Trends in Scholarly Publishing

OPEN ACCESS

AI in Peer Review: Tool or Threat to Editorial Integrity?

Mingfang Lu
Executive President, AiScholar, China

Corresponding Author: Mingfang Lu (lumingfang@hotmail.com)

ABSTRACT

The rise of Artificial Intelligence (AI) in scholarly publishing is rapidly reshaping the peer review process. Once limited to tasks like plagiarism checks and manuscript triage, AI tools now assist in reviewer selection, ethical screening, and even pre-evaluation. While these innovations offer unmatched efficiency and scalability, particularly beneficial for under-resourced journals, they also pose serious concerns regarding transparency, editorial independence, and bias. This perspective explores the evolving role of AI in peer review, highlighting its dual potential as both a powerful tool and a source of ethical risk. Drawing on recent studies and global publishing practices, we argue that AI should enhance, not replace, human judgment in peer evaluation. The article proposes practical best practices for responsible integration, emphasizing transparency, bias mitigation, and global accessibility. It concludes with a call for balanced and inclusive approaches that protect editorial integrity while leveraging the benefits of AI.

INTRODUCTION

We're witnessing a profound shift in how peer review is conducted. AI, once confined to manuscript screening and plagiarism checks, is now infiltrating deeper layers of editorial decision-making[1]. It's fast, efficient, and smart, but is it also trustworthy, fair, and accountable? Will it transform scholarly peer review from triage to reviewer selection, preliminary evaluation, and decision-making regarding ethical violations, bringing unmatched efficiency? While these tools offer efficiency, they may also threaten the scholarly principles of trust, transparency, and human judgment. Even though these were quickly adopted, they also raise concerns about reducing human oversight, compromising editorial independence, and risking academic integrity[2,3]. In scholarly publishing and its peer review process, generative AI tools should only be used as supportive instruments to assist human editors, reviewers, and other stakeholders (e.g., Editorial Board Members and Editors in Chief) in managing manuscript workflows. This article examines AI's dual role as a revolutionary tool and a possible threat in peer review and provides suggestions for ethical and effective use.

RISING ROLE OF AI IN PEER REVIEW

The AI technologies are now experimentally integrated into various stages of the peer review process. It underlies supporting manuscript triage, reviewer selection through Natural Language Processing (NLP), and even automated feedback. Editorial management systems use AI to assist desk rejections, evaluate manuscript scope, and detect plagiarism or image manipulation. A 2024 synthesis of Large Language Model (LLM) tools in peer review found that while they assist efficiency, they fall short in assessing novelty

ISSN: 2789-9446 (Online)
ISSN: 2789-9438 (Print)
https://doi.org/10.21124/tsp.2025.49.53

Received: 20 May, 2025
Accepted: 11 Jul. 2025
Published: 15 Aug. 2025

and scientific rigor[1]. This January, 2025 study surveys 79 papers and a case study of the RIPPLE tool, presenting a six-part framework for AI-enhanced peer assessment, including tasks like assigning assessors, improving individual reviews, and deriving grades[2]. A joint perspective by the European Association of Science Editors (EASE) and the Asian Council of Science Editors (ACSE) further emphasizes that while AI enhances efficiency, it must be integrated responsibly to protect editorial independence and human oversight[3].

These advancements greatly decrease the editorial workload, shorten turnaround times, and improve the scalability of journals. Other notable examples include:

- Manuscript screening tools such as iThenticate and ImageChecker. AI tools can help tailor iThenticate subject specific. For example, in disciplines that frequently use long, standardized nomenclature or chemical formulas, such phrases should not be mistakenly flagged as intentional duplication, particularly in concise versions like abstracts
- Reviewer suggestion engines based on publication history and citation networks
- Language correction tools like Grammarly or Writefull to improve readability before review

Artificial Intelligence tools can assist with technical checks of a manuscript, such as assessing alignment with journal scope or detecting possible plagiarism, including AI generated text and images, either during pre screening or the formal peer review stage. However, all editorial decisions must ultimately be made by a human editor or authorized editorial staff. AI tools should never be used to make editorial decisions, as they cannot assume responsibility or accountability. This is consistent with the principle that AI tools cannot be listed as authors on research papers (as noted by COPE[4]. Besides, AI models trained on existing knowledge cannot reliably assess the scientific merit of a research paper that presents novel findings.

OPPORTUNITIES

Efficiency, accuracy, and scalability: The AI significantly boosts peer review efficiency by processing manuscripts quickly and consistently. It offers a level of consistency and objectivity that can reduce human biases. Algorithms can cover a pre-review of hundreds of submissions in minutes and screening for misconduct, allowing editorial teams to handle growing content volumes without losing speed. Tools like iThenticate and NLP-based triage systems help detect plagiarism, check scope, and suggest reviewers, reducing bias and improving transparency through documented decision trails. Automated tools also improve transparency by recording decision-making processes and making sure ethical checks are consistently applied[5].

For under-resourced journals, especially in the Global South, AI can act as a vital equalizer, offering affordable access to editorial tools that would otherwise be out of reach[6]. Projects like the ReviewFlow system are a concrete example. In a controlled study, it allowed novice reviewers to produce more insightful and comprehensive reviews compared to those working without AI assistance. It provided in-context annotations, novelty checks, and structured outline guidance, enhancing review quality without replacing human insight[7]. The AI tools used in scholarly publishing and peer review should be strictly limited to secure, localized models that meet high standards for information security for example, models operating in isolated environments without internet access or data retention capabilities, such as Mistral or Mixtral. When implemented ethically, AI supports faster, fairer, and more scalable publishing.

THREATS

Ethical risks and oversight challenges: Despite its benefits, AI in peer review introduces several risks:

- **Opaque decision-making:** Many AI tools operate as black boxes. Editors using them may be unable to explain why certain manuscripts are rejected, or reviewers may select a transparency gap that runs counter to academic values. This finding that AI tools remain inherently opaque is highlighted by a systematic review[1]. Moreover, another study reviewed explainable AI methods, concluding that understanding and debugging "black-box" systems remains technically and practically challenging[8]

- **Bias reinforcement:** If AI is trained on historically biased data, it can reinforce inequities in topic selection, author representation, or language fluency. A survey highlights that LLM evaluations often reflect biases, such as favoritism toward English or dominant viewpoints, meaning AI trained on skewed historical data can unintentionally reinforce inequities in topic selection, author diversity, and language fluency[9]
- **Reviewer replacement concerns:** Experiments with AI-generated peer reviews are underway, yet these reviews lack the nuance, critical inquiry, and disciplinary context that human scholars contribute. The ReviewEval framework compared AI-generated reviews against human ones and found that many "AI" reviews miss subtlety, provide less actionable feedback, and require refinement to approach expert-level critique[10]
- **Overdependence on automation:** There's a risk that editors and reviewers, pressed for time, may begin to treat AI-generated outputs as final judgments rather than preliminary guides. This could erode the editorial discernment essential to maintaining scholarly standards. A recent comprehensive review of AI in peer review published in the Journal of Korean Medical Science warns that "if editors and reviewers become too dependent on AI-generated suggestions, the quality of peer review could decline"[1]

EDITORIAL INTEGRITY AT A CROSSROADS

Peer review is a highly intelligent task rather than a mechanical one, it is a scholarly responsibility grounded in trust, expertise, and judgment. While AI can filter and summarize, it cannot assess the originality of a theoretical contribution, the ethical ramifications of a study, or the broader implications for a field. A recent systematic study in Ethics in AI demonstrated that AI-driven peer review systems, when overly trusted, can produce misleading or biased assessments that sideline meaningful human evaluation. This poses a serious risk to scholarly integrity when AI outputs are accepted as final judgments[11]. The Council of Science Editors (CSE) experts argue that while AI automation streamlines workflows, it may erode personalized reviewer feedback and trust if not balanced with transparent governance and human oversight[12].

Editorial decisions require the kind of contextual intelligence and ethical reflexivity that no algorithm can replicate. Another study underscores this concern, warning that over-reliance on AI may erode editorial accountability, marginalize human expertise, and introduce systemic biases if not critically governed[13].

The uncritical adoption of AI risks shifting decision-making power away from trained editors toward unaccountable systems, thereby weakening the integrity of the peer review process. Without clear boundaries and responsible oversight, we risk transforming peer review from a critical scholarly dialogue into a computational transaction.

At present, all major scholarly publishers prohibit editors and reviewers from using or uploading manuscript content to unauthorized AI tools during manuscript evaluation or peer review. Doing so risks compromising the intellectual property, originality, confidentiality, and integrity of unpublished research. Though these policies may evolve as AI technology advances, in order to harness its benefits[14].

BEST PRACTICES FOR RESPONSIBLE AI INTEGRATION

To preserve integrity while harnessing AI's benefits, publishers and editorial boards must establish clear ethical framework[13-16]:

- **Transparency:** Ensure that editors and reviewers understand how AI tools function, what data they use, and what limitations they carry
- **Human oversight:** AI should be an assistant, not an arbiter. Final decisions must rest with accountable human editors

- **Bias mitigation:** Regular audits should identify and correct systemic biases in training data or algorithmic behavior
- **Reviewer empowerment:** Provide reviewers with AI-generated support (e.g., reference checks, summaries) without compromising their autonomy
- **Ethical infrastructure:** Safeguard data privacy, consent, and ensure tools comply with global ethical standards

Journal policies should be updated to reflect these principles, and editorial board training should include AI literacy as a core componsent.

GLOBAL DISPARITIES IN AI ACCESS

The benefits of AI are unevenly distributed. While top-tier journals and publishers can afford sophisticated AI systems, many journals in underdeveloped regions remain excluded from these advancements. If left unchecked, this disparity could exacerbate the North-South divide in academic publishing, marginalizing voices from developing countries. Open-source AI tools, collaborative platforms, and capacity-building initiatives are critical to ensuring that AI becomes a force for inclusion rather than exclusion.

BALANCED PATH FORWARD

The AI is neither a cure-all nor a threat to be feared; it is a tool that, if used responsibly, can strengthen peer review systems. The goal should be to build symbiotic relationships where AI handles the repetitive, mechanical aspects of peer review, allowing human editors and reviewers to focus on ethical judgment, scholarly quality, and strategic insight. If we treat AI as a co-editor rather than a decision-maker, it can relieve administrative burdens, enhance consistency, and expand access, without displacing the human insight that defines rigorous peer evaluation.

What's at stake is not just workflow efficiency, but the trustworthiness of the scholarly record. Editorial integrity must remain the cornerstone of scholarly publishing. The path forward demands balance: automation where appropriate, human judgment where essential. With thoughtful implementation, transparent governance, and inclusive access, AI can become a trusted ally in upholding the credibility and rigor of the peer review process.

CONCLUSION

The integration of AI into peer review presents both transformative opportunities and significant ethical challenges. While AI tools enhance efficiency, scalability, and consistency in manuscript evaluation, their unchecked use risks undermining editorial integrity through opaque decision-making, embedded biases, and over-reliance on automated systems. The key lies in striking a balance, leveraging AI as a supportive tool while preserving the irreplaceable role of human expertise and ethical judgment. By adopting best practices such as transparency in AI usage, bias mitigation strategies, and inclusive global standards, the scholarly publishing community can harness AI's potential without compromising the credibility and fairness of peer review.

SIGNIFICANCE STATEMENT

This perspective underscores the urgent need for a structured and ethical approach to AI adoption in peer review. As AI becomes increasingly embedded in scholarly publishing, its impact on editorial integrity cannot be ignored. The discussion highlights the dual nature of AI, as both a facilitator of efficiency and a potential threat to accountability, and calls for proactive measures to ensure responsible implementation. By fostering collaboration among researchers, editors, and AI developers, the academic community can safeguard the peer review process while embracing technological advancements that enhance, rather than erode, the foundations of scholarly rigor and trust.

REFERENCES

1. Doskaliuk, B., O. Zimba, M. Yessirkepov, I. Klishch and R. Yatsyshyn, 2025. Artificial intelligence in peer review: Enhancing efficiency while preserving integrity. J. Korean Med. Sci., Vol. 40. 10.3346/jkms.2025.40.e92.
2. Topping, K.J., E. Gehringer, H. Khosravi, S. Gudipati, K. Jadhav and S. Susarla, 2025. Enhancing peer assessment with artificial intelligence. Int. J. Educ. Technol. Higher Educ., Vol. 22. 10.1186/s41239-024-00501-1.
3. Sayab, M., M. Sarwar and D. Nicholas, 2019. 5th peer review week: Perspectives from European Association of Science Editors and the Asian Council of Science Editors. Eur. Sci. Ed., 45: 82-83.
4. van Quaquebeke, N., S. Tonidandel and G.C. Banks, 2025. Beyond efficiency: How artificial intelligence (AI) will reshape scientific inquiry and the publication process. Leadersh. Q., Vol. 36. 10.1016/j.leaqua.2025.101895.
5. Bravo, G., F. Grimaldo, E. López-Iñesta, B. Mehmani and F. Squazzoni, 2019. The effect of publishing peer review reports on referee behavior in five scholarly journals. Nat. Commun., Vol. 10. 10.1038/s41467-018-08250-2.
6. Perlin, M.S. T. Imasato and D. Borenstein, 2018. Is predatory publishing a real threat? Evidence from a large database study. Scientometrics, 116: 255-273.
7. Hassija, V., V. Chamola, A. Mahapatra, A. Singal and D. Goel *et al*., 2024. Interpreting black-box models: A review on explainable artificial intelligence. Cognit. Comput., 16: 45-74.
8. Chang, Y., X. Wang, J. Wang, Y. Wu and L. Yang *et al*., 2024. A survey on evaluation of large language models. ACM Trans. Intell. Syst. Technol., Vol. 15. 10.1145/3641289.
9. Maturo, F., A. Porreca and A. Porreca, 2025. The risks of artificial intelligence in research: Ethical and methodological challenges in the peer review process. AI Ethics, 10.1007/s43681-025-00775-9.
10. Sarwar, M., M. Machado, J. Robens, G. Dyke and M. Sayab, 2025. Bridging tradition and technology: Expert insights on the future of innovation in peer review. Sci. Editor, Vol. 48. 10.36591/SE-4801-01.
11. Romeo, G. and D. Conti, 2025. Exploring automation bias in human-AI collaboration: A review and implications for explainable AI. AI Soc., 10.1007/s00146-025-02422-7.
12. Lubowitz, J.H., 2024. Guidelines for the use of generative artificial intelligence tools for biomedical journal authors and reviewers. Arthroscopy: J. Arthroscopic Relat. Surg., 40: 651-652.
13. Kocak, B., M.R. Onur, S.H. Park, P. Baltzer and M. Dietzel, 2025. Ensuring peer review integrity in the era of large language models: A critical stocktaking of challenges, red flags, and recommendations. Eur. J. Radiol. Artif. Intell., Vol. 2. 10.1016/j.ejrai.2025.100018.
14. Ebadi, S., H. Nejadghanbar, A.R. Salman and H. Khosravi, 2025. Exploring the impact of generative AI on peer review: Insights from journal reviewers. J. Acad. Ethics, 10.1007/s10805-025-09604-4.
15. Seghier, M.L., 2025. AI-powered peer review needs human supervision. J. Inf. Commun. Ethics Soc., 23: 104-116.
16. Alnaimat, F., A.R.F. AlSamhori, O. Hamdan, B. Seiil and A.B. Qumar, 2025. Perspectives of artificial intelligence use for in-house ethics checks of journal submissions. J. Korean Med. Sci., Vol. 40. 10.3346/jkms.2025.40.e170.