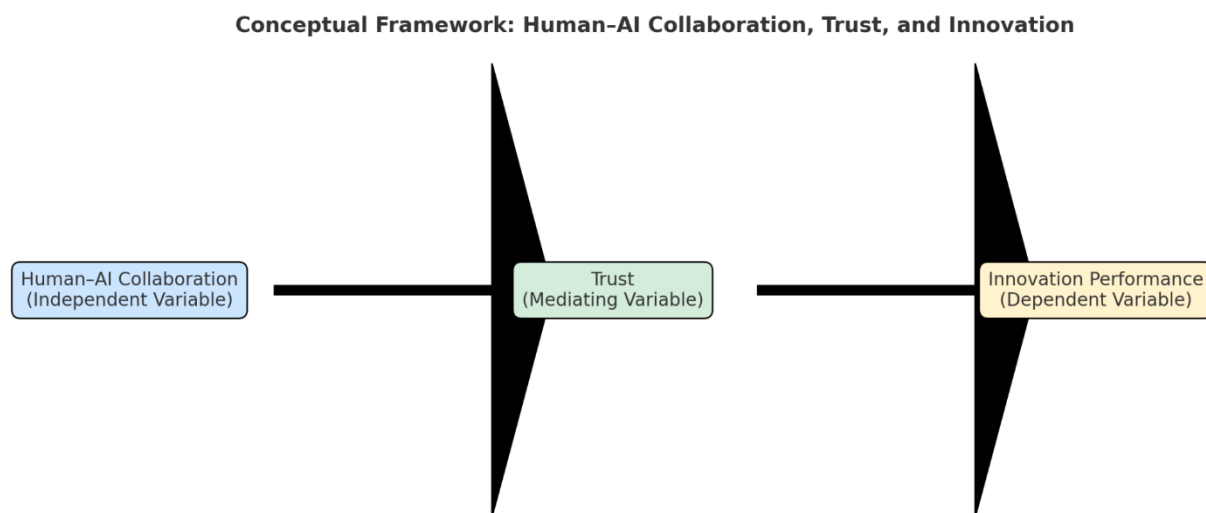


Figure 2: Theoretical Framework Of This Study

Proposition Development

Human-AI Collaboration affect Trust

Human–AI Collaboration (HAIC) in Industry 5.0 fundamentally reconfigures industrial work by integrating intelligent systems that augment human cognition, creativity, and decision-making—enabling adaptive manufacturing, personalized production, and novel business models (Dewangan & Kumar, 2024; Pawar et al., 2025). However, the innovative potential of HAIC is not automatic; it is critically contingent upon the level of trust that human operators place in AI partners. Trust in this context is not passive acceptance, but an active, dynamic belief shaped by system transparency, explainability of AI decisions, reliability in volatile environments, and adherence to ethical principles such as fairness and accountability (Zywiołek, 2024; Ahangar et al., 2025). When AI systems are designed using human-centric approaches like Trust by Design which embeds ethical safeguards throughout the AI lifecycle operators are more likely to engage in collaborative risk-taking, knowledge exchange, and experimental behaviors that drive innovation (Merchán-Cruz et al., 2025). Conversely, opaque “black-box” systems, algorithmic biases, and poor user experiences erode trust, leading to disengagement or misuse, thereby stifling co-creative potential (Ramírez-Gordillo et al., 2026). Thus, trust functions as the essential psychological and operational bridge that transforms technical collaboration into meaningful innovation.

Proposition 1: In Industry 5.0, trust mediates the positive relationship between Human–AI Collaboration and innovation, such that higher levels of calibrated, context-aware trust amplify the innovative outcomes of human–AI partnerships.

Trust effect Innovation Performance

In the human-centric ecosystem of Industry 5.0, trust serves as a critical enabler of innovation performance, transforming reliable human–AI interactions into tangible outcomes such as novel product development, adaptive process redesign, and sustainable business model innovation. When human operators trust AI systems perceiving them as transparent, explainable, fair, and aligned with ethical and operational goals they are more likely to engage in exploratory behaviors, share tacit knowledge, and co-create solutions with intelligent agents (Zywiołek, 2024; Merchán-Cruz et al., 2025). This calibrated trust reduces cognitive load and psychological resistance, freeing cognitive resources for higher-order creative tasks that drive innovation (Ahangar et al., 2025). Conversely, in low-trust environments characterized by opaque algorithms, inconsistent AI behavior, or ethical ambiguities workers may override, ignore, or misuse AI recommendations, leading to suboptimal decisions and stifled innovation potential (Ramírez-Gordillo et al., 2026). Empirical studies further confirm that real-time trust assessment through behavioral indicators (e.g., gaze patterns, task delegation frequency) correlates strongly with collaborative efficiency and innovation output in industrial settings (Campagna et al., 2024, 2025). Thus, trust is not merely a relational outcome but a performance catalyst that directly shapes the quality, speed, and sustainability of innovation in Industry 5.0.

Proposition 2: In Industry 5.0, higher levels of trust in AI systems are positively associated with enhanced innovation performance, as trust facilitates cognitive engagement, risk-taking, and co-creative collaboration between humans and intelligent technologies.

Human-AI Collaboration effect Innovation Performance

Human–AI Collaboration (HAIC) lies at the heart of Industry 5.0’s vision for a human-centric, resilient, and sustainable industrial future, directly catalyzing innovation performance through the synergistic integration of human creativity and artificial intelligence. By augmenting human cognitive and physical capabilities, AI systems enable workers to engage in higher-order tasks such as strategic decision-making, adaptive problem-solving, and co-design of intelligent workflows (Dewangan & Kumar, 2024; Pawar et al., 2025). Collaborative robots (cobots), AI-driven analytics, and immersive technologies like AR/VR enhance operational efficiency, accelerate prototyping cycles, and facilitate the development of personalized products and circular business models key indicators of innovation performance in advanced manufacturing (Pawar et al., 2025; Liu et al., 2025). Moreover, HAIC fosters a dynamic feedback loop where human intuition guides AI learning, and AI insights refine human judgment, creating a continuous cycle of iterative improvement and novelty (Katuk et al., 2024). However, the full innovative potential of HAIC is only realized when systems are designed with human needs at the core ensuring usability, contextual relevance, and ethical alignment (Ahangar et al., 2025). Empirical evidence from smart manufacturing environments confirms that effective HAIC correlates with faster time-to-market, increased product customization, and enhanced process adaptability (Shabur et al., 2025).

Proposition 3: In Industry 5.0, Human–AI Collaboration is positively associated with innovation performance, as the integration of human expertise with AI capabilities enables adaptive, responsive, and human-centric innovation across products, processes, and business models.

Mediating Role of Trust on The Relationship Between Human-AI Collaboration and Innovation Performance

While Human–AI Collaboration (HAIC) provides the structural foundation for innovation in Industry 5.0, its transformative impact is not direct but contingent upon the mediating role of trust a dynamic, context-sensitive belief that enables humans to fully engage with AI as a collaborative partner rather than a mere tool. Without sufficient trust, even the most advanced AI systems may be met with skepticism, underutilization, or active resistance, thereby limiting their capacity to augment human creativity and drive innovation (Zywiołek, 2024; Ahangar et al., 2025). Trust acts as the psychological and operational conduit through which HAIC translates into innovation performance: when operators perceive AI systems as transparent, explainable, reliable, and ethically aligned principles embedded in frameworks like Trust by Design (Merchán-Cruz et al., 2025) they are more likely to share insights, delegate complex tasks, and co-create novel solutions, fostering adaptive and human-centric innovation (Ramírez-Gordillo et al., 2026). Empirical studies in industrial settings confirm that real-time trust calibration, informed by behavioral indicators and system responsiveness, significantly enhances collaborative efficiency and innovation output (Campagna et al., 2024, 2025). Thus, trust does not merely moderate but actively mediates the pathway from collaboration to innovation, transforming technical integration into meaningful co-creation.

Proposition 4: In Industry 5.0, trust mediates the positive relationship between Human–AI Collaboration and innovation performance, such that the effect of HAIC on innovation is realized primarily through the development of calibrated, context-aware trust between humans and AI systems.

Conclusion

This study advances our understanding of how innovation is cultivated in the emerging landscape of Industry 5.0 by proposing and elaborating a conceptual framework centered on the mediating role of trust in Human–AI Collaboration. The key finding is that trust is not merely a supportive condition but an essential mechanism through which human–AI partnerships translate into tangible innovation outcomes. Specifically, when AI systems are designed to be transparent, explainable, reliable, and ethically aligned, they foster calibrated trust that empowers human workers to engage in co-creation, experimentation, and adaptive problem-solving. Without this trust, even the most sophisticated AI integrations risk remaining underutilized or misaligned with human intent, thereby limiting their innovative potential.

Theoretically, this research contributes by bridging fragmented literatures on socio-technical systems, trust in automation, and innovation management into a unified model that positions trust as the dynamic linchpin between collaboration and performance. It reframes trust from a static attitude to an ongoing, context-sensitive process that evolves with interaction quality and system behavior. Practically, the framework offers actionable guidance for industrial leaders, AI developers, and policymakers: to unlock innovation in Industry 5.0, organizations must prioritize human-centric AI design that embeds ethical principles such as fairness, accountability, and user control throughout the system lifecycle. Real-time trust assessment tools, informed by behavioral cues and explainable interfaces, can further enable adaptive AI that responds to human needs and maintains collaborative synergy.

Despite its contributions, this study has limitations inherent to its conceptual nature. As a narrative review, it does not provide empirical validation of the proposed relationships, and its insights are derived primarily from recent, emerging literature that may evolve as Industry 5.0 matures. Additionally, the framework focuses on industrial manufacturing contexts and may not fully capture nuances in service-oriented or non-Western settings where cultural, regulatory, or workforce dynamics differ significantly. The operationalization of “innovation performance” also remains broad, encompassing product, process, and business model dimensions that may require distinct trust mechanisms.

Future research should seek to empirically test the proposed mediation model through mixed-methods studies in real-world Industry 5.0 environments, using both qualitative insights from workers and quantitative metrics of innovation output. Longitudinal designs could examine how trust evolves over time as humans and AI systems co-adapt. Furthermore, cross-cultural comparisons and sector-specific investigations such as in healthcare, energy, or logistics would enhance the framework’s generalizability. Finally, developing standardized, real-time trust calibration protocols that integrate behavioral sensing, explainable AI, and ethical safeguards represents a critical next step toward making trustworthy human–AI collaboration a scalable reality in the factories and workplaces of tomorrow.

-
- Acknowledgements:** The authors would like to express their sincere gratitude to the Kedah State Research Committee, UiTM Kedah Branch, for the generous funding provided under the Tabung Penyelidikan Am. This support was crucial in facilitating the research and ensuring the successful publication of this article.
- Funding Statement:** No Funding
- Conflict of Interest Statement:** The authors declare that there is no conflict of interest regarding the publication of this paper. All authors have contributed to this work and approved the final version of the manuscript for submission to the International Journal of Innovation and Industrial Revolution (IJIREV).
- Ethics Statement:** This study did not involve any human participants, animals, or sensitive data requiring ethical approval. The authors confirm that the research was conducted in accordance with accepted academic integrity and ethical publishing standards.
- Author Contribution Statement:** All authors contributed significantly to the development of this manuscript. All authors read and approved the final version of the manuscript prior to submission.
-

References

- Ahangar, M. N., Farhat, Z. A., & Sivanathan, A. (2025). AI trustworthiness in manufacturing: Challenges, toolkits, and the path to Industry 5.0. *Sensors*, 25(14), 4357. <https://doi.org/10.3390/s25144357>
- Bunte, A., Wunderlich, P., Moriz, N., Li, P., Mankowski, A., Rogalla, A., & Niggemann, O. (2019). Why symbolic ai is a key technology for self-adaption in the context of cpps. *2019 24th IEEE International Conference On Emerging Technologies*.
- Campagna, G., Chrysostomou, D., & Rehm, M. (2024). Analysis of facial features for trust evaluation in industrial human-robot collaboration. *In 2024 IEEE International Conference on Advanced Robotics and its Social Impacts (ARSO) (pp. 1–6)*. IEEE. <https://doi.org/10.1109/ARSO60199.2024.10557748>
- Campagna, G., Dadgostar, M., Chrysostomou, D., & Rehm, M. (2024). A data-driven approach utilizing body motion data for trust evaluation in industrial human-robot collaboration. *In 2024 33rd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN) (pp. 1984–1990)*. IEEE. <https://doi.org/10.1109/RO-MAN60168.2024.10731437>
- Campagna, G., Lagomarsino, M., Lorenzini, M., Chrysostomou, D., Rehm, M., & Ajoudani, A. (2025). Estimating trust in human-robot collaboration through behavioral indicators and explainability. *IEEE Robotics and Automation Letters*, 10(10), 10218–10225. <https://doi.org/10.1109/LRA.2025.3600170>
- Charalambous, G., Fletcher, S., & Webb, P. (2016). The development of a scale to evaluate trust in industrial human-robot collaboration. *International Journal of Social Robotics*, 8(2), 193–209. <https://doi.org/10.1007/s12369-015-0333-8>
- Chesbrough, H. &. (2014). Explicating open innovation: clarifying an emerging paradigm for understanding innovation. *New Frontiers in Open Innovation*. Oxford: Oxford University Press, Forthcoming, 3–28.
- Golovianko, M., Terziyan, V., Branytskyi, V., & Malyk, D. (2022). Industry 4.0 vs. Industry 5.0: Co-existence, transition, or a hybrid. *Procedia Computer Science*, 217, 102–113. <https://doi.org/10.1016/j.procs.2022.12.206>
- Katuk, N., Gilal, A. R., & Safdar, S. (2024). Artificial intelligence in Industry 5.0: Revolutionising human–computer synchronisation. *In The future of human–computer integration: Industry 5.0 technology, tools, and algorithms (pp. 12–24)*. CRC Press. <https://doi.org/10.1201/9781003479727-2>
- Nahavandi, S. (2019). Industry 5.0—A human-centric solution. *Sustainability*, 11(16), 4371. <https://doi.org/10.3390/su11164371>
- Pawar, P. P., Hemamalini, V., Kumari, S., & Tyagi, A. K. (2025). Industry 5.0: A new industrial revolution and technological applications. *In Computational intelligence for analysis of trends in Industry 4.0 and 5.0 (pp. 426–457)*. CRC Press. <https://doi.org/10.1201/9781003533023-17>
- Rožanec, J. M., Novalija, I., Zajec, P., Kenda, K., Tavakoli Ghinani, H., Suh, S., Veliou, E., Papamartzivanos, D., Giannetsos, T., Menesidou, S. A., Alonso, R., Cauli, N., Meloni, A., Recupero, D. R., Kyriazis, D., Sofianidis, G., Theodoropoulos, S., Fortuna, B., Mladenčić, D., & Soldatos, J. (2023). Human-centric artificial intelligence architecture for Industry 5.0 applications. *International Journal of Production Research*, 61(20), 6847–6872. <https://doi.org/10.1080/00207543.2022.2138611>
- Shabur, M. A., Shahriar, A., & Ara, M. A. (2025). From automation to collaboration: Exploring the impact of Industry 5.0 on sustainable manufacturing. *Discover Sustainability*, 6(1), 341. <https://doi.org/10.1007/s43621-025-01201-0>

- Xu, X., Lu, Y., Vogel-Heuser, B., & Wang, L. (2021). Industry 4.0 and Industry 5.0—Inception, conception and perception. *Journal of Manufacturing Systems*, 61, 530–535. <https://doi.org/10.1016/j.jmsy.2021.10.006>
- Yanytska, L. (2025). The rise of human-centric manufacturing in the Industry 5.0 era. *International Journal of Advanced Manufacturing Technology*, 139(9–10), 5067–5077. <https://doi.org/10.1007/s00170-025-16192-5>
- Zywiołek, J. (2024). Trust-building in AI–human partnerships within Industry 5.0. *System Safety: Human – Technical Facility – Environment*, 6(1), 89–98. <https://doi.org/10.2478/czoto-2024-0011>