Revista Información Científica

ISSN 1028-9933
Universidad de Ciencias Médicas de Guantánamo

July 2025 Volume 104 e5053
DOI: https://doi.org/10.5281/zenodo.15707725

EDITORIAL

Scientific review and the ethical challenge of artificial intelligence

La revisión científica frente al reto ético de la inteligencia artificial

Revisão científica e o desafio ético da inteligência artificial

Javier Pérez-Capdevila

Universidad de Guantánamo. Delegación Territorial de Ciencia, Tecnología y Medio Ambiente. Guantánamo, Cuba.

***Corresponding author:** jvpzcp@gmail.com

Received: 20-06-2025 **Accepted:** 03-07-2025 **Published:** 06-07-2025

How to cite this article:
Pérez-Capdevila J. Scientific review and the ethical challenge of artificial intelligence. Rev Inf Cient[Internet]. 2025[cited Access date]; 104:e5053. Available at: http://www.revinfcientifica.sld.cu/index.php/ric/article/view/5053

Advances in artificial intelligence (AI) technology have revolutionized multiple fields, especially in the generation of academic content, optimizing processes such as writing, editing, and interpreting documents. However, their unregulated application creates dilemmas surrounding ethics and intellectual authorship, prompting the creation of systems capable of distinguishing between human texts and those produced by algorithms.

These mechanisms analyze stylistic and probabilistic features typical of automated language generators. Although valuable, their effectiveness depends on factors such as the complexity of the material being reviewed. Therefore, specialized publications have begun to use them as part of their verification protocols.

The academic publishing sector must reconcile technological progress with the principles of research rigor and integrate these digital solutions into the peer-review process. Evidence of this is the development and implementation of editorial policies related to the use of AI in many journals. However, their role is auxiliary, as the final assessment falls to experts who interpret the findings from a critical perspective.

From this point on, the use of artificial intelligence (AI) detectors by editors and referees in the review of scientific articles, from a very different perspective than usual, could be analyzed as a potentially unethical practice that undermines the fundamental principles of academic research: transparency, fairness, trust, and evaluation based on scientific merit. Some possible reasons for this approach are presented below.

AI detectors generate false positives and harm academic autonomy because they lack absolute accuracy, and often generate false alarms; they stigmatize legitimate work as "non-human." This threatens the autonomy of authors whose carefully written texts could be rejected without objective grounds. A study rejected due to suspected AI (without concrete evidence of plagiarism or fraud) violates a researcher's right to be evaluated for their methodological rigor, not their writing style. According to Bilgiç[(1)], relying on automated detectors distracts from central problems such as fraud or poor human oversight.

The algorithms behind these detectors are often black boxes, whose criteria are not accessible to either authors or reviewers, creating algorithmic opacity and a lack of due process. Basing editorial decisions on opaque systems contradicts the principle of transparency that should govern science. If an editor cannot explain how a text was determined to be "AI-generated," how can a rejection be ethically justified? The lack of clear appeals against these automated accusations negates the right to defense. Current AI-generated text detectors are ill-equipped to adequately identify content created by language models in the peer-review process, raising concerns about the fairness and reliability of using them to evaluate scientific manuscripts.[(2)]

AI detectors are often trained with predominantly English-speaking and Western data, penalizing the discursive styles of non-native researchers or researchers from different academic cultures. This reinforces global inequalities by excluding marginalized voices whose linguistic patterns differ from the algorithmic "norm." Science loses intellectual diversity when biased tools act as gatekeepers of legitimacy.

Furthermore, many AI detectors rely on commercial platforms (privatization of scientific evaluation) whose economic interests can collide with ethical ones. Subordinating refereeing to private systems whose mechanisms and purposes are not public erodes the collective integrity of peer review, traditionally managed by the academic community. Furthermore, submitting texts to third parties (detectors store information) poses privacy risks, especially if sensitive data is stored without consent.

Science is built through critical dialogue between humans. Replacing this dialogue with automated verification prioritizes efficiency over expert judgment, reducing review to a mere technical formality and dehumanizing the scientific process. Is it ethical for an algorithm to decide which ideas deserve discussion while referees avoid delving into the originality or real impact of the work?

The answer to the question at the end of the previous paragraph is categorically negative. Although some AI detectors report accuracy rates of up to 99.9% in controlled environments. their performance drops drastically in real-life contexts, calling into question their practical usefulness[3]. Free AI detection tools achieve an average accuracy of 48%, while paid ones barely exceed 64%, demonstrating a high error rate in educational and publishing contexts[4]. It is also worth considering that the OpenAI detector (considered to date the most influential and advanced player in artificial intelligence) was withdrawn after achieving only 26% accuracy in detecting AI-generated texts, with a 74% false positive rate in human texts[5].

Finally, it could be recommended that, instead of pursuing dubious surveillance tools, journals should strengthen traditional mechanisms such as verified plagiarism detectors and open reviews, which prioritize the intrinsic quality of research. If AI is used in writing, this should be declared and openly discussed, without criminalizing its use a priori. Scientific ethics demands that human judgment, not imperfect algorithms, define what knowledge advances.

Institutional mistrust of authors, normalized by AI detectors, hinders academic collaboration. The answer is not technological surveillance, but rather an editorial culture that values transparency, corrects biases, and trusts human discernment as an indispensable pillar of science.

BIBLIOGRAPHIC REFERENCES

1. Bilgiç ET. Pasquale, Frank, New laws of robotics: defending human expertise in the age of AI. Massachusetts: The Belknap Press of Harvard University Press; 2020. 330 p. DOI: https://doi.org/10.33630/ausbf.1439964
2. Yu S, Luo M, Madusu A, Lal V, Howard, P. (2025). Is your paper being reviewed by an LLM? A new benchmark dataset and approach for detecting AI text in peer review. arXiv:2502.19614[Preimpresión]. 2025[cited 14 Jun 2025]. DOI: https://doi.org/10.48550/arXiv.2502.19614
3. Gritsai G, Voznyuk A, Grabovoy A, Chekhovich Y. Are AI detectors good enough? A survey on quality of datasets with machine-generated texts. arXiv:2410.14677 [Preimpresión]. 2025 [cited 10 Jun 2025]. DOI: https://doi.org/10.48550/arXiv.2410.14677
4. Borell N. How reliable are AI detection tools? [Internet]. 4sysops, Dic 2024 [cited 14 Jun 2025]. Available at: https://4sysops.com/archives/how-reliable-are-ai-detection-tools/
5. AI Busted Team. How reliable are AI detectors? An in-depth analysis[Internet].AI Busted Blog, Abr 2025. [cited 14 Jun 2025]. Available at: https://blog.aibusted.com/how-reliable-are-ai-detectors/

Conflict of Interest:
The author declares no conflicts of interest.

Financing:
The author did not receive any funding for the development of this article.