

## HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL EN LA INVESTIGACIÓN ACADÉMICA Y CIENTÍFICA: NORMATIVAS, DESAFÍOS Y PRINCIPIOS ÉTICOS

GLEND A ERNST, PABLO YOUNG

Departamento de Docencia e Investigación, Hospital Británico de Buenos Aires, Argentina

E-mail: pabloyoung2003@yahoo.com.ar

El artículo del Dr. Eduardo L. De Vito publicado en *MEDICINA* debate la realidad acerca de la inteligencia artificial (IA) y el ChatGPT<sup>1</sup>.

El ChatGPT es un modelo lingüístico destacado (MLD) desarrollado por OpenAI, disponible como un chatbot de IA de acceso público. Fundamentado en la arquitectura del Transformador Generativo Preentrenado (GPT), tanto GPT-3.5 como GPT-4 sirven como pilares de esta tecnología. Dichos modelos han sido entrenados con un amplio conjunto de datos, lo que les capacita para llevar a cabo conversaciones dinámicas en múltiples idiomas. Aunque su entrenamiento se basó en datos hasta enero de 2022, el ChatGPT tiene la capacidad de integrar información en tiempo real o basada en conocimiento mediante *plugins* o complementos. El lanzamiento de ChatGPT ha marcado el surgimiento de varios MLD similares, entre los cuales se incluyen Google Bard, GenAi, entre otros<sup>2-5</sup>.

Respecto de la generación de contenidos médicos, la confiabilidad y rendimiento del ChatGPT, ha sido evaluado mediante el *United States Medical Licensing Exam* (USMLE), que representa un conjunto de tres pruebas estandarizadas de conocimientos a nivel de experto que se exigen para obtener la licencia médica en EE. UU. Sus resultados comprobaron que ChatGPT alcanza o se aproxima al umbral de aprobación del 60% de precisión, mostrando un razonamiento comprensible y conocimientos clínicos válidos,

lo que aumenta la confianza y la capacidad de explicación<sup>3</sup>. Más aún, también ha sido evaluado el ChatGPT, en pruebas de admisión estandarizadas en el Reino Unido, incluidas la Prueba de Admisión Biomédica (BMAT), la Prueba de Matemáticas para la Admisión a la Universidad (TMUA), la Prueba Nacional de Aptitud en Derecho (LNAT) y la Evaluación de Habilidades de Pensamiento (TSA), para comprender su potencial como herramienta innovadora para la educación y la preparación de exámenes. Los hallazgos de este estudio mostraron que el ChatGPT resulta una herramienta complementaria para asignaturas y formatos de examen que evalúan la aptitud, la resolución de problemas y el pensamiento crítico, y la comprensión lectora<sup>4</sup>.

Estas tecnologías de IA, son cada vez más utilizadas en los entornos académicos, por lo que es necesario examinar cómo estas herramientas interactúan con cuestiones de la autoría, la integridad académica y las metodologías de investigación.

De Vito resume la importancia de integrar tanto las oportunidades como las preocupaciones que surgen con el uso de esta dinámica tecnología, y la necesidad de establecer políticas éticas y transparentes en su aplicación en la producción de contenido científico y académico<sup>1</sup>.

Las reglamentaciones respecto de la utilización de las herramientas de IA han comenzado

a identificarse por parte de las editoriales más conocidas. Por ejemplo, la Revista *Nature* ha publicado dentro su normativa que no aceptará ningún modelo de lenguaje como autor acreditado en un artículo de investigación y que los investigadores que utilicen modelos de lenguaje deben documentar este uso en las secciones de métodos o agradecimientos<sup>6</sup>.

Estas reglas buscan promover el uso ético de los modelos de lenguaje en la investigación académica y garantizar la transparencia y la integridad en el proceso de generación de conocimiento.

Baskwill y col. enuncian conceptos claves respecto de la utilización de las herramientas de IA para generar conocimiento científico y destacan la importancia del entrenamiento en datos para que resulten diversos y representativos, ya que es crucial garantizar que los conjuntos de datos utilizados para entrenar los modelos de IA generativa sean libres de sesgos para generar resultados inclusivos e imparciales. Fomentar herramientas de detección y mitigación de sesgos en los datos de entrenamiento y en el propio modelo de IA. Y finalmente la transparencia en el funcionamiento de los sistemas de IA generativa y ofrecer explicaciones sobre sus resultados puede ayudar a los usuarios a comprender el razonamiento que subyace a los contenidos generados<sup>7</sup>.

La Organización Mundial de la Salud, enuncia seis principios claves para garantizar que la IA se utilice con normativas comunes, aumentando al máximo las oportunidades de la utilización de la IA en el ámbito de la salud y reglamentando su utilización<sup>8</sup>:

1. *Preservar la autonomía del ser humano*: Los seres humanos deberían seguir siendo partícipes de su atención en salud y formando parte de las decisiones médicas. Se debe preservar la privacidad y la confidencialidad, y los pacientes deben dar su consentimiento informado y válido por medio de marcos jurídicos adecuados para la protección de datos.

2. *Promover el bienestar y la seguridad de las personas y el interés público*: Los diseñadores de tecnologías de IA deben cumplir los requisitos normativos en materia de seguridad, precisión y eficacia para indicaciones o usos bien definidos. Se deben instaurar medidas de control de la ca-

lidad en la práctica y de mejora de la calidad en la utilización de la IA.

3. *Garantizar la transparencia, la claridad y la inteligibilidad*: La transparencia exige que se publique o documente información suficiente antes de la concepción o el despliegue de una tecnología de IA. Esa información debe ser fácilmente accesible y facilitar consultas y debates provechosos sobre la concepción de la tecnología y sobre el uso que se debería hacer o no de esta.

4. *Promover la responsabilidad*: Las tecnologías de IA permiten realizar tareas específicas; y deben ser utilizadas en condiciones apropiadas y por personas debidamente formadas.

5. *Garantizar la inclusividad y la equidad*: La inclusividad requiere que la IA aplicada a la salud sea concebida de manera que aliente la utilización y el acceso equitativos en la mayor medida de lo posible, con independencia de la edad, sexo, género, raza, origen étnico, orientación sexual, la capacidad u otras características amparadas por los códigos de derechos humanos.

6. *Promover una IA con capacidad de respuesta y sostenible*: Los diseñadores, desarrolladores y usuarios deberían evaluar de forma continua y transparente las aplicaciones de la IA en situación real a fin de determinar si esta responde de manera adecuada y apropiada a las expectativas y las necesidades.

Las herramientas de IA se encuentran en plena fase de expansión y su utilización en los diferentes ámbitos médicos como científicos se ven beneficiados. Las normativas actualmente planteadas, refuerzan el principio básico de transparencia y reproducibilidad de los resultados.

Las personas seguimos siendo responsables de nuestras acciones, los autores de nuestros manuscritos y debemos detallar cuales herramientas de IA fueron utilizadas para redacción, búsqueda o análisis de los datos publicados, de modo de garantizar la reproducibilidad de nuestro trabajo. Para concluir, la inteligencia artificial y específicamente el ChatGPT, han demostrado ser herramientas valiosas y prometedoras en diversos campos, incluyendo la medicina y la educación. Sin embargo, su rápida evolución y expansión plantean desafíos éticos y metodológicos que deben ser abordados de manera proactiva y transparente. Es fundamental que como usuarios, investigadores y profesionales

comprendamos las implicaciones y limitaciones de estas tecnologías, y que las empleemos de manera responsable y criteriosa. Esto incluye el cumplimiento de las normativas y directrices establecidas por instituciones académicas y editoriales, así como la promoción de la transparencia, la integridad y la reproducibilidad en la generación y difusión del conocimiento.

En última instancia, es responsabilidad de todos nosotros asegurarnos de que la IA se de-

sarrolle y se utilice de manera que beneficie a las personas, respetando los derechos humanos y las libertades fundamentales. Como autores y usuarios de estas herramientas, debemos comprometernos con la transparencia y la integridad, documentando y divulgando adecuadamente cualquier uso de IA en nuestro trabajo. Solo así podremos garantizar la confiabilidad y la validez de los resultados obtenidos con la ayuda de estas nuevas tecnologías

## Bibliografía

1. De Vito EL. Inteligencia artificial y chatGPT. ¿Usted leería a un autor artificial? *Medicina (B Aires)* 2023; 83: 329-32.
2. Strong E, DiGiammarino A, Weng Y, et al. Chatbot vs Medical Student Performance on Free-Response Clinical Reasoning Examinations. *JAMA Intern Med* 2023; 183: 1028-30.
3. Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digit Health* 2023; 2: e0000198.
4. Giannos P, Delardas O. Performance of ChatGPT on UK Standardized Admission Tests: Insights from the BMAT, TMUA, LNAT, and TSA examinations. *JMIR Med Educ* 2023; 9: e47737.
5. Wang X, Sanders HM, Liu Y, et al. ChatGPT: promise and challenges for deployment in low- and middle-income countries. *Lancet Reg Health West Pac* 2023; 41: 100905.
6. Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. *Nature* 2023; 613: 612.
7. Baskwill A. Navigating Generative AI: Opportunities, Limitations, and Ethical Considerations in Massage Therapy and Beyond. *Int J Ther Massage Bodywork* 2023; 16: 1-4.
8. La OMS publica el primer informe mundial sobre inteligencia artificial (IA) aplicada a la salud y seis principios rectores relativos a su concepción y utilización. 28 de junio de 2021 Comunicado de prensa Ginebra. En: <https://www.who.int/es/news/item/28-06-2021-who-issues-first-global-report-on-ai-in-health-and-six-guiding-principles-for-its-design-and-use>; consultado marzo 2024.