

¿Cómo detectar textos científicos generados con inteligencia artificial?

How to detect scientific texts generated with artificial intelligence?

Omar Chávez-Martínez^{1a}

Resumen

Los modelos de lenguaje de gran escala han transformado la escritura científica, lo cual facilita la redacción y la revisión de textos, pero a su vez genera riesgos éticos y epistemológicos. Aunque su uso promueve la equidad lingüística, la falta de transparencia y la manipulación de información amenazan la integridad académica. Los detectores de inteligencia artificial (IA), como Originality.ai, ZeroGPT o Turnitin muestran eficacia variable y no ofrecen resultados concluyentes, especialmente ante los “humanizadores de texto”. Los textos generados por IA se caracterizan por su coherencia formal, pero también por su previsibilidad y uniformidad estilística. Por ello, la detección debe combinarse con la evaluación ética y crítica hecha por humanos, y hay que entender que la auténtica integridad científica depende del juicio intelectual y no de la automatización tecnológica.

Abstract

Large language models have transformed scientific writing, which facilitates text drafting and revision, but at the same time introduces ethical and epistemological risks. Even though their use promotes linguistic equity, the lack of transparency and the manipulation of information threaten academic integrity. AI detectors —such as Originality.ai, ZeroGPT, or Turnitin— show variable effectiveness and do not provide conclusive results, especially against “text humanizers.” AI-generated texts are characterized by formal coherence, but also by predictability and stylistic uniformity. Therefore, detection must be combined with ethical and critical evaluation made by humans, and it must be understood that true scientific integrity depends on intellectual judgment rather than technological automation.

¹Instituto Mexicano del Seguro Social, Coordinación de Investigación en Salud, División de Investigación Clínica. Ciudad de México, México

ORCID: 0000-0003-2633-1898^a

Palabras clave
Inteligencia Artificial Generativa
Escritura Médica
Ética en Investigación

Keywords
Generative Artificial Intelligence
Medical Writing
Ethics, Research

En los últimos años, los modelos de lenguaje de gran escala (LLM, por sus siglas en inglés) han transformado profundamente las prácticas académicas, lo cual ha reconfigurado los modos de producción, revisión y difusión del conocimiento. Estas herramientas han introducido una nueva dinámica en la escritura científica que posibilita una notable aceleración de procesos tradicionalmente laboriosos, como la redacción de manuscritos, la revisión bibliográfica o la corrección gramatical. Actualmente se estima que

aproximadamente el 76% de los investigadores emplea tecnologías de inteligencia artificial (IA) para optimizar su escritura y reducir las desigualdades lingüísticas entre hablantes nativos y no nativos del inglés.¹

Esta expansión del uso de herramientas basadas en IA en la investigación representa una oportunidad significativa para promover la equidad y la eficiencia en la comunicación científica, ya que permite a muchos autores mejorar la clari-

Comunicación con:
Omar Chávez Martínez
 omar.chavezm@imss.gob.mx
 55 5627 6900, extensión 22129

Cómo citar este artículo: Chávez-Martínez O. ¿Cómo detectar textos científicos generados con inteligencia artificial? Rev Med Inst Mex Seguro Soc. 2026;64(1):e6952. doi: 10.5281/zenodo.17477984

dad técnica de sus manuscritos. Sin embargo, este avance conlleva una serie de riesgos que comprometen la integridad del ecosistema académico y científico, como la falta de transparencia en el uso de estas herramientas y la posible manipulación de información. Estudios recientes advierten que entre 2021 y 2023 se identificó un aumento de entre 1.79 y 2.37 veces más de resúmenes de oncología que probablemente incluyen contenido generado por IA, lo que sugiere una adopción acelerada y, en muchos casos, poco controlada. Este fenómeno se ve agravado por la capacidad de modelos como ChatGPT para producir textos científicamente verosímiles, pero con datos inventados o tergiversados, conocidos como “alucinaciones”. Tales prácticas, cuando no son detectadas, pueden derivar en la aceptación de artículos manipulados con IA en revistas indexadas, lo cual mina la confianza en la literatura científica y diluye los estándares de rigor académico.²

En este contexto, la identificación de textos generados por IA se convierte en una tarea crucial, pero más allá de considerarse como un mero ejercicio técnico, es importante que se adopte una estrategia ética y epistemológica orientada a preservar la autenticidad de la producción científica. Esto implica analizar los indicios lingüísticos, estilísticos y semánticos que revelan la intervención de modelos generativos.

Actualmente existen herramientas destinadas a identificar textos producidos por modelos de lenguaje, los cuales se han desarrollado con rapidez debido a la creciente preocupación ante el uso indiscriminado de IA en la escritura académica. No obstante, los estudios más recientes evidencian una variabilidad considerable en los resultados de los detectores, lo que cuestiona su fiabilidad como mecanismos de control editorial o ético.

Por ejemplo, en términos comparativos, Originality.ai se ubica entre los sistemas más eficaces, al detectar el 100% de los textos reescritos por IA, mientras que ZeroGPT alcanzó un 88% de aciertos y Turnitin apenas un 30%. Aun más limitado fue BrandWell AI Writer and Editor Review, que solo asignó un 48% de probabilidad de autoría artificial a un texto comprobadamente generado por IA. Estas diferencias reflejan una deficiencia en la consistencia de la verificación de escritura humana, lo que ha ocasionado altas tasas de falsos positivos y negativos que pueden derivar tanto en acusaciones injustas como en la omisión de casos reales de escritura automatizada.^{3,4,5}

El panorama se complica con la aparición de los llamados “humanizadores de texto”, que son programas diseñados para modificar textos generados por IA con la finalidad de volverlos indistinguibles de los escritos humanos. Estas herramientas afectan la eficacia de muchos detectores y agravan los dilemas éticos asociados a su uso.⁶

Es por lo anterior que los detectores deben usarse como herramientas que apoyan la evaluación de textos y no como generadoras de pruebas concluyentes. Confiar exclusivamente en sus resultados es arriesgado, ya que en lugar de convertir la vigilancia tecnológica en el eje de la integridad académica, resulta necesario comenzar a construir marcos regulatorios integrales basados en la transparencia, la educación ética y la evaluación crítica del contenido, en donde la detección sea solo un componente dentro de una política más amplia de preservación de la confianza científica.

También resulta importante destacar que los textos producidos por IA se distinguen por su verosimilitud, corrección gramatical y coherencia interna, lo que les otorga una apariencia de rigor académico, solidez formal que suele ocultar una uniformidad lingüística y discursiva que la diferencia de la escritura humana. En este sentido, suelen presentar baja perplejidad, es decir, una alta previsibilidad estadística: las palabras, las frases y las transiciones siguen patrones regulares que generan fluidez, pero también monotonía.⁵

Por ejemplo, ChatGPT revela ciertos patrones discursivos recurrentes que, aunque no afectan su corrección formal, sí delatan su origen automatizado. Uno de los más notorios es el uso frecuente de la raya tipográfica (—) en lugar de comas o paréntesis tradicionales. Esta preferencia tipográfica, empleada para introducir aclaraciones o enfatizar ideas, confiere al texto una cadencia artificiosa, más propia de un sistema que busca estructurar con precisión lógica cada fragmento del discurso. Asimismo, ChatGPT tiende a utilizar construcciones paralelas del tipo “no solo... sino también”, que aparecen con una regularidad mayor que en la escritura humana, lo que genera una sensación de equilibrio mecánico y previsibilidad argumentativa.⁷

Otro rasgo distintivo es la presencia reiterada de marcadores de cierre, como “finalmente”, “en conclusión” o “en última instancia”, incluso antes de que el texto llegue efectivamente a su conclusión, hábito que produce una clausura anticipada del discurso y crea la impresión de que el modelo intenta cumplir con una secuencia discursiva preestablecida más que desarrollar una reflexión orgánica.^{8,9}

En el plano léxico, destaca el empleo de términos genéricos o semánticamente inadecuados según el contexto idiomático y disciplinar, como el uso de “sanitario” en lugar de “hospitalario”, o de expresiones amplias que diluyen la precisión terminológica propia del lenguaje académico. A ello se suma la tendencia al uso excesivo de viñetas o listados, con el fin de estructurar la información de manera esquemática.

Precisamente, estas regularidades léxicas, sintácticas y estructurales constituyen la base sobre la cual operan los

detectores automáticos de IA, diseñados para reconocer los patrones distintivos de la escritura generada por modelos de lenguaje. No obstante, la revisión ética y el juicio humano deben mantenerse como componentes esenciales para salvaguardar la integridad y autenticidad de la producción científica asistida por IA. Si bien estos detectores resultan útiles como herramientas de cribado preliminar, su alcance sigue siendo limitado y no permite establecer conclusiones definitivas sobre la autoría o la naturaleza del texto.

En consecuencia, más que depositar una confianza ciega en los algoritmos de detección, el verdadero desafío consiste en integrar su empleo con criterios humanos de evaluación lingüística, ética y metodológica, a fin de preservar la autenticidad, la responsabilidad intelectual y el rigor académico en la escritura científica contemporánea.⁶

de apoyo que complemente en vez de reemplazar, y que ayude a mejorar formal y lingüísticamente los textos sin desplazar el juicio académico ni la deliberación humana, elementos que dan legitimidad a la producción científica.⁶

Desde esta perspectiva, es pertinente recordar que la Ley Federal del Derecho de Autor mexicana establece que el derecho de autor corresponde exclusivamente al creador de la obra y precisa que solo las personas físicas pueden ser reconocidas como autores. En concordancia con ello, el Tribunal Federal de Justicia Administrativa, en la tesis IX-CASE-PI-3, determinó que las obras generadas por inteligencia artificial no son susceptibles de protección jurídica bajo dicha ley, dado que la creatividad humana es un requisito esencial para que una obra sea considerada original y, por tanto, protegida. La resolución señala que, aun

Según la Ley Federal de Derecho de Autor de México, el derecho de autor le atañe exclusivamente al creador de la obra y únicamente pueden ser reconocidos como autores las personas físicas. Por lo tanto, las obras producidas totalmente con sistemas de inteligencia artificial no pueden ser registradas ni tener derechos de autor

Desde una perspectiva ética es esencial determinar que la IA no debe ser considerada autora de un trabajo académico, ya que carece de responsabilidad moral, intención cognitiva y capacidad de juicio crítico. La autoría científica implica una relación directa con la veracidad, la interpretación y la responsabilidad sobre el contenido, atributos que solo pueden ser asumidos por personas. Por ello, los investigadores deben declarar de forma explícita cualquier uso de herramientas de IA en la redacción, edición o revisión de sus textos, garantizando la transparencia y la trazabilidad del proceso de producción.

Asimismo, la ética científica de autoría requiere que el investigador conserve el control intelectual sobre lo que ha escrito y compruebe la coherencia del argumento y la precisión de los hechos antes de hacer pública la obra. En este contexto, la IA debe concebirse como una herramienta

cuando una persona proporcione instrucciones o parámetros a un sistema algorítmico, el resultado final proviene de la ejecución automática de los algoritmos y constituye un contenido artificial, carente del elemento creativo humano indispensable. En consecuencia, las obras producidas de manera totalmente autónoma por sistemas de IA no pueden registrarse ni gozar de derechos de autor, al no cumplir con los requisitos establecidos en la legislación mexicana.^{10,11}

Por lo tanto, es imprescindible reforzar la capacitación sistemática de revisores, editores e investigadores en alfabetización de IA y en ética digital. Únicamente con este esfuerzo será factible combinar la innovación tecnológica con los valores fundamentales de la investigación científica: el rigor académico, la responsabilidad compartida y la honestidad intelectual.

Referencias

1. Gautam N. Regulating the unseen hand: AI, authorship, and trust in medical science. Ann Med Surg (Lond). 2025;87(8): 5353-5. doi: 10.1097/MS9.000000000003498
2. Howard FM, Li A, Riffon MF, et al. Characterizing the Increase in Artificial Intelligence Content Detection in Oncology Scientific Abstracts From 2021 to 2023. JCO Clin Cancer Inform. 2024;8:e2400077. doi: 10.1200/CC.24.00077
3. Gao CA, Howard FM, Markov NS, et al. Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. NPJ Digit Med. 2023;6

- (1):75. doi: 10.1038/s41746-023-00819-6
- 4. Odri GA, Ji Yun Yoon D. Detecting generative artificial intelligence in scientific articles: Evasion techniques and implications for scientific integrity. *Orthop Traumatol Surg Res.* 2023; 109(8):103706. doi: 10.1016/j.otsr.2023.103706
 - 5. Erol G, Ergen A, Gülsen Erol B, et al. Can we trust academic AI detective? Accuracy and limitations of AI-output detectors. *Acta Neurochir (Wien).* 2025;167(1):214. doi: 10.1007/s00701-025-06622-4
 - 6. Pellegrina D, Helmy M. AI for scientific integrity: detecting ethical breaches, errors, and misconduct in manuscripts. *Front Artif Intell.* 2025;8:1644098. doi: 10.3389/frai.2025.1644098
 - 7. Shi R. A comparative study on the meta-functions of artificial intelligence and human discourse markers. *J Lit Art Stud.* 2024;14(10):819-829. doi:10.17265/2159-5836/2024.10.008
 - 8. Muñoz-Ortiz A, Gómez-Rodríguez C, Vilares D. Contrasting Linguistic Patterns in Human and LLM-Generated News Text. *Artif Intell Rev.* 2024;57(10):265. doi: 10.1007/s10462-024-10903-2
 - 9. Reinhart A, Markey B, Laudenbach M, et al. Do LLMs write like humans? Variation in grammatical and rhetorical styles. *Proc Natl Acad Sci U S A.* 2025;122(8):e2422455122. doi: 10.1073/pnas.2422455122
 - 10. Ley Federal del Derecho de Autor (Última reforma publicada el 01-07-2020) [Internet]. México: Cámara de Diputados del H. Congreso de la Unión; 1996 [consultado el 1 de noviembre de 2025]. Disponible en: <https://www.diputados.gob.mx/LeyesBiblio/pdf/LFDA.pdf>
 - 11. Tribunal Federal de Justicia Administrativa. Derechos de autor. Las obras creadas por inteligencia artificial no son sujetas de protección de conformidad con la Ley Federal del Derecho de Autor. R.T.F.J.A. Novena Época. Año III, No. 36, diciembre 2024, p. 388. Publicada el 9 de diciembre de 2024 [Internet]. México: TFJA; 2024 [consultado el 1 de noviembre de 2025]. Disponible en: <https://www.tfja.gob.mx/cesmdfa/sctj/tesis-pdf-detalle/47734/>