

CHAT GPT: PELIGRO DE CIENCIA-FICCIÓN

GPT CHAT: SCIENCE-FICTION DANGER

VIRGINIA INÉS SIMÓN ¹



Resumen

Hace poco tiempo en la red interna de comunicación de RENICS (Red Nacional de Información en Ciencias de la Salud), se conversó sobre ChatGTP y referencias bibliográficas fantasmas, generando una serie de testimonios con respecto a la detección de este fenómeno, junto con alucinaciones. Aquí comenzamos a entender el origen del título de este escrito que alerta sobre los nuevos peligros que acechan a la ciencia y que finalmente fueron el puntapié para imaginar el peligro de la ciencia-ficción.

Abstract

Recently, in the internal communication network of RENICS (National Health Sciences Information Network), there was discussion about ChatGTP and phantom bibliographic references, generating a series of testimonies regarding the detection of this phenomenon, along with hallucinations. Here we begin to understand the origin of the title of this writing that warns about the new dangers that threaten science and that finally were the kickstart to imagine the danger of science fiction.

Palabras clave: Inteligencia artificial

Keywords: Artificial Intelligence

¹ Virginia Inés Simón. Biblioteca del Centro Médico de Mar del Plata. Lic. en Bibliotecología y Documentación, Universidad Nacional de Mar del Plata. Diplomada en Discapacidad y Derechos Humanos, Pontificia Universidad Católica de Perú. Egresada CopyrightX Harvard Law School, the HarvardX. E-mail: biblioteca@centromedicomdp.org.ar

Fecha de recepción: 24/08/2023

Fecha de aceptación: 01/09/2023

Hace poco tiempo en RENICS (Red Nacional de Información en Ciencias de la Salud), en la red interna de comunicación, se compartió la publicación del artículo de Enrique Orduña-Malea y Alvaro Cabezas-Clavido titulado “*Chat GPT and the potential growing of ghost bibliographic references*”². A raíz de este artículo se generaron una serie de testimonios con respecto a la detección de alucinaciones y fantasmas. Aquí comenzamos a entender el origen del título de este escrito. Profesionales de información de distintos puntos de la provincia de Buenos Aires compartieron testimonios sobre la detección de artículos y citas falsas solicitadas por personal médico y de investigación de las instituciones, de alcance nacional, con las que trabajan. Se mencionaron casos en la biblioteca FLENI, Biblioteca del Hospital Italiano, Biblioteca del Instituto Universitario CEMIC y Biblioteca del Centro Médico de Mar del Plata. Como conclusión de estas conversaciones surgió una catarsis (y este escrito) de alerta sobre los nuevos peligros que acechan a la ciencia y que finalmente fueron el puntapié para imaginar el peligro de la ciencia-ficción.

ChatGPT (Chat Generative Pre-Trained Transformer) es una aplicación de [chatbot](#) de [inteligencia artificial](#) desarrollada en 2022 por [OpenAI](#)³. Se trata de un programa de computación que utiliza inteligencia artificial y procesamiento de lenguaje natural para entender preguntas y automatizar respuestas⁴ simulando una conversación humana. Sus principales ventajas radican en la posibilidad de uso en cualquier momento, prescindiendo de atención humana real, ahorrando tiempos y costos en atención al cliente⁵.

De alguna manera, los modelos de chatbot son instrucciones programadas de comunicación automatizada, que prevén una serie determinada de respuestas a preguntas que se realizan. Esta lógica aplicada a los modelos de búsqueda de información basadas en lenguaje natural permite

²https://link.springer.com/epdf/10.1007/s11192-023-04804-4?sharing_token=NSELq2jZ7AdV4hiYA_Vs7fe4RwlQNchNByi7wbcMAY4QCijpDp2vF80NYaK4T03ifomMVpemoMverw8D9GFdv03qa1GPPvfypfdJ_njl8xxdev-FPGJOTIQpbipnkndnuOetJFZvWxaWcUT4kW-o0u_tsghgNRyBRdj5bQ0zFFhQ%3D

³<https://openai.com/chatgpt>

⁴<https://www.ibm.com/topics/chatbots>

⁵<https://www.forbes.com/advisor/business/software/what-is-a-chatbot/>

imaginar cómo funciona ChatGPT, interpretando por entrenamiento y repetición de términos, las palabras claves que conforman una pregunta expresada en término coloquial o lenguaje libre.

A veces pareciera que suponemos sobre la inteligencia artificial, tecnologías que escriben por nosotros, investigan por nosotros, deciden por nosotros. Tecnologías que nos reemplazan. Esta primera impresión puede ser un espejismo engañoso para el futuro de la investigación científica.

La inteligencia artificial (IA) se refiere a aplicaciones que realizan tareas complejas para las que antes eran necesaria la intervención humana, como la comunicación en línea con clientes⁶. En forma más compleja, se trata de un campo de la ciencia que estudia la creación de pensamiento o razonamiento automatizado a través de las tecnologías. Se pueden identificar diferentes tipos de inteligencia artificial y aprendizaje. Este escrito solo aborda definiciones generales para dar un contexto a las reflexiones por delante.

¿Cómo se ha llegado a pensar que un tipo de tecnología puede responder a las preguntas elementales de una investigación o resolver un escrito? ¿Cuándo renunciamos al juicio ético, a la moral, al criterio de originalidad y al acto creativo que implican los procesos de toda escritura? Desde la aparición de Chat GPT y similares, se ha popularizado la idea de resolución de búsquedas de información a través de la modalidad de conversación con chat-bot. Si bien en principio los resultados parecen coherentes para preguntas sencillas, sobre por ejemplo ubicaciones o recetas, consultas más complejas arrojan resultados que sin conocimientos sobre el tema, son pasibles de una gran cadena de desinformación. Los errores recurrentes pueden corresponder a entradas o preguntas mal formuladas, pero también al exceso de creatividad de la propia aplicación, que genera lo que se denomina como “alucinaciones”, un fenómeno de respuestas realistas pero falsas, tanto en el contenido de respuesta en texto como en las citas bibliográficas que puede generar.

Hace algunos meses se presentaba en The Scholarly Kitchen, un texto titulado “Academic Publishers Are Missing the Point on ChatGPT⁷”. En él, Avi Staiman presenta el contexto del monopolio editorial con respecto a la irrupción de la inteligencia artificial en la ciencia. Con respecto a la posibilidad de modelos de inteligencia artificial como autores o coautores de trabajos académicos, los principales monopolios editoriales se manifestaron rápido excluyendo cualquier

⁶ <https://www.oracle.com/ar/artificial-intelligence/what-is-ai/>

⁷ <https://scholarlykitchen.sspnet.org/2023/03/31/guest-post-academic-publishers-are-missing-the-point-on-chatgpt/>

mecanismo de inteligencia artificial como parte de la identificación de autores en un equipo de trabajo⁸. Así por ejemplo, Science indica que

“El texto generado a partir de IA, aprendizaje automático o herramientas algorítmicas similares no se puede utilizar en artículos publicados en revistas científicas, ni las figuras, imágenes o gráficos adjuntos pueden ser productos de dichas herramientas, sin el permiso explícito de los editores. Además, un programa de IA no puede ser autor de un artículo de revista científica. Una violación de esta política constituye una mala conducta científica.”⁹

Es interesante el caso de ELSEVIER, que intenta adaptarse a la problemática al tiempo de desligarse de las responsabilidades que pueden acarrear los errores de este tipo de tecnologías. Así se interpreta en algunas de las siguientes afirmaciones:

“La revisión por pares está en el corazón del ecosistema científico y Elsevier cumple con los más altos estándares de integridad en este proceso. Gestionar la evaluación editorial de un manuscrito científico implica responsabilidades que solo pueden ser atribuidas a humanos. Los editores no deben utilizar tecnologías de IA generativa o asistidas por IA para ayudar en el proceso de evaluación o toma de decisiones de un manuscrito, ya que el pensamiento crítico y la evaluación original necesarios para este trabajo están fuera del alcance de esta tecnología...”

Con respecto a la posibilidad de escribir con tecnologías basadas en inteligencia artificial:

“...La política de autor de IA de Elsevier¹⁰ establece que los autores pueden usar IA generativa y tecnologías asistidas por IA en el proceso de escritura antes del envío, pero solo para mejorar el lenguaje y la legibilidad de su artículo y con la divulgación adecuada, según nuestras instrucciones en la Guía de Elsevier para Autores...”

En su guía para autores¹¹ se detalla que:

“Declaración de IA generativa en la escritura científica

La siguiente guía solo se refiere al proceso de redacción y no al uso de herramientas de IA para analizar y extraer información de los datos como parte del proceso de investigación...

...La IA y las tecnologías asistidas por IA no deben figurar como autor o coautor, ni citarse como autor. La autoría implica responsabilidades y tareas que solo

⁸ <https://scholarlykitchen.sspnet.org/2023/03/31/guest-post-academic-publishers-are-missing-the-point-on-chatgpt/>

⁹ Traducción personal.

¹⁰ <https://www.elsevier.com/about/policies/publishing-ethics>

¹¹ <https://www.elsevier.com/journals/journal-of-biotechnology/0168-1656/guide-for-authors>

pueden ser atribuidas y realizadas por humanos, como se describe en la política de IA de Elsevier para autores...”

“...Los autores deben revelar en su manuscrito el uso de IA y tecnologías asistidas por IA en el proceso de escritura siguiendo las instrucciones a continuación. Una declaración aparecerá en el trabajo publicado. Tenga en cuenta que los autores son, en última instancia, responsables del contenido del trabajo...”

Nature ha sido inflexible en su posición, afirmando este mismo año que herramientas como ChatGPT amenazan la ciencia transparente¹². Entre sus posiciones puede leerse

“La gran preocupación en la comunidad de investigación es que los estudiantes y los científicos puedan hacer pasar engañosamente el texto escrito por LLM como propio, o usar LLM de manera simplista (como para realizar una revisión de literatura incompleta) y producir un trabajo que no sea confiable. Varios preprints y artículos publicados ya han acreditado ChatGPT con autoría formal.”

...

“Nature, junto con todas las revistas de Springer Nature, ha formulado los siguientes dos principios, que se han agregado a nuestra guía existente para autores (ver go.nature.com/3j1jxsw)...”

“En primer lugar, no se aceptará ninguna herramienta LLM como autor acreditado en un trabajo de investigación. Esto se debe a que cualquier atribución de autoría conlleva responsabilidad por el trabajo, y las herramientas de IA no pueden asumir tal responsabilidad.

En segundo lugar, los investigadores que utilizan herramientas LLM deben documentar este uso en las secciones de métodos o agradecimientos...”

La mayoría de los editores coinciden en que las aplicaciones tecnológicas de inteligencia artificial no pueden tener ningún tipo de asignación de autoría, principalmente en razón de no ser personas humanas y por ello, no ser pasibles de asumir las consecuencias y responsabilidades del sistema de propiedad intelectual. En efecto, el hecho de no tener autoconciencia sobre la producción de contenidos implica que tampoco tienen la capacidad, ni el derecho, de decidir sobre los límites de sus creaciones en cuanto a la asignación consciente de licencias de derecho de autor en toda su diversidad. Valga decir que la aparición de una aplicación de inteligencia artificial en la mención de autoría de un escrito no es una autodeterminación de la propia aplicación sino la decisión arbitraria de un ser humano, que restringe en forma sistematizada el derecho a elegir, a la determinación del alcance de los derechos sobre la obra, a esta nueva entidad autoral, que no tiene

¹² <https://www.nature.com/articles/d41586-023-00191-1>

la capacidad de decidir por sí misma qué tipo de licencia de derecho de autor quisiera en su contribución.

Investigar y escribir textos académicos implican sin dudas un acto creativo por sí mismo. La producción científica conlleva los criterios de creación original, verificabilidad y contribución científica. El tiempo que se emplea en la redacción del escrito, entendido este como el mecanismo de afirmación de conocimiento y divulgación de la ciencia, no debería ser considerado una pérdida de tiempo para permitir que una tecnología reemplace el ingenio humano en su producción.

En la medida en que el uso de las tecnologías basadas en inteligencia artificial pueda apoyar las prácticas investigativas sin reemplazar a la investigación y la escritura como actos en sí mismos, su uso puede ser beneficioso. Sin embargo, en tiempos de sobre exposición digital, masificación de información no verificada y generación constante de contenidos falsos, debe ponderarse un criterio juicioso y ético sobre el uso versus abuso de tecnologías para la producción académica. De lo contrario, nos alejaremos poco a poco del concepto de ciencia para confluir en una ciencia-ficción.

Referencias

El Cronista. (2023, June 8). *Los creadores de ChatGPT advierten sobre el mayor riesgo de la Inteligencia Artificial: a qué le temen*. El Cronista. Retrieved August 24, 2023, from <https://www.cronista.com/infotechnology/actualidad/los-creadores-de-chatgpt-advierten-sobre-el-mayor-riesgo-de-la-inteligencia-artificial/>

Elsevier. (2023). *Guide for authors - Journal of Biotechnology - ISSN 0168-1656*. Elsevier. Retrieved August 24, 2023, from <https://www.elsevier.com/journals/journal-of-biotechnology/0168-1656/guide-for-authors>

Elsevier. (2023). *Home About Policies Publishing ethics*. Elsevier. Retrieved August 24, 2023, from <https://www.elsevier.com/about/policies/publishing-ethics>

IBM. (n.d.). *What is a chatbot?* IBM. Retrieved August 11, 2023, from <https://www.ibm.com/topics/chatbots>

Nature. (2023, Enero 24). *Tools such as ChatGPT threaten transparent science; here are our ground rules for their use*. Nature. Retrieved Agosto 24, 2023, from <https://www.nature.com/articles/d41586-023-00191-1>

OpenAI. (n.d.). *ChatGPT*. OpenAI. Retrieved August 11, 2023, from <https://openai.com/chatgpt>

Oracle. (n.d.). *¿Qué es la inteligencia artificial (IA)?* Oracle. Retrieved August 24, 2023, from <https://www.oracle.com/ar/artificial-intelligence/what-is-ai/>

Orduña-Malea, E., & Cabezas-Clavijo, Á. (2023). ChatGPT and the potential growing of ghost bibliographic references. *Scientometrics*, 128, 5351-5355. doi.org/10.1007/s11192-023-04804-4

Staiman, A., & Willis, M. (2023, March 31). *Guest Post - Academic Publishers Are Missing the Point on ChatGPT - The Scholarly Kitchen*. The Scholarly Kitchen. Retrieved August 24, 2023, from <https://scholarlykitchen.sspnet.org/2023/03/31/guest-post-academic-publishers-are-missing-the-point-on-chatgpt/>

Watts, R., & Shweta, K. M. (2022, August 21). *What Is A Chatbot? Everything You Need To Know*. Forbes. Retrieved August 24, 2023, from <https://www.forbes.com/advisor/business/software/what-is-a-chatbot/>