

La revisión científica frente al reto ético de la inteligencia artificial

Scientific review and the ethical challenge of artificial intelligence

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.

Para la correspondencia: jvpzcp@gmail.com

Recibido: 20-06-2025 Aprobado: 03-07-2025 Publicado: 06-07-2025

Cómo citar este artículo:

Pérez-Capdevila J. La revisión científica frente al reto ético de la inteligencia artificial. Rev Inf Cient [Internet]. 2025 [citado Fecha de acceso]; 104:e5053. Disponible en: <http://www.revinfcientifica.sld.cu/index.php/ric/article/view/5053>

Los avances en tecnología sobre inteligencia artificial (IA) han revolucionado múltiples campos, especialmente en la generación de contenidos académicos, donde optimizan procesos como la escritura, corrección e interpretación de documentos. No obstante, su aplicación sin regulación genera dilemas en torno a la ética y la autoría intelectual, lo que motiva la creación de sistemas capaces de discernir entre textos humanos y aquellos producidos por algoritmos.

Estos mecanismos analizan rasgos estilísticos y probabilísticos propios de los generadores de lenguaje automatizado. Aunque resultan valiosos, su eficacia depende de factores como la complejidad del material examinado. Por ello, las publicaciones especializadas han comenzado a emplearlos como parte de sus protocolos de verificación.

El sector editorial académico debe conciliar el progreso tecnológico con los principios de rigor investigativo e integrar estas soluciones digitales en las fases de arbitraje. Prueba de ello es la elaboración y aplicación de políticas editoriales relacionadas con el uso de la IA en muchas revistas. Sin embargo, su función es auxiliar, ya que la valoración final recae en expertos que interpretan los hallazgos con perspectiva crítica.

A partir de este punto, el empleo de detectores de inteligencia artificial (IA) por parte de editores y árbitros en la revisión de artículos científicos, desde una perspectiva muy diferente a la habitual, pudiese analizarse como una posible práctica antiética que socava los principios fundamentales de la investigación académica: transparencia, equidad, confianza y evaluación basada en el mérito científico. A continuación, se exponen algunas posibles razones que fundamentan este planteamiento.

Los detectores de IA generan falsos positivos y daño a la autonomía académica ya que carecen de precisión absoluta y suelen generar falsas alarmas; estigmatiza trabajos legítimos como "no humanos". Esto amenaza la autonomía de autores cuyos textos, redactados con esmero, podrían ser rechazados sin fundamento objetivo. Un estudio rechazado por sospecha de IA (sin pruebas concretas de plagio o fraude) viola el derecho del investigador a ser evaluado por su rigor metodológico, no por su estilo de escritura. Según Bilgiç⁽¹⁾, confiar en detectores automatizados distrae de problemas centrales como el fraude o la supervisión humana deficiente.

Los algoritmos detrás de estos detectores suelen ser cajas negras, cuyos criterios no son accesibles ni para autores ni para revisores, lo cual crea opacidad algorítmica y ausencia de debido proceso. Basar decisiones editoriales en sistemas opacos contradice el principio de transparencia que debe regir la ciencia. Si un editor no puede explicar cómo se determinó que un texto es "IA-generado", ¿cómo se justifica éticamente un rechazo? La falta de apelación clara ante estas acusaciones automatizadas niega el derecho a la defensa. Los detectores actuales de textos generados por IA no están preparados para identificar adecuadamente contenidos creados por modelos de lenguaje en el proceso de revisión por pares, lo que genera inquietudes sobre la equidad y confiabilidad al usarlos para evaluar manuscritos científicos.⁽²⁾

Los detectores de IA suelen entrenarse con datos predominantemente anglófonos y occidentales, penalizando estilos discursivos de investigadores no nativos o de culturas académicas distintas. Esto refuerza desigualdades globales, al excluir voces marginadas cuyos patrones lingüísticos difieren de la "norma" algorítmica. La ciencia pierde diversidad intelectual cuando herramientas sesgadas actúan como guardianes de la legitimidad.

Por otra parte, muchos detectores de IA dependen de plataformas comerciales (privatización de la evaluación científica) cuyos intereses económicos pueden colisionar con los éticos. Subordinar el arbitraje a sistemas privados cuyos mecanismos y fines no son públicos, erosiona la integridad colectiva de la revisión por pares, tradicionalmente gestionada por la comunidad académica. Además, someter textos a terceros (los detectores guardan información) plantea riesgos de privacidad, especialmente si se almacenan datos sensibles sin consentimiento.

La ciencia se construye mediante el diálogo crítico entre humanos. Reemplazar este diálogo con verificación automatizada prioriza la eficiencia sobre el juicio experto, reduciendo la revisión a un mero trámite técnico deshumanizando del proceso científico. ¿Es ético que un algoritmo decida qué ideas merecen ser discutidas mientras árbitros evitan profundizar en la originalidad o el impacto real del trabajo?



La respuesta a la interrogante que finaliza el párrafo anterior es categóricamente negativa. Aunque algunos detectores de IA reportan precisiones de hasta el 99,9 % en entornos controlados, su rendimiento cae drásticamente en contextos reales, lo que cuestiona su utilidad práctica.⁽³⁾ Las herramientas gratuitas de detección de IA alcanzan una precisión promedio del 48 %, mientras que las de pago apenas superan el 64 %, lo que demuestra una alta tasa de error en contextos educativos y editoriales⁽⁴⁾ y, es muy valioso considerar que el detector de OpenAI (considerado hasta hoy como el actor más influyente y avanzado en la inteligencia artificial) fue retirado tras alcanzar solo un 26 % de precisión en la detección de textos generados por IA, con un 74 % de falsos positivos en textos humanos.⁽⁵⁾

Para finalizar se pudiese recomendar que, en lugar de perseguir herramientas dudosas de vigilancia, las revistas deben fortalecer mecanismos tradicionales como detectores de plagio verificados y revisiones abiertas, que prioricen la calidad intrínseca de la investigación. Si la IA se usa en la escritura, ello debe declararse y debatirse abiertamente, sin criminalizar *a priori* su empleo. La ética científica exige que el criterio humano, no algoritmos imperfectos, defina qué conocimiento avanza.

La desconfianza institucional hacia autores, normalizada mediante detectores de IA, entorpece la colaboración académica. La respuesta no es la vigilancia tecnológica, sino una cultura editorial que valore la transparencia, corrija sesgos y confíe en el discernimiento humano como pilar irrenunciable de la ciencia.

REFERENCIAS BIBLIOGRÁFICAS

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 [citado 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 [citado 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 [citado 14 Jun 2025]. Disponible en: <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. [citado 14 Jun 2025]. Disponible en: <https://blog.aibusted.com/how-reliable-are-ai-detectors/>

Declaración de conflicto de intereses:

El autor declara que no existen conflictos de intereses.

Financiación:

El autor no recibió financiación para el desarrollo de este artículo.

