Editorial

Disclosing AI Use in Scientific Publications: Between Formal Transparency and the Methodological Void

La declaración del uso de IA en publicaciones científicas: entre la transparencia formal y el vacío metodológico

Reis Ríos[1]

How to Cite

Ríos, R. (2026). Disclosing AI Use in Scientific Publications: Between Formal Transparency and the Methodological Void. *Trilogía Ciencia Tecnología Sociedad, 18*(38), art. 4038. https://doi.org/10.22430/21457778.4038

Within the "police regimes" that reproduce the normative structure of science (Merton, 1977), scientific journals serve as epistemic institutions that certify valid knowledge, validate appropriate research practices, and define the minimum standards for manuscripts to become part of the scientific record. In this sense, Artificial Intelligence (AI) disclosure statements offer an observable indicator of editorial rigor. According to Queiroz et al. (2025), 70.9% of dentistry journals indexed in Web of Science have established policies governing the use of AI in manuscript preparation, and all require some form of explicit disclosure. However, there is no standardized guidance on what should be disclosed, where the disclosure should appear, how much detail should be provided, or what methodological criteria should be applied. Ganjavi et al. (2024) conducted a bibliometric analysis of the author guidelines of the top 100 academic publishers and journals across multiple disciplines, confirming that this regulatory fragmentation is not confined to a single field. Only 17% of leading publishers provided guidance on generative AI use, and such guidance varied substantially in both specificity and scope.

That is the problem that this editorial seeks to describe. In the current scientific publishing ecosystem, disclosing the use of AI has become a formal requirement rather than a guarantee of methodological rigor. Disclosure is a necessary condition for transparency, but not a sufficient condition for rigor. It does not specify when AI was used in the research process, for what purpose, under what level of human oversight, or to what extent its documented limitations were considered. Knowing that a researcher employed a Large Language Model (LLM) does not indicate whether it was

[1] Universidad de La Sabana, Chía, Colombia, reisri@unisabana.edu.co

CC BY NC SA OPEN ACCESS

used appropriately. Without methodological criteria, transparency is, at best, incomplete; at worst, it creates a false sense of control over a problem whose scope has yet to be clearly delineated.

Beginning in November 2022, the widespread adoption of tools such as ChatGPT prompted scientific journals to require authors to disclose whether they had used AI, identify the tools employed, and describe the purpose of their use. Organizations such as the International Committee of Medical Journal Editors (ICMJE) and the Committee on Publication Ethics (COPE) established that AI cannot be listed as an author because it cannot assume intellectual or legal responsibility for the content it generates (De Cassai et al., 2026; Vahedian-Azimi et al., 2026). However, these guidelines are not binding on journals outside their affiliated networks, making compliance a voluntary editorial decision without formal consequences for nonadherence. Carlson et al. (2026) analyzed a sample of neurosurgery journals and found that, although all required AI disclosure, the mandatory content, designated section, and restrictions on use varied considerably. Notably, the Delphi RULE-AP consensus explicitly recommends that its own guidance be continuously updated (De Cassai et al., 2026), thereby acknowledging its provisional nature despite representing one of the most comprehensive frameworks.

Consequently, the lack of consensus in editorial policies undermines the consistent preservation of scientific integrity. Although current guidelines on AI use are relatively clear, their implementation faces a challenge that regulatory frameworks have yet to resolve. Disclosure establishes that AI was used, but not whether that use was methodologically sound. A researcher may disclose using an LLM to retrieve scientific information without verifying whether the references actually exist or whether the cited data accurately reflect the source content. Similarly, another may report using AI solely to improve writing without reviewing the logical structure of the arguments. In both cases, the disclosure is formally correct, yet the methodological rigor of that use remains unevaluated.

Even more concerning is the fact that the practice of disclosure has little to do with scientific transparency. An international study of authors publishing in orthopedic journals found that only 43.1% consistently disclosed AI use, while 66.7% feared being perceived negatively when doing so (de Geofroy et al., 2026). Among doctoral students, 85.2% chose not to disclose AI use, arguing that its contribution had been minimal or that disclosure might undermine their credibility with evaluation committees and prospective employers (Hoomanfard & Shamsi, 2025). In these cases, nondisclosure is driven not by methodological considerations but by perceived reputational risk. Moreover, given that current editorial policies lack verifiable mechanisms for penalizing nondisclosure, many researchers regard silence as the least costly option. This situation is further compounded by the poor reliability of AI-generated text detectors (Epelde, 2026; Hosseini et al., 2023), which have disproportionately disadvantaged authors whose first language is not English.

To understand why disclosure alone is insufficient as a guarantee of rigor, it is necessary to examine the evidence supporting AI use at each stage of the research process. A review of the available empirical literature reveals a marked concentration of research in only a few stages. The most visible stages of the research process—literature review, writing, and data analysis—account for most published studies, whereas methodologically sensitive stages, such as instrument design, data collection, and interpretation of findings, have received far less attention. This distribution constitutes

an epistemological warning. If a researcher discloses AI use at stages for which empirical evidence remains limited, it cannot be determined with sufficient certainty whether AI was used appropriately, what systematic errors may have occurred, or what level of human oversight was required. A similar situation exists even in the most extensively studied stages. In literature review, for example, the outputs of retrieval-augmented systems depend heavily on the balance of the evaluation *corpus*, and validation protocols are not yet standardized (Scherbakov et al., 2025). Likewise, in scientific writing, documented risks include inadvertent plagiarism, the generation of nonexistent references, and challenges to authorship criteria (Gutiérrez-Cirlos et al., 2023; Owusu Sarfo, 2023; Salvagno et al., 2023).

In conclusion, this is not merely a matter of authors' willingness to disclose AI use, because journals, as epistemic institutions, cannot limit themselves to recording a fact. Rather, the fundamental issue is that current editorial policies provide no effective methodological criteria for evaluating AI use. If journals are satisfied with requiring only a formal disclosure statement, they overlook the central question: *Is the reported use of AI methodologically valid?* Empirical evidence on the appropriate use of AI varies dramatically across the stages of the research process. In this context, failing to move beyond the minimum threshold of disclosure is an editorial decision with direct consequences for the quality of the scientific knowledge ultimately published.

REFERENCES

Carlson, B., Laffaye, T., Gray, L., Bathini, A. R., Dubey, A., & Patra, D. P. (2026). The use of artificial intelligence in neurosurgical manuscript writing: Journal specific policies and their implementation. *Journal of Clinical Neuroscience*, *144*, art. 111819. https://doi.org/10.1016/j.jocn.2025.111819

De Cassai, A., Dost, B., Augoustides, J. G., Azamfirei, L., Alanoğlu, Z., Azi, L. M., Calvache, J. A., Cerny, V., De Hert, S., Eldawlatly, A., Farber, M. K., Sobreira-Fernandes, D., Fettiplace, M. R., Galante, D., Garg, R., Goldstein, H. V., Abad-Gurumeta, A., Gupta, L., Hemmings, H. C., ... Zdanowski, S. (2026). Responsible use of large language models in manuscript authorship, peer review, and editorial processes: A Delphi consensus among editors-in-chief of anaesthesia and pain medicine journals (RULE-AP). *British Journal of Anaesthesia*, *136*(5), 1625–1633. https://doi.org/10.1016/j.bja.2026.01.029

de Geofroy, B., Gérard, A., Lupon, É., Desrayaud, C., Choufani, C., & Micicoi, G. (2026). Adoption and perception of artificial intelligence by corresponding authors in orthopaedic research: A global cross-sectional study. *Orthopaedics & Traumatology: Surgery & Research*, art. 104653. https://doi.org/10.1016/j.otsr.2026.104653

Epelde, F. (2026). At the frontiers – and perils – of AI-text detectors in the screening of medical manuscripts. *Medicina Clínica*, *166*(3), art. 107300. https://doi.org/10.1016/j.medcli.2025.107300

Ganjavi, C., Eppler, M. B., Pekcan, A., Biedermann, B., Abreu, A., Collins, G. S., Gill, I. S., & Cacciamani, G. E. (2024). Publishers' and journals' instructions to authors on use of generative artificial intelligence in academic and scientific publishing: Bibliometric analysis. *BMJ*, *384*, art. e077192. https://doi.org/10.1136/bmj-2023-077192

Gutiérrez-Cirlos, C., Carrillo-Pérez, D. L., Bermúdez-González, J. L., Hidrogo-Montemayor, I., Carrillo-Esper, R., & Sánchez-Mendiola, M. (2023). ChatGPT: Opportunities and risks in the fields of medical care, teaching, and research. *Gaceta Médica de México*, *159*(5), 372–379. https://doi.org/10.24875/GMM.M23000811

Hoomanfard, M., & Shamsi, Y. (2025). Generative AI in dissertation writing: L2 doctoral students' self-reported use, AI-giarism, and perceived training needs. *Journal of English for Academic Purposes*, *78*, art. 101570. https://doi.org/10.1016/j.jeap.2025.101570

Hosseini, M., Resnik, D. B., & Holmes, K. (2023). The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts. *Research Ethics*, *19*(4), 449–465. https://doi.org/10.1177/17470161231180449

Merton, R. K. (1977). La estructura normativa de la ciencia. In *La sociología de la ciencia, 2. Investigaciones teóricas y empíricas* (pp. 355–368). Alianza Editorial.

Owusu Sarfo, J. (2023). Artificial Intelligence Chatbot – ChatGPT and High-Tech Plagiarism Concerns in a Digital Age: Is Detection Possible? *Journal of Advocacy, Research and Education*, *10*(2), 55–58. https://doi.org/10.13187/jare.2023.2.55

Queiroz, A. B. L., Sartori, L. R. M., Lima, G. da S., & Moraes, R. R. (2025). Editorial policies for use and acknowledgment of artificial intelligence in dental journals. *Journal of Dentistry*, *161*, art. 105923. https://doi.org/10.1016/j.jdent.2025.105923

Salvagno, M., Taccone, F. S., & Gerli, A. G. (2023). Can artificial intelligence help for scientific writing? *Critical Care*, *27*(1), art. 75. https://doi.org/10.1186/s13054-023-04380-2

Scherbakov, D., Hubig, N., Jansari, V., Bakumenko, A., & Lenert, L. A. (2025). The emergence of large language models as tools in literature reviews: a large language model-assisted systematic review. *Journal of the American Medical Informatics Association*, *32*(6), 1071–1086. https://doi.org/10.1093/jamia/ocaf063

Vahedian-Azimi, A., Arabfard, M., & Blot, S. (2026). Authorship integrity and artificial intelligence. *Intensive & Critical Care Nursing*, *93*, art. 104283. https://doi.org/10.1016/j.iccn.2025.104283