Review Article

The Role of Artificial Intelligence in the Lifecycle of Scientific Manuscripts: Authoring, Reviewing, and Editorial Selection

Jose L. Domingo[1]

1. School of Medicine, Universitat Rovira i Virgili, Spain

The integration of Large Language Models (LLMs) and Artificial Intelligence (AI) into scientific publishing is accelerating, driven by systemic crises in peer review and academic economic pressures. This manuscript provides a critical three-part analysis: a) AI as a co-authoring tool, balancing its democratizing potential against risks like citation fabrication; b) AI's proficiency in technical review versus its inability to assess novelty; and c) the risk of amplified bias in AI-driven editorial decisions. Recent evidence confirms that citation hallucination remains a persistent threat. Furthermore, 2025 surveys indicate over 50% of researchers utilize AI during peer review, often violating existing policies. This shift occurs within an exploitative model that relies on unpaid labor while charging substantial Article Processing Charges (APCs). To address these challenges, this paper proposes a sustainable, human-centered framework. A model is proposed in which AI is restricted to technical verification and efficiency, while judgments on scientific merit, ethics, and paradigm-shifting research are reserved for compensated human experts. Maintaining scientific credibility requires both the ethical integration of AI and a fundamental reform of the economic structures governing research communication.

Corresponding author: Jose L. Domingo, joseluis.domingo@urv.cat

1. Introduction

The world of academic publishing is currently facing a major crisis: a massive increase in new journals, many papers, a growing shortage of willing peer reviewers, and a business model that profits from researchers' hard work without giving much back[1][2][3][4]. By early 2026, this unsustainable system is being simultaneously challenged and complicated by the rapid integration of Artificial Intelligence (AI). Large Language Models (LLMs) like ChatGPT, Claude, Deep Seek, Grok, Gemini, and many more are now ubiquitous in labs and editorial offices[5]. A comprehensive survey of 1,600 academics in 2025 found that more than 50% used AI tools while peer-reviewing manuscripts[6]. This undisciplined adoption creates a paradox. While AI promises to alleviate burdens like linguistic inequality and technical error detection, it introduces existential threats to data integrity through fabricated citations and the potential automation of scientific conservatism[7][8].

This manuscript presents a holistic examination of the role of AI across the publishing triad: authorship, review, and editorial decision. It expands upon previous critiques of publishing economics[9], which highlighted the ethical contradiction of charging authors article processing charges (APCs) while reviewers work for free. Here, the potential of AI to address logistical crises is juxtaposed with its inherent limitations and dangers. The point is the following: can AI tools truly solve a peer-review crisis that is, at its core, an economic and systemic failure? In this paper, recent empirical studies are analyzed, the complementary and conflicting capabilities of human versus AI reviewers are compared, and a governance framework that prioritizes scientific integrity over mere efficiency is proposed. The analysis draws upon the latest guidelines from the Committee on Publication Ethics[10], which emphasize transparency, human accountability, and mandatory disclosure of AI use across all stages of the publication process. Unlike studies that focus narrowly on AI's technical capabilities in manuscript screening or language editing[11], this analysis situates AI adoption within the broader political economy of scholarly publishing, arguing that technological fixes cannot resolve a crisis rooted in labor exploitation and profit-driven incentives. The main argument is that unless the financial model of publishing is fundamentally changed, using AI could further frustrate scientists and damage the credibility of research.

2. The Roots of the Crisis: Unsustainable Economics and the Exploitation of Peer Reviewers

2.1. The broken social contract of peer review

Peer review operates on a fragile gift economy. Authors receive career-advancing credit for publication, while reviewers receive only intangible academic capital. The Global State of Peer Review report[12] revealed that most researchers felt overburdened[13][14], a situation that has intensified in recent years. Reviewer acceptance rates have significantly decreased, and editors now routinely invite numerous colleagues to secure two to three reviews, creating a massive hidden workload[15][16][17]. This reviewer fatigue is not merely a logistical problem. It is, rather, a symptom of a system that undervalues critical scholarly labor.

2.2. The gold open access paradox and the rise of the "Greedy Publisher"

The transition to digital and Open Access (OA) publishing was announced as a democratizing force. Instead, it has meant a lucrative new business model. As noted in Domingo[9], "*Authors who want to see their paper published as Open Access must pay tremendous amounts (APCs > 3,000-4,000 Euros)... In contrast, interestingly, the reviewers of that paper do not receive a single euro/dollar for their work.*" This creates a fundamental ethical breach; if publication is a paid service, why is its quality control unpaid labor? This model has been perfected by so-called "predatory" journals but has also been eagerly adopted by legacy publishers, blurring ethical lines[9].

Modern science is increasingly treated like a commercial product, where the goal is to move papers through the system as quickly as possible rather than carefully maintaining the quality of knowledge. This profit-driven atmosphere is the real reason AI is being introduced to the peer-review process. Supporters claim AI will solve the shortage of human reviewers, but this solution hides a deeper problem: the system's reliance on the free labor and generosity of experts. By using AI to handle the massive influx of papers, we might simply be treating the symptoms of the volume crisis while allowing the actual disease, the exploitation of researchers, to become even more deeply rooted.

3. AI in Manuscript Preparation

3.1. Democratization and enhanced efficiency

AI writing assistants offer tangible benefits, particularly in promoting linguistic equity. For non-native English speakers, these tools can reduce the linguistic tax, improving readability, grammar, and adherence to academic conventions, which may positively influence editors′ first impressions[18]. More than just editing, AI can help structure complex arguments, draft technical methodology sections from structured data, and assist in initial literature synthesis, potentially increasing researcher productivity[19].

3.2. The hallucination epidemic and the erosion of verifiability

The most severe and well-documented threat from the use of AI to write manuscripts is citation hallucination; that is, the generation of plausible but entirely fictitious references. This is not an occasional mistake but a fundamental feature of LLMs, which generate text based on statistical patterns. Recent (2023–2025) empirical evidence reveals the alarming scope of this problem across various models and disciplines (Table 1). Various studies have consistently demonstrated that while newer models show incremental improvements, fabrication and error rates remain unacceptably high for scientific work. Walters and Wilder[8] reported alarming rates. While GPT-4 improved upon GPT-3.5's 55% fabrication rate, it still hallucinated 18% of citations. A comprehensive 2025 study found that ChatGPT GPT-4o fabricated approximately 20% of academic citations, introducing errors in 45% of real references. Among fabricated citations that included DOIs, 64% linked to real but completely unrelated papers, making detection extremely difficult without careful verification[20][21]. A recent analysis showed that ChatGPT fabricated 20% of academic citations and introduced errors in 45% of real references, posing obvious risks for researchers[20][21]. Cheng et al.[22] reported that 38% of ChatGPT-generated references contained incorrect or fabricated DOIs. Only a low percentage (7%) corresponded to fully accurate DOIs. Similarly, in a study conducted by Kim et al.[23] using four LLMs, it was found that they frequently fabricated DOIs in academic citations, which was especially notable for lower-income countries (e.g., India and Bangladesh). Error rates even exceeded 80%, while hallucinations were found to be most common in recent publications, varying by model and highlighting geographic and economic biases. A 2025 cross-model study revealed that only 26.5% of AI-generated references were entirely correct, with nearly 40% being erroneous or completely fabricated, remarking on the persistent unreliability across different language models[24]. AI errors increase significantly in specialized fields, where finding an expert to check the facts is most difficult[25].

The reliability of published research is under serious threat as fake references become increasingly common. Buchanan et al.[26] showed that these artificial citations are deceptively realistic, often attributing fake findings to actual researchers within credible-looking journal templates. Despite advances in Retrieval-Augmented Generation (RAG) systems aimed at reducing hallucinations, these tools face fundamental challenges, including source conflicts and data poisoning, and they cannot substitute for a true understanding of scientific literature. This places the entire responsibility for fraud on authors, reviewers, and editors, a task that is becoming nearly impossible in an age of data saturation. Although the adoption of RAG tools offers a partial shield, improving accuracy, it cannot replace true comprehension, leaving the fundamental risk of AI-generated misinformation unresolved.

Study	Year	AI Model(s) Tested	Sample Size / Methodology	Key Finding: Fabrication / Error Rate	Additional Notes
Walters and Wilder[8]	2023	GPT-3.5, GPT-4	Comparative analysis	GPT-3.5: 55% fabrication; GPT-4: 18% fabrication	Early evidence of improvement in newer models, but persistent hallucination risk
Linardon et al.[20]	2025	ChatGPT (GPT-4o)	Topic-specific mental health citations	20% fabrication; 45% contain errors in real references	Hallucination rates spike for niche topics where expert verification is difficult
Cabezas-Clavijo and Sidorenko-Bautista[24]	2025	Various LLMs (ChatGPT, Claude, Copilot, Gemini, Perplexity, Grok, DeepSeek, others)	400 references across 5 academic disciplines	26.5% entirely correct; 39.8% erroneous or completely fabricated; 33.8% real but with partial errors	Cross-model comparison; Grok and DeepSeek performed best; Copilot, Perplexity, Claude had highest hallucination rates
Cheng et al.[22]	2025	ChatGPT	DOI-specific analysis	38% of generated references contained incorrect or fabricated DOIs; only 7% fully accurate DOIs	DOI fabrication is particularly deceptive as it appears authentic
Kim et al.[23]	2025	4 LLMs (unspecified)	Geographic bias analysis (India, Bangladesh, other low-income countries)	Error rates exceeded 80% for some models in certain regions; geographic and economic biases evident	Hallucinations most common in recent publications; bias varies by model and geographic context
Buchanan et al.[26]	2024	ChatGPT	Detection of fabricated citations	Fabricated citations attributable to real	Demonstrates deceptive realism of hallucinations;

Study	Year	AI Model(s) Tested	Sample Size / Methodology	Key Finding: Fabrication / Error Rate	Additional Notes
			in economics	researchers in credible-looking journal templates	difficult to detect without verification

Table 1. Summary of empirical evidence on AI citation fabrication rates and error patterns (2023–2025)

4. Human Reviewers vs. AI Chatbot Reviewers

4.1. AI as a technical auditor: strengths in verification

Where AI shows genuine promise is in augmenting the technical aspect of peer review. Automated systems can perform with superhuman consistency in specific, rule-based tasks:

a. Statistical and Mathematical Check: AI can detect inconsistencies in p-values, degrees of freedom, and mathematical derivations with over 90% accuracy[7].
b. Image Manipulation Detection: AI algorithms can screen for duplicated, spliced, or otherwise manipulated images more thoroughly than the human eye.
c. Plagiarism and Text Recycling: LLMs can cross-check text against massive databases more efficiently than standard software.
d. Compliance Checking: Ensuring adherence to reporting guidelines (e.g., CONSORT, PRISMA) and ethical statements.

These capabilities address known weaknesses in human review, where statistical errors are frequently undetected[27]. AI can act as a first-line technical auditor, freeing human reviewers from tedious verification work.

4.2. The irreplaceable human roles: judgment, novelty, and context

This is where the comparison becomes stark and highlights the non-fungible value of the human reviewer. The core responsibilities of peer review extend far beyond technical checking:

doi.org/10.32388/9AOB4V

a. Judging Novelty and Significance: Can the work change thinking in the field? Because contemporary AI systems are trained predominantly on existing scientific literature, they tend to reproduce established patterns of thought, which may inadvertently discourage submissions that challenge prevailing paradigms or pursue high-risk, high-impact avenues of inquiry[28]. It lacks the intuition to recognize a brilliant flaw or a new revolutionary idea.
b. Contextual and Ethical Reasoning: Does the work fit into the broader landscape? Are its ethical implications sound? Human reviewers draw on deep domain knowledge, understanding of societal impact, and professional ethics.
c. Constructive Dialogue: Peer review is an academic dialogue. A human reviewer can suggest alternative experiments, propose collaborative connections, or engage in a dialectic that improves the science. AI generates feedback based on patterns, not understanding or creativity.

As Domingo[9] emphasized, the reviewer is *"the basic point of the chain of quality"*, which is unquestionable. Replacing human judgment with computer-generated scores would change peer review from a careful evaluation into a routine filtering process. This shift could strengthen existing prejudices and prevent new ideas from emerging. The fear is not that AI will be wrong but that it will be consistently and invisibly biased toward the conventional.

4.3. The hybrid review model

A sustainable future might lie in a hybrid model that benefits from the strengths of both:

1. Role of AI: Technical pre-screening. Perform initial checks for statistics, plagiarism, image integrity, and formatting. Flag potential issues for human attention. Summarize manuscript content for editor triage.
2. Human's Role: Evaluate novelty, significance, creativity, and broader impact. Interpret AI-generated flags with contextual wisdom. Make final recommendations on acceptance/rejection. Engage in constructive dialogue with authors.

This model respects humans as the referees of science while using AI as a powerful, precise tool. Figure 1 illustrates this complementary framework, showing how AI functions as a technical auditor, whereas humans retain authority over scientific judgment across the entire manuscript lifecycle.

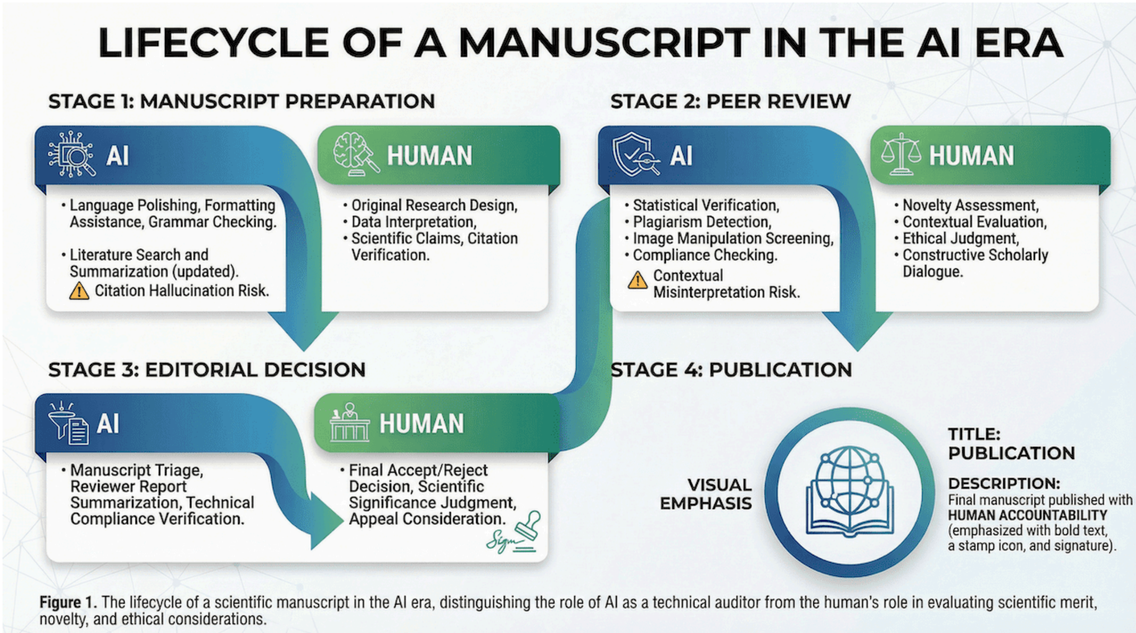

Figure 1. The lifecycle of a scientific manuscript in the AI era, distinguishing the role of AI as a technical auditor from the human's role in evaluating scientific merit, novelty, and ethical considerations.

Figure 1. The lifecycle of a scientific manuscript in the AI era, distinguishing the role of AI as a technical auditor from the human's role in evaluating scientific merit, novelty, and ethical considerations.

4.4. Comparative analysis: a structured overview

To synthesize the arguments developed in Sections 4.1–4.3, Table 2 provides a structured comparison of the main capabilities, limitations, and implications of human reviewers versus AI (LLM) reviewers across key assessment dimensions in scientific peer review. This overview highlights that, while AI excels in speed and technical verification, it remains fundamentally inadequate for evaluating novelty, contextual, and ethical issues. This also ensures the delivery of truly constructive academic feedback, which emphasizes the essential role of responsible human judgment.

doi.org/10.32388/9AOB4V

Assessment Dimension	Human Reviewer	AI (LLM) Reviewer	Implications for Peer Review
Technical Verification	Prone to fatigue, oversight, and variable standards. High expertise but inconsistent.	Exceptional speed, consistency, and recall for defined tasks (statistics, plagiarism detection).	AI is well suited for initial objective technical screening, reducing human workload and error rates.
Novelty and Significance	Core strength. Uses intuition, experience, and field foresight to judge paradigm-shifting potential.	Biased toward incremental science that conforms to patterns in training data.	AI cannot reliably identify high-risk, transformative research; exclusive reliance may homogenize science.
Contextual and Ethical Reasoning	Integrates societal impact, ethical nuance, and interdisciplinary context.	Limited to text-based pattern recognition; lacks genuine ethical understanding or situational awareness.	Human oversight remains essential for evaluating ethical compliance, dual-use concerns, and societal implications.
Constructive Feedback	Provides creative suggestions, proposes new experiments, and engages in scholarly dialogue.	Produces structured but often generic feedback lacking original insight.	The iterative and developmental nature of peer review depends on human intellectual contribution.
Bias	Subject to conscious and unconscious biases related to affiliation, gender, nationality, or academic networks.	Inherits and may amplify systemic biases embedded in training data.	Both systems are bias-prone. However, AI bias is systemic, opaque, and highly scalable.
Economic Model	Relies on unpaid academic labor; sustainability is increasingly compromised by reviewer fatigue.	Operates at near-zero marginal cost once deployed.	Economic incentives favor AI adoption but risk devaluing human expertise and undermining review quality.
Accountability	Professionally and ethically accountable for evaluations and recommendations.	Lacks legal, ethical, or professional accountability; operates as a non-transparent system.	Responsibility for the integrity of the scientific record must remain with accountable human reviewers.

Table 2. Comparative capabilities of human reviewers vs. AI (LLM) reviewers in scientific peer review

5. To Pay the Human or Replace with AI

The advancement of capable AI-based peer-review systems places publishers at a decisive economic juncture. They must determine whether to continue relying on uncompensated expert reviewers, introduce mechanisms for fair remuneration, or shift toward inexpensive algorithmic review processes. This represents more than a simple operational shift. It requires a deep ethical and philosophical discussion about the fundamental value of human expertise in a digital world. Table 3 presents four distinct scenarios for the future of peer review economics, contrasting the status quo, full AI automation, a fair hybrid model, and community-led reclamation. This structured overview makes explicit how different combinations of AI use and reviewer compensation may shape costs, incentives, and ultimately, the likely outcomes for the quality and credibility of science.

Scenario	Model	Pros	Cons	Likely Outcome for Science
Status Quo Exploitation	Continued reliance on unpaid human reviewers while maintaining high article processing charges (APCs).	Low direct financial cost for publishers; preserves elements of human judgment.	Increasing reviewer fatigue; ethical concerns; progressive decline in review quality and system sustainability.	Systemic collapse. Progressive degradation of peer review quality leading to erosion of journal credibility.
Full AI Automation	Replacement of human reviewers with fully automated AI-based review systems, while APCs remain unchanged.	Near-zero marginal cost per review; extremely rapid editorial turnaround.	Inability to assess true novelty; amplification of algorithmic bias; absence of accountability.	Homogenized stagnation: A technically consistent but incremental and uninspired scientific literature.
Fair Hybrid Model	AI-based tools conduct technical screening, complemented by compensated human reviewers for intellectual judgment.	Sustainable labor model; combines AI efficiency with human expertise; preserves scientific creativity.	Higher operational costs for publishers; requires structural reform of current publishing economics.	Sustainable integrity that maintains scientific quality, supports innovation, and ethically values expert labor.
Community Reclamation	Scientific- or society-owned journals with reviewer compensation funded through transparent, non-profit budgets.	Alignment of incentives with academic community values; elimination of profit-driven distortions.	Requires coordinated collective action and development of new editorial infrastructure.	Renewed trust: A scholar-led publishing ecosystem prioritizing quality, openness, and knowledge dissemination.

Table 3. Scenarios for the future of peer review economics

5.1. The case for compensating human reviewers

The argument for payment is grounded in fairness, sustainability, and quality preservation.

Fairness in a for-profit system: As Domingo[9] argues, the current model is ethically unjustifiable. If the publication process generates important benefits, its most critical quality-control component deserves financial recognition.

Sustainable labor model: The gift economy is collapsing. Payment, whether direct stipends (e.g., €200–€500 per review, which must be absorbed by publisher profits, not with higher APCs) or review credits redeemable for APCs, creates a sustainable incentive structure. A 2025 pilot by the Journal of Clinical Medicine offering modest honoraria increased reviewer acceptance by 40%, with a 25% decrease in review time[7].

Preserving quality and commitment: Compensation professionalizes the role. Paid reviewers are more likely to accept invitations within their expertise, meet deadlines, and provide thorough, constructive feedback. It formally recognizes reviewing as essential scholarly work, not a secondary activity[15][29].

Counteracting commercialization: Fair compensation is a step toward rebalancing power. It acknowledges that the research community provides the core product, challenging the role of publishers as mere intermediaries.

5.2. The publisher temptation: "zero-cost"

For profit-driven publishers, the economic appeal of AI is irresistible: a one-time software investment that eliminates the recurring costs (in time and effort) of managing a volunteer reviewer pool. AI promises instant, 24/7 review, reducing time-to-first-decision and eliminating the reviewer chase. This maximizes throughput, a key metric in high-volume, APC-driven models. Marketing AI as an unbiased tool (despite evidence to the contrary) could be used to deflect criticism about editorial decisions, creating a tag of technical neutrality. In a competitive market, the first major publisher to fully automate peer review for certain article types could undercut others on cost and speed, forcing widespread adoption regardless of quality concerns[30].

5.3. Replacing human reviewers with AI

The replacement of human reviewers with AI to save costs would be a catastrophic trade-off, sacrificing scientific progress for corporate efficiency. Scientific progress relies on an essential basis of mutual trust within the scientific community, based on the belief that human experts conduct thorough and impartial evaluations. The adoption of automated peer-review technologies risks undermining this implicit social agreement, as opaque algorithmic decision processes may significantly reduce confidence in the legitimacy and integrity of the scientific record[31][32]. The community trusts that peers have fairly evaluated their work. Replacing them with inscrutable algorithms would fundamentally break this contract, delegitimizing published science. AI systems remain unable to reliably identify true novelty, contextual significance, or ethical complexity, which makes their exclusive use in peer review a direct threat to scientific progress. A fully automated system would create an incremental science filter, publishing only those papers that confirm existing paradigms. Innovative work from unknown labs or on emerging topics would be systematically rejected. Moreover, who is responsible when an AI reviewer misses a fatal flaw, endorses plagiarized work, or rejects a brilliant idea? The publisher would hide behind the black box, leaving authors with no recourse. Replacing human reviewers with AI merely externalizes the cost of quality onto the scientific community, which will suffer from a stagnant, unreliable literature. The cost savings achieved come at the expense of an investment in the integrity of the scientific record itself[33].

6. Algorithmic Bias in Editorial Decisions

Editors are increasingly using AI-driven predictive analytics to select manuscripts, estimating scientific interest based on historical citation data, author prestige, and keyword trends[34]. This poses a profound danger of algorithmic gatekeeping. An AI trained on past successful papers will inherently favor research that resembles past success, systematically disadvantaging interdisciplinary work, studies from less prestigious institutions, and truly novel paradigms. It creates a vicious, self-reinforcing cycle that could homogenize scientific output.

As noted in Domingo[9], the over-reliance on journal-level metrics (Impact Factor, Cite Score, etc.) already distorts evaluation. AI-powered interest predictors risk automating and exacerbating this bias, making it more opaque and harder to challenge. The definition of interesting science must remain a dynamic, human social construct, responsive to new challenges and perspectives[35].

7. A Framework for Ethical Integration and Systemic Reform

7.1. Governance, transparency, and mandatory disclosure

Ad hoc policies are insufficient. According to the guidelines from COPE[10], authors must detail how AI was used (e.g., for language polishing, for literature search, for generating figures, etc.), with vague statements being unacceptable. In addition, all final publication decisions must have explicit human approval, while AI output cannot be the sole basis for rejection. Moreover, journals using AI for review or triage must disclose this to authors and provide a mechanism to appeal or request human assessment. Moving toward open review reports can increase accountability for both human and AI-assisted reviews. Major publishers have now established comprehensive AI policies that universally prohibit AI authorship and mandate transparency regarding AI tool usage, while emphasizing that authors retain full responsibility for all content accuracy and integrity.

7.2. Addressing the economic problems

If AI integration proceeds without economic reform, it will merely make a profitable but exploitative system more efficient. Incentives must be redesigned. Thus, if APCs are charged, a significant portion must fund reviewer honoraria or systematic rewards (e.g., review credits redeemable for APCs at partner journals)[29]. In turn, the scientific community should promote and progressively adopt society-owned, diamond open-access, and cooperative publishing models[9]. In any case, the formal recognition of peer review in tenure, promotion, and funding decisions is essential.

7.3. Training the next generation of scientists

Academic institutions and research centers must implement compulsory AI literacy programs that provide scientists with a comprehensive understanding of machine learning capabilities, inherent constraints, and ethical vulnerabilities. A key part of this training must be the development of rigorous verification protocols. Specifically, researchers require the skills to systematically audit AI-generated technical content and validate the authenticity of provided citations. Training programs should specifically address the identification of hallucinated citations, proper citation verification workflows using databases (CrossRef, PubMed, Scopus, etc.), and the responsible use of RAG-enhanced tools while recognizing their limitations. Furthermore, establishing a robust ethical framework is fundamental to distinguishing between legitimate computational assistance and academic dishonesty.

8. Conclusions

The integration of AI into scientific publishing is not a simple question of adoption. It is a complex negotiation at a time of systemic crisis. AI offers powerful tools to enhance technical rigor and manage scale, particularly in addressing the overwhelming burden contributing to reviewer fatigue. However, it simultaneously introduces severe risks to the veracity of the scientific record through hallucinations and to the diversity of scientific thought through algorithmic bias.

The economic crossroads is clear. The path of least resistance for publishers, replacing unpaid human labor with unpaid machine labor, is a direct threat to the progress of science. The only ethically and scientifically defensible path is Scenario 3: The Fair Hybrid Model (Table 3). This requires viewing AI not as a replacement, but as a tool that handles technically arduous work, thereby making the valuable time of human experts more efficient and sustainable. It must be paired with fair compensation for those experts, reforming the economic model that has brought us to this crisis point. The goal of fully automating research must be rejected; instead, we must adopt a human-centered approach, where AI serves only to assist and enhance human work.

In that model, AI handles the technical, the repetitive, and the scalable, while human experts (adequately valued and compensated) retain authority over judgments of significance, novelty, ethics, and creativity. The future of a vibrant, credible scientific literature depends on choosing a model that invests in human expertise, not one that seeks to reduce or eliminate costs.

Statements and Declarations

Conflicts of Interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

There is no funding source.

Acknowledgements

Artificial Intelligence (Skywork AI Smart Slide Tool) was used to assist in the graphic design of Figure 1.

References

1. ^Clark AD, Myers TC, Steury TD, Krzton A, Yanes J, Barber A, Barry J, Barua S, Eaton K, Gosavi D, Nance R, Pervaiz Z, Ugochukwu C, Hartman P, Stevison LS (2024). *"Does It Pay to Pay? A Comparison of the Benefits of Open-Access Publishing Across Various Sub-Fields in Biology." PeerJ. 12:e16824. doi:10.7717/peerj.16824.*
2. ^Clark LM, Wang DK, Adkins BD, Fitzhugh VA, Walker PD, Khan SS, Fadare O, Stephens LD, Coogan AC, Booth GS, Jacobs JW (2024). *"Paying to Publish: A Cross-Sectional Analysis of Article Processing Charges and Journal Characteristics Among 87 Pathology Journals." Acad Pathol. 11(4):100153. doi:10.1016/j.acpath.2024.100153.*
3. ^Horta H, Jung J (2024). *"The Crisis of Peer Review: Part of the Evolution of Science." High Educ Q. 78(4):e12511. doi:10.1111/hequ.12511.*
4. ^Routledge D, Pariente N (2025). *"On Improving the Sustainability of Peer Review." PLoS Biol. 23(3):e3003127. doi:10.1371/journal.pbio.3003127.*
5. ^Hosseini M, Eisenman D, Riddle J, Pyle S, Peters A, Cobb N, Holmes AK (2026). *"Guidelines Needed for the Use of AI in the Preparation or Review of IRB, IBC, and IACUC Applications." Account Res. 2612564. doi:10.1080/08989621.2025.2612564.*
6. ^Naddaf M (2025). *"More Than Half of Researchers Now Use AI for Peer Review - Often Against Guidance." Nature. doi:10.1038/d41586-025-04066-5.*
7. a, b, cDoskaliuk B, Zimba O, Yessirkepov M, Klishch I, Yatsyshyn R (2025). *"Artificial Intelligence in Peer Review: Enhancing Efficiency While Preserving Integrity." J Korean Med Sci. 40(7):e92. doi:10.3346/jkms.2025.40.e92.*
8. a, b, cWalters WH, Wilder EI (2023). *"Fabrication and Errors in the Bibliographic Citations Generated by ChatGPT." Sci Rep. 13(1):14045. doi:10.1038/s41598-023-41032-5.*
9. a, b, c, d, e, f, gDomingo JL (2024). *"To Publish Scientific Journals: For Some, the Big Business of the Century." Qeios. 6(2):YTFEF2.2. doi:10.32388/YTFEF2.2.*
10. a, bCOPE, Committee on Publication Ethics (2025). *"Position Statement on Authorship and AI Tools." Committee on Publication Ethics. https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools (accessed December 20, 2026).*
11. ^Carobene A, Padoan A, Cabitza F, Banfi G, Plebani M (2023). *"Rising Adoption of Artificial Intelligence in Scientific Publishing: Evaluating the Role, Risks, and Ethical Implications in Paper Drafting and Review Process." Clin Chem Lab Med. 62(5):835–843. doi:10.1515/cclm-2023-1136.*

12. ^Publons (2018). "Global State of Peer Review." Clarivate Analytics. https://clarivate.com/academia-government/lp/global-state-of-peer-review-report/ (accessed December 15, 2025).
13. ^Breuning M, Backstrom J, Brannon J, Gross BI, Widmeier M (2015). "Reviewer Fatigue? Why Scholars Decline to Review their Peers' Work." PS: Polit Sci Polit. 48(4):595–600. doi:10.1017/S1049096515000827.
14. ^Fox CW, Albert AY, Vines TH (2017). "Recruitment of Reviewers Is Becoming Harder at Some Journals: A Test of the Influence of Reviewer Fatigue at Six Journals in Ecology and Evolution." Res Int Peer Rev. 2(1):3. doi:10.1186/s41073-017-0027-x.
15. a, bAczel B, Barwich AS, Diekman AB, Fishbach A, Goldstone RL, Gomez P, Gundersen OE, von Hippel PT, Holcombe AO, Lewandowsky S, Nozari N, Pestilli F, Ioannidis JPA (2025). "The Present and Future of Peer Review: Ideas, Interventions, and Evidence." Proc Natl Acad Sci U S A. 122(5):e2401232121. doi:10.1073/pnas.2401232121.
16. ^Bauchner H, Rivara FP (2024). "Use of Artificial Intelligence and the Future of Peer Review." Health Aff Sch. 2(5):qxae058. doi:10.1093/haschl/qxae058.
17. ^Malički M, Mehmani B (2024). "Structured Peer Review: Pilot Results from 23 Elsevier Journals." PeerJ. 12:e17514. doi:10.7717/peerj.17514.
18. ^Khalifa AA, Ibrahim MA (2024). "Artificial Intelligence (AI) and ChatGPT Involvement in Scientific and Medical Writing, a New Concern for Researchers. A Scoping Review." Arab Gulf J Sci Res. 42(4):1770–1787. doi:10.1108/AGJSR-09-2023-0423.
19. ^Liang W, Izzo Z, Zhang Y, Lepp H, Cao H, Zhao X, Chen L, Ye H, Liu S, Huang Z, McFarland DA, Zou JY (2024). "Monitoring AI-Modified Content at Scale: A Case Study on the Impact of ChatGPT on AI Conference Peer Reviews." arXiv preprint. doi:10.48550/arXiv.2403.07183.
20. a, b, cLinardon J, Jarman HK, McClure Z, Anderson C, Liu C, Messer M (2025). "Influence of Topic Familiarity and Prompt Specificity on Citation Fabrication in Mental Health Research Using Large Language Models: Experimental Study." JMIR Ment Health. 12:e80371. doi:10.2196/80371.
21. a, bStudyFinds Analysis (2025). "ChatGPT's Hallucination Problem: Study Finds More Than Half Of AI's References Are Fabricated or Contain Errors in Model GPT-4o." StudyFinds Analysis. https://studyfinds.org/chatgpts-hallucination-problem-fabricated-references/ (accessed December 19, 2025).
22. a, bCheng A, Calhoun A, Reedy G (2025). "Artificial Intelligence-Assisted Academic Writing: Recommendations for Ethical Use." Adv Simul (Lond). 10(1):22. doi:10.1186/s41077-025-00350-6.
23. a, bKim E, Kipchumba F, Min S (2025). "Geographic Variation in LLM DOI Fabrication: Cross-Country Analysis of Citation Accuracy Across Four Large Language Models." Publications. 13(4):49. doi:10.3390/publication

s13040049.

24. ^a, b^*Cabezas-Clavijo Á, Sidorenko-Bautista P (2025). "Assessing the Performance of 8 AI Chatbots in Bibliographic Reference Retrieval: Grok and DeepSeek Outperform ChatGPT, but None Are Fully Accurate." arXiv (preprint). doi:10.48550/arXiv.2505.18059.*
25. ^*Chelli M, Descamps J, Lavoué V, Trojani C, Azar M, Deckert M, Raynier JL, Clowez G, Boileau P, Ruetsch-Chelli C (2024). "Hallucination Rates and Reference Accuracy of ChatGPT and Bard for Systematic Reviews: Comparative Analysis." J Med Internet Res. 26:e53164. doi:10.2196/53164.*
26. ^a, b^*Buchanan J, Hill S, Shapoval O (2024). "ChatGPT Hallucinates Non-existent Citations: Evidence from Economics." Am Econ. 69(1):80–87. doi:10.1177/05694345231218454.*
27. ^*Lee J, Lee J, Yoo JJ (2025). "The Role of Large Language Models in the Peer-Review Process: Opportunities and Challenges for Medical Journal Reviewers and Editors." J Educ Eval Health Prof. 22:4. doi:10.3352/jeehp.2025.22.4.*
28. ^*Maddi A, Sapinho D (2022). "Article Processing Charges, Altmetrics and Citation Impact: Is There an Economic Rationale?" Scientometrics. 127(12):7351–7368. doi:10.1007/s11192-022-04284-y.*
29. ^a, b^*Cheah PY, Piasecki J (2022). "Should Peer Reviewers Be Paid to Review Academic Papers?" Lancet. 399(10335):1601. doi:10.1016/S0140-6736(21)02804-X.*
30. ^*Drozdz JA, Ladomery MR (2024). "The Peer Review Process: Past, Present, and Future." Br J Biomed Sci. 81:12054. doi:10.3389/bjbs.2024.12054.*
31. ^*Kemal Ö (2025). "Artificial Intelligence in Peer Review: Ethical Risks and Practical Limits." Turk Arch Otorhinolaryngol. 63(3):108–109. doi:10.4274/tao.2025.2025-8-12.*
32. ^*Schintler LA, McNeely CL, Witte J (2023). "A Critical Examination of the Ethics of AI-Mediated Peer Review." arXiv Preprint. doi:10.48550/arXiv.2309.12356.*
33. ^*Resnik DB, Hosseini M (2025). "The Ethics of Using Artificial Intelligence in Scientific Research: New Guidance Needed for a New Tool." AI Ethics. 5(2):1499–1521. doi:10.1007/s43681-024-00493-8.*
34. ^*Peddi S, Manoharan G (2026). "Ethical Guidelines for the Thoughtful Implementation of AI in Higher Education." EthAIca. 5:409. doi:10.56294/ai2026409.*
35. ^*Park W, Cullinane A, Gandolfi H, Alameh S, Mesci G (2023). "Innovations, Challenges and Future Directions in Nature of Science Research: Reflections from Early Career Academics." Res Sci Educ. 1–22. doi:10.1007/s11165-023-10102-z.*

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.

doi.org/10.32388/9AOB4V