

Type of
Contribution:

Research Paper
Review Paper
Case Study

ENERGY: JURNAL ILMIAH

ILMU-ILMU TEKNIK

Special Issue 2025 pp 345-361

DOI: 10.51747/energy.si2025.251



E-ISSN: 2962-2565

This article
contributes to:



4 QUALITY
EDUCATION



A Conceptual Framework for Human AI Collaboration: Ontological and Epistemological Perspectives

Meyti Eka Apriyani^{1*}, Syaad Patmanthara¹

¹ Department of Electrical Engineering and Informatics, State University of Malang, 65145, Indonesia

*meyti.eka.2505349@students.um.ac.id

Abstract

Collaboration between humans and artificial intelligence (AI) has become a pivotal phenomenon in the evolution of information systems, yet its philosophical foundations remain underexplored. This study develops an integrative conceptual framework that combines ontological and epistemological perspectives to examine how human–AI collaboration shapes knowledge creation and decision-making within sociotechnical contexts. The proposed framework identifies five ontological levels of AI agency and four epistemological processes underlying hybrid knowledge formation. It further integrates six interrelated dimensions—ontological, epistemological, technical, ethical, social, and organizational—that collectively define the dynamics of human–AI collaboration. The findings contribute to the theoretical discourse by introducing the constructs of quasi-epistemic entities and hybrid epistemology, which reconceptualize AI not merely as a computational artifact but as a participant in epistemic processes, thereby extending existing theories of distributed cognition and epistemic accountability beyond instrumental human–machine models. Practically, the framework informs the design of transparent, adaptive, and ethically aligned human–AI systems within information-intensive environments.

Keywords: Epistemology; Ontology; Human–AI Collaboration; Information Systems; Knowledge Co-Creation

1. Introduction

The rapid development of artificial intelligence (AI) technology has profoundly changed the structural paradigm of contemporary information systems,

Article Info

Submitted:

2025-10-25

Revised:

2025-12-21

Accepted:

2025-12-25

Published:

2025-12-30



This work is
licensed under a
Creative
Commons
Attribution-
NonCommercial
4.0 International
License

Publisher

Universitas
Panca Marga

raising significant philosophical questions related to the interaction between human and machine agents in organizational contexts. AI has shifted from being a mere computational instrument to an entity endowed with capabilities for self-learning, autonomous decision-making, and proactive participation in knowledge application processes [1]. These developments challenge traditional notions of human agency, cognitive mechanisms, and the delineation between human and non-human entities [2]. Consequently, the essence of collaboration between humans and AI has emerged as a critical issue, requiring a thorough examination encompassing not only technical aspects but also philosophical dilemmas, particularly in the domains of ontology and epistemology [3].

Despite the proliferation of research on human–AI collaboration, the current literature remains fragmented. Most studies emphasize technical issues such as explainability [4], [5] or ethical concerns such as algorithmic bias [6], but few systematically integrate philosophical foundations with practical system design. This lack of integration can lead to ineffective or ethically problematic systems. Empirical evidence further reveals paradoxes such as automation bias where humans overly rely on AI outputs even when erroneous [7] and cases where excessive transparency paradoxically diminishes user trust [8]. In the Indonesian context, AI adoption has grown rapidly across sectors, yet scholarly exploration of its ontological and epistemological implications remains limited. To address these challenges, this study identifies four fundamental problems: (1) ontological ambiguity regarding the status and agency of AI in human collaboration; (2) epistemological complexity in understanding how hybrid systems form and validate knowledge; (3) theoretical fragmentation across interdisciplinary studies; and (4) the gap between abstract philosophical discourse and practical implementation. Based on these issues, four research questions are proposed: (a) How does the ontological perspective explain the role and status of AI in collaboration (b) How does the epistemological perspective explain knowledge formation in hybrid systems (c) How can both perspectives be integrated into a coherent framework and (d) What theoretical and practical implications arise from this integration.

This scholarly investigation aims to establish a comprehensive conceptual framework designed to clarify the dynamics of human-AI collaboration through the lenses of ontological and epistemological inquiry. Specifically, the study seeks to: (1) articulate and analyze ontological perspectives that define the transformative role of AI—from a passive tool to an active collaborator; (2) develop a hybrid epistemological model that illuminates the genesis of knowledge arising from interactive human-AI engagements; (3) integrate philosophical foundations with technical, ethical, social, and organizational dimensions; (4) outline theoretical advancements to the current body of literature on information systems; and (5)

propose actionable guidelines for the design and governance of ethically responsible AI systems. Theoretically, the study contributes to information systems scholarship by introducing the concepts of quasi-epistemic entities and hybrid epistemology, which offer new insights into distributed cognition and collaboration. Practically, it provides design principles, evaluative frameworks, and policy recommendations to foster transparent, accountable, and adaptive AI-driven systems. Unlike previous research focusing solely on technical or ethical aspects, this study unifies ontological and epistemological perspectives to advance a holistic understanding of human–AI collaboration.

2. Related Work

Research on human–AI collaboration has expanded across multiple domains of information systems, emphasizing diverse dimensions such as technical integration, organizational impact, and ethical governance. However, despite the growing volume of studies, conceptual fragmentation persists, particularly concerning the ontological and epistemological foundations that underlie collaborative intelligence. Technical and functional perspectives have dominated recent research, focusing on optimizing task performance and human oversight. Frenette [5] and Seufert & Meier [6] describe hybrid intelligence systems where AI functions as a cognitive partner that enhances decision quality through machine-assisted reasoning. Similarly, Przegalińska et al. [7] emphasize the organizational value of AI–human teaming within resource-based and task–technology fit frameworks. These studies, however, tend to treat AI as an instrumental agent, overlooking questions of agency, intentionality, and epistemic status. From an ethical and socio-technical viewpoint, other scholars have explored fairness, transparency, and accountability. Li et al. [2] highlight the challenges of algorithmic bias, while Cabitza et al. [16] call for moving beyond the notion of agential AI toward more empirically grounded perspectives of human–machine collaboration. Although these works contribute to responsible AI discourse, they seldom examine the ontological implications of AI autonomy or the epistemological consequences of distributed cognition between humans and machines. Philosophical and epistemic perspectives on human–AI interaction remain relatively scarce within the information systems discipline. Yang & Ma [3] proposed early classifications of epistemic relationships in hybrid systems, while Lin [5] introduced the concept of *Cognitio Emergens* to capture the co-creation of knowledge in human–AI contexts. Nonetheless, these frameworks have not yet been synthesized into a coherent model that unites ontological and epistemological reasoning. Existing literature tends to treat ontology and epistemology separately, leading to theoretical gaps in explaining how knowledge is constructed, validated, and shared across human–AI systems. In

summary, prior research has illuminated valuable insights into the technical, ethical, and managerial aspects of human–AI collaboration but has not yet articulated a philosophically grounded conceptual framework. This study addresses that gap by integrating ontological and epistemological perspectives into a unified theoretical model that captures both the nature of AI’s agency and the dynamics of hybrid knowledge formation in sociotechnical environments.

3.Method

3.1 Conceptual Framework for Human AI Collaboration

This study adopts a theory-building conceptual methodology following established guidelines for conceptual research in information systems. Specifically, it employs an integrative conceptual synthesis approach, combining philosophical analysis, structured literature review, and abductive reasoning to develop a coherent framework for human–AI collaboration.

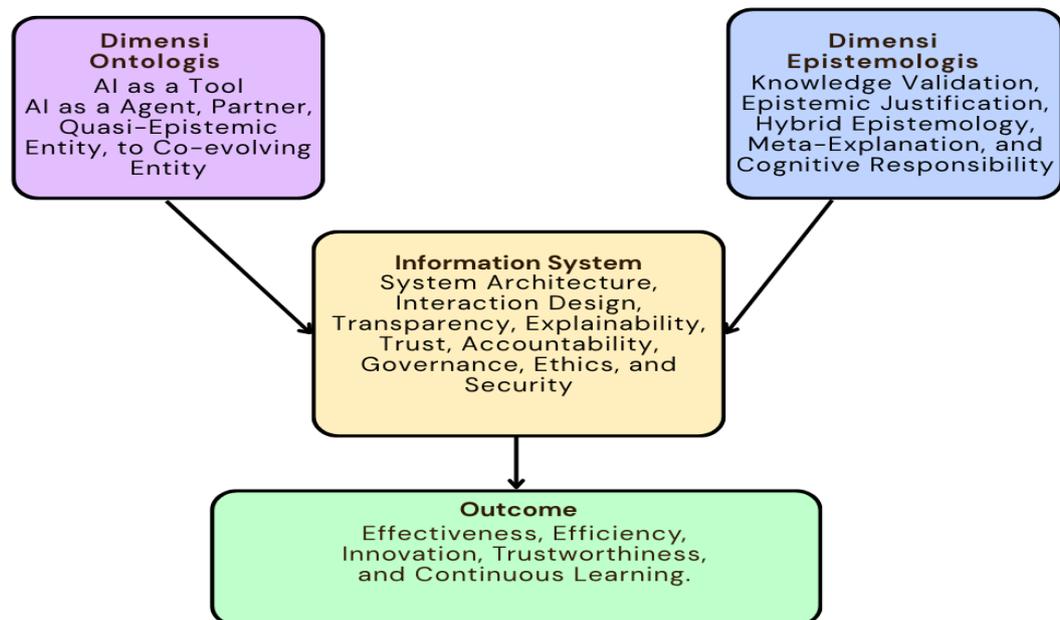


Figure 1. Conceptual Framework for Human–AI Collaboration

As shown in **Figure 1**, this conceptual framework consists of three main interrelated layers the Ontological Dimension, the Epistemological Dimension, and the Information System Context. These three layers work in an integrated manner to produce effective collaboration outcomes between humans and AI.

This framework is built on the premise that effective human-AI collaboration depends not only on the technical capabilities of the system, but also on a deep understanding of three fundamental questions: (1) what is the nature of AI and its ontological role in the system (ontological questions), (2) how humans and AI interact and jointly construct knowledge (epistemological questions), and (3) how

information systems can effectively facilitate this interaction (practical-technical questions). These three dimensions do not exist in isolation but rather influence each other and form a holistic whole that supports the creation of adaptive and meaningful human-AI collaboration.

3.2 Ontological Perspective: The Status and Role of AI in Collaboration

Analysis of the literature shows that the ontological status of artificial intelligence (AI) in collaboration with humans cannot be understood as a single category, but rather as a continuum representing varying degrees of agency and autonomy [3], [10]. This spectrum illustrates the evolution of AI's role from an entity functioning as a passive tool to an active co-creator in collaborative knowledge creation and innovation in **Figure 2**.

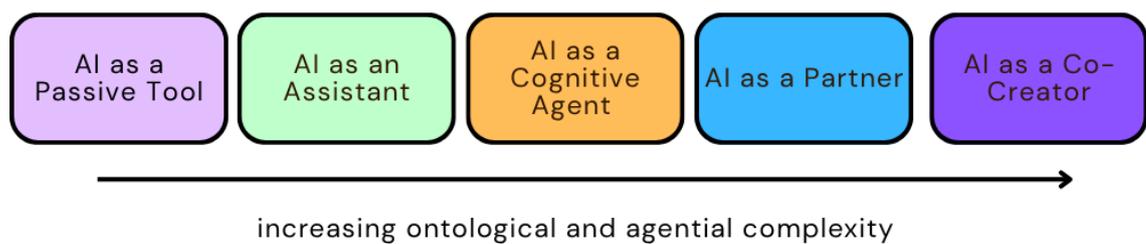


Figure 2. Spectrum Illustrates The Evolution

As shown in **Figure 2**, AI acts as a passive tool entirely controlled by humans. At this stage, AI has no agency or intentionality its function is limited to executing explicit commands from the user [2]. Relevant examples include calculators, number processing systems, and spreadsheet software. The relationship between humans and AI in this context is instrumental, with full responsibility for the results remaining with the human operator. Then, in Stage 2, AI transforms into an intelligent assistant with limited agency. At this stage, AI is capable of providing recommendations or suggestions based on data analysis, such as in recommendation systems, chatbots, or decision support systems [11]. Although the final decision remains with humans, AI begins to play a significant role as an advisor, enriching the decision-making process. Next, in Stage 3, AI functions as a cognitive agent with autonomous decision-making capabilities in specific domains [12]. Real-world examples include autonomous vehicles, trading algorithms, and machine learning-based medical diagnostic systems.

AI at this level demonstrates substantive agency through its ability to learn from experience and adapt behavior based on environmental feedback. Responsibility for the results becomes distributed between humans and the AI system. Then, at stage 4, AI is positioned as a collaborative partner on an equal footing with humans in carrying out shared tasks [13]. The relationship between humans and AI is dialogic and symmetrical, with both negotiating to determine work goals and strategies. Examples of applications include collaborative robots (cobots), co-creative systems, and advanced knowledge management systems. In this context, AI plays an active role in coordination and shared responsibility.

Finally, at level 5, AI is positioned as a co-creator with the ability to contribute creatively and innovatively to the creation of new knowledge [1].

Generative AI, automated scientific discovery systems, and creative AI are representative of this stage. Collaboration between humans and AI produces a form of "hybrid creativity" in which the boundaries between human and machine contributions become blurred. In this configuration, AI possesses generative agency and participates substantively in epistemic processes, while simultaneously raising new questions about the distribution of responsibility and ethical values in the outcomes of collaboration. Importantly, each ontological level conditions the epistemological possibilities of collaboration. As AI transitions from a passive tool to a co-creator, the epistemic processes shift from unidirectional validation toward reciprocal justification, negotiation, and meta-epistemic reflection. Thus, ontological status is not merely descriptive but constitutive of how knowledge is produced, validated, and distributed within human–AI systems.

Tabel 1. Ontological spectrum of ai collaboration

Level	Ontological Status	Key Characteristics	Example Systems	Human–AI Relationship	Responsibility Distribution
1	Passive Tool	Executes explicit commands	Spreadsheet, calculator	Instrumental	Full human responsibility
2	Intelligent Assistant	Provides recommendations	Chatbot, recommender system	Advisory	Shared (human-led)
3	Cognitive Agent	Autonomous decision-making	Trading algorithm, diagnostic AI	Delegative	Distributed (by domain)
4	Collaborative Partner	Dialogic and co-decision	Cobots, co-design AI	Symmetrical	Joint accountability
5	Co-Creator	Generates new knowledge	Generative AI, creative AI	Co-evolutionary	Shared epistemic responsibility

As shown in **Tabel 1**, This table delineates the ontological continuum of AI in collaborative contexts, emphasizing the progressive shift from instrumental to co-creative agency. Understanding the ontological spectrum of AI has several important implications for information system design. An ontological alignment is necessary between the system design and the expected role of AI. A system designed for AI as an assistant will be very different from one that accommodates AI as a collaborative partner [14]. A mismatch between technical design and ontological status can lead to inefficiencies and collaborative dysfunction. The principle of ontological transparency is crucial. Users must clearly understand the ontological status of the AI they are using whether it functions as a tool, assistant, agent, or partner [15]. Ambiguity about this status can create unrealistic expectations and lead to misinterpretation of the AI's work. Systems need ontological flexibility, the ability

to adapt the AI's role to the user's context and needs [16]. In some situations, AI may simply function as an assistant, but in other contexts it may evolve into a collaborative partner. This flexibility allows for better adaptation to the dynamics of the work environment and the complexity of the tasks at hand. Thus, the ontological perspective not only provides a philosophical framework for understanding the position of AI, but also contributes directly to more reflective, adaptive, and ethical information system design practices, in line with the principles of human-AI collaboration that are oriented towards shared values and responsibilities.

3.3 Epistemological Perspective: Knowledge Formation in Human–AI Collaboration

Hybrid Epistemological Model

Human–AI collaboration raises a fundamental epistemological question: how is knowledge formed, validated, and justified when cognitive processes involve non-human entities operating with architectures, representations, and logics that differ from those of humans? To address this, this study proposes a hybrid epistemological model that integrates human and AI epistemic processes within a coherent framework [17]. Rather than simply summing “human knowledge” and “AI output,” this model views knowledge generated by human–AI systems as emergent knowledge that emerges from dynamic interactions, iterative feedback, and a complementary division of cognitive roles. To illustrate the comparative epistemic characteristics of human and AI agents within the hybrid epistemological model, **Tabel 2.** presents the main features and their integrative outcomes.

This table summarizes the complementary epistemic roles of humans and AI, forming the basis of hybrid knowledge through synergy, mutual validation, and distributed cognition. The epistemic interplay summarized in **Tabel 2.** underpins the dynamic processes of validation, justification, negotiation, and reflection

Epistemic Processes in Human–AI Collaboration

The formation of hybrid knowledge relies on a series of interrelated epistemic processes. First, knowledge validation is carried out through dual mechanisms: technical validation to check the accuracy, consistency, and reliability of AI output using quantitative metrics; and contextual validation by humans to assess relevance, appropriateness, and meaning in a specific context. Both are strengthened by cross-validation—using AI output to check human judgment and vice versa [18]. Second, epistemic justification addresses the opacity of AI reasoning (the black box problem) with an explainable AI (XAI) approach to increase transparency [19], meta-explanation to explain how AI explanations are generated [20], and hybrid justification that combines algorithmic argumentation with contextual and normative human reasoning [21]. Third, epistemic negotiation is necessary when conflicts arise between human judgment and AI output: identifying the source of the

discrepancy (data, assumptions, or evaluation criteria), assessing the relative strengths (domains where AI is reliable vs. where humans excel), and synthesizing solutions that integrate insights from both [22]. Fourth, meta-epistemic reflection ensures adaptive collaboration: recognizing each agent's limitations and blind spots, calibrating trust in AI based on track record and context, and building continuous epistemic learning about how humans and AI best work together [23].

Tabel 2. Comparative epistemic characteristics in the hybrid epistemological model

Aspect	Human Epistemic Characteristics	AI Epistemic Characteristics	Hybrid / Emergent Knowledge Outcome
Nature of Cognition	Based on experience, intuition, and contextual understanding	Based on algorithmic inference and data pattern recognition	Complementary reasoning that combines context sensitivity with computational precision
Knowledge Source	Tacit knowledge, social interaction, creativity, ethical reasoning	Explicit data, structured models, statistical learning	Co-created insights integrating tacit and explicit dimensions
Validation Mechanism	Contextual relevance and human judgment	Quantitative accuracy and consistency metrics	Dual validation: cross-verification between human and AI
Learning Process	Reflective and adaptive based on experience	Iterative training through feedback and data	Co-evolution through continuous mutual adaptation
Representation of Knowledge	Semantic, narrative, and symbolic	Numeric, logical, and probabilistic	Hybrid representation linking symbols and patterns
Epistemic Limitation	Prone to cognitive bias, limited scalability	Prone to data bias, lack of contextual understanding	Balanced hybrid cognition minimizing both weaknesses

Epistemological Challenges

While promising, human–AI collaboration poses several epistemological challenges. First, the black box problem: many modern models (e.g., deep networks) are difficult to inspect, creating a dilemma – can knowledge whose processes cannot be understood still be accepted as a basis for action? Responses in the literature include strengthening XAI, implementing the “trust but verify” principle through independent verification, and limiting the use of opaque models to low-risk domains [24]. Second, epistemic responsibility: when decisions are made by hybrid systems, responsibility for knowledge is not singular, but distributed among designers, operators, and the AI system itself, requiring governance designs that explicitly allocate roles, authority, and mechanisms for epistemic accountability [25]. Third, epistemic bias: AI can inherit or amplify biases in training data, thus undetectably compromising the quality of knowledge. This requires detection of bias in data and

models, debiasing techniques, and transparency about limitations and potential biases as part of epistemic documentation [26]. Fourth, overreliance and automation bias: the human tendency to rely too much on AI advice can dull critical reasoning and degrade epistemic quality. Therefore, systems need to be designed to encourage appropriate reliance a level of trust calibrated to context, uncertainty, and the model's performance track record [27]. By systematically addressing these challenges, hybrid epistemological models provide a conceptual foundation for producing knowledge that is more reliable, accountable, and useful in practice.

Framework Integration: Dimensions of Human-AI Collaboration

A comprehensive conceptual framework for human-AI collaboration requires a harmonious integration of ontological and epistemological perspectives with other practical dimensions that influence the interaction and implementation of these systems. This research identifies six key interrelated dimensions that together form the basis for effective collaboration between humans and AI. These dimensions include ontological, epistemological, technical, ethical, social, and organizational dimensions, each with its own role in supporting productive and sustainable interactions between humans and machines.

The ontological dimension addresses fundamental questions regarding the status, agency, and relationship between humans and AI. It highlights AI's position in collaboration, from being a mere passive tool to functioning as an active co-creator in the knowledge creation process. Understanding this ontological aspect is crucial for determining the expectations and limitations of AI's role in collaborative systems. Meanwhile, the epistemological dimension highlights how knowledge is formulated, validated, and justified in hybrid systems. This process ensures that the resulting knowledge is valid, accountable, and relevant to the shared goals of the collaboration, including mechanisms for validating AI output and justifying its knowledge. From a technical perspective, this dimension encompasses implementation elements such as system architecture, interface design, algorithms, and AI models to support collaborative tasks, data management, and security and privacy aspects to protect user data.

The ethical dimension highlights fundamental issues related to morality, such as establishing accountability for errors, process transparency, fair distribution of benefits and risks, respect for human autonomy, and the principles of beneficence and non-maleficence to maximize the benefits and minimize the negative impacts of AI use. The social aspect encompasses an analysis of social capital and the dynamics of human-AI interactions, including trust building, the influence of organizational culture, factors of technology adoption and resistance, team collaboration dynamics, and the changing roles of humans driven by the integration of AI within organizations and society. The organizational dimension emphasizes the

organization's internal structures and processes, such as changes in organizational structure and governance, the integration of AI into business processes, strategic harmonization, and change management to ensure a successful transformation toward an AI-based organization. By understanding and integrating these six dimensions, human-AI collaboration systems can be designed that are not only technically effective but also adaptable to the demands of sustainability, ethics, and the dynamics of social and organizational change.

4. Results and Discussion

To capture the distinctive epistemic role of AI in collaborative systems, this study introduces the concept of quasi-epistemic entities (QEEs), defined as non-human agents that participate in epistemic processes—such as inference, validation, and knowledge generation—without possessing full epistemic autonomy or intentionality traditionally attributed to human knowers. Based on the multi-dimensional analysis outlined previously, this study proposes a layered implementation framework for human AI collaboration in information systems. This framework is designed not only to ensure effective technical integration of AI but also to ensure sustainability, adaptability, and ethical compliance across various organizational processes. The framework consists of five main layers that interact with each other and contribute to guiding the design and implementation of optimal collaborative systems.

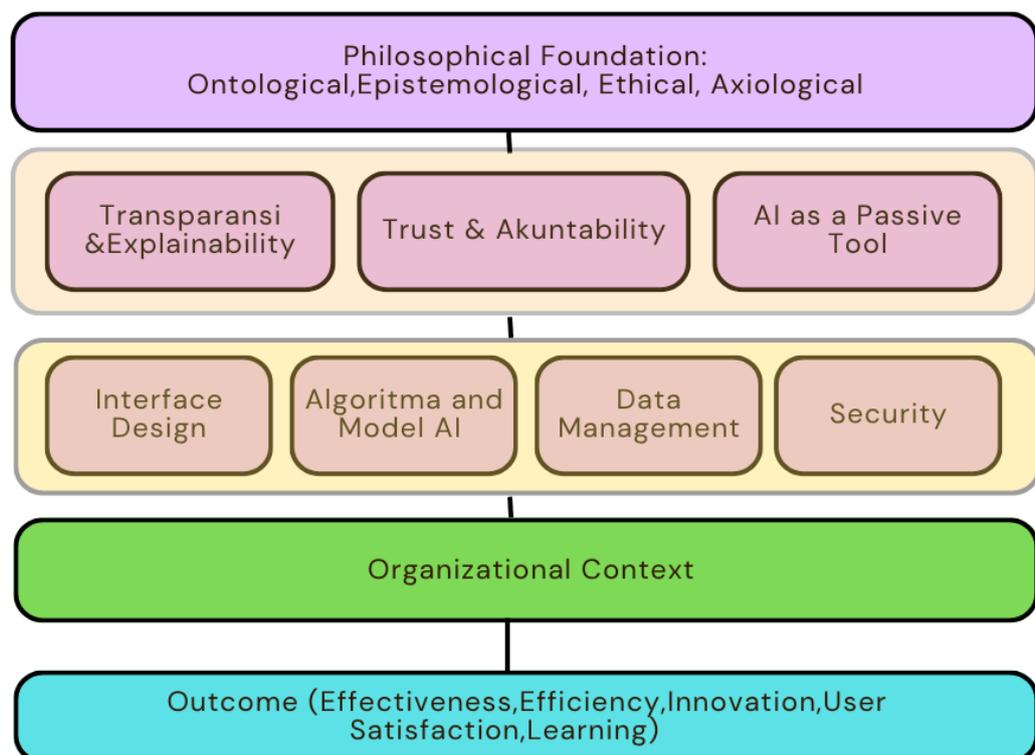


Figure 3. Framework for human AI

As shown in **Figure 3**, the first layer, the philosophical foundation, serves as the foundation for the entire implementation process. This foundation contains a set of ontological, epistemological, ethical, and axiological principles that serve as a reference for interpreting the status of AI, the mechanisms for knowledge formation and validation, and the upholding of moral values throughout the collaboration process [36], [39], [44]. With a solid foundation, the relationship between humans and AI can be built proportionally while minimizing various epistemic and ethical risks. The second layer, the operational design principles, derive from translating philosophical aspects into concrete implementation guidelines. Principles such as transparency and explainability ensure every process can be explained, trust and accountability build confidence and establish clear responsibilities, and adaptability and learning affirm the system's ability to adapt and continuously improve through user input and new data.

The third layer of the framework focuses on the technical aspects of implementation, including user-based interface design so that interactions between humans and AI are intuitive and efficient; the development of relevant AI algorithms and models to generate informative and accurate decisions; robust data governance to manage the volume and quality of collaborative data [31], [32]; and the implementation of stringent security measures to protect data and maintain user privacy. The fourth layer emphasizes the broader organizational context, considering the prevailing culture, structure, business processes, policies, and governance. The implemented system must be able to adapt to the dynamics of the organization and support the achievement of overall strategic goals. The fifth layer, outcomes, serves as indicators of success oriented toward the expected end results. Outcomes include the effectiveness of achieving organizational goals, operational efficiency, innovation resulting from human-AI collaboration, user satisfaction levels, and a continuous learning process that allows the system to evolve and improve performance over time. This framework is expected to optimize the human-AI collaboration process in various information system sectors, providing a comprehensive and applicable framework with a strong foundation of principles, in order to produce optimal value for users and organizations. This scholarly investigation has adeptly formulated a comprehensive conceptual framework aimed at elucidating human-AI collaboration within information systems, particularly through ontological and epistemological lenses.

The findings of this research indicate that the ontological status of AI in collaborative contexts does not conform to a singular category; rather, it manifests as a continuum that spans from AI functioning as a passive tool to its role as an active co-creator. Five distinct ontological tiers have been delineated: AI as a passive tool, assistant, cognitive agent, partner, and co-creator. In order to encapsulate the

distinctive function of AI within epistemic processes, the notion of quasi-epistemic entities (QEEs) is introduced as a novel ontological classification. From an epistemological standpoint, knowledge within human-AI systems transcends mere aggregation of human knowledge and AI-generated output; it is instead an emergent construct arising from the dynamic interplay between the two entities. The hybrid epistemological model elucidates the mechanisms of knowledge validation, epistemic justification, epistemic negotiation, and meta-epistemic reflection that transpire during collaborative interactions. Nonetheless, several critical epistemological obstacles were identified, including the black box problem, distributed epistemic responsibility, epistemic bias, and automation bias, all of which warrant careful consideration in the advancement of human-AI-centric information systems.

The multidimensional framework produced by this research underscores the imperative to amalgamate six pivotal dimensions—ontological, epistemological, technical, ethical, social, and organizational—to realize optimal human-AI collaboration. The layered implementation framework devised herein comprises philosophical underpinnings, design principles, technical components, organizational context, and anticipated outcomes to steer comprehensive system design and execution. Theoretically, this framework contributes novel expansions to ontological classifications, fortifying distributed epistemology, redefining the human-computer interaction paradigm, and proposing a theory of distributed responsibility. Practically, this framework furnishes explicit guidance for the design of transparent and accountable information systems, collaborative governance mechanisms, evaluation metrics, and serves as a reference point for cultivating human resource competencies through adaptive training programs [37].

In summation, this research yields substantial contributions from both theoretical and practical standpoints. The synthesis of fragmented philosophical discourses enriches the existing body of literature in information systems, while the resulting conceptual framework offers a coherent analytical structure alongside new conceptual terminologies such as QEEs and hybrid epistemology. In the realm of implementation, this framework equips organizations and information systems designers with guiding principles, evaluation metrics, governance frameworks, and a collaborative competency training framework, thereby addressing the necessity for adaptation in the increasingly intricate and dynamic landscape of human-AI collaboration.

5. Conclusion

Collaboration between humans and artificial intelligence constitutes an emergent and transformative area within the discipline of information systems,

prompting profound philosophical inquiries regarding the essence of agency, processes of knowledge formation, and accountability within sociotechnical ecosystems. This investigation has developed a conceptual framework that synthesizes ontological and epistemological perspectives to furnish an analytical basis for comprehending human–AI collaboration across various strata of the information technology milieu. The results underscore that the efficacy of human–AI collaboration transcends mere technical performance, necessitating a profound understanding of AI’s ontological standing, the epistemic processes that underpin hybrid knowledge construction, and the comprehensive integration of ethical, social, and organizational factors. This paradigm shift—from perceiving AI solely as a tool to acknowledging it as an epistemic collaborator—carries extensive ramifications for the conceptualization and governance of information systems in the contemporary digital landscape.

Theoretically, this investigation propels the dialogue surrounding human–AI collaboration by amalgamating philosophical exploration with the formulation and stewardship of intelligent systems. It broadens the comprehension of distributed epistemology and presents innovative constructs such as quasi-epistemic entities and hybrid epistemology, thereby facilitating the advancement of a cohesive framework for interpreting the interdependence between humans and AI. Practically, the proposed framework yields pragmatic insights for system architects, organizational executives, and policymakers. It emphasizes the imperative of integrating philosophical contemplation into the lifecycle of system development—ensuring that principles of transparency, accountability, and equity are woven into each phase of technological deployment. Moreover, it guides the formulation of governance structures and educational initiatives aimed at enhancing AI literacy and fostering ethical co-creation between human agents and intelligent systems.

Nevertheless, this framework remains bounded by its philosophical scope. It does not claim that AI possesses consciousness or moral agency, nor does it resolve debates on strong AI. Instead, it offers a pragmatically grounded epistemological account of collaboration, applicable to contemporary sociotechnical systems where human oversight and value alignment remain central.

Acknowledgement

The author would like to express sincere gratitude for the financial support received through the PDDI Scholarship 2025. This research would not have been possible without the guidance and support of Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbudristek). The authors acknowledge the use of AI-assisted tools (ChatGPT).

Authors' Declaration

Authors' contributions and responsibilities - The authors made substantial contributions to the conception and design of the study. The authors took responsibility for data analysis, interpretation, and discussion of results. The authors read and approved the final manuscript.

Funding - No funding information from the authors.

Availability of data and materials - All data is available from the authors.

Competing interests - The authors declare no competing interest.

Additional information - No additional information from the authors.

References

- [1] Human-AI Co-Scholarship: Reframing AI as an Epistemic Contributor to Knowledge Creation. Social Science Research Network. DOI: 10.2139/ssrn.5310319. Diakses dari https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5310319.
- [2] Constructivist Mixed Human-AI Approaches Overcome Epistemic Limitations of LLMs: A Cognitive Insight from Socio-Technical Research. AIS Electronic Library. Diakses dari <https://aisel.aisnet.org/oisiworkshop2025/11/>
- [3] Symbiotic Epistemology: Quasi-Epistemological Entities and the Philosophy of Human-AI Cognitive Partnership. PhilPapers. Diakses dari <https://philpapers.org/rec/KAPSEQ>.
- [4] Co-evolutionary Intelligence: Rethinking Human-AI Interaction. PhilPapers. Diakses dari <https://philpapers.org/rec/MOCCIR>.
- [5] Jaakkola, E. (2020). Designing conceptual articles: four approaches. *AMS Review*, 10(1-2), 18-26.
- [6] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- [7] Popay, J., et al. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. *ESRC Methods Programme*, 15(1), 47-71.
- [8] Williamson, T. (2020). *Philosophical Method: A Very Short Introduction*. Oxford University Press.
- [9] Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, 14(4), 490-495.
- [10] Smirnov, A., & Ponomarev, A. (2023). Collaborative Decision Support with Ontology-Based Neuro-Symbolic Artificial Intelligence: Challenges and Conceptual Model. DOI: 10.1007/978-3-031-19620-1_6. Diakses dari <https://scispace.com/papers/collaborative-decision-support-with-ontology-based-neuro-16zc7mby>.
- [11] Smirnov, A., & Ponomarev, A. (2023). Ontology-Based Explanations of Neural Networks for Collaborative Human-AI Decision Support Systems. DOI: 10.1007/978-3-031-43789-2_33. Diakses dari <https://scispace.com/papers/ontology-based-explanations-of-neural->

- networks-for-2p0fqcit67
- [12] Fabri, L., Häckel, B., & Oberländer, A. M. (2023). Disentangling Human-AI Hybrids. *Business & Information Systems Engineering*. DOI: 10.1007/s12599-023-00810-1. Diakses dari <https://scispace.com/papers/disentangling-human-ai-hybrids-1z7exvwo>.
 - [13] Patil, S. R., Sharma, S., & Mahdavi, M. (2025). From Loop to Partnership: A Framework for Understanding the Evolving Paradigms of Human-AI Collaboration. DOI: 10.22541/au.175647875.57157328/v1. Diakses dari <https://scispace.com/papers/from-loop-to-partnership-a-framework-for-understanding-the-9earzrva1o16>.
 - [14] Seeber, I., et al. (2020). Machines as teammates: A research agenda on AI in team collaboration. *Information & Management*, 57(2), 103174.
 - [15] Bansal, G., et al. (2021). Does the whole exceed its parts? The effect of AI explanations on complementary team performance. *CHI Conference on Human Factors in Computing Systems*, 1-16.
 - [16] Dellermann, D., et al. (2019). Hybrid intelligence. *Business & Information Systems Engineering*, 61(5), 637-643.
 - [17] Adhnouss, F., El-Asfour, H., & McIsaac, K. (2023). A Hybrid Approach to Representing Shared Conceptualization in Decentralized AI Systems: Integrating Epistemology, Ontology, and Epistemic Logic. *AppliedMath*. DOI: 10.3390/appliedmath3030032. Diakses dari <https://scispace.com/papers/a-hybrid-approach-to-representing-shared-conceptualization-2bj5bubne1>
 - [18] Lebovitz, S., Levina, N., & Lifshitz-Assaf, H. (2022). To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. *Organization Science*, 33(1), 126-148.
 - [19] Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE Access*, 6, 52138-52160.
 - [20] Younas, A. (2024). A Philosophical Inquiry into AI-Inclusive Epistemology. *Social Science Research Network*. DOI: 10.2139/ssrn.4822881. Diakses dari <https://scispace.com/papers/a-philosophical-inquiry-into-ai-inclusive-epistemology-5319yp4afn>
 - [21] Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1-38.
 - [22] Wilder, B., Horvitz, E., & Kamar, E. (2021). Learning to complement humans. *International Joint Conference on Artificial Intelligence*, 1526-1533.
 - [23] Lai, V., & Tan, C. (2019). On human predictions with explanations and predictions of machine learning models: A case study on deception detection. *FAT Conference**, 29-38.*
 - [24] Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 2053951715622512.
 - [25] Coeckelbergh, M. (2020). Artificial intelligence, responsibility attribution, and a relational justification of explainability. *Science and Engineering Ethics*, 26(4), 2051-2068.
 - [26] Ntoutsis, E., et al. (2020). Bias in data-driven artificial intelligence systems – An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and*

- Knowledge Discovery, 10(3), e1356.
- [27] Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: a systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1), 121-127.
- [28] Abbass, H. A. (2019). Social integration of artificial intelligence: Functions, automation allocation logic and human-autonomy trust. *Cognitive Computation*, 11(2), 159-171.
- [29] Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389-399.
- [30] Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, 46(1), 192-210.
- [31] Benbya, H., Davenport, T. H., & Pachidi, S. (2020). Artificial intelligence in organizations: Current state and future opportunities. *MIS Quarterly Executive*, 19(4), 9.
- [32] Amershi, S., et al. (2019). Guidelines for human-AI interaction. *CHI Conference on Human Factors in Computing Systems*, 1-13.
- [33] Mingers, J., & Willcocks, L. (2017). An integrative semiotic methodology for IS research. *Information and Organization*, 27(1), 17-36.
- [34] Design Principles for Human-AI Collaborative Knowledge Service Systems. DOI: 10.1007/978-3-031-95901-1_1. Diakses dari https://link.springer.com/chapter/10.1007/978-3-031-95901-1_1
- [35] Kahr, P., et al. (2023). Transparent machine learning in hospitality: Designing explanation types for hotel revenue management forecasting. *International Journal of Hospitality Management*, 108, 103357.
- [36] Dignum, V. (2019). *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*. Springer.
- [37] Burton, J. W., et al. (2020). How to integrate artificial intelligence into the classroom. *Nature Machine Intelligence*, 2(11), 631-636.
- [38] Smirnov, A., & Ponomarev, A. (2022). Collaborative Decision Support with Ontology-Based Neuro-Symbolic Artificial Intelligence. DOI: 10.1007/978-3-031-19620-1_6
- [39] Cath, C. (2018). Governing artificial intelligence: ethical, legal and technical opportunities and challenges. *Philosophical Transactions of the Royal Society A*, 376(2133), 20180080.
- [40] Grover, V., & Lyytinen, K. (2015). New state of play in information systems research: The push to the edges. *MIS Quarterly*, 39(2), 271-296.
- [41] Faraj, S., et al. (2018). Working on and with algorithmic systems. *Information and Organization*, 28(1), 62-70.
- [42] Rai, A., Constantinides, P., & Sarker, S. (2019). Editor's comments: Next-generation digital platforms. *MIS Quarterly*, 43(1), iii-ix.
- [43] Durán, J. M., & Jongasma, K. R. (2021). Who is afraid of black box algorithms? On the epistemological and ethical basis of trust in medical AI. *Journal of Medical Ethics*, 47(5), 329-335.
- [44] Cummings, M. L. (2006). Integrating ethics in design through the value-

sensitive design approach. *Science and Engineering Ethics*, 12(4), 701-715.