



## El uso de la inteligencia artificial y las publicaciones científicas en la Revista Colombiana de Obstetricia y Ginecología

*Hernando Gaitán-Duarte, MD, MSc<sup>1</sup>.*

En la actualidad hay una gran atención por lo que representa la inteligencia artificial (IA) para nuestra sociedad, y las expectativas para el futuro a corto y mediano plazo son aún mayores. La Real Academia Española define la inteligencia como la capacidad de entender, comprender, o de resolver problemas. A su vez, la inteligencia artificial (IA) ha sido definida como “la ciencia y la ingeniería de fabricar máquinas inteligentes, especialmente programas informáticos inteligentes” (1). Entonces, podríamos decir que la IA es la ciencia y la ingeniería de fabricar máquinas que entienden, comprenden y resuelven problemas. Más recientemente, la International Business Machines Corporation (IBM), citando a Russell y Norvik (2), indica que la definición de IA tiene dos enfoques que se clasifican en función de la racionalidad y el pensamiento frente a la acción. Uno es el enfoque humano, del cual hacen parte los sistemas que piensan y actúan como humanos. El segundo, el enfoque ideal, en el que se ubican los sistemas que piensan y actúan racionalmente.

Entre las tecnologías más conocidas de la inteligencia artificial tenemos a los robots, que ya están disponibles no solo en los hospitales (3), sino que ya también se pueden comprar para uso domiciliario en sitios de ventas virtuales (4). Otros ejemplos de inteligencia artificial son los procesadores de voz como Siri o Google, o el programa ChatGPT de la casa de innovación tecnológica OpenAI® (5), un chatbot (se comunica por interacciones escritas),

que ha significado una revolución en términos de lograr que la máquina pueda no solo producir textos simples ante preguntas, sino que, además, pareciera que aprende rápidamente a partir de las solicitudes y de las respuestas previamente producidas. Ante estas novedades tecnológicas y la velocidad con la que se desarrolla la IA, hay todo tipo de reacciones que varían entre las expectativas ilimitadas sobre lo que viene, así como la negación, la incredulidad y el rechazo de estas nuevas máquinas. Lo importante es que ya están aquí y llegaron para quedarse, y que debemos conocer sus fortalezas y debilidades.

Para el mundo académico y científico, el desarrollo de ChatGPT ha significado un desafío, ya que el uso de esta aplicación podría ir en contra de la integridad científica, pues su uso podría determinar la presencia de plagio, bajo su más amplia definición. Según la American Psychologist Association (APA) (secciones 8.2 y 8.3) (6) el plagio es “el acto de presentar las palabras, ideas o imágenes de otro como propias; el negar a los autores que son creadores de contenidos el crédito que les corresponde”. Vamos a intentar explicar, desde el punto de vista de un neófito en estas tecnologías, algunos elementos que nos permitirán entender de dónde provienen los riesgos que representa esta aplicación y su manejo, según las recomendaciones del Comité Internacional de Editores de Revistas Médicas (International Committee of Medical Journal Editors - ICJME).

GPT es la sigla de Generative Pre-training Transformer. Es un modelo de procesamiento de lenguaje natural de aprendizaje automático (Large

1. Revista Colombiana de Obstetricia y Ginecología (RCOG), Bogotá (Colombia).

Language Model - LLM). Los LLM digieren grandes cantidades de datos de texto e infieren relaciones entre palabras dentro este. Se lanzó en 2019 y permite codificar, en algoritmos computacionales, textos de entrada y decodificar estos algoritmos para producir otros textos de salida. En su interior, la aplicación permite al modelo ponderar de manera diferencial partes de la secuencia de un texto para inferir significado y contexto. El transformador (una o varias redes neuronales) intenta comprender el modelado de lenguaje para “acertar” la relación entre palabras y producir respuestas más claras. En la versión lanzada en 2022 (GPT-3), el transformador se surte de libros, artículos, imágenes y páginas web, entre otros. Por otra parte, hace el proceso de selección de las piezas de texto de ingreso (llamadas token) que “considera” importantes para entender el contexto con base en ponderaciones de la importancia de los textos en el orden de salida, esto de manera iterativa (*self attention mechanisms*), con lo que logra hacer predicciones sobre el texto de salida, sin embargo, tiene limitaciones para alinear este texto con las necesidades del usuario (7). Actualmente se dispone como una aplicación (app).

El ChatGPT es un modo de aprendizaje profundo de máquinas, tiene la ventaja sobre sus predecesores de poder incorporar comentarios humanos en el proceso de capacitación para alinear mejor los resultados del modelo con la intención del usuario (*reinforcement learning*). Es decir, la máquina aprende a dar las respuestas más apropiadas según la especificidades del usuario y la calificación o retroalimentación que el humano da a estas respuestas. Sin embargo, estos programas no pueden generar algún tipo de razonamiento intelectual o modelo mental (8).

Ya para el caso que nos ocupa –las tareas académicas o la publicación de textos científicos– surgen varios problemas con el uso de ChatGPT, estos son:

- No es posible identificar qué parte de la información que emerge de los algoritmos es válida y cuál no.

- La máquina incorpora, además de las preferencias, los sesgos de los usuarios, como también puede producir textos sin sentido (alucinaciones).
- Además, dado que toma textos de libros o artículos y los mezcla, no identifica la fuente original de la información y, por lo tanto, no puede dar el crédito respectivo.
- Por otra parte, los textos de las solicitudes al ChatGPT pueden ser incorporados al texto emergente, es decir, pasan a ser parte de las ideas de dominio público.
- En el texto generado por el ChatGPT es imposible determinar si todo, o en qué magnitud, el contenido fue desarrollado como un modelo de texto y cuánto fue realizado por el ser humano; de esta manera es difícil decidir cuándo se debe dar el crédito al ChatGPT.

Estos últimos problemas llevan a la pérdida de los derechos de autoría, que son muy importantes para un investigador que presenta un hecho o una idea, la cual se incorpora en un campo del conocimiento. Como se mencionó, esto constituye un plagio y es un hecho que atenta contra la integridad científica. Actualmente se están desarrollando herramientas para detectar qué partes son escritas por el chat GPT, sin embargo su exactitud es aún incierta (9); por otra parte, hay otros chatbot, como Google Bard, que produce textos con referencias y las fuentes o que permite a la persona incluir las referencias en el texto. Sin embargo, la máquina no hace la verificación si la fuente de la cita es primaria o secundaria, la cual debe ser hecha por el ser humano (10).

Para evitar este problema, en la *Revista Colombiana de Obstetricia y Ginecología* (RCOG) explicamos a los autores la relevancia del uso adecuado de las citas directas (citas textuales), y de las citas indirectas (parfraseo, pero citando al autor original de la idea) (11), para otorgar el crédito requerido a los autores que originalmente presentan los hechos mencionados. Además, se prioriza el uso de las fuentes primarias (fuente original, donde por primera vez se expuso el hecho o idea), sobre las secundarias (citación a partir de documentos resumidos por otros), en las cuales, en

caso de ser utilizadas, se deberá referenciar la fuente original de la información (12).

En este sentido, y para evitar estas nuevas formas de plagio, a partir del próximo número la RCOG se sumará a la incitativa del ICJME (13), en cuanto a que se insta a los autores a declarar la utilización de la inteligencia artificial en el manuscrito presentado, ya que es el ser humano el responsable de la utilización de la IA en el documento. Además, los autores deben revisar cuidadosamente y editar el texto final, porque la IA puede generar resultados que parecen adecuados y que pueden ser incorrectos, incompletos o sesgados. Esta información debe ir tanto en la carta de presentación (página titular) como en la sección de materiales y métodos en el trabajo presentado. Los GPT y el ChatGPT no deben ser listados como autores o coautores ya que no cumplen con el criterio de responder por la integridad, validez y originalidad del trabajo, tampoco deben citar estas aplicaciones de IA en la bibliografía respectiva. Los autores deben poder asegurar que no hay plagio en el texto, en las citas o en las imágenes, y que hay una atribución apropiada de todo el material citado, incluidas citas completas.

Esperamos con esto hacer un aporte para que nuestros investigadores y estudiantes comprendan la responsabilidad que implica ser autores de un documento académico o científico, las limitaciones de estas máquinas y su adecuada utilización, en beneficio de la validez de la información que se brinda al lector y de la integridad científica.

## REFERENCIAS

1. McCarthy J. From here to human-level AI. *Artif Intell.* 2007;171(18):1174-82. <https://doi.org/10.1016/j.artint.2007.10.009>
2. International Business Machines Corporation. What is artificial intelligence (AI)? [Internet]. Disponible en: <https://www.ibm.com/topics/artificial-intelligence>
3. Morrell AL, Morrell Jr AC, Morrell AG, Freitas J, Tustumi F, De-Oliveira L, et al. The history of robotic surgery and its evolution: When illusion becomes reality. *Rev Col Bras Cir.* 2021;48:e20202798. <https://doi.org/10.1590/0100-6991e-20202798>
4. Alibaba.com. [Internet]. Disponible en: <https://www.alibaba.com/showroom/humanoid-robot.html>
5. Open AI. [Internet]. Disponible en: <https://chat.openai.com/auth/login>
6. American Psychological Association. APA Style. 7th ed. APA [Internet]. 2022. Disponible en: <https://apastyle.apa.org/products/publication-manual-7th-edition>
7. Ruby M. How ChatGPT works: The model behind the bot [Internet]. *Towards Data Science*; 2023. Disponible en: <https://towardsdatascience.com/how-chatgpt-works-the-models-behind-the-bot-1ce5fca96286>
8. Dien J. Editorial: Generative artificial intelligence as a plagiarism problem. *Biol Psychol.* 2023;181:108621. <https://doi.org/10.1016/j.biopsycho.2023.108621>
9. Demers, T. 16 of the best AI and ChatGPT content detectors compared. *Search Engine Land* [Internet]. (2023, April 25). Disponible en: <https://searchengineland.com/aichatgpt-content-detectors-395957>
10. Google Bard [Internet]. Disponible en: <https://bard.google.com/?hl=es>
11. Universidad de Puerto Rico. Sistema de bibliotecas. Manual APA 7a edición. Citas y referencias [Internet]. 2023. Disponible: <https://uprrp.libguides.com/c.php?g=985694&p=7256246>
12. Universidad de Guadalajara. Clasificación general de las fuentes de información. [Internet]. 2023. Disponible: <http://biblioteca.udgvirtual.udg.mx/portal/clasificacion-general-de-las-fuentes-de-informacion>
13. International Committee for Medical Journal Editors - ICMJE. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals [Internet]. 2023. Disponible en: <https://www.icmje.org/icmje-recommendations.pdf>