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 9

2024

Peer Review in the Age of Generative AI

Atreyi Kankanhalli
National University of Singapore, atreyi@comp.nus.edu.sg

Recommended Citation

Kankanhalli, Atreyi (2024) "Peer Review in the Age of Generative AI," *Journal of the Association for Information Systems*, 25(1), 76-84.
DOI: 10.17705/1jais.00865
Available at: https://aisel.aisnet.org/jais/vol25/iss1/9

Journal of the Association for Information Systems (2024) **25**(1), 76-84
doi: 10.17705/1jais.00865

EDITORIAL

ISSN 1536-9323

Peer Review in the Age of Generative AI

Atreyi Kankanhalli[1]
[1]National University of Singapore, Singapore, atreyi@comp.nus.edu.sg

Abstract

Rapid advances in artificial intelligence (AI), including recent generative forms, are significantly impacting our lives and work. A key aspect of our work as IS researchers is the publishing of research articles, for which peer review serves as the primary means of quality control. While there have been debates about whether and to what extent AI can replace researchers in various domains, including IS, we lack an in-depth understanding of how AI can impact the peer review process. Considering the high volume of submissions and limited reviewer resources, there is a pressing need to use AI to augment the review process. At the same time, advances in AI have been accompanied by concerns about biases introduced by AI tools and the ethics of using them, among other issues such as hallucinations. Thus, critical issues to understand are: how can AI augment and potentially automate the review process, what are the pitfalls in doing so, and what er the implications for IS research and peer review practice. I will offer my views on these issues in this opinion piece.

Keywords: Artificial Intelligence (AI), IS Research, Peer Review Process, Augment, Bias, Ethics, Generative AI, Hallucinations

Dov Te'Eni and David Schwartz were the accepting senior editors. This paper was submitted on June 17, 2023 and underwent two revisions. This paper is part of the Special Issue on The Future Impact of AI on Academic Journals and the Editorial Process.

1 Introduction

Rapid advances in artificial intelligence (AI) are significantly impacting our lives and work. AI refers to computing technologies for performing tasks that normally require human intelligence, such as perception, planning, reasoning, and learning (Russell & Norvig, 2016). Most recently, generative AI technologies such as ChatGPT and its variants have further enhanced the power of AI. These AI techniques, such as large language and diffusion models, are capable of generating novel content including text, images, audio, product designs and software code from training data. The widespread adoption of generative AI is currently disrupting the way we work, learn, and create.

Even prior to generative AI, there have been numerous predictions of how AI will replace jobs (e.g., Lee, 2017) and research on how this will impact different professions (e.g., Acemoglu, 2021). A critical issue for IS researchers is the extent and ways in which AI can replace or augment our research activities (Loebbecke et al., 2020). A key and arguably the most salient aspect of our work is the publishing of research articles, for which peer review is the primary means of quality control (Iivari, 2016). Submissions to peer reviewed journals have seen relentless growth due to increasing publication pressure and burgeoning numbers of active researchers (Bornmann et al., 2021; Publons, 2018). This has placed a tremendous strain on the publication process, with increasing demands on peer reviewers and their time (Petrescu & Krishen, 2022; Publons, 2018). Insufficient reviewer capacity is seen as a salient reason for longer publication cycle times and the compromised quality of reviews in IS (Avital, 2018).

Thus, there is a vital need for techniques that can improve the efficacy of the peer review process, which serves as the gatekeeper for research quality. To address problems with the peer review process, various solutions have been proposed, including providing rewards and feedback to reviewers (Avital, 2018; Iivari, 2016). In terms of AI solutions, tools have been developed for review tasks such as format checks, reviewer-manuscript matching, and readability checks (Checco et al., 2021). Though some benefits of these tools have been reported, there are also concerns that the use of AI may add to biases or ethical and other issues in peer review (Chubb et al., 2022). Given the magnitude of the peer review problem (Petrescu & Krishen, 2022; Avital, 2018), it is useful to discuss the promise and perils of AI to support the peer review process, which is what I aim to do in this opinion piece. Thus, I am interested in examining: (1) How can AI augment and potentially automate the peer review process for IS research? (2) What are the challenges in doing so? (3) What are the implications of AI for the peer review of IS research going forward?

The first question is important since AI tools are able to perform some tasks in the peer review process, such as routine checking, but not necessarily all parts of the process (Chubb et al., 2022). Thus, it is useful to understand where in the review process efficiency gains can be obtained and which tasks have yet to be addressed by AI. The second question is salient since AI tools may encounter major issues in performing complex intellectual activities, such as peer review, where the ground truth is not necessarily apparent (Shah, 2022). These challenges need to be identified and acknowledged in order to be addressed. The third question is important because AI and its recent developments are likely to impact the peer review of IS research in complex ways.

In response to the first question, I will outline the key tasks performed during the review process and discuss how AI tools are able to augment and potentially automate them. Regarding the second question, I will identify the challenges associated with using AI tools to augment peer review tasks, including problems of bias, ethics, and hallucination. Finally, I will offer my views on the research implications of these tools and provide suggestions to enhance their use for peer review of IS research. Overall, this opinion piece contributes to the discussion on how developments in AI implicate the peer review process, which is the primary means of quality control for our research.

2 Related IS Research

A number of IS scholars have raised concerns about deterioration in the quality of reviews (Iivari, 2016) and long review cycle times in IS journals (Stafford, 2016). This has resulted in repeated calls to improve the peer review process (Avital, 2018). In response, some recommendations have focused on reviewers, for example, by providing them guidelines on how to write good reviews (Kohli & Straub, 2011; Leidner et al., 2022), including developmental reviews (Saunders, 2005). Premium journals such as *MIS Quarterly* have organized workshops for reviewer development (Rai, 2019). Others have suggested editorial approaches, such as editors providing systematic feedback to reviewers, rewarding good reviews, and making reviewers more accountable by revealing their identity to authors in certain conditions (Iivari, 2016).

Related to this opinion piece, digital and other solutions have also been proposed to support the review process—for example, by using wikis (Kane & Fichman, 2009) or blockchain technology (Avital, 2018), through structured communication (Mandviwalla et al., 2008), or by employing an open knowledge approach (Hardaway & Scamell, 2012). Presciently, Ein-Dor suggested an open review process where a community of peers serve as reviewers (Peffers et al., 2003), which is now being adopted in some journals and conferences. As AI technology has advanced, recent editorial articles have begun to discuss the impact of generative AI on all stakeholders of their journals more broadly (Burton-Jones, 2023; Shmueli et al., 2023; Susarla et al., 2023). However, we continue to lack an in-depth understanding of the implications of AI tools (both pre-generative and generative AI) for the IS peer review process. Aligning with Susarla et al.'s (2023) call for such work, this opinion piece aims to provide such understanding.

3 AI Augmentation and Automation of Peer Review

In this section, I discuss how AI tools could augment and potentially automate different tasks of the peer review process, which is a pressing need. Peer review tasks can be divided into two stages i.e., pre-peer review screening, and peer review (Spezi et al., 2018). *Pre-peer review screening* consists of a number of checks, including *format checks*, *plagiarism detection*, and *checking of language quality*. In many cases, if a paper does not satisfy these checks, or fails in the initial editorial screening, it will be "desk rejected." If the paper passes this stage, reviewers will be matched to the manuscript to enable peer review. For manuscript screening, AI tools have made good progress in automating tasks such as *format checks*. Journals and conferences impose various format rules for submissions—e.g., page limits, paper size, margins, font size, line spacing, paper structure, as well as section heading, figure, table, and reference styles—whose compliance needs to be checked. For instance, tools like Penelope.ai[1] offer multiple features for checking a manuscript's structure, styles, references, and metadata that can be configured to satisfy the requirements of different publication outlets. These checks can be largely automated, thus offering high potential for AI automation. Typically, this task does not require generative AI or other advanced AI techniques, as the format rules are usually unambiguous to detect.

Plagiarism detection entails checking that a submission does not "copy" from other sources without proper attribution, which then helps to determine if the authors are trying to claim other's ideas, text, and data as their own. To do so, many major publishers make use of services like CrossCheck powered by iThenticate[2] software, which transforms each manuscript into a form of digital fingerprint for comparison with existing sources in order to compute similarity scores. This can not only detect verbatim text matches but also fuzzy matches, e.g., synonyms, word substitutions, paraphrasing, and translation. Such tools have thus far been fairly effective, though some human judgment is required to interpret similarity scores and determine the nature and extent of plagiarism (Anson et al., 2020). However, the ability to detect plagiarism has been upended by the advent of generative AI tools such as ChatGPT, which creates plausible content by synthesizing content from extant sources (Lund et al., 2023). Would a plagiarism detector be able to check if authors used ChatGPT for writing their submission, which combines others' work in a complex manner? We currently do not have the means to reliably do so or to know what sources ChatGPT has learned from. While tools like ZeroGPT[3] claim to be able to detect if a text has been produced by AI versus humans, there is much skepticism about their accuracy (Sadasivan et al., 2023). Indeed, plagiarism detection has become highly challenging in the era of generative AI, offering low potential for AI automation.

The *language quality* of submissions includes the basics such as correctness of spelling and grammar, as well as writing style and organization. Spelling and grammar checks in review tools are common, e.g., Enago Grammar Checker, and are now routinely seen in word processing software as well. Additionally, language quality is often assessed by readability metrics such as the Flesch-Kincaid Grade Level (Kincaid et al., 1975), which captures simple features of words (e.g., characters per word) and sentences (i.e., words per sentence). However, these metrics do not explicitly capture writing style and organization, and other language-quality aspects like cohesion and logic (Crossley et al., 2023). Generative AI-based methods offer promise in terms of detecting academic writing style (Crossley et al., 2023). However, it is still challenging to assess the cohesive organization of ideas and logic in writing, which tools like UNSILO[4] aim to do. Thus, I rate this task as overall having medium potential for AI automation, both with pre-generative and generative AI.

If a manuscript passes the pre-review screening, it will be sent for review. Current peer review systems for IS journals usually allow editors to search for reviewers based on the match between the keywords of the manuscript and the expertise areas provided by reviewers. However, this approach suffers from limitations since information about reviewers' expertise areas is often incomplete or not detailed enough to correspond to manuscript keywords, entailing considerable effort from editors to find appropriate reviewers for a submission (Price & Flach, 2017). For reviewer-manuscript matching, disciplines like computer science have developed AI tools such as the Toronto Paper Matching System (TPMS) (Charlin & Zemel, 2013). This system creates reviewer profiles not only by using reviewer-provided expertise areas but also by mining the web to find the reviewer's publications and extract representative keywords and topics from them. Conference review systems like Microsoft CMT use a combination of suggestions from editors, reviewer bids, and TPMS recommendations to match reviewers to manuscripts (Price & Flach, 2017). Systems like TPMS can address the limitations of simpler keyword-based reviewer matching tools, thus offering high potential for AI automation.

The second stage, i.e., *peer review*, typically involves the assessment of four main criteria i.e., relevance or scope, rigor or soundness, novelty or originality, and significance or importance. This is followed by writing the review report (for reviewers) and integrating reports (for editors). Assessment of a submission's scope and relevance can be assisted through text summarization tools such as UNSILO and more recently GPT-4.[5] However, these still require editorial judgment to review the text summaries against the publication outlets' scope and domain (Susarla et al., 2023), thus offering medium potential for AI automation. With regards to the soundness and rigor of a submission, tools such as Enago[6] state that they can

[1] https://www.penelope.ai/
[2] https://www.ithenticate.com/
[3] https://www.zerogpt.com/
[4] https://discovery.researcher.life/publisher
[5] https://openai.com/gpt-4
[6] https://www.enago.com/publisher.htm

verify the soundness of claims and arguments made by a manuscript by checking the supporting references. GPT-4 is able to identify relevant references on a topic (Agrawal et al., 2023) but can summarize their content wrongly due to hallucinations, i.e., false information generated by large language models (Ji et al., 2023). Statcheck (Nuijten & Polanin, 2020) and StatReviewer (Shanahan, 2016) aim to check the soundness of statistical tests reported in manuscripts. However, doubts have been expressed about their reliability (e.g., Schmidt, 2017), indicating that human intervention may be needed to verify their results. Additionally, such methods do not apply to qualitative research studies. The combined capabilities of these tools imply a medium-low level of AI automation. Last, the novelty and significance of submissions have been the most challenging aspects to assess, with a low potential for AI automation. Tools like ReviewAdviser[7] are said to specifically assist in evaluating the contribution of a submission based on novelty and significance, though limited details and evidence are available to support the claims. Recent editorial articles have highlighted the need for caution in using generative AI tools to augment reviewing and editing activities, stating that human verification is always required due to the possibility of hallucinations (Shmueli et al., 2023; Susarla et al., 2023). Table 1 below summarizes the potential of AI tools to automate the different tasks of the peer review process.

Other than the above peer review tasks, reproducibility checks are becoming prevalent in related disciplines like data science (Shmueli et al., 2023) and economics (Vilhuber, 2020). This requires the authors of conditionally accepted papers to provide their data and code to journals' data editors to verify the reproducibility of their results. In IS journals, we are seeing a move toward greater research transparency (Burton-Jones et al., 2021). Conceivably, we could see similar reproducibility requirements become a part of the peer review process for quantitative IS studies. If so, AI tools could aid data editors by performing reproducibility checks i.e., this task has a high potential for AI augmentation. Overall, the above discussion indicates the ability of AI tools (both pre-generative and generative AI) to augment tasks of the review process but a somewhat limited ability to fully automate the tasks.

Table 1. Automation Potential of AI for the Peer Review Process

Task	AI automation potential	**Example of AI tool**
Format check: checking that the manuscript follows publication outlet format rule of structure, styles, references, and metadata	High	Penelope.ai
Plagiarism detection: identifying the extent and nature of copying from other sources without source attribution	Low	iThenticate, ZeroGPT
Language quality: assessing audience-appropriate readability, cohesion, logic	Medium	UNSILO
Manuscript-reviewer matching: finding suitable reviewers for a manuscript using reviewer profiling	High	TPMS
Scope/relevance: assessing fit with the scope of the publication outlet	Medium	UNSILO, GPT-4
Soundness/rigor: checking that study methodology and analysis are rigorous and robust	Medium-Low	Enago, StatCheck, StatReviewer
Novelty: newness or departure from the existing body of knowledge	Low	ReviewAdviser
Significance: importance of the phenomenon that the manuscript is focusing on	Low	ReviewAdviser
Writing and integrating reviews: reviewer and editor roles of writing and integrating reports	Medium	GPT-4
Reproducibility check: author code and data analysis checks for reproducibility of findings	High	GPT-4

[7] https://github.com/neulab/ReviewAdvisor

4 Challenges in AI Use for Peer Review

A major challenge in peer reviewing is to avoid *biases* that can cause unfairness, some of which are hidden. Indeed, many sociocultural biases are present in peer review (Lee et al., 2013) and could propagate to AI systems through algorithmic bias (Mittelstadt et al., 2016; Zarsky, 2016). Bias in the review process can take different forms. This includes first-impression bias, ideological/theoretical orientation, language, social identity, and prestige biases (Lee et al., 2013; Siler et al., 2015). Challenges remain in modeling these biases, although some studies are attempting to do so (e.g., Feliciani et al., 2019).

In particular, machine-learning AI techniques are inherently conservative, as they are trained with data from the past. This could lead to bias when such tools are used to inform current review decisions. For example, papers from countries that have historically been under-represented in the literature could have a higher rejection rate using AI methods since the reviews may not adequately account for the rising quality of submissions from such regions over time (Checco et al., 2021). Biases can also be introduced by the fact that editors have mostly selected reviewers from developed regions in the past (Publons, 2018). Such bias applies to both pre-generative and generative AI techniques that are tuned with human feedback (Ferrara, 2023).

Furthermore, using AI tools to flag problematic papers to reviewers could introduce additional bias i.e., bias amplification (Lloyd, 2018). This could result in the opposite issue to the above i.e., the way the model interprets the manuscript could propagate to the reviewer, resulting in unintended biased outcomes. For example, if the AI model identifies as potential issues: (1) the presence of typos, (2) the citing of references from under-represented regions, or (3) the use of techniques that have been employed in previously rejected papers, this could increase the salience of such factors in the mind of the reviewers and influence their judgment. With the increasing focus on diversity, equity, and inclusion in IS research (Burton-Jones & Sarker, 2021; Wright et al., 2023) it would be imperative to identify and mitigate such biases.

Additionally, there are *ethical concerns* arising from the use of AI tools in the peer review process. In particular, two salient concerns come to mind. First is the issue of explainability (Checco et al., 2021), which impacts the transparency of review decisions. When the relationship between the original data and the AI model prediction is unclear, there is a lack of algorithm *explainability*, which in turn leads to mistrust toward the AI tool. An author will not trust a review if there is no transparency about the rationale for the decision. If AI tools are used to assist in peer review, it is crucial to ensure transparency about how the models work to justify the decisions made. Second is the issue of *accountability* for peer reviews when AI assists in the process. Review decisions have important implications for authors' careers. Leaving such crucial decisions to AI tools could indicate that editors are avoiding responsibility and accountability. Thus, these applications need to be treated as decision support systems rather than decision makers.

Other concerns arise with the use of AI tools, particularly for writing review reports. Popular generative AI tools like ChatGPT could be used by reviewers and editors to summarize important points of a submission and carry out verification checks on it, followed by creating reports. There are several issues with such use, the most important being *hallucination* i.e., producing incorrect or imaginary output, stated with as much certainty as correct output (Ji et al., 2023). Hallucination in peer review reports compromises the quality of the report. Another issue relates to the *privacy* of confidential information e.g., parts of a submission that reviewers upload into generative AI tools to help prepare their review reports. This information would then be used to further train the AI tool and could appear in future responses to the tool's users, compromising the confidentiality and potentially the intellectual property of the authors. Yet another issue is the possibility of convergence of style and the content of review reports when many reviewers/editors use generative AI tools to create their reports over time.

The above concerns need to be considered and managed carefully when designing and deploying AI tools for peer review. As can be seen from the above discussions, AI technologies can serve as a double-edged sword in the review process. Continued research on their development and use is essential to help ensure that AI tools play a positive role in the peer review process.

5 Implications for IS Research and Practice

After discussing the opportunities and challenges of using AI tools for peer review, I now outline the research and practical implications of these tools for the peer review of IS research. But before I do that, I would like to highlight what is unique about the IS discipline in terms of peer review. IS is a multidisciplinary field that spans computer science (CS) and social sciences like management, economics, psychology, and sociology (Tarafdar & Davison, 2018). Its links to CS as a reference discipline are especially important for this discussion since the development of AI tools for peer review has most frequently been undertaken by CS researchers for the peer review of their research e.g., the TPMS. Thus, it makes sense to discuss how IS research differs from CS research to understand how these tools could be adapted to our discipline.

Two salient differences that implicate peer review processes come to mind. First, CS research largely concerns the formulation of new processes/methods/ algorithms, mostly using mathematically based analyses and not referring to other disciplines (Glass et al., 2004). For example, a new algorithm could be designed for recognizing specific activities in a video sequence and then benchmarked against other recent algorithms on public datasets to show its superior performance. This kind of research, in some sense, simplifies the automation of the peer review process since the scope is relatively narrow and comparisons to prior work are relatively objective—making it easier (compared to most IS research) to identify the body of related studies and build AI tools to carry out tasks like assessing the novelty and significance of the work.

Second, much of CS research is published in conference proceedings, where papers are shorter in length and have shorter and fewer review cycles—as compared to IS research where journal publications play a much greater role. These differences have implications for the peer review of IS research: there could be a more pressing need for AI automation in order to speed up review cycles, but the tasks of automation may be more challenging. CS researchers have also built AI tools to perform automatic format checks of conference papers (e.g., SIGSOFT Submission Checker), which are missing in IS conferences like ICIS and PACIS (which have fairly complex format rules). At the same time, IS research differs from its social science reference disciplines in that IS editors and reviewers are likely to be more open to and familiar with the use of AI tools for peer review, compared to these groups in the other disciplines. Thus, the unique characteristics of the IS field offer both advantages and drawbacks for the use of AI tools for peer review as compared to related disciplines.

The IS field's characteristics also inform the research implications of the use of AI tools for the peer review of IS research. Particularly, while IS researchers may not be at the forefront of developing such AI tools (in contrast to CS researchers), our unique expertise in understanding the adoption and consequences of IS allows us to examine important questions around the adoption, effectiveness, and broader impacts of these tools e.g., the disruptive consequences of generative AI tools for peer review. Further, while pre-generative AI tools (with some recently incorporating generative techniques) are being adopted by publishers and editors of journals and could save reviewer resources, rigorous evaluation of their costs and benefits is needed, including for different IS research paradigms. Again, IS researchers are in a position to do so, with their research spanning quantitative and qualitative research paradigms and employing design, econometrics, and behavioral methods, among others.

With respect to the practical implications for the peer review of IS research, I offer several suggestions for editors and reviewers. Recently there has been much debate about how IS journals should respond to the challenges posed by the use of generative AI by authors, editors, and reviewers. All major IS journals are in the process of developing policies for regulating the use of these tools by authors, editors, and reviewers (e.g., Susarla et al., 2023; Burton-Jones, 2023). In the meantime, however, it is still important to understand which tasks of the review process are more amenable to automation and to ensure that IS research as a discipline is up to date in its use of such tools for the peer review of conference and journal submissions. For example, such tools can be used for format checks, reviewer-manuscript matching, and reproducibility checks (as per Table 1). The editors and publishers of IS journals and conferences could then target the peer review tasks with medium automation potential, like language-quality assessment, and potentially work with CS researchers and developers to build tools for these tasks. For soundness and rigor checks, it is important to note that the ease of automation is likely to be higher for quantitative IS research studies with statistical or econometrics tests than for qualitative IS studies, where the coding of data may be subjective. Finally, there are some tasks in the review process (such as writing, integrating reviews, plagiarism detection, or assessing novelty and significance) where editors and publishers would need to wait for advances in AI techniques (e.g., systems that use knowledge graphs to verify the outputs of generative AI methods, to prevent hallucinations) to automate them. However, beyond such technological advances, the use of these tools requires the development of legal and policy frameworks that protect the intellectual property of sources and ensure appropriate source attribution by generative AI and its users.

6 Conclusion

This opinion piece discussed how AI tools could be used to obtain efficiencies in the quality control and peer review processes for IS research. This is especially true for the more routine aspects of the review process, which require less intellectual input or domain expertise e.g., format and reproducibility checks, language/readability assessments, reviewer matching, and improving the writing of review reports. At the same time, there are important negative implications of using these tools in terms of biases, hallucinations, ethics, and other issues. Thus, editors and reviewers need to be cognizant of these challenges when employing AI to gain efficiencies in the review process. Going forward, more research is needed to mitigate these challenges and realize the full potential of AI augmentation and even automation for the peer review of IS research.

References

Acemoglu, D. Ed. (2021). *Redesigning artificial intelligence*. MIT Press.

Agrawal, A., Suzgun, M., Mackey, L., & Kalai, A. T. (2023). *Do language models know when they're hallucinating references?* Available at https://arxiv.org/abs/2305.18248

Anson, C. M., Hall, S., Pemberton, M., & Moskovitz, C. (2020). Reuse in STEM research writing: Rhetorical and practical considerations and challenges. *AILA Review, 33*(1), 120-135.

Avital, M. (2018). Peer review: Toward a blockchain-enabled market-based ecosystem. *Communications of the Association for Information Systems*, *42*, 646-653.

Bornmann, L., Haunschild, R., & Mutz, R. (2021) Growth rates of modern science: A latent piecewise growth curve approach to model publication numbers from established and new literature databases. *Humanities and Social Sciences Communication*, *8*, 1-15

Burton-Jones, A. (2023). Editing Innovation in MIS Quarterly. *MIS Quarterly*, *47*(3), i-xii.

Burton-Jones, A., Boh, W. F., Oborn, E., & Padmanabhan, B. (2021). Advancing research transparency at MIS Quarterly: A pluralistic approach. *MIS Quarterly, 45*(2), iii-xviii.

Burton-Jones, A., & Sarker, S. (2021). Editor's comments: Creating our editorial board position statement on diversity, equity, and inclusion (DEI). *MIS Quarterly, 45*(4), i-x.

Charlin, L., & Zemel, R.S. (2013). The Toronto Paper Matching System: An automated paper-reviewer assignment system. *Proceedings of 30th International Conference on Machine Learning*.

Checco, A., Bracciale, L., Loreti, P., Pinfield, S., & Bianchi, G. (2021). AI-assisted peer review. *Humanities and Social Sciences Communications*, *8*(1), 1-11.

Chubb, J., Cowling, P., & Reed, D. (2022). Speeding up to keep up: Exploring the use of AI in the research process. *AI & Society*, *37*(4), 1439-1457.

Crossley, S., Heintz, A., Choi, J. S., Batchelor, J., Karimi, M., & Malatinszky, A. (2023). A large-scaled corpus for assessing text readability. *Behavior Research Methods, 55*(2), 491-507.

Feliciani, T., Luo, J., Ma, L., Lucas, P., Squazzoni, F., Marušić, A., & Shankar, K. (2019). A scoping review of simulation models of peer review. *Scientometrics*, *121*, 555-594.

Ferrara, E. (2023). *Should ChatGPT be biased? Challenges and risks of bias in large language models.* Available at https://arxiv.org/abs/2304.03738

Glass, R. L., Ramesh, V., & Vessey, I. (2004). An analysis of research in computing disciplines. *Communications of the ACM, 47*(6), 89-94.

Hardaway, D. E., & Scamell, R. W. (2012). Open knowledge creation: Bringing transparency and inclusiveness to the peer review process. *MIS Quarterly*, *36*(2), 339-346.

Iivari, J. (2016). How to improve the quality of peer reviews? Three suggestions for system-level changes. *Communications of the Association for Information Systems*, *38*, 263-273.

Ji Z, Lee N, Frieske R, Yu T, Su D, Xu Y, Ishii E, et al. (2023) Survey of hallucination in natural language generation. *ACM Computing Surveys*, *55*(12), 1–38.

Kane, G. C., & Fichman, R. G. (2009). The shoemaker's children: Using wikis for information systems teaching, research, and publication. *MIS Quarterly*, *33*(1), 1-17.

Kincaid, J.P., Fishburne, R.P., Rogers, R.L., & Chissom, B.S. (1975). *Derivation of new readability formulas for Navy enlisted personnel* (Research Branch Report 8-75). Naval Technical Training. Available at https://stars.library.ucf.edu/cgi/viewcontent.cgi?article=1055&context=istlibrary

Kohli, R., & Straub, D. (2011). How reviews shape MIS Quarterly: A primer for reviewers and editors. *MIS Quarterly*, *35*(3), iii-vii.

Lee, C. J., Sugimoto, C. R., Zhang, G., & Cronin, B. (2013). Bias in peer review. *Journal of the American Society for Information Science and Technology*, *64*(1), 2-17.

Lee, K.-F. (2017). The real threat of Artificial Intelligence. *The New York Times*. https://www.nytimes.com/2017/06/24/opinion/sunday/artificial-intelligence-economic-inequality.html

Leidner, D. E., Carte, T., Chatterjee, S., Chen, D., Jones, M., & Preston, D. (2022). On civil critique: reviewing for JAIS. *Journal of the Association for Information Systems*, *23*(1), 1-12.

Lloyd, K. (2018). *Bias amplification in artificial intelligence systems.* ArXiv. https://arxiv.org/abs/1809.07842

Loebbecke, C., El Sawy, O. A., Kankanhalli, A., Markus, M. L., & Te'eni, D. (2020). Artificial intelligence meets IS researchers: Can it replace us? *Communications of the Association for Information Systems*, *47*(1), 273-283.

Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology, 74*(5), 570-581.

Mandaviwalla, M., Patnayakuni, R., & Schuff, D. (2008). Improving the peer review process with information technology. *Decision Support Systems, 41*(4), 29-40.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), 1-21.

Nuijten, M. B., & Polanin, J. R. (2020). "Statcheck": Automatically detect statistical reporting inconsistencies to increase reproducibility of meta-analyses. *Research Synthesis Methods*, *11*(5), 574-579.

Peffers, K., Ein-Dor, P., & Zmud, R. W. (2003). Scholarly products in IS: Will advances in electronic media promote evolutionary or radical change? *Communications of the Association for Information Systems*, *11*, 498-512.

Petrescu, M., & Krishen, A. S. (2022). The evolving crisis of the peer-review process. *Journal of Marketing Analytics*, *10*(3), 185-186.

Price, S., & Flach, P. A. (2017). Computational support for academic peer review: A perspective from artificial intelligence. *Communications of the ACM*, *60*(3), 70-79.

Publons (2018) *Global state of peer review 2018*. https://publons.com/static/Publons-Global-State-Of-Peer-Review-2018.pdf

Rai, A. (2019) Developing virtuous reviewers, *MIS Quarterly, 43*(4), iii-iv.

Russell, S. J., & Norvig, P. (2016) *Artificial intelligence: A modern approach*. Pearson.

Sadasivan, V. S., Kumar, A., Balasubramanian, S., Wang, W., & Feizi, S. (2023). *Can AI-generated text be reliably detected?* ArXiv. https://arxiv.org/abs/2303.11156

Saunders, C. (2005). From the trenches: Thoughts on developmental reviewing. *MIS Quarterly, 29*(2), iii.

Schmidt, T. (2017). *Statcheck does not work: All the numbers. Reply to Nuijten et al. (2017).* Available at http://psyarxiv.com/hr6qy

Shah, N. B. (2022). Challenges, experiments, and computational solutions in peer review. *Communications of the ACM, 65*(6), 76-87.

Shanahan, D. (2016, May 23). A peerless review? Automating methodological and statistical review. *Research in Progress Blog*. https://blogs.biomedcentral.com/bmcblog/2016/05/23/peerless-review-automating-methodological-statistical-review/

Shmueli, G., Maria Colosimo, B., Martens, D., Padman, R., Saar-Tsechansky, M., Liu Sheng, O. R., Street, N.W., and Tsui, K. L. (2023). How can IJDS authors, reviewers, and editors use (and misuse) generative AI? *INFORMS: Journal on Data Science, 2*(1), 1-9.

Siler, K., Lee, K., & Bero, L. (2015). Measuring the effectiveness of scientific gatekeeping. *Proceedings of the National Academy of Sciences*, *112*(2), 360-365.

Spezi, V., Wakeling, S., Pinfield, S., Fry, J., Creaser, C., & Willett, P. (2018). "Let the community decide"? The vision and reality of soundness-only peer review in open-access mega-journals. *Journal of Documentation*, *74*(1), 137-161.

Stafford, T.F. (2016). Optimizing the business side of science: Publication review cycles and process management considerations. *Communications of the Association for Information Systems, 38,* 296-306.

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.

Tarafdar, M., & Davison, R.M. (2018). Research in information systems: Intra-disciplinary and inter-disciplinary approaches. *Journal of the Association for Information Systems*, *19*(6), 523-551.

Vilhuber, L. (2020). Reproducibility and replicability in economics. *Harvard Data Science Review, 2*(4), 1-39.

Wright, R. T., Saunders, C., Sarker, S., Kankanhalli, A., Whitley, E. A., Tuunainen, V. K., & Henriksen, H. Z. (2023). The scholarly divide: Insights from the AIS well-being project. *Communications of the Association for Information Systems, 52*(1), 777-797.

Zarsky, T. (2016). The trouble with algorithmic decisions: An analytic road map to examine efficiency and fairness in automated and opaque decision making. *Science, Technology, & Human Values*, *41*(1), 118-132.

About the Authors

Atreyi Kankanhalli is Provost's Chair Professor at the Department of Information Systems and Analytics, National University of Singapore. She served as assistant dean for Research and for Undergraduate Studies. She obtained her B.Tech. from IIT Delhi and Ph.D. from NUS. She has visited at London School of Economics, the University of California Berkeley, and ESSEC Business School. She conducts research in online communities and digital (including human-AI) collaboration, digital innovation, and transformation, such as for healthcare. Her work has appeared in premium journals such as *MIS Quarterly*, *Information Systems Research*, *Research Policy*, *Journal of Management Information Systems*, *Proceedings of the National Academy of Sciences*, and *Journal of the Association for Information Systems*. She has over 200 publications, with more than 20,000 citations. She is a senior editor at *Information Systems Research* and at *AIS Transactions on TRR* and *Human-Computer Interaction*. She has been a program chair for conferences like PACIS 2015 and ICIS 2021. She has received a number of awards including AIS Fellow, IBM Faculty Award, and Singapore 100 Women in Tech.