

INTELIGENCIA ARTIFICIAL Y CHAT-GPT. ¿USTED LEERÍA A UN AUTOR ARTIFICIAL?

La inteligencia artificial (IA) está cada vez más presente en muchos ámbitos de nuestra vida cotidiana y no deja de sorprender e inquietar. *Revista Medicina* se ha ocupado recientemente del tema (1,2). Uno de los máximos exponentes de esta herramienta es el ChatGPT (3), merecedora de nutridos espacios publicitarios y que podría hacerle pasar un mal rato a Google y sorprender a sus usuarios con respuestas a preguntas existenciales como *¿Cuál es el sentido de la vida?* y mucho más (4).

No es posible soslayar el impacto de la estrategia promocional de esta herramienta, y que no pocos artículos periodísticos de difusión general que abordan la IA la ilustran con imágenes de robots. Esta antropomorfización de la IA se enlaza erróneamente con la idea según la cual esta tiene el potencial de actuar como un agente en el mundo real. Pero mantengamos la calma, el fin del mundo está aún muy lejano.

¿Qué es el ChatGPT?

El ChatGPT (Generative Pre-trained Transformer 3), conocido mundialmente como GPT3, es un modelo de aprendizaje utilizado para generar textos similares a los escritos por humanos (5). A poco de su lanzamiento, en noviembre de 2022, ChatGPT se convirtió en una sensación cultural. Se trata de una tecnología conversacional de libre acceso (hasta un determinado nivel) a través de un portal web creado por la empresa OpenAI, fundada en 2015 en EE.UU. por Sam Altman, en la que participó Elon Musk. Microsoft acaba de comprar el 49% de las acciones de ChatGPT, por lo cual no nos debería extrañar que la empresa incorpore un nuevo ícono a las herramientas de oficina ya conocidas.

Se trata de un programa que crea un texto automáticamente en función de las indicaciones ingresadas. Está preentrenado en un gran conjunto de datos, por lo que puede comprender y generar textos sobre una amplia gama de temas con alta precisión y una naturalidad similar a la escritura humana, en unos 100 idiomas. Además de ser utilizado como un portal de juegos, existen inquietantes implicaciones para el ámbito académico y científico.

¿Qué puede hacer el ChatGPT?

Desde su lanzamiento, ha sorprendido a millones de usuarios al recitar poemas, cuentos, ensayos y consejos personales. Es capaz de crear una partitura musical e incluso un ensayo académico comparando dos teorías (5). El programa también puede producir resúmenes científicos que tienen la calidad similar a la de un humano y podría engañar a revisores. La herramienta ChatGPT ya figura como coautora en un puñado de artículos (6). Desde su lanzamiento, se ha hecho evidente que este tipo de tecnología de IA podrá tener influencias en la forma en que trabajan las universidades y los investigadores. Estos son algunos de los motivos del presente editorial.

Defensores y críticos del ChatGPT

El ChatGPT ha recibido críticas positivas y negativas (5). Estas últimas llegaron del ámbito de la educación y de la ciencia: por ejemplo, en la Universidad de Estrasburgo, Francia, descubrieron que 20 estudiantes realizaron un examen a distancia usando ChatGPT como asistente, y deberán hacer el examen de manera presencial (7). El Departamento de Educación de la ciudad de Nueva York ha restringido el acceso a ChatGPT desde internet y los dispositivos en sus escuelas públicas. De manera similar, la universidad Sciences Po de París prohibió a sus alumnos emplearla, en especial sin reconocimiento de haber recurrido a esa tecnología, y ocho universidades australianas han reconocido que su uso por los estudiantes podría considerarse como *copiar* (8).

En el ámbito de la ciencia, las posiciones sobre ChatGPT están más divididas. Algunos editores han prohibido o restringido su utilización en la preparación de envíos, mientras que otros ven su adopción como algo inevitable. El editor en jefe de la revista estadounidense *Science* actualizó su política editorial prohibiendo el uso de textos de ChatGPT y aclarando que el programa no podía figurar como autor (6). Springer-Nature, que publica entre otras *Scientific American* y *Nature*, no lo ha prohibido por completo, pero esta herramienta y similares por ahora pueden ser usadas en la preparación de documentos, siempre que se revelen todos los detalles en el manuscrito. Elsevier, que tiene en su portafolio a *Cell* y *The Lancet*, ha adoptado una postura similar.

La editora en jefe de *Nature* en Londres cree, por otra parte, que con las medidas de seguridad adecuadas, ChatGPT y herramientas de IA similares podrían ser beneficiosas para la ciencia, sobre todo para nivelar el campo de juego para los hablantes no nativos de inglés que podrían usar programas de IA para hacer que el idioma en sus trabajos sea más fluido.

Otra forma de ver ChatGPT es considerarlo como una herramienta más dentro del conjunto de herramientas en constante crecimiento, disponible para escribir un artículo. Uno podría, por ejemplo, implementarlo como un motor de búsqueda que responda preguntas directamente en lugar de sólo referirse a fuentes donde uno mismo debe buscar las respuestas (9).

El ChatGPT, coautor de artículos científicos

Las principales revistas científicas exigen que los autores firmen un formulario en el que declaran que son responsables de su contribución al trabajo. ChatGPT no puede hacer esto (9); sin embargo, es uno de los 12 autores en una preimpresión sobre el uso de la herramienta para la educación médica, publicado en el repositorio médico medRxiv (10).

Una de las pautas editoriales es que un coautor debe hacer una “contribución académica significativa” al artículo, lo que podría ser posible con herramientas como ChatGPT. Pero también debe tener la capacidad de aceptar ser coautor y asumir la responsabilidad de un estudio o, al menos, la parte a la que contribuyó. “Es realmente esa segunda parte en la que la idea de otorgar la coautoría de una herramienta de IA realmente se topa con un obstáculo” (10).

¿Es posible detectar un texto científico en el que intervino ChatGPT?

Más allá de definir si está permitido o no, parece razonable precisar la utilización de algo que ya se está utilizando y cuya mayor difusión parece inevitable. En este punto podría apelarse, al menos por ahora, a que los autores sean sinceros y describan cómo se usó, y que asuman la responsabilidad de sus resultados.

Los resúmenes generados por ChatGPT fueron puestos a prueba a través de verificadores de plagio (*plagiarism*): la mediana de puntuación de originalidad fue del 100%, lo que indica que no detectaron ningún plagio (11). Estos programas, cuando identifican repeticiones, ofrecen un servicio de *paraphrasing*, a efectos de escamotear la copia.

El profesor de Informática de la Universidad de Princeton Arvind Narayanan dice que las soluciones a estos problemas no deberían centrarse en el chatbot (hablar con un robot) en sí mismo, “sino en los incentivos perversos que conducen a este comportamiento, como que las universidades realicen revisiones de contratación y promoción contando trabajos

sin tener en cuenta su calidad o impacto”. Narayanan ha calificado a ChatGPT como un generador de tonterías (*bullshit generator*). En su opinión, este tipo de IA está entrenada para producir un texto plausible. Es muy bueno para ser persuasivo, pero no es capaz de producir declaraciones verdaderas. A menudo produce afirmaciones verdaderas como efecto secundario de ser plausible y persuasivo, pero ese no es el objetivo. Es un discurso que pretende persuadir sin tener en cuenta la verdad. En otros términos, el software funciona simplemente haciendo predicciones sobre qué decir en lugar de sintetizar ideas como hacen los cerebros humanos (12).

En línea con esta idea se encuentra lo expresado por Van Dis y colaboradores en *Nature* (13). Los autores pusieron a prueba la interacción de ellos con el ChatGPT en temas científicos específicos. Esto arrojó imprecisiones no menores (13). Parece que los mismos sesgos que a menudo desvían a los humanos, como los sesgos de disponibilidad, selección y confirmación, se reproducen y, a menudo, incluso se amplifican.

Entendemos que en campos donde la información falsa puede poner en peligro la seguridad de las personas, como la medicina, las revistas tendrán que tomar un enfoque más riguroso para verificar que la información sea precisa.

Entonces, ¿usted leería a un autor artificial?

La respuesta de este editorialista es sí, pero sabiendo que ChatGPT ha contribuido a su realización. El trabajo científico es una de las empresas más gratificantes. Amalgama la curiosidad, la imaginación y el descubrimiento. La IA tiene consecuencias de largo alcance para la ciencia y la sociedad, y es probable que el desarrollo de la modalidad conversacional modifique las prácticas de docencia, de investigación y de publicaciones, creando tanto oportunidades como preocupaciones.

En mi opinión, el uso de esta tecnología es inevitable, por lo tanto, prohibirla no funcionará. La atención debe centrarse en aprovechar la oportunidad y gestionar los riesgos. Un camino no suficientemente explorado (o al menos difundido) es la idea de priorizar el desarrollo y la implementación de tecnología de IA de código abierto (13). Esto ayudaría a desarrollar tecnologías de IA avanzadas de código abierto, transparentes y controladas democráticamente.

No es posible poner en duda la importancia de los avances científicos y tecnológicos y su impacto en la calidad de vida de los seres humanos. El mal uso de estos avances no son las “desventajas de la ciencia” sino las consecuencias del mal uso del conocimiento científico o de las nuevas tecnologías por parte de los humanos, que son los mismos que programan la IA y le asignan tareas que claramente no deberían estar en manos de la misma. Como ejemplo de esto, Narayanan menciona los programas para la identificación de personas que corren riesgo de abandonar la universidad,

para la detección de riesgos criminales y para la moderación de contenido en las redes sociales. Estos programas han tenido fallas notables. Las redes sociales se han utilizado para incitar a la violencia, incluso quizás a la violencia genocida, en muchas partes del mundo, como Myanmar, Sri Lanka y Etiopía. Todas estas fueron fallas en la moderación de los contenidos, incluida la IA como moderadora de esos contenidos (12).

Indiscutiblemente, como en otros ámbitos, creemos que los seres humanos siempre deben ser responsables de la práctica científica. Esto incluye las estructuras que los albergan y se desarrollan, las instituciones de investigación, las universidades, los comités de ética de investigación, etc. Los editores de revistas científicas deben adoptar políticas explícitas que generen conciencia y exijan transparencia sobre el uso de la IA conversacional en la preparación del material que formará parte de una publicación. Cada autor debe

ser éticamente responsable de verificar cuidadosamente su texto, resultados, datos, códigos y referencias.

Apenas estamos en los primeros capítulos de la primera temporada, que promete ser apasionante. Por ahora, sugerimos continuar leyendo los artículos de *Revista Medicina* en las redes, todos los cuales, hasta donde sabemos, han sido escritos por autores humanos.

Eduardo L. de Vito

Profesor titular de Medicina Interna, PhD

Instituto de Investigaciones Médicas Alfredo Lanari, Buenos Aires, Argentina
eldevito@gmail.com

Esta nota editorial se reproduce con la autorización de la Revista *Medicina* (Buenos Aires).

Referencias bibliográficas

1. Lüthy IA. Inteligencia artificial y aprendizaje de máquina en diagnóstico y tratamiento del cáncer. *Medicina (B Aires)* 2022;82:798-800
2. Alcañiz M, Chicchi Giglioli IA, Sirera M y col. Biomarcadores del trastorno del espectro autista basados en bioseñales, realidad virtual e inteligencia artificial. *Medicina (B Aires)* 2020;80:31-6
3. Chat GPT. Chat GPT Pro is here. Disponible en: <https://chat.openai.com/auth/login>. [Último acceso: diciembre de 2022.]
4. 20 bits. ¿Cuál es el sentido de la vida? La inteligencia artificial sorprende con lo contundente de su respuesta. Disponible en: <https://www.20minutos.es/tecnologia/cual-es-sentido-vida-inteligencia-artificial-sorprende-respuesta-5083405/>. [Último acceso: diciembre de 2022.]
5. ChatGPT. En: <https://en.wikipedia.org/wiki/ChatGPT>. [Último acceso: enero de 2023.]
6. Sample I. Science journals ban listing of ChatGPT as co-author on papers. *The Guardian* 26 de enero de 2023. Disponible en: <https://www.theguardian.com/science/2023/jan/26/science-journals-ban-listing-of-chatgpt-as-co-author-on-papers>
7. Machete. *Página/12*. Disponible en: <https://www.pagina12.com.ar/pirulo/521247>. [Último acceso: enero de 2023.]
8. Sociedad. Inteligencia artificial: una universidad francesa prohibió el uso de ChatGPT. *Página/12* 28 de enero de 2023
9. Gordijn B, Have HT. ChatGPT: evolution or revolution? *Med Health Care and Philos* 2023. Disponible en: <https://doi.org/10.1007/s11019-023-10136-0>.
10. Stokel-Walker C. ChatGPT listed as author on research papers: many scientists disapprove. *Nature* 2023;613:620-1
11. Else H. Abstracts written by ChatGPT fool scientists. *Nature* 2023;613:423
12. Angwin J. Decoding the hype about AI. The Marck-up. Disponible en: <https://themarkup.org/hello-world/2023/01/28/decoding-the-hype-about-ai>. [Último acceso: enero de 2023.]
13. Van Dis EAM, Bollen J, Zuidema W, et al. ChatGPT: five priorities for research. *Nature* 2023;614:224-6