Journal of the Association for Information Systems

Volume 25
Issue 1 *Special Issue: The Future Impact of AI on Academic Journals and the Editorial Process (pp. 37-181)*

Article 2

2024

Democratizing Knowledge Creation Through Human-AI Collaboration in Academic Peer Review

Suprateek Sarker
University of Virginia, sarkers@virginia.edu

Anjana Susarla
Michigan State University, asusarla@broad.msu.edu

Ram Gopal
University of Warwick, ram.gopal@wbs.ac.uk

Jason Bennett Thatcher
University of Colorado / University of Manchester, jason.thatcher@colorado.edu

Recommended Citation

Sarker, Suprateek; Susarla, Anjana; Gopal, Ram; and Thatcher, Jason Bennett (2024) "Democratizing Knowledge Creation Through Human-AI Collaboration in Academic Peer Review," *Journal of the Association for Information Systems*, 25(1), 158-171.
DOI: 10.17705/1jais.00872
Available at: https://aisel.aisnet.org/jais/vol25/iss1/2

Journal of the Association for Information Systems (2024) **25**(1), 158-171
doi: 10.17705/1jais.00872

EDITORIAL

ISSN 1536-9323

Democratizing Knowledge Creation Through Human-AI Collaboration in Academic Peer Review

Suprateek Sarker,[1] Anjana Susarla,[2] Ram Gopal,[3] Jason Bennett Thatcher[4]

[1]University of Virginia, USA, sarkers@virginia.edu
[2]Michigan State University, USA, asusarla@broad.msu.edu
[3]University of Warwick, UK, ram.gopal@wbs.ac.uk
[4]University of Colorado, USA / Alliance Manchester Business School, UK, Jason.Thatcher@colorado.edu

Abstract

In the rapidly evolving landscape of academic research, artificial intelligence (AI) is poised to revolutionize traditional academic peer review processes and knowledge evaluation systems. We believe that the growing collaboration between humans and AI will disrupt how academics assess scholarly manuscripts and disseminate published works in a way that facilitates the closing of gaps among diverse scholars as well as competing scholarly traditions. Such human-AI collaboration is not a distant reality but is unfolding before us, in part, through the development, application, and actual use of AI, including language learning models (LLMs). This opinion piece focuses on the academic peer review process. It offers preliminary ideas on how human-AI collaboration will likely change the peer review process, highlights the benefits, identifies possible bottlenecks, and underscores the potential for democratizing academic culture worldwide.

Keywords: Peer Review, Language Learning Models, Artificial Intelligence

David G. Schwartz was the accepting senior editor. This paper was submitted on June 19, 2023 and underwent two revisions. It is part of the Special Issue on The Future Impact of AI on Academic Journals and the Editorial Process. All authors contributed equally.

1 Introduction

Academic journals play a crucial role in developing, validating, and distributing knowledge. They help create and sustain knowledge communities by offering a platform for individuals with similar interests to share their research (Lamont, 2009). For scholars, authoring papers that appear in highly regarded journals enhances their wider exposure and prestige, creates opportunities for recognition in their discipline, and boosts possibilities for career advancement.

To maintain the integrity of journals that serve as platforms for knowledge dissemination, journal editors implement rigorous and fair peer review processes that afford equitable access to all interested authors. While peer review dates to the 18th century ("Scholarly Peer Review," 2023), modern peer review processes date to the end of World War II, when a technological change—the advent of the photocopy machine—made it easy to create copies of submissions to send to peer reviewers. Even as the growth of the internet has made electronic processing of manuscript submissions and distribution of published papers possible, the basic mechanics of peer review (Csiszar, 2016), notably the recording of reviewer evaluations/comments, the aggregation of numeric scores, and the maintenance of manuscripts' history, have largely remained unchanged.

The *Journal of the Association for Information Systems* (JAIS), under the visionary leadership of the late Professor Philip Ein-Dor, was founded to fulfill the mission of equal opportunity to publish the highest possible quality research by the global information systems community. JAIS was designed to be an "electronic journal" (or e-journal), with no physical copy of the issues/articles, and by lowering costs, it sought to democratize access to contribute and share knowledge across the IS community. This was *pathbreaking* for its time, and despite the scope and efficiencies that electronic publishing made possible, JAIS initially struggled to gain recognition at many business schools. Indeed, Palmer et al. (2000, p. 2) noted that "acceptance of electronic journals among business school faculty has two hurdles to overcome: technological and, more challenging, garnering legitimacy within the academic community. A survey targeted at business school faculty in the United States … suggest[s] that at the time of publication, electronic publications were seen as less desirable than paper counterparts for tenure and review.[1]"

While the prejudice against e-journals appears to have come down substantially, recent advances in AI present new challenges—and opportunities—for JAIS. Tools offering AI-powered cognition have prompted many scholars to reflect on the role of journals and the nature of the review process (Checco et al. 2021, Grimaldi et al. 2023). Are journals and papers published therein, as we know them, going to remain relevant in a world of generative AI? If so, do we retain the classical "human expert-driven" peer review model in selecting manuscripts for publication? Or, is it possible to infuse AI to transform peer review processes, thereby enhancing support for the core values of rigor, fairness, and access? Also, will such an AI-infused review process create challenges for democratization or enable it?

In this opinion piece, we assume that (1) for researchers worldwide, securing journal publications will remain relevant to succeed in academic life, and (2) AI, particularly generative AI, will spark transformative changes to scholarship in journals, especially in authorial and peer review processes. The following pages offer our thoughts regarding likely changes, possible challenges, and potential impacts.

We first provide conceptual foundations, including capabilities that AI technologies bring to collaboration with humans with respect to peer review and core values that journals should continue to represent. Next, we outline, for the near term and for the longer term, possibilities for changes in peer review processes and their impacts. We conclude by reflecting on these issues and calling for scholars to be forward-looking but vigilant in this journey toward human-AI-infused review processes.

2 Key Capabilities of AI and Core Values of The Academic Peer Review Process

The emerging classes of predictive and generative AI have much potential to transform how journal submissions are evaluated. Such transformation can occur as a result of automating some elements of the review process (e.g., plagiarism checking), thereby reducing uncertainty (e.g., by verifying the accuracy of references) and cycle times (e.g., through text-analysis-informed reviewer recommendations) associated with evaluating journal submissions. They can also facilitate fact-checking of the literature (e.g., verifying the accuracy of quotations), offer insight into alternative interpretations of results, and help draft reviews and decision letters that support the editorial processes. Indeed, from the perspective of techno-optimists, AI will soon have the ability to mimic essential scholarly activities.

Herbert Simon (2019), in "The Sciences of the Artificial," distinguished between natural phenomena and artificial phenomena, such as technological artifacts, a category that encompasses AI systems. In particular, we believe that predictive and generative AI systems are closely related to Herbert Simon and Allen Newell's (1956) classification of thinking machines that emulate human decisions such that the more sophisticated large language models may be able to automate perceptual and reasoning tasks traditionally assigned to editors and reviewers. That is, the current generation of LLMs is showing the potential to synthesize reasoning and facilitate the recombination of human thoughts in ways that emulate essential actions in the peer review process, such as assessing the quality of prose or the internal consistency of arguments (Drori et al., 2022).

With the growth of AI's capabilities, generative AI can already undertake some tasks in evaluating scholarship that earlier required human judgment. For example, AI is already being widely employed to screen papers for plagiarism. AI can also inform editorial tasks, such as detecting novelty by identifying similarities between submitted papers and published work in journals or other outlets (e.g., Bauersfeld et al., 2023). These are examples of how AI has helped speed up the peer review processes.

What is uncertain pertains to the specific roles that AI-infused review processes will play—whether they will complement the capabilities of human peer reviewers

[1] One of the authors of this paper, a former EIC of JAIS, recalls the advice of some Advisory Board members of JAIS to format papers in such a way that, when printed, the hard copy would look as if it were from a traditional printed journal. JAIS actually undertook a major initiative to reformat the papers it publishes to address this sentiment.

and editors or eliminate the need for humans altogether by automating writing and editorial work. Such possibilities for augmentation and automation by AI have yielded contradictory positions and narratives in scholarly communities. While some might welcome how it eases the work of human decision-making, others might fear that the need for human decision makers will be eliminated altogether (Wang et al., 2019). As the broader discourse on AI has become more sophisticated, it has moved from expressing fear about AI replacing or competing with humans to expressing hope for human-AI collaboration[2] (Fugener et al., 2021, Jain et al., 2021, Jussupow et al., 2021, Raisch & Krakowski, 2021) because participants are increasingly realizing that human-AI augmentation and automation roles are intertwined and will change together over time.

While there are numerous ways to describe the potential collaboration between humans and AI, this brief opinion piece focuses on *the relative dominance in contributions of humans and AI with respect to each other* as part of the peer review process. The notion of dominance here not only includes the proportion of the tasks undertaken but also the level of control over the process and its outcomes. For example, in the case of human dominance, one can assume that the human is generally aware of and in control of the entire review process; the human *utilizes* AI for specific tasks within the process. However, if AI becomes dominant, humans may participate without an overall understanding or control over the process. Further, as suggested in the cocreation literature, collaboration can be *additive* (i.e., involving AI doing part of the process and humans doing a part of the process independently, with the work being put together unproblematically) or it can involve the *amalgamation or synergistic integration* of the capabilities and outputs of the human and AI components (Sarker et al., 2012; Sahaym et al., 2023). The latter is obviously emergent, far more complex, and hence harder to explain or control.

We are hopeful that a more cogent, broader vision for human-AI collaboration as part of scholarly knowledge production and peer review will emerge from the ongoing broader discourse on generative AI. We believe this is essential because, in contrast to predictive AI, which readily augments and automates some human decisions, generative AI has the potential to fundamentally change a broad range of activities long viewed as the exclusive domain of humans, such as assessing idea generation and supporting creative tasks in many disciplines (e.g., Susarla et al., 2023). Because generative AI can assist in evaluating the novelty of content, differentiating between good and bad ideas, and assessing the outputs of scholarly work, it holds the potential to transform AI-human collaboration from prediction—where, for example, we label a paper as simply plagiarized or piecework by detecting patterns of similarity to existing work as an outcome—to generation, where we partner with AI to create synthetic reports that evaluate the quality of basic or applied science or, in some cases, ask the AI to help create suggestions for how to improve the quality and significance of intellectual artifacts (e.g., journal articles).

If humans can effectively collaborate with AI as part of the peer review process, we envision a scholarly community where knowledge is created and shared far more effectively and efficiently, consistent with the democratization vision. For example, if enabled by AI, stakeholders with constraints from different parts of the world (e.g., language competency, limited access to published research, and social networks) will not necessarily be held back from contributing to the highest-level journals if they have valuable and original ideas. Effectively taking advantage of human-AI collaboration will require scholarly communities to be purposeful as they update peer review processes.

Table 1 outlines potential scenarios for human-AI collaboration in the peer review process. As Table 1 shows, we see Type I as the state of synergy between humans and AI, where we would like to end up, given our understanding of the review process and AI capabilities today.

Type III is not particularly valuable, beyond basic experimentation that is not well supervised, since it would involve the suboptimal use of AI without sufficient human agency or control. As a result, we believe that the journey towards AI-human collaboration should ideally commence with Type II, where AI assists human effort. Type IV involves substituting humans with AI in significant aspects of the review process. While attractive to AI optimists, the Type IV scenario, characterized by an overreliance on AI for decision-making, can lead to dysfunctional outcomes due to automation bias. To avoid automation bias, we maintain that human judgment, at least in the near future, is indispensable for evaluating the novelty and contributions in a fair and transparent manner.

[2] In this paper, the term human-AI collaboration is used broadly, denoting that AI and humans participate in a process in order to contribute to a common goal.

Table 1. Modes of Human-AI Collaboration in the Peer Review Process

	AI dominant	AI subordinate
Human dominant	**Type I (human-AI synergistic):** Complementarity of human and AI capabilities being synergistically harnessed to implement a reimagined review process **Path B**	**Type II (AI-assisted):** AI augmenting the human capabilities in some aspects of peer review; humans in control **Path A (recommended)**
Human subordinate	**Type IV (AI-led):** AI substituting for the human in many aspects of peer review; some reimagination of roles but, notably, human control is limited. At least in the foreseeable future, we risk automation bias with a large proportion of review processes handed over to AI	**Type III (ambiguity in humans and AI roles):** This state may involve suboptimal use of AI without sufficient human agency or control. *Not a particularly desirable or valuable state*, beyond some basic experimentation.

Our vision is that a transition towards Type I will benefit scholarly communities the most. This approach will require reimagining review/editorial tasks and harnessing a harmonious and creative amalgamation of human and AI elements to accomplish tasks. It calls for a collective vision from the community regarding the desired outcomes of this synergistic review process. Furthermore, it leverages the domain expertise of reviewers to infuse the emergent capabilities into the AI system. For us, the ideal path would be Path A (from Type II to Type I). While Path B (from Type IV to Type I) may be attractive to some, it is inherently riskier, given the possible automation bias of Type IV and the resulting undesirable path dependencies created. As mentioned earlier, we do not feel that the review process in line with Type IV, with loosely coordinated, unsupervised actors, will be viable in the near future. We suspect that journals experimenting with a Type III initiative will quickly move to Type IV or, ideally, to Type II, before transitioning to Type I.

Absent the purposeful construction of human-AI collaboration in peer review processes, we believe there is a pressing risk that widespread and uncritical integration of AI into peer review processes could fuel even greater inequality (Acemoglu, 2021) in academic communities, which contradicts the founding vision of JAIS, including our yearning for the democratization of knowledge. We also worry that the rapid infusion of AI in reviewing and editorial tasks might widen gaps in inclusion and access between the Global North and South and reinforce preexisting divides in representation and diversity across scholarly communities (Reidtpath & Allotey, 2019). For instance, training datasets for AI systems often draw heavily from journals in the Global North, marginalizing innovative methodologies/theories more common in the Global South. An AI trained solely on such datasets may discount high-quality manuscripts employing alternative approaches. Similarly, AI trained on accepted papers in top journals may exhibit gender bias if their peer review processes have historically favored male authors. For instance, the AI might downgrade papers using qualitative methods more common among female scholars, thus perpetuating disparities in representation. It could do so by inadvertently reinforcing biased norms and standards within the existing corpus of knowledge found in the "top journals" used to train AI, which too often excludes ideas or issues of interest to scholars outside of elite institutions (Wanderley et al., 2021).

3 Conceptualization of a Value System for Human-AI Collaboration in Reviewing

To ensure that the design of human-AI collaboration is purposeful, academic communities must carefully consider how they integrate AI into the review and editorial processes. This necessitates reflection on the potential roles AI can play, the values inherent in the routines assigned to AI, and the methods by which AI is used to create value for authors, reviewers, and editors.

First, it is important to recognize that the review process contains paradoxical elements. On the one hand, as "gatekeepers," editors and their teams must critically evaluate the publishability of potential manuscripts, which requires surfacing weaknesses and challenges in a paper submission's design and arguments. On the other hand, as "diamond cutters" and "champions," editors and their teams must help develop and sharpen ideas (e.g., Saunders, 2005; Sarker et al., 2015), which requires creatively identifying solutions and opportunities for strengthening a submission's analysis and narrative. Understanding that this paradox exists is important because it frames the choices that we, in scholarly communities, collectively make about how to design and integrate AI as partners in the peer review process.

Second, in light of the paradoxes posed by gatekeeping and developmental objectives, it is imperative that journals transparently instantiate the traditional academic values of rigor, fairness, and access in any model of human-AI peer review. Absent the instantiation of these values, we fear that academic communities will reject such review or feel suspicious about peer review resulting from human-AI collaboration. However, doing so will require labeling these values in terms understood by both academics and system developers: notably, values such as explainability, trustworthiness, and transparency that are highlighted in the literature on responsible AI (e.g., Dignum, 2019) and aligning them with community norms and feedback enshrined in the academic peer review process in the discipline.

Third, given current generative AI models, which are trained using sometimes opaque techniques and extremely large data sets that integrate billions of parameters, it is imperative that journals partner with scholarly communities to ensure explainability, transparency, and responsibility to an adequate degree in collaborative human-AI peer review processes. To do so, it will be helpful for AI designers and academics to draw on the European Commission's ethical guidelines for trustworthy AI (EU, 2020; EU, 2023) with direction to (1) human agency and oversight; (2) technical robustness and safety; (3) privacy and data governance; (4) transparency, diversity, non-discrimination and fairness; (5) societal and environmental well-being; and (6) accountability.

We contend that effectively instantiating such guidelines in human-AI peer review processes will require scholarly communities to consider three interlinked sets of issues.

3.1 Algorithmic Opacity, Transparency, and Need for Explanation

Given the black-box nature of AI models and peer review processes, introducing AI to peer review may make already opaque editorial models even more confusing for authors and reviewers. Because authors have certain expectations with respect to the outcomes of these emergent human-AI processes, we need to be especially vigilant about such interpretability and explainability concerns. This is especially the case because evidence suggests that traditional review processes have traditionally harmed under-represented and historically disadvantaged subcommunities of authors as well as subjects (Silbiger & Stubler), and the involvement of AI increases the possibility of additional harm.

To make human-AI peer review processes explainable, scholarly communities must engage in a sensemaking process that encourages "public reason." Rawls argues that public reason "is characteristic of a democratic people: it is the reason of its citizens, of those sharing the status of equal citizenship" (Rawls, 1993, p. 213). Public reason incorporates "standards of correctness" and "criteria of justification" (Rawls, 1993, p. 220). AI ethicists have suggested that an important first step toward establishing human-AI collaborations is addressing a deficit of public reason in the Rawlsian sense (MacClure, 2021), that is, developing shared explanations and standards of the role of AI in peer review.

Developing transparent and explainable AI-human peer review processes will be a complex task. It will require us to address existing issues of equity and transparency in traditional peer review, which is inherently challenging. Moreover, we will need to incorporate these considerations into the design principles for AI-led, AI-assisted, or AI-synergistic peer review systems, as well as in how we evaluate outcomes from human-AI systems (e.g., Guidotti et al., 2018). This endeavor calls for inclusive dialogue within the community, thoughtful design, and constant testing. We must acknowledge that the complexity is compounded by the fact that we don't yet have a universally accepted standard for what constitutes a "good" explanation (Lipton, 2018), especially in the context of decisions made through human-AI collaborations.

3.2 Trustworthiness

The success of AI-human peer review processes hinges significantly on authors' trust in the fairness of these systems. Peer review, in essence, operates as a social contract. Editors (and reviewers), who are trusted and well-versed in community norms, apply (or at least, try to apply) rules consistently to all authors. Given the emergent nature of peer review, trust is a vital component of the sociotechnical system that supports the review process.

Research in AI ethics suggests that for a human-AI peer review system to be deemed trustworthy, users must have confidence that it will meet its stated requirements regarding expertise and predictability in evaluation. Moreover, it should provide verifiable evidence that it operates as intended (ISO, 2020). A human-AI peer review model must not only incorporate the scholarly norms of rigor, fairness, and equity but must also demonstrate that these norms are upheld through procedural safeguards in the peer review processes (Kaur et al., 2022).

Setting aside the issue of "hallucination" in the review process for now (see Susarla et al., 2023, for example), it may be feasible to design systems that encode norms of trust and safeguard the social capital derived from credible peer review processes. However, it remains less clear how to design AI-human routines and roles that effectively merge human elements such as empathy, cognition, and intuition. The goal is to achieve manuscript assessments that are critical and incisive while also embodying charity and constructiveness—characteristics expected in high-quality, developmental peer review processes.

3.3 Norms and Community Building

Each journal's peer review process undeniably carries elements of the sociocultural aspect of its discipline. Journals such as JAIS have made significant strides in fostering open, inclusive cultures that adopt a broad perspective on information systems consistent with the value system of the Association for Information Systems. Journals propagate these values by cultivating a sense of community among reviewers, editors, and authors that support norms such as ensuring charitable and constructive comments, even when the assessment is critical. Reviewers and editors perform their roles with a service-oriented mindset, fostering increased responsibility, accountability, and trustworthiness. Community is also important because, within the connectionist school of knowledge management in particular, knowledge is not universal: rather than being understood and valued identically across communities, it is deeply contextual (Joshi et al., 2007; Kogut & Zander, 1992). As we move toward AI-augmented peer review, the challenge lies in designing processes that preserve these very human considerations, norms, and sentiments while also taking advantage of the speed and precision offered by AI.

Human-computer interaction theories provide some guidance on constructing human-AI peer review systems that accommodate diverse peer review paradigms and reasoning systems (e.g., Amershi et al., 2019). However, the field of human-AI collaboration is still in its infancy, and there is much to discover about such collaboration. We discuss some of the relevant issues in the subsequent sections.

4 The Path Forward with an AI-Human Peer Review Process: Envisioning Editorial Roles and Transformational Capabilities That Bridge Human Cognition With AI

Transitioning from pre-AI norms and practices to a new era of AI-infused peer review will require scholarly communities to adapt their routines, taking into consideration the new affordances and constraints (e.g., Leonardi, 2011). This shift will necessitate that community members, including authors, reviewers, editors, and readers, gain a nuanced understanding of the institutional learning modes that underpin peer review (Beane, 2019). Without this understanding, there is a risk that AI could compromise the professional judgment of community members when assessing paper quality (e.g., Lebovitz et al., 2021).

We envision the future of peer review as a symbiotic, synergistic relationship between human intellect and contextualizing capabilities and AI's computational capabilities. To outline a path toward this goal, we explore the evolution of the peer review process in two stages (refer to Tables 2, 3, and 4).

4.1 Stage 1: Collaboration within Existing/Near-Term AI Capabilities

The first stage could use AI technologies to enhance the review process. Human judgment could be augmented and, when appropriate, reimagined and automated by AI in the following ways:

As shown in Table 2, AI could assist in the peer review process by identifying issues with methods, statistical analyses, or logical inconsistencies. This could help reduce the burden on human reviewers to identify fundamental issues in design and logic. AI could also *detect plagiarism,* both in simple forms, such as copying text without attribution, as well as more complex forms, such as mosaic plagiarism or patchwriting, where one borrows ideas without attribution. It could also identify instances of *self-plagiarism* or *duplicate publication*.

Enabling greater transformation, AI might *enable the formulation of new and more relevant criteria* for acceptance (Table 3). For example, transparency and reproducibility might become important factors in the acceptance of research. As a result, reporting how AI was used as part of the research process might also become a requirement for consideration of a manuscript for publication. Furthermore, the transparency and reproducibility of the emergent criteria used in the review process require careful scrutiny. There is a real danger that the automated screening of manuscripts might result in conformity to accepted research criteria and reduce the role of serendipity and the potentially enriching inspiration derived from reviewer comments.

As AI becomes more integrated into the review process, *new ethical considerations* might arise. For example, how should AI-generated research be credited, and should such crediting be validated in the review process *to ensure that intellectual property rights are not violated*? How can the review and editorial process guard against *the bias in AI training* and *hallucinations* in the reviews constructed with the aid of AI? Broadly speaking, what technological and human safeguards should be in place to prevent the misuse of AI in research, reviewing, and editing?

Table 2. Editorial Roles and Transformational Capabilities with AI in the Short Term

Review activities	As currently done	In collaboration with AI	Impacts	Possible pitfalls (including violation of core values)	Implementation guidelines to ensure explainability, trustworthiness, and community building
Peer review process	Basic IT systems for transactional aspects and human cognition for the actual assessment	AI can prescreen papers and identify potential issues with the method, statistical analysis, or logical inconsistencies	May reduce the load on human reviewers; may improve quality of reviews and research. This may change the human review process from a gate-keeping to a diamond- cutting and championing role	(1) Depending on the training datasets used, AI may not be able to spot some types of inconsistencies (see, for example, Susarla et al., 2023) (2) the AI may be opaque and not be perceived as trustworthy, or AI may make or prompt judgments without a nuanced understanding of the community's perspective on an issue. (3) Overreliance on AI may result in the depletion of stocks of human expertise on specific topics and methods in scholarly communities.	Ensuring the primacy of human scholarly interpretation and the development of domain knowledge of the phenomenon, methods, and disciplinary area (in line with the connectionist epistemology in knowledge management), ensuring norms of trustworthiness, explainability, and community acceptance
Plagiarism Detection	Both AI and human cognition	AI can be used to detect different types of plagiarism, such as self-plagiarism or cleverly disguised plagiarism.	May reduce the burden on human reviewers and could lead to more original research; it may enhance the diamond-cutting and championing role of reviewers and editors	(1) With AI-generated content widely prevalent, plagiarism detectors may not be able to distinguish between human-generated and AI-generated content, which can undermine trust in the human-AI peer review system. (2) Overreliance on AI may result in superficially similar patterns of word choice to describe discrete concepts, resulting in false positives.	(1) Scholars must have a clear understanding of what is plagiarism, especially as our understanding evolves from "direct" or "cut and paste" plagiarism to include "mosaic" or "patchwriting" plagiarism, where one steals the unique structure of arguments and ideas. (2) Scholars must learn to write in partnership with AI while preserving their own voice. (3) There is also a danger of lack of human-AI alignment where AI-based screening in the review process results in published research that is primed for the algorithm rather than directed toward questions of deeper significance to the field.
Criteria for acceptance (please see Table 3)	Human cognition	Emergent criteria	This may reduce the load on reviewers and lead to a more fair evaluation process	Transparency and reproducibility in the formulation of and application of criteria in the reviews	We need greater clarity on how we assess transparency and replicability. Such assessments would have to be perceived as trustworthy and in accordance with community norms.

Table 3. New Criteria and Considerations for Manuscripts in the Short Term

Review activities	**As currently done**	**Expected positive impacts on norms and community values**	**Possible issues and pitfalls (including violation of core values)**	**Implementation guidelines to ensure explainability, trustworthiness, and community building**
Research accessibility	Human cognition	AI can be used in translation into multiple languages; can enhance the participation of researchers worldwide and engage a broader reviewer pool; can also help in the dissemination of knowledge to laypersons	Biases in AI may widen some of the existing disparities in research access, reducing transparency	It is necessary to have criteria for inclusion and representation in building AI-augmented review processes. Such norms also need to be legitimated in the community through attention to explainability and trustworthiness, and building/enacting community-centric norms.
Criteria for acceptance	Human cognition	Emergent criteria such as ensuring that human cognition and reasoning are respected during the review process. May reduce the load on reviewers and further accentuate the role of reviewers and editors as “diamond-cutters” and champions	Disentangling human and AI contributions may prove challenging. Additionally, should AI- and human-created content be evaluated separately with different criteria? This would require infusing the AI with peer review norms in specific research communities	We need continuous review and greater direction on specifying how AI has been used in the research and the review process. We also need maturity and AI alignment in augmenting AI with human reviewers to realize the promise of AI-synergistic review processes.

In the short term, AI may lack the capabilities to help guide an AI-infused review process that satisfies the norms of explainability, trustworthiness, and adherence to community norms. As a result, the AI-assisted review process (Type II in Table 1) is recommended initially, with careful attention to ensuring that norms of equity, empathy, and fairness are enacted and that attention is paid to developing the experience in the community needed to collaborate on the development of more advanced systems in the future.

4.2 Stage 2: Longer Term—The Path toward an AI-Human Synergistic Review Process

Moving from an AI-assisted (Type II) to human-AI synergistic (Type I) collaboration could involve reimagining the review process and building or fine-tuning a language model using the wealth of data from *relevant, credible, and diverse sources.* This would allow the knowledge to be more sensitive to the context of the disciplinary community. The ideal state would be a model of an AI-human synergistic review process where the human and AI bring complementary capabilities that are blended in ways that harness the strengths while avoiding some of the pitfalls discussed above.

A properly tuned AI could support peer review processes in several ways. For example, it could be used to prescreen papers by providing an initial set of similar papers, assessing similarity to published work, and suggesting possible peer reviewers with expertise on the content (e.g., topic or method), thereby reducing the workload for editors. It could also identify common issues, suggest improvements, and point to areas (e.g., novelty) that require more careful inspection by human reviewers. By doing so, it could speed up the prescreening process and reduce the workload for reviewers. It could also check for more easily detected problems such as plagiarism, inconsistencies in the methodology, or conclusions that do not map to the results.

A properly tuned AI could also help inform broad editorial policy, such as board composition, emerging topics, and opportunities for community building. Editors could employ AI to analyze trends in the field. By analyzing the topics and methodologies of submitted papers, the AI could identify emerging trends and gaps in the current body of knowledge and inform opportunities for selecting new editorial board members. It could also be used to inform the training of new reviewers. By analyzing the feedback provided by experienced reviewers on successful papers or rejected papers, AI could provide guidance and examples for new reviewers to learn from. AI could also identify opportunities for helping new reviewers learn how to strengthen their reviews. While human editors or reviewers would need to identify situations where AI is providing incorrect feedback, a properly tuned AI could allow the editor and editorial board members to focus on identifying opportunities, developing papers, and thinking about strategies for advancing their discipline. This, in turn, would free up resources for editors to reimagine the future trajectory of the discipline and could redirect their own efforts toward building discipline-specific norms and acculturating new reviewers into the values of the academic community.

This aspect of AI infusion into the peer review process could involve authors interacting with a bespoke journal-specific (or community-specific) AI to develop their work. The AI-infused review portal would leverage an LLM trained on papers from the journal (or a set of relevant sources for the community) that prioritizes its values to automate the initial review process, provide constructive feedback to authors, and assign a quality rating to each paper. This system could significantly improve the efficiency, transparency, and accessibility of the academic publishing process. However, such changes could potentially reify existing values, reinforce echo chambers, and encourage the growth of normal science. To avoid such outcomes, humans must be attuned to the possibilities, be mindful of falling into the trap of overreliance on AI, and cultivate an awareness of the conditions in which such outcomes are likely.

Although a futuristic scenario at present, it presents the broad outline of how we envisage the contours of an AI-infused review process unfolding. When an author submits a paper to the portal, the AI would first assess whether the paper meets the minimum standards for academic writing. This could involve checking the basic elements like an abstract, introduction, methodology, results, and discussion, as well as more sophisticated elements like the clarity of the writing and the soundness of the argument. If the paper meets minimum standards, the AI would then generate an automated review.[3] This could include an assessment of the paper's clarity, the methodology's validity, the interpretation of the results, and the overall coherence of the argument. Human reviewers would then assess and engage with the AI's review and add their own comments. The human and AI comments would provide a more robust foundation for editors to make decisions. The authors would receive this combined feedback and have the opportunity to revise their paper accordingly. If this process works well, not only would the cycle times be reduced, but the review process would be freed of many of the idiosyncratic elements that we experience in human review systems and are likely to see in AI-infused review systems that have not sufficiently matured, i.e., not moved to Type I. Table 4 provides a summary of issues related to this stage.

Table 4. Editorial Roles and Transformational Capabilities with AI-Human Synergistic Peer Review Processes

Scholarly activities	As currently done	In collaboration with AI	Nature of collaboration	Expected positive impacts	Possible pitfalls (including violation of core values)	Implementation guidelines
Human-AI synergistic research portals	These do not currently exist	Journals could deploy AI to rapidly check for fit and quality	Editorial staff can use AI to check basic elements like the presence of an abstract, introduction, methodology, results, and conclusion, as well as assess the clarity of the writing and the soundness of the argument	On reviewers: Editorial staff can use AI to generate automated reviews, which human reviewers can then check and add nuances to	The quality of the review would largely depend on the quality of the AI used	We need stringent criteria in building and evaluating outputs from bespoke LLMs and other AI models that can be used to automate the review process. This is likely to impose a nontrivial burden on journals. Human participation cannot be removed; it is necessary to ensure explainability, transparency, and accordance with community norms.
Recognizing reviewers' contributions	This is currently done by humans.	AI can help in recognizing quality of human reviews.	Reviewers could be rated based on the quality of their feedback.	Reviewers who consistently provide high-quality, constructive feedback could be recognized.	We need to ensure standards of review are consistent across different subdisciplines.	We need AI that is trained on datasets that are sufficiently representative of the diversity of research in the field.

[3] This is similar to how Drori et al. (2022) envisaged the process of LLMs involved in synthesizing and reasoning tasks that were once considered the purview of humans.

A bespoke AI could ensure that authors receive clear and consistent feedback to improve their papers (even before the formal review process) and increase their chances of publication. Such screening could be used without prejudice, allowing authors to revise and resubmit their work until they are satisfied that they have addressed the concerns identified in their work by the AI. This could be particularly beneficial for authors who do not have access to extensive academic networks or resources—for example, those in the Global South or outside of elite institutions.

Once submitted, the bespoke AI could quickly assess whether a paper meets the minimum standards and provide an immediate assessment for editors and reviewers to work with. This could significantly reduce the time it takes for papers to be reviewed and published, easing pressure for early career faculty and allowing new research to be disseminated more quickly. Given that surveys of researchers have surfaced gender disparities in submission behavior among top journals—i.e., women are less likely to submit to top journals, thereby lowering their likelihood of eventual acceptance (Basson et al., 2023)—such AI-assisted prescreening might lower such disparities in publication.

A bespoke AI could also aggregate and summarize information on the quality of reviews and reviewers. When combined with qualitative assessments of reviews, such a system could be used as a reputation mechanism for reviewers. This could provide recognition for reviewers who consistently provide high-quality, constructive feedback or demonstrate the big-picture thinking needed to serve on editorial boards.

By automating prescreening as a complement to the peer review process, while ensuring that core values are not violated and humans remain engaged with the process, a bespoke AI system could help editors quickly handle a larger volume of papers, thus removing human bottlenecks and making academic publishing more accessible.

Although this system has many potential benefits, we need to be vigilant about four potential bottlenecks and challenges:

First, the system's effectiveness would largely depend on the quality of the review that the AI system provides. If an AI cannot accurately assess the paper's quality or provide useful feedback, the system may not be effective. Thus, it would be necessary for a journal and its supporting scholarly community to provide the data and knowledge necessary to provide bespoke training for the AI. Careful testing and ongoing monitoring would be required.

Second, while AI could reduce the workload for human reviewers, they would still need to assess and engage with the AI's review and add their own comments. This back and forth with AI could still create a significant workload, especially if the system increases the number of papers submitted to the journal. It would also require providing additional training to avoid complacency regarding any bias in the system that may creep in.

Third, AI could introduce serious risks of bias and fairness into the review process, particularly if the AI system is trained on a dataset and instantiates values not representative of the diversity of research traditions. This would require going beyond the dominant scholarly community associated with a journal to include voices from emerging and adjacent communities as well as voices within the discipline that have been drowned out in the past. Such balancing between including diverse knowledge sources and being sensitive to the disciplinary context must be done in a deliberate manner. Furthermore, human review teams must continue to monitor potentially biased AI behaviors.

Fourth, while we know that AI-human synergistic collaboration needs to ensure that norms of equity, empathy, and fairness are preserved in the review process, we lack rubrics for how to assess whether and how they are preserved. This will require creativity and astute leadership from the involved editors.

If these concerns are not addressed effectively, authors will find it difficult to trust the output of a human-AI system (particularly as AI begins to take on more significant roles) to the same extent that they would trust a human reviewer.

5 Concluding Remarks

JAIS was founded with the vision of equal opportunity to publish the highest possible quality research by the global information systems community, and the contribution of Professor Ein-Dor in implementing this vision cannot be overstated. Subsequent editors have done their bit, and the journal has progressed to being widely regarded as one of the top four journals in the discipline. Now, we are faced with the opportunity to guide the journal to the next level in the era of AI—to help reimagine how the journal should operate to achieve a true democratization of knowledge—*this would be a fitting tribute to Professor Ein-Dor.*

We have outlined a phased pathway for journals to introduce AI into the peer review process that emphasizes transparency, rigor, and inclusion. In the near term, AI can augment human capabilities in discrete tasks like plagiarism detection and reviewer recommendations. However, humans must remain dominant in peer review processes and avoid overreliance on imperfect AI systems. In the longer term, bespoke AI trained on community knowledge and norms (reflecting the diverse IS community) could collaborate synergistically with human reviewers to increase efficiency and access. AI designers and journal editors will need to partner with scholarly communities to ensure that the review process reflects the diverse perspectives in our discipline and withstands ethical scrutiny. If carefully co-constructed, AI-infused review processes could significantly enhance efficiencies while also broadening opportunities for marginalized voices to contribute high-quality manuscripts.

Ultimately, realizing the democratizing potential of AI-human collaboration requires centering humanistic academic values. JAIS can lead this agenda by piloting initiatives that transparently assess AI impacts on equity and inclusion. If guided by a humanistic compass, the future portends an era in which AI elevates rather than undermines journals' founding visions to advance knowledge for the betterment of society. We urge JAIS to commit to this journey.

References

Acemoglu, D. (2021). *Harms of AI* (NBER Working Paper No. 29247). https://doi.org/10.3386/w29247

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for human-AI interaction. *Proceedings of the CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3290605.3300233

Basson, I., Ni, C., Badia, G., Tufenkji, N., Sugimoto, C. R., & Larivière, V. (2023). *Gender differences in submission behavior exacerbate publication disparities in elite journals.* bioRxiv. at https://www.biorxiv.org/content/10.1101/2023.08.21.554192v1

Bauersfeld, L, Romero, A., Muglikar, M., & Scaramuzza, D. (2023). Cracking double-blind review: Authorship attribution with deep learning. *PLOS ONE, 18*(6), Article e0287611.

Beane, M. (2019). Shadow learning: Building robotic surgical skill when approved means fail. Administrative Science Quarterly, *64*(1), 87-123.

Checco, A., Bracciale, L., Loreti, P., Pinfield, S., & Bianchini, C. (2021). AI-assisted peer review. *Humanities and Social Sciences Communications, 8*, Article 25.

Csiszar, A. (2016). Peer review: Troubled from the start. Nature, 532(7599), 306-308.

Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way*. Springer.

Drori, I., Zhang, S., Shuttleworth, R., Tang, L., Lu, A., Ke, E., Liu, K., Krishnan, R., Chang, Y., Vijayaraghavan, A., Goodman, E., & Roy, D. (2022). A neural network solves, explains, and generates university math problems by program synthesis and few-shot learning at human level. *PNAS 119*(32), Article e2123433119.

European Commission. (2023). Ethics guidelines for trustworthy AI. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

European Commission. (2020). *White paper on artificial intelligence: A European approach to excellence and trust.* https://ec.europa.eu/digital-single-market/en/news/white-paper-artificial-intelligence-european-approach-excellence-and-trust

Fugener, A., Grahl, J., Gupta, A., & Ketter, W. (2021). Cognitive challenges in human-artificial intelligence collaboration: Investigating the path toward productive delegation. *Information Systems Research, 33*(2), 678-696.

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys, 51*(5), Article 93.

Jain, H., Padmanabhan, B., Pavlou, P. A., & Raghu, T. S. (2021). Editorial for the special section on humans, algorithms, and augmented intelligence: The future of work, organizations, and society. *Information Systems Research, 32*(3), 675-687. https://doi.org/10.1287/isre.2021.1046

Joshi, K. D., Sarker, S., & Sarker, S. (2007). Knowledge transfer among face-to-face information systems development team members. *Decision Support Systems, 43*(2), 322-335.

Jussupow, E., Spohrer, K., Heinzl, A., & Gawlitza, J. (2021). Augmenting medical diagnosis decisions? An investigation into physicians' decision-making process with artificial intelligence. *Information Systems Research, 32*(3), 713-735.

Kaur, D., Uslu, S., Rittichier, K. J., & Durresi, A. (2022). Trustworthy artificial intelligence: A review. *ACM Computing Surveys, 55*(2), Article 39.

Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science, 3*(3), 383-397.

Lamont, M. (2009). *How professors think: Inside the curious world of academic judgment.* Harvard University Press.

Lebovitz, S., Levina, N., & Lifshitz-Assaf, H. (2021). Is AI ground truth really "true"? The dangers of training and evaluating AI tools based on experts' know-what. *MIS Quarterly, 45*(3), 1501-1525.

Leonardi, P. M. (2011). When flexible routines meet flexible technologies: Affordance, constraint, and the imbrication of human and material agencies. *MIS Quarterly, 35*(1), 147-167.

Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue, 16*(3), 31-57.

Maclure, J. (2021). AI, explainability and public reason: The argument from the limitations of the human mind. *Minds and Machines, 31*(3), 421-438.

Newell, A., & Simon, H. A. (1956). The logic theory machine: A complex information processing system. *IRE Transactions on information theory, 2*(3), 61-79.

Palmer, J. W., Speier, C., Wren, D., & Hahn, S. (2000). Electronic journals in business schools: Legitimacy, acceptance, and use. *Journal of the Association for Information Systems, 1*, Article 2.

S., & Krakowski, S. (2021). Artificial intelligence and management: The automation-augmentation paradox. *Academy of Management Review, 46*(1), 192-210.

Rawls, J. (1993). *Political liberalism.* Columbia University Press.

Reidpath, D. D., & Allotey, P. (2019). The problem of 'trickle-down science' from the Global North to the Global South. *BMJ Global Health, 4*(2), Article e001719.

Sahaym, A., Vithayathil, J., Sarker, S., & Bjørn-Andersen, N. (2023). Value destruction in information technology ecosystems: A mixed-method investigation with interpretive case study and analytical modeling. *Information Systems Research, 34*(2), 508-531.

Sarker, S., Sarker, S., Sahaym, A., & Bjørn-Andersen, N. (2012). Exploring value cocreation in relationships between an ERP vendor and its partners: A revelatory case study. *MIS Quarterly, 36*(1), 317-338.

Sarker, S., Agarwal, R., Goes, P., Gregor, S., Henfridsson, O., Saunders, C., & Tan, B. (2015). Roles and responsibilities of a senior editor. *Journal of the Association for Information Systems, 16*(9), 707-728.

Saunders, C. (2005). Editor's comments: Looking for diamond cutters. *MIS Quarterly, 29*(1), iii-viii. https://doi.org/10.2307/25148664

Scholarly Peer Review. (2023, October 25). In Wikipedia. https://en.wikipedia.org/wiki/Scholarly_peer_review

Silbiger, N. J., & Stubler, A. D. (2019). Unprofessional peer reviews disproportionately harm underrepresented groups in STEM. *PeerJ, 7,* Article e8247.

Simon, H. A. (2019). *The sciences of the artificial* (3rd ed.). MIT Press.

Susarla, A., Gopal, R. D., 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*. Advance online publication. https://doi.org/10.1287/isre.2023.ed.v34.n2

Teye, M. T., Missah, Y. M., Ahene, E., & Frimpong, T. (2022). Evaluation of conversational agents: understanding culture, context and environment in emotion detection. *IEEE Access, 10*, 24976-24984.

Wanderley, S., Alcadipani, R., & Barros, A. (2021). Recentering the Global South in the Making of Business School Histories: Dependency Ambiguity in Action. *Academy of Management Learning & Education, 20*(3), 361-381.

Wang, D., Weisz, J. D., Muller, M., Ram, P., Geyer, W., Dugan, C., Wallach, H., & Gray, M. (2019). Human-AI collaboration in data science: Exploring data scientists' perceptions of automated AI. *Proceedings of the ACM on Human-Computer Interaction, 3*(CSCW), Article 211.

About the Authors

Suprateek Sarker Suprateek Sarker is the Rolls-Royce Commonwealth Commerce Professor at the McIntire School of Commerce, University of Virginia, USA. Professor Sarker has held short-term visiting professorships, and visiting scholar positions at many institutions around the globe and currently holds a visiting professorship at the London School of Economics and Political Science (LSE), UK, and an honorary professorship at the University of Sydney, Australia. He has published articles in many leading journals and his work has been funded by agencies such as NSF. He serves on (or has served on) the editorial boards of a number of leading academic journals, including *Information Systems Research* (as editor-in-chief), *Journal of Management Information Systems* (board of editors), *Journal of the Association for Information Systems* (past editor-in-chief), and *MIS Quarterly* (as senior editor emeritus). He is also a past senior editor of *Decision Sciences Journal*. Currently, Suprateek is serving as the immediate past president of the Association for Information Systems.

Anjana Susarla is the Omura-Saxena Professor of Responsible AI at the Eli Broad College of Business, Michigan State University. Her research interests include the economics of information systems and artificial intelligence. Her research has been published in *Information Systems Research*, *Knowledge Discovery and Data Mining*, *Journal of Management Information Systems*, *Management Science*, and *MIS Quarterly* and has received funding from agencies such as the National Institute of Health.

Ram Gopal is the Information Systems Society's Distinguished Fellow and a professor of information systems and management at the Warwick Business School. He has a diverse portfolio of research that spans big data analytics, health informatics, financial technologies, information security, privacy and valuation, intellectual property rights, online market design and business impacts of technology. His research has appeared in *Management Science*, *Management Information Systems Quarterly*, *Operations Research*, *INFORMS Journal on Computing*, and *Information Systems Research*.

Jason Bennett Thatcher is a professor of information systems at the Leeds School of Business at the University of Colorado Boulder. He is also a professor of management science at the Alliance Manchester Business School at the University of Manchester. He is a generalist, broadly interested in the impact of information technology on the performance of technology users, organizations, and societies. He has published in the *Journal of the Association for Information Systems, MIS Quarterly*, *Information Systems Research, Journal of Applied Psychology*, and other outlets. Currently, Jason is learning to deep fry tofu and perfect Korean-style ramen.