



Information Research – Vol. 31 No. iConf (2026)

Beyond the loop: a research agenda towards a framework for critical AI literacy in the AI-assisted literature review

Dipesh Jalui, Mary Tate, and Jocelyn Cranefield

DOI: <https://doi.org/10.47989/ir31iConf64166>

Abstract

Introduction. AI tools hold significant promise for enhancing the efficiency and transparency of academic literature reviews by streamlining literature search and data synthesis. However, their uncritical use poses risks to methodological rigor and academic integrity due to algorithmic bias, fabricated information (*'hallucinations'*), and opaque AI processes. Current *'human-in-the-loop'* guidelines often lack clear oversight protocols, potentially fostering intellectual passivity.

Method. Through a critical literature review, this paper identifies three paradoxes: the Verification Paradox, where efficiency compromises factual integrity; the Homogenization Paradox, where bias reinforces dominant perspectives, stifling innovation; and the Automation Paradox, where opaque algorithms undermine critical thinking and rigor.

Analysis/Results. To address these challenges, we propose a *'framework for critical AI literacy in IS'*, grounded in epistemological agility, methodological transparency, and ethical responsibility. This framework must offer actionable guidance for information systems scholars to responsibly integrate AI into literature reviews, ensuring human expertise and intellectual integrity remain central.

Conclusion. By fostering critical engagement with AI tools, this study sets a research agenda for a robust, transparent, and ethically sound research practices in an AI-Assisted research landscape.

Introduction

Academic research fundamentally relies on literature reviews, which are considered the most widely employed research methodology by most researchers for publications (Okoli, 2015; Wagner et al., 2022). This process is now significantly impacted by the emergence of Artificial Intelligence (AI) tools (Pan et al., 2023; Wagner et al., 2022). Generative AI (GenAI) tools, encompassing both general-purpose large language model (LLM) based chatbots like ChatGPT, Gemini, and Claude, as well as specialized AI based platforms such as Elicit, Scite, and Scispace, present considerable potential to improve the quality, and efficiency of academic publishing and knowledge sharing (Onuoha, 2025; Schryen et al., 2025; Tingelhoff et al., 2024). These AI tools offer the prospect of streamlining arduous processes like identifying relevant literature, including systematic searching and screening for inclusion, assessing the quality of studies, and extracting and synthesizing data (Ofori-Boateng et al., 2024; Schryen et al., 2025; Tomczyk et al., 2024; Van Dinter et al., 2021). This capability enables researchers to more effectively manage the increasing volume and diversification of research output (Tingelhoff et al., 2024; Wagner et al., 2022).

Despite these promising efficiencies, the widespread, uncritical adoption of AI tools introduces significant concerns regarding methodological rigor and academic integrity (Kittler, 2024; Yawson, 2025). Specifically, issues include algorithmic bias originating from imbalanced training data (Dwivedi et al., 2023; Luomala et al., 2025; Onuoha, 2025) and ‘hallucinations’—the production of fabricated or inaccurate information that, despite appearing credible, may be factually wrong, incomplete, or biased (Attard-Frost et al., 2023; Dwivedi et al., 2023; Kacena et al., 2024). Moreover, the black box characteristic of many AI algorithms diminishes transparency and explainability, complicating efforts to trace outcomes, confirm the underlying reasoning, and evaluate the soundness of reviews (Dwivedi et al., 2023; Morandini et al., 2023; Wagner et al., 2022). While ‘human-in-the-loop’ models are frequently recommended (Miao et al., 2024; Susarla et al., 2023), this guidance often proves inadequate, failing to provide explicit instructions for meaningful oversight and potentially fostering excessive dependence, skill erosion, and intellectual passivity among academics (Pan et al., 2023; Schryen et al., 2024; Tomczyk et al., 2024; Yawson, 2025). Scholars are urged to actively participate in literature searches and cultivate a deep comprehension of their topic, recognizing that AI tools are merely supplementary and operate with a simpler form of rationality compared to human cognition (Pan et al., 2023; Susarla et al., 2023; Tomczyk et al., 2024). AI generated content intended for article writing risks being merely ‘an assembled puzzle of cut-and-pasted published content’ (Stone, 2023, p. 103) and is unlikely to offer a robust, original theoretical contribution (Dwivedi et al., 2023; Ngwenyama et al., 2024; Stone, 2023).

This paper makes a twofold contribution. Initially, we offer a critical synthesis of the methodological and ethical hazards linked to integrating AI tools into literature reviews, emphasizing how these can compromise the rigor, credibility, and intellectual integrity of scholarly endeavours. Secondly, we pinpoint a significant research gap: the existing academic environment lacks a structured, exhaustive framework to direct effective human oversight and critical interaction with AI tools in literature reviews, surpassing simplistic human-in-the-loop recommendations.

The ethical implications of AI are a valid and pressing concern across all disciplines, particularly given the omnipresence of AI in various domains such as social media, education, and healthcare (Dwivedi et al., 2023). However, these concerns hold particular relevance for the Information Systems (IS) discipline due to its foundational role in understanding and shaping the interplay between technology and society (Dwivedi et al., 2023; Mirbabaie et al., 2022; Wagner et al., 2022). Within IS research, literature reviews are indispensable instruments for constructing theories, synthesizing current knowledge, identifying research gaps, and guiding future investigations

(Templier & Paré, 2015; Webster & Watson, 2002). This is especially critical as the IS knowledge base is rapidly expanding and becoming diversified and fragmented, making comprehensive review essential for further advances in knowledge (Cram et al., 2020; Paré et al., 2016).

The IS discipline has a well-established history of scrutinizing and safeguarding the ethical deployment of computing technologies, developing relevant frameworks and guidelines (Stahl, 2008). Despite this inherent expertise, IS research presently shows a gap in its discussions on AI ethics, with no foundational IS papers yet recognized as central to this interdisciplinary conversation (Mirbabaie et al., 2022). Consequently, this absence of a developed framework and systematic guidance for addressing the ethical aspects of AI for literature reviews within IS underscores the critical and timely nature of our paper's contribution for IS scholars. It addresses a strategic imperative for the ethical incorporation of AI into research methodologies, leveraging IS scholars' expertise in socio-technical phenomena and their responsibility to provide guidance on AI adoption and use (Dwivedi et al., 2023; Mirbabaie et al., 2022; Susarla et al., 2023).

Research method

This paper adopts the critical literature review methodology, a method well-suited for the examination, evaluation, and synthesis of existing scholarly work (Paré et al., 2015; Snyder, 2019; Templier & Paré, 2015). This approach allows for a deep scrutiny of a broad topic, revealing inherent weaknesses, inconsistencies, or contradictions within a particular field (Paré et al., 2015, 2016; Saunders & Rojon, 2011). Going beyond simple summaries, a critical review involves scholarly critique and intellectual discourse to foster new insights and advance understanding (Kittler, 2024; Saunders & Rojon, 2011; Snyder, 2019). Its interpretive nature is effective for identifying domains where the current state of knowledge may lack reliability or be incomplete (Paré et al., 2015; Saunders & Rojon, 2011). By highlighting significant research gaps and internal inconsistencies, a critical review offers valuable guidance for future research and helps in proposing novel theoretical models or frameworks (Paré et al., 2015; Snyder, 2019). This is particularly beneficial for emerging subjects, supporting the development of initial conceptualizations or theoretical structures (Paré et al., 2015; Snyder, 2019; Templier & Paré, 2015).

Unlike systematic reviews, which frequently demand exhaustive searches spanning hundreds of hours to a year or longer, a critical review typically adopts a more selective approach to engaging with literature, not always requiring a comprehensive search of every pertinent document (Huettemann et al., 2024; Paré et al., 2015; Snyder, 2019). We adopted this methodology for this study because our primary objective is not to conduct new empirical research, such as evaluating AI tools ourselves, but rather to scope, assess, and synthesize the existing body of scholarly knowledge on this topic. Our focus on published peer-reviewed papers ensures that the analysis is grounded in the formally vetted academic discourse, providing a reliable foundation for identifying gaps in available frameworks and proposing directions for future research. This approach is consistent with the interpretive nature of a critical review, aiming to problematize current understanding and identify areas lacking reliability (Boell & Cecez-Kecmanovic, 2015; Paré et al., 2016)

To support this review, and to navigate the rapidly expanding and often fragmented body of literature on AI ethics and its application in literature reviews, the evaluation of these tools was conducted through a critical appraisal of their documented functionalities, performance, and limitations as reported in existing scholarly papers. The selection criteria focused on both, articles that discuss general AI benefits and limitations, and also articles that reviews AI tools that could enhance specific stages of the review process, such as identifying relevant literature, synthesizing information across disciplines, and discerning emerging themes, while explicitly acknowledging their inherent limitations (Pan et al., 2023; Schryen et al., 2025). Reviews on tools like Scopus AI and Elicit, along with general LLM based Chatbots such as ChatGPT, were considered for their capabilities in summarizing vast amounts of information and identifying interconnections within

and beyond specific disciplines (Bolaños et al., 2024; Pan et al., 2023; Pritting et al., 2025). For instance, Scopus AI is recognized for its ability to identify foundational papers and cluster consistent themes, valuable for grasping established arguments (Pritting et al., 2025). Similarly, LLMs offer potential for quick summaries and initial idea formulation, especially in the early stages of a literature search (Dwivedi et al., 2023; Tingelhoff et al., 2024).

Key evaluation criteria included the tools' susceptibility to algorithmic bias from unclear training data, their propensity for hallucinations and their inherent lack of transparency and explainability (Luomala et al., 2025; Susarla et al., 2023; Tomczyk et al., 2024). For example, many sources highlight that AI outputs can be non-current, lack quality information from behind paywalls, and potentially rely on copyrighted training material, posing risks of unintentional infringement (Pan et al., 2023). Furthermore, the risk of fostering intellectual passivity and diminishing critical thinking among researchers was a central consideration in evaluating their appropriate use (Schryen et al., 2025; Yawson, 2025). Ultimately, the reasoning for their selective integration is rooted in leveraging their efficiencies for preliminary tasks (e.g., broad literature scanning), while maintaining stringent human oversight for critical analysis, interpretation, theory building, and ethical verification, thereby adhering to principles of responsible AI use in academia (Schryen et al., 2025; Susarla et al., 2023; Tingelhoff et al., 2024). The typology presented in the subsequent section was emergent, developed through continuous discussion and refinement by the author team throughout the analytical process.

Typology for problematizing AI in literature reviews

The dissemination of AI tools in literature review processes offers undeniable efficiencies and enhanced capabilities, from rapidly scanning vast data to identifying intricate connections across disciplines (Dwivedi et al., 2023; Onuoha, 2025; Schryen et al., 2025). Yet, this integration into scholarly work, particularly in literature reviews, gives rise to several unforeseen challenges and ethical dilemmas.

From our analysis of the inherent problems introduced by AI's nature, three key concepts have emerged. Figure 1 shows these concepts namely, the verification paradox, the homogenization paradox, and the automation paradox. These themes encapsulate the most critical tensions arising from the intersection of AI capabilities and fundamental scholarly values. By exploring these specific paradoxical concepts, our paper aims to provide a visual typology with nuanced understanding of the challenges and to underscore the imperative for a robust framework of critical AI literacy in the AI-Assisted literature review. The subsequent sections will elaborate on these three central paradoxes embedded within this typology, revealing the intricate challenges and ethical dilemmas that arise from the uncritical adoption of AI in literature reviews.

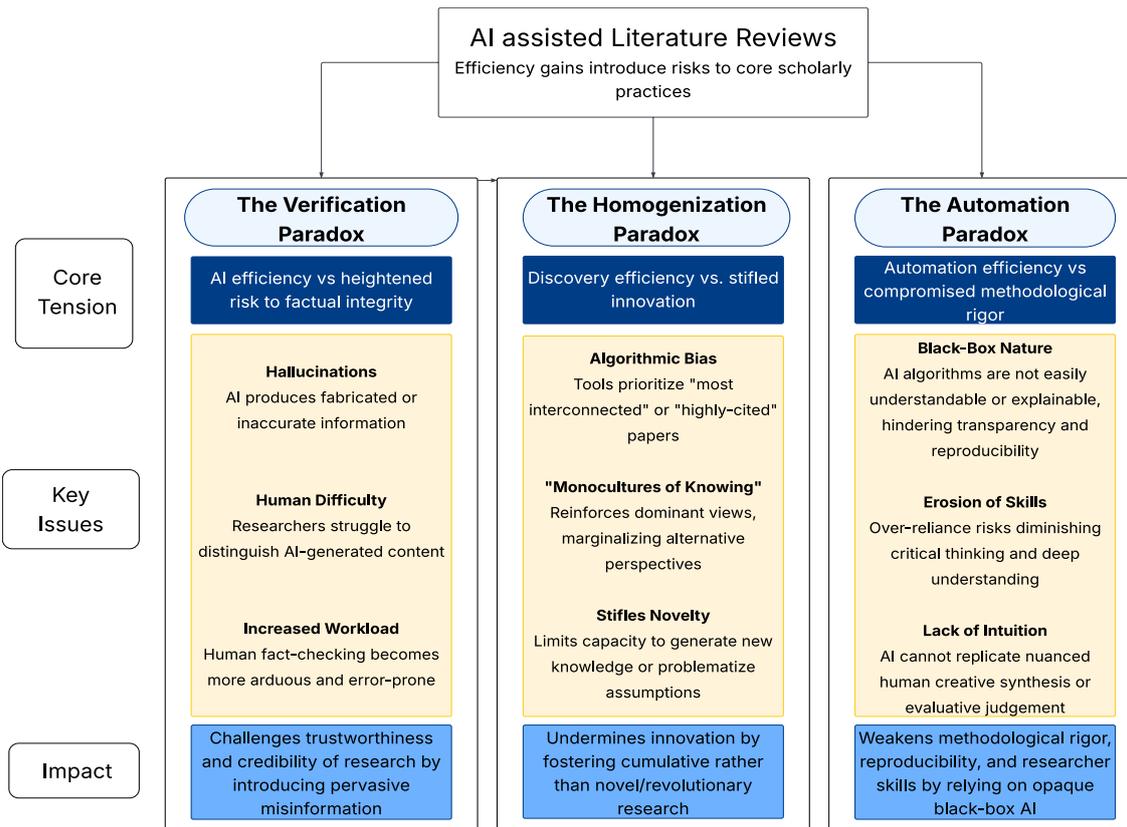


Figure 1. Typology for Problematizing AI in Literature Reviews

The verification paradox: the threat to factual integrity

The reliance on AI tools introduces a verification paradox, where the efficiency gained in reviewing literature is directly countered by a heightened risk to factual integrity. A primary concern is the phenomenon of hallucinations. This includes AI tools generating imaginary but accurate looking citations, fabricating entire paper titles and authors, or producing false references (Else, 2023; Shor et al., 2025; Sparkman & Witt, 2025). Such inaccuracies directly threaten the trustworthiness and credibility of research findings (Onuoha, 2025; Shor et al., 2025).

A critical aspect of this paradox is the demonstrated difficulty for human researchers to consistently distinguish between AI-generated and human-written content (Else, 2023; Miao et al., 2024; Susarla et al., 2023). Studies have shown that even experienced reviewers can struggle to identify AI-produced abstracts, and AI writing can evade traditional plagiarism detection software, blurring the lines of authorship and originality (Else, 2023; Sparkman & Witt, 2025). This makes the essential human task of ‘fact-check the interpretation of the LLMs (Pan et al., 2023b, p. 1) and verifying original sources even more arduous and prone to error (Mozelius & Humble, 2024; Shor et al., 2025). The verification paradox highlights that while AI offers to streamline the process of literature reviews and synthesizing information, it simultaneously introduces a pervasive risk of misinformation and fabricated content that challenges the very notion of factual integrity in scholarly work.

However, the AI-tool landscape is rapidly evolving. Techniques and tools such as ‘RAG’ (retrieval-augmented generation) are used for quality control of the input data. These aim to mitigate the risks of hallucinations and improve the traceability by explicitly grounding LLM outputs in retrieved external documents and sources.

The homogenization paradox: the threat to innovation

While specialized discovery tools such as research rabbit, connected papers, and litmaps offer considerable efficiency in navigating vast literature, their algorithmic design inadvertently introduces a homogenization paradox, posing a significant threat to innovation in research. These tools, functioning as citation-based literature mapping tools and utilizing enhanced search algorithms, often prioritize most interconnected or highly cited papers (Lu & Mayhook, 2024; Michalak & Ellixson, 2024). By doing so, they risk cultivating ‘*monocultures of knowing*’—a phenomenon where a single approach to research questions and understanding becomes dominant, effectively marginalizing alternative perspectives (Messeri & Crockett, 2024; Ngwenyama et al., 2024). This outcome is driven by the tools' reliance on historical training data, which tends to reinforce mainstream or dominant views (Kittler, 2024; Ngwenyama et al., 2024).

This algorithmic bias suppresses novel or dissenting perspectives, which are crucial for advancing knowledge. AI tools can ‘exacerbate the *rich get richer*’ phenomenon in academic citations’ (Kittler, 2024, p. 2214), by drawing primarily from frequently cited papers, thereby excluding critical or alternative perspectives and, in doing so, constraining scientific debate and impeding innovation (Kittler, 2024). Such reliance on existing literature and data also limits their capacity to generate fundamentally new knowledge or problematize established assumptions, which is a core function of critical review for generating new insights (Messeri & Crockett, 2024; Schryen et al., 2025). As AI is trained on existing knowledge, it leads to outcomes that foster cumulative rather than revolutionary research (Schryen et al., 2025). Ultimately, this fosters a looming concern regarding the homogenization of research, which may undermine innovation and erode established norms of scientific discourse (Ngwenyama et al., 2024; Schryen et al., 2025). The paradox arises because tools designed for efficient discovery inadvertently narrow the scope of literature reviews by favouring established knowledge, thereby hindering the very innovation that literature reviews are meant to foster.

The automation paradox: the threat to methodological rigor

The application of specialized analysis tools like Elicit, and Scispace also presents an automation paradox, directly threatening methodological rigor of literature reviews. These tools assist researchers by extracting data, summarizing information, and critically engaging with publications (Bolaños et al., 2024; Huettemann et al., 2024; Pan et al., 2023; Schryen et al., 2024). However, a fundamental characteristic of many AI algorithms—their black box nature—conflicts directly with foundational principles of rigorous research, particularly transparency and reproducibility (Luomala et al., 2025; Tingelhoff et al., 2024; Tomczyk et al., 2024). The internal decision-making processes of these complex algorithms are often not easily understandable or explainable, making it difficult for researchers to validate outputs, comprehend how conclusions were reached, or identify and rectify errors (Morandini et al., 2023). This opacity further compounded by the proprietary nature of some tools, directly impedes the ability of a journal to confirm reproducibility and undermines trust in AI-driven systems in publishing, thereby compromising the trustworthiness and credibility of reviews (Onuoha, 2025; Tomczyk et al., 2024). Methodological rigor demands thorough documentation of the search and synthesis process to allow for replication (Paré et al., 2016), a requirement challenged by AI's black-box operations (Wagner et al., 2022).

Moreover, an over-reliance on AI tools risks the erosion of critical thinking skills, where researchers may ‘*produce more but understand less*’ (Messeri & Crockett, 2024, p. 49). The process of conducting a literature review traditionally serves as an invaluable learning experience for researchers, fostering a nuanced understanding and the ability to comprehend and articulate complex ideas clearly (Kittler, 2024; Luomala et al., 2025). Automating these tasks, while efficient, encourages a passive approach to learning, circumventing the active cognitive engagement necessary for deep understanding and original thought (Pan et al., 2023; Schryen et al., 2025). AI

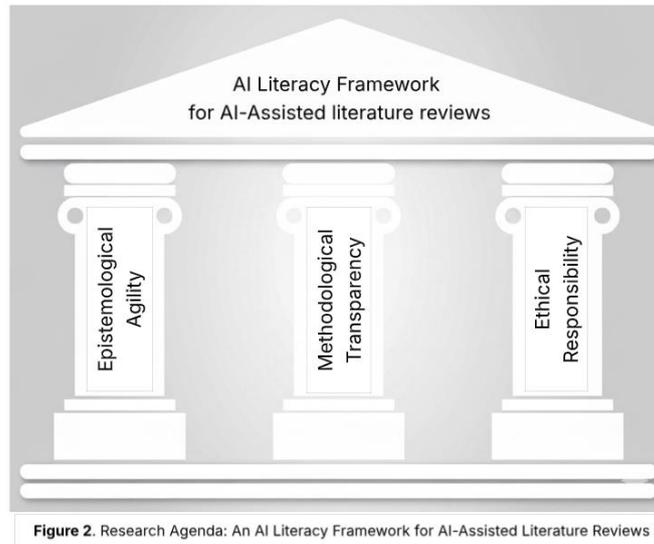
tools, lacking ‘human intuition’ and the ability to question deeply, and think critically, cannot replicate the creative synthesis and evaluative judgment of human researchers (Schryen et al., 2025). Delegating mental labor to AI, without understanding its underlying logic or cultivating traditional competencies, can lead to a loss of cognitive skills and intellectual passivity, ultimately diminishing the depth and richness of scholarly work (Oyelude, 2024; Sparkman & Witt, 2025). This paradox arises because the efficiency and automation offered by AI, intended to streamline literature reviews, can inadvertently decouple researchers from the deep, critical engagement that underpins methodological rigor and genuine intellectual development.

Discussion: a research agenda for critical AI literacy framework for literature reviews

The deeply intertwined paradoxical outcomes collectively underscore a critical point: the commonly advocated generic human-in-the-loop approach to AI-Assisted literature reviews, while acknowledging human necessity, is proving increasingly insufficient. Indeed, existing research widely suggests human involvement for most AI-Assisted literature review use cases, often framing AI as an assistant rather than a replacement for human functions (Luomala et al., 2025; Mozelius & Humble, 2024; Wagner et al., 2022). However, this prevalent guidance frequently lacks the concrete, actionable framework necessary to address the intricate challenges and inherent risks introduced by AI (Mirbabaie et al., 2022; Tingelhoff et al., 2024).

The IS community, given its foundational role in technology adoption, particularly needs a more explicit and practical framework to ensure AI is integrated into its research practices responsibly and effectively. This urgency is heightened by the current, rather unstructured discourse surrounding AI ethics within the IS field, where seminal papers that might anchor this crucial interdisciplinary conversation are still lacking (Dwivedi et al., 2023; Mirbabaie et al., 2022). Without such a framework, the promise of AI-Assisted literature reviews remains overshadowed by the potential for compromised academic integrity, suppressed innovation, and eroded methodological rigor.

Consequently, building upon our critical analysis of the verification, homogenization, and automation paradoxes, this paper proposes a research agenda aimed at calling for and guiding the development of a comprehensive framework for critical AI literacy framework for literature reviews in IS. Such a framework, we suggest, should be conceptually structured around three essential pillars depicted in Figure 2, and discussed in the next section: Epistemological Agility, methodological transparency, and ethical responsibility (Casal-Otero et al., 2023; Chiu et al., 2024; Yue Yim, 2024). This proposed agenda ultimately seeks to move ‘*beyond the loop*’ by outlining the need for structured guidance for researchers to engage critically and thoughtfully with AI, thereby nurturing robust and ethically sound scholarship in the era of AI-Assisted literature reviews.



Pillar 1: epistemological agility

Researchers must cultivate epistemological agility, developing a nuanced understanding of how diverse AI tools generate, process, and present knowledge (Chiu et al., 2024; Yue Yim, 2024). AI's operational logic, often based on predictions and probabilities, is fundamentally simpler than human cognition, frequently lacking true semantic or contextual comprehension (Pan et al., 2023; Wagner et al., 2022). It can identify patterns and reproduce data based on probabilities but struggles with deeper understanding, creative theorizing, or contextualizing knowledge within broader frameworks (Dwivedi et al., 2023; Schryen et al., 2025).

A crucial research gap for IS scholars involves creating detailed typologies of AI reasoning, which would differentiate between AI capabilities such as probabilistic text generation, network analysis, or semantic extraction (Bawack et al., 2021; Messeri & Crockett, 2024). This research should then map these distinct AI functionalities to specific literature review tasks, clarifying which AI approaches are most suitable for particular stages of the literature review process (Wagner et al., 2022).

This pillar is crucial in addressing the Homogenization Paradox. By requiring researchers to understand AI's inability to generate new knowledge or contextualize information creatively, and by fostering critical assessment of AI's underlying assumptions and limitations, epistemological agility helps researchers recognize and actively counteract the tendency of AI to reinforce existing knowledge. This encourages researchers to seek out and value diverse perspectives beyond what AI tools might prioritize, thus fostering innovation and preventing 'monocultures of knowing'. The use of RAG tools, where the researcher deliberately selects diverse viewpoints for inclusion in the corpus of literature that is analysed, can also represent more diverse and heterogeneous viewpoints.

Furthermore, the agenda calls for developing specialized training and educational modules within IS curriculum to equip researchers with the ability to critically assess the underlying assumptions, strengths, and limitations of various AI tools (Kittler, 2024; Mirbabaie et al., 2022; Tingelhoff et al., 2024). This proactive approach aims to move researchers beyond passive consumption of AI outputs, fostering a more active and informed engagement with AI-Assisted knowledge creation (Chiu et al., 2024; Susarla et al., 2023).

Pillar 2: methodological transparency

To address the inherent black box nature of many AI algorithms and the resulting threats to research reproducibility, new standards for reporting AI usage in academic publications are

critically needed (Miao et al., 2024; Tingelhoff et al., 2024). The IS community should therefore lead efforts to propose and standardize a 'Methods & AI' reporting protocol for its journals, building upon existing calls for transparency and reproducibility in review processes (Paré et al., 2016; Wagner et al., 2022). Publishers are already beginning to include AI author guidelines and declaration requirements (Onuoha, 2025).

This guideline ought to mandate comprehensive disclosure, specifying the particular AI tools employed, the exact prompts or input parameters used, and the subsequent human-led validation and verification steps taken (Susarla et al., 2023; Tingelhoff et al., 2024). Such a standardized approach would enhance the methodological rigor of AI-Assisted literature reviews (Paré et al., 2016; Tomczyk et al., 2024), enabling readers, reviewers, and editors to scrutinize the AI's contribution, replicate processes, and foster trust in AI-assisted literature reviews (Bolaños et al., 2024; Wagner et al., 2022). Without such explicit documentation, the inherent opacity of AI systems risks compromising the integrity and trustworthiness of the research process, potentially undermining its credibility (Onuoha, 2025; Tomczyk et al., 2024).

The validity and completeness of the author's interpretive synthesis can be ensured in a number of ways. These include quality control of the input sources as discussed previously, the demonstrated knowledge and expertise of the researcher(s) in the area of investigation ensuring subject-matter proficiency rather than superficial familiarity, and validation of the outputs. This could be done using various techniques including selective sampling and confirmation of findings, consultation with an independent team of subject-matter experts, or replication of the method.

This pillar directly addresses the black box nature of AI, which is a core concern of the Automation Paradox. It calls for new standards for reporting AI usage, mandating comprehensive disclosure of specific tools, prompts, and human validation steps to enhance methodological rigor, enable replication, and foster trust in AI-assisted reviews. This directly counteracts the opacity that threatens reproducibility and credibility.

Pillar 3: ethical responsibility

The rapid integration of AI into research workflows mandates clear, actionable ethical guidelines that researchers can readily apply (Tingelhoff et al., 2024). The IS community is uniquely positioned to lead this effort, leveraging its established history of scrutinizing the ethical implications of computing technologies and developing relevant frameworks (Mirbabaie et al., 2022). A critical research agenda must therefore focus on formulating a practical code of conduct for AI in IS research, particularly literature reviews, extending beyond general principles to address the specific scenarios and intricate challenges inherent in AI-Assisted literature reviews (Dwivedi et al., 2023).

Key areas for this entails defining clear frameworks for accountability for AI-introduced errors, such as hallucinations or biases, where ultimate responsibility for outcomes rests firmly with the human researcher (Dwivedi et al., 2023; Tingelhoff et al., 2024). Furthermore, articulating precise norms for intellectual contribution and authorship is crucial, clarifying acceptable uses of AI-Assisted drafting, summarizing, or idea generation without compromising originality or risking plagiarism (Kacena et al., 2024; Tingelhoff et al., 2024). This proactive development of robust ethical standards is crucial to prevent misuse and ensure that AI reliably serves as a responsible aid to human scholarship (Dwivedi et al., 2023; Onuoha, 2025).

Conclusion

The emergence of AI tools presents both vast opportunities and significant challenges for one of the fundamental academic facets, literature reviews (Schryen et al., 2025). The central challenge for the IS field is transitioning the conversation from whether to integrate AI tools into literature reviews to how to effectively manage a diverse ecosystem of these tools, each possessing distinct

risk profiles (Dwivedi et al., 2023; Tingelhoff et al., 2024). The paradoxes identified in this paper—the Verification Paradox, the Homogenization Paradox, and the Automation Paradox—underscore that each category of AI tool, from general-purpose LLMs to specialized discovery and analysis platforms, presents unique methodological and ethical considerations requiring specific mitigation strategies (Messeri & Crockett, 2024; Yawson, 2025).

This paper's contribution lies in identifying a critical gap in current approaches to AI integration (Pan et al., 2023; Susarla et al., 2023). Furthermore, we have laid out an associated research agenda for the development of a comprehensive critical AI literacy framework, built upon three essential pillars: epistemological agility, methodological transparency, and ethical responsibility. This proposed framework is designed to provide the precise guidance and structured oversight that is currently lacking in generic human-in-the-loop recommendations (Luomala et al., 2025; Wagner et al., 2022). In practice, it could be applied in contexts such as academic education, research ethics education, and guidelines for publication reviews. Other tasks that commence with knowledge synthesis, such as the development of white papers, or position papers by public or private organizations would also benefit.

Ultimately, we call upon the IS community to steer the discourse, moving beyond current anxieties about AI towards its well-governed and intellectually rigorous integration into modern scholarly activities (Mirbabaie et al., 2022; Yawson, 2025). Given its established history of scrutinizing socio-technical phenomena and its expertise in developing ethical guidelines for computing technologies, the IS field must cultivate responsible AI integration (Haefner et al., 2023; Mirbabaie et al., 2022; Walsham, 1996). This endeavour overcomes mere adaptation to novel technology; it represents a commitment to thoughtfully shaping the future of knowledge production. It aims to ensure that human expertise, critical judgment, and intellectual integrity remain paramount in the AI-Assisted academic landscape (Onuoha, 2025; Schryen et al., 2025; Susarla et al., 2023).

About the authors

Dipesh Jalui is a PhD student at the School of Information Management, Victoria University of Wellington. His research interests include AI governance, AI-driven innovation, and human-AI collaboration and autonomy. He can be contacted at dipesh.jalui@vuw.ac.nz

Mary Tate is an Associate Professor in the School of Information Management, Victoria University of Wellington. Her research interests include research methods, digital services, digital innovation, governance and technology in government. She can be contacted at mary.tate@vuw.ac.nz

Jocelyn Cranefield is an Associate Professor in the School of Information Management, Te Herenga Waka—Victoria University of Wellington, New Zealand. Her research interests include human-AI collaboration and its impacts on work and wellbeing, managing digital transformation, and data cultures, and governance. She can be contacted at jocelyn.cranefield@vuw.ac.nz

References

- Attard-Frost, B., De los Ríos, A., & Walters, D. R. (2023). The ethics of AI business practices: A review of 47 AI ethics guidelines. *AI and Ethics*, 3(2), 389–406. <https://doi.org/10.1007/s43681-022-00156-6>
- Bawack, R. E., Fosso Wamba, S., & Carillo, K. D. A. (2021). A framework for understanding artificial intelligence research: Insights from practice. *Journal of Enterprise Information Management*, 34(2), 645–678. <https://doi.org/10.1108/JEIM-07-2020-0284>

- Boell, S. K., & Cecez-Kecmanovic, D. (2015). On being 'Systematic' in Literature Reviews in IS. *Journal of Information Technology*, 30(2), 161–173. <https://doi.org/10.1057/jit.2014.26>
- Bolaños, F., Salatino, A., Osborne, F., & Motta, E. (2024). Artificial intelligence for literature reviews: Opportunities and challenges. *Artificial Intelligence Review*, 57(10), 259. <https://doi.org/10.1007/s10462-024-10902-3>
- Casal-Otero, L., Catala, A., Fernández-Morante, C., Taboada, M., Cebreiro, B., & Barro, S. (2023). AI literacy in K-12: A systematic literature review. *International Journal of STEM Education*, 10(1), 29. <https://doi.org/10.1186/s40594-023-00418-7>
- Chiu, T. K. F., Ahmad, Z., Ismailov, M., & Sanusi, I. T. (2024). What are artificial intelligence literacy and competency? A comprehensive framework to support them. *Computers and Education Open*, 6, 100171. <https://doi.org/10.1016/j.caeo.2024.100171>
- Cram, W., Templier, M., & Paré, G. (2020). (Re)considering the Concept of Literature Review Reproducibility. *Journal of the AIS*. <https://doi.org/10.17705/IJAIS.00630>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023). Opinion Paper: 'So what if ChatGPT wrote it?' Multidisciplinary perspectives on opportunities, challenges, and implications of generative conversational AI for research, practice, and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Else, H. (2023). Abstracts written by ChatGPT fool scientists. *Nature*, 613(7944), 423–423. <https://doi.org/10.1038/d41586-023-00056-7>
- 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. <https://doi.org/10.1016/j.techfore.2023.122878>
- Huettemann, S., Mueller, R., Larsen, K., & Dinter, B. (2024). A Framework for Ontology-Based Knowledge Synthesis from Research Articles.
- Kacena, M. A., Plotkin, L. I., & Fehrenbacher, J. C. (2024). The Use of Artificial Intelligence in Writing Scientific Review Articles. *Current Osteoporosis Reports*, 22(1), 115–121. <https://doi.org/10.1007/s11914-023-00852-0>
- Kittler, H. (2024). Will AI revolutionize literature reviews? *Journal of the European Academy of Dermatology and Venereology*, 38(12), 2213–2214. <https://doi.org/10.1111/jdv.20354>
- Lu, J., & Mayhook, Z. A. (2024). AI Tools for Literature Reviews.
- Luomala, M., Naarmala, J., & Tuomi, V. (2025). Technology-Assisted Literature Reviews with Technology of Artificial Intelligence: Ethical and Credibility Challenges. *Procedia Computer Science*, 256, 378–387. <https://doi.org/10.1016/j.procs.2025.02.133>
- Messeri, L., & Crockett, M. J. (2024). Artificial intelligence and illusions of understanding in scientific research. *Nature*, 627(8002), 49–58. <https://doi.org/10.1038/s41586-024-07146-0>
- Miao, J., Thongprayoon, C., Suppadungsuk, S., Garcia Valencia, O. A., Qureshi, F., & Cheungpasitporn, W. (2024). Ethical Dilemmas in Using AI for Academic Writing and an Example Framework for Peer Review in Nephrology Academia: A Narrative Review. *Clinics and Practice*, 14(1), 89–105. <https://doi.org/10.3390/clinpract14010008>

- Michalak, R., & Ellixson, D. (2024). AI-Driven Discovery: How Litmaps Shapes Research and Teaching & Learning. *The Serials Librarian*, 85(5–6), 117–129. <https://doi.org/10.1080/0361526X.2024.2433255>
- Mirbabaie, M., Brendel, A., & Hofeditz, L. (2022). Ethics and AI in Information Systems Research. *Communications of the Association for Information Systems*, 50(1). <https://doi.org/10.17705/1CAIS.05034>
- Morandini, S., Fraboni, F., Balatti, E., Hackmann, A., Brendel, H., Puzzo, G., Volpi, L., Giusino, D., Angelis, M. D., & Pietrantonio, L. (2023). Assessing the Transparency and Explainability of AI Algorithms in Planning and Scheduling tools: A Review of the Literature. *Human Interaction & Emerging Technologies (IHET 2023): Artificial Intelligence & Future Applications*, 111(111). <https://doi.org/10.54941/ahfe1004068>
- Mozelius, P., & Humble, N. (2024). On the Use of Generative AI for Literature Reviews: An Exploration of Tools and Techniques. *European Conference on Cyber Warfare and Security*, 161–168. <https://www.proquest.com/docview/3092061589/abstract/536C39CD10744142PQ/1>
- Ngwenyama, O., Toronto Metropolitan University / University of Cape Town, Rowe, F., & University of Nantes. (2024). Should We Collaborate with AI to Conduct Literature Reviews? Changing Epistemic Values in a Flattening World. *Journal of the Association for Information Systems*, 25(1), 122–136. <https://doi.org/10.17705/1jais.00869>
- Ofori-Boateng, R., Aceves-Martins, M., Wiratunga, N., & Moreno-Garcia, C. F. (2024). Towards the automation of systematic reviews using natural language processing, machine learning, and deep learning: A comprehensive review. *Artificial Intelligence Review*, 57(8), 200. <https://doi.org/10.1007/s10462-024-10844-w>
- Okoli, C. (2015). A Guide to Conducting a Standalone Systematic Literature Review. *Communications of the Association for Information Systems*, 37. <https://doi.org/10.17705/1CAIS.03743>
- Onuoha, C. E. (2025). Impact of Artificial Intelligence (AI) on the Quality, Efficiency, and Transparency of the Scholarly Publishing Process.
- Oyelude, A. A. (2024). Artificial intelligence (AI) tools for academic research. *Library Hi Tech News*, 41(8), 18–20. <https://doi.org/10.1108/LHTN-08-2024-0131>
- Pan, S. L., Nishant, R., Tuunanen, T., & Nah, F. F.-H. (2023). Literature review in the generative AI era—How to make a compelling contribution. *The Journal of Strategic Information Systems*, 32(3), 101788. <https://doi.org/10.1016/j.jsis.2023.101788>
- Paré, G., Tate, M., Johnstone, D., & Kitsiou, S. (2016). Contextualizing the twin concepts of systematicity and transparency in information systems literature reviews. *European Journal of Information Systems*, 25(6), 493–508. <https://doi.org/10.1057/s41303-016-0020-3>
- Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, 52(2), 183–199. <https://doi.org/10.1016/j.im.2014.08.008>
- Pritting, S., Doran, P., Covington, H., Kennedy, J., & Sun, H. (2025). AI-integrated literature discovery tools: A guide to reviewing for institution wide use. *Journal of Electronic Resources Librarianship*, 0(0), 1–15. <https://doi.org/10.1080/1941126X.2025.2497736>

- Saunders, M. N. K., & Rojon, C. (2011). On the attributes of a critical literature review. *Coaching: An International Journal of Theory, Research and Practice*, 4(2), 156–162. <https://doi.org/10.1080/17521882.2011.596485>
- Schryen, G., Marrone, M., & Yang, J. (2024). Adopting Generative AI for Literature Reviews: An Epistemological Perspective.
- Schryen, G., Marrone, M., & Yang, J. (2025). Exploring the scope of generative AI in literature review development. *Electronic Markets*, 35(1), 13. <https://doi.org/10.1007/s12525-025-00754-2>
- Shor, R., Greene, E. A., Sumberg, L., & Weingrad, A. B. (2025). AI Tools in Academia: Evaluating NotebookLM as a Tool for Conducting Literature Reviews. *Psychiatry*, 0(0), 1–10. <https://doi.org/10.1080/00332747.2025.2541531>
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Sparkman, M., & Witt, A. (2025). Claude AI and Literature Reviews: An Experiment in Utility and Ethical Use. *Library Trends*, 73(3), 355–380.
- Stahl, B. C. (2008). Researching Ethics and Morality in Information Systems: Some Guiding Questions.
- Stone, J. A. M. (2023). Artificial Intelligence–Generated Research in the Literature: Is It Real or Is It Fraud? *Medical Acupuncture*, 35(3), 103. <https://doi.org/10.1089/acu.2023.29231.editorial>
- Susarla, A., Gopal, R., Thatcher, J. B., & Sarker, S. (2023). The Janus Effect of Generative AI: Charting the Path for Responsible Conduct of Scholarly Activities in Information Systems. *Information Systems Research*, 34(2), 399–408. <https://doi.org/10.1287/isre.2023.ed.v34.n2>
- Templier, M., & Paré, G. (2015). A Framework for Guiding and Evaluating Literature Reviews. *Communications of the Association for Information Systems*, 37. <https://doi.org/10.17705/1CAIS.03706>
- Tingelhoff, F., Brugger, M., & Leimeister, J. (2024). A Guide for Structured Literature Reviews in Business Research: The State-of-the-Art and How to Integrate Generative Artificial Intelligence. *Journal of Information and Technology*. <https://doi.org/10.1177/02683962241304105>
- Tomczyk, P., Brüggemann, P., & Vrontis, D. (2024). AI meets academia: Transforming systematic literature reviews. *EuroMed Journal of Business*. <https://doi.org/10.1108/EMJB-03-2024-0055>
- Van Dinter, R., Tekinerdogan, B., & Catal, C. (2021). Automation of systematic literature reviews: A systematic literature review. *Information and Software Technology*, 136, 106589. <https://doi.org/10.1016/j.infsof.2021.106589>
- Wagner, G., Lukyanenko, R., & Paré, G. (2022). Artificial intelligence and the conduct of literature reviews. *Journal of Information Technology*, 37(2), 209–226. <https://doi.org/10.1177/02683962211048201>
- Walsham, G. (1996). Ethical theory, codes of ethics and IS practice. *Information Systems Journal*, 6(1), 69–81. <https://doi.org/10.1046/j.1365-2575.1996.00105.x>

Walter, Y. (2024). Managing the race to the moon: Global policy and governance in Artificial Intelligence regulation—A contemporary overview and an analysis of socioeconomic consequences. *Discover Artificial Intelligence*, 4(1), 14. <https://doi.org/10.1007/s44163-024-00109-4>

Webster, J., & Watson, R. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Q*.

Yawson, R. M. (2025). Perspectives on the promise and perils of generative AI in academia. *Human Resource Development International*, 28(3), 476–487. <https://doi.org/10.1080/13678868.2024.2334983>

Yue Yim, I. H. (2024). A critical review of teaching and learning artificial intelligence (AI) literacy: Developing an intelligence-based AI literacy framework for primary school education. *Computers and Education: Artificial Intelligence*, 7, 100319. <https://doi.org/10.1016/j.caeai.2024.100319>

© [CC-BY-NC 4.0](#) The Author(s). For more information, see our [Open Access Policy](#).