

Citas fantasmas en artículos científicos: problemática creciente ante el uso de la inteligencia artificial

Ghost citations in scientific articles: a growing problem in the face of the use of artificial intelligence

Luis Ernesto Paz-Enrique^{1*}  <https://orcid.org/0000-0001-9214-3057>

¹ Universidad Nacional Autónoma de México. Ciudad de México, México.

* Autor para la correspondencia: luisernestopazenrique@gmail.com

Recibido: 22/09/2023.

Aceptado: 28/09/2023.

En la última década, el campo de la inteligencia artificial (IA) ha experimentado avances significativos. Uno de los desarrollos más notables es el modelo de lenguaje Generative Pre-trained Transformer (GPT), que ha revolucionado la forma en que las máquinas pueden interactuar con los humanos. Su capacidad para comprender, generar y responder texto ha llevado a su aplicación en una amplia variedad de campos, incluyendo la redacción de artículos y documentos científico-técnicos.

Antes de la llegada de los sistemas de ChatGPT, la redacción de artículos científicos y documentos técnicos solía ser un proceso laborioso y lento. Los investigadores y expertos debían invertir mucho tiempo y esfuerzo en la redacción y revisión de estos textos. Sin embargo, con el advenimiento de los sistemas de chat basados en GPT, se ha vuelto posible automatizar gran parte de este proceso. Los algoritmos de IA pueden extraer y procesar información de múltiples fuentes, como revistas científicas, bases de datos y libros, para proporcionar una visión general de un tema en particular.⁽¹⁾ Además, pueden identificar patrones y tendencias en los datos, lo que puede resultar



útil para los investigadores al analizar resultados experimentales o realizar revisiones sistemáticas.

Los modelos de lenguaje GPT están entrenados con grandes cantidades de texto de diferentes dominios, lo que les permite entender y generar argumentos en lenguaje natural. Esto significa que pueden ser programados para redactar expresiones científicas y técnicas con relativa facilidad. Al entrenar el modelo con corpus científicos relevantes, los chats pueden generar contenido coherente y bien estructurado, siguiendo las convenciones y el estilo de escritura requeridos en determinadas áreas, dominios y campos del conocimiento. Otra ventaja de utilizar la IA en la redacción científica es su habilidad para eliminar sesgos y errores humanos como imprecisiones en la representación de los datos o la omisión de información relevante. La IA puede ayudar a minimizar estos errores al aplicar algoritmos de verificación y corrección automática.

Una de las principales ventajas del uso de ChatGPT en la redacción científica y técnica es la capacidad de sistematizar mucha información. Estos sistemas tienen acceso a vastas bases de conocimiento; sin embargo, no es así para proporcionar referencias y datos relevantes para respaldar sus argumentos.⁽²⁾ También son capaces de seguir rigurosamente las convenciones de formato requeridas para presentar información científica de manera clara y concisa. Asimismo, la IA puede ayudar a superar bloqueos creativos o dificultades en la expresión de ideas. Al proporcionar sugerencias y opciones cuando se les solicita, estos sistemas pueden ser una herramienta valiosa para estimular nuevas ideas y enfoques en la escritura científica.

Pese a las ventajas mencionadas, es importante destacar que el uso de ChatGPT en la redacción científica y técnica tiene muchos desafíos. Ningún modelo puede garantizar la precisión de los datos e información, debido a que su principal función es triangular la producción científica. En consecuencia, cada texto debe ser revisado y editado por expertos humanos antes de ser publicados o presentados. Para redactar artículos y material científico, la IA no reemplaza a los investigadores o escritores humanos. Más bien, se trata de una herramienta complementaria que puede ayudar a agilizar el proceso de redacción y mejorar la calidad de los documentos resultantes.

Otro de los principales desafíos, es la necesidad de entrenar adecuadamente los algoritmos de IA para garantizar que generen contenido de calidad y con mayor precisión. Los algoritmos deben aprender a distinguir entre la información confiable y la que no lo es, así como a evaluar la validez y relevancia de los resultados obtenidos.⁽¹⁾ Esto requiere una gran cantidad de datos de entrenamiento y la participación de expertos en el campo. Se adiciona a la problemática la pérdida de la subjetividad y la creatividad inherentes a la redacción humana. Aunque la IA puede generar contenido científicamente con relativa precisión, carece de la capacidad para comprender completamente el contexto social o cultural en el que se realiza la investigación. Esto puede resultar en la falta de sutilezas o interpretaciones subjetivas que los investigadores pueden aportar a la redacción científica.

Las citas y referencias en los artículos científicos desempeñan un papel crucial en el proceso de creación de conocimiento. Actúan como una forma de respaldo y validación de los argumentos presentados por los investigadores, permitiendo que otros científicos puedan rastrear la información utilizada y construir sobre ella.⁽³⁾ Son una forma de reconocer y dar crédito a los trabajos previos que han servido como base para el artículo en cuestión. Al incluir estas referencias, los autores indican que se han



basado en investigaciones anteriores y que están contribuyendo a la acumulación de conocimiento científico de la comunidad. Esto no solo es un gesto de cortesía hacia los autores anteriores, sino que también ayuda a los investigadores a crear un diálogo continuo y a respetar la autoría intelectual.

Las citas y referencias brindan a los lectores la oportunidad de examinar las fuentes utilizadas en el artículo y evaluar la solidez del argumento presentado. Al proporcionar referencias, los investigadores permiten que otros puedan investigar más a fondo y confirmar o refutar los hallazgos presentados, por lo que tiene función de verificabilidad.⁽⁴⁾ Esto fomenta la transparencia y la confianza en el proceso científico, ya que otros investigadores pueden replicar los experimentos o análisis utilizados, lo que ayuda a garantizar la rigurosidad y validez de los resultados.

De igual forma, ayudan a los investigadores a construir una red de conocimiento científico interconectado. Al incluir referencias a investigaciones previas, los autores pueden mostrar cómo su estudio se relaciona con las teorías existentes, identificar las lagunas en el conocimiento actual y sugerir nuevas direcciones de investigación. Esto permite a la comunidad científica avanzar y refinar sus teorías, al tiempo que evita la duplicación innecesaria de esfuerzos de investigación.

La construcción de artículos científicos requiere de una extensa investigación, análisis crítico y síntesis de información relevante. Uno de los aspectos fundamentales en la construcción de estos artículos es la citación adecuada de fuentes. Esto es esencial para respaldar las afirmaciones y resultados presentados, al igual que para dar crédito a otros autores cuyos trabajos han sido utilizados como base para desarrollar nuevos conocimientos.

Las citas fantasmas en la comunicación científica se refieren a una problemática cada vez más común en la actualidad. Esto ocurre cuando los autores incluyen referencias bibliográficas en sus artículos científicos, pero sin haber consultado realmente los trabajos a los que hacen referencia. Esto puede tener consecuencias negativas en la calidad y veracidad de la información científica presentada.

Una de las principales causas de las citas fantasmas es la falta de tiempo o recursos por parte de los autores para revisar exhaustivamente la literatura existente.⁽⁵⁾ En muchos casos, los investigadores se sienten presionados para publicar rápidamente y cumplir con los estándares de productividad académica, lo que puede llevarlos a incluir citas sin haber verificado adecuadamente su contenido. La citación incorrecta puede llevar a la difusión de información errónea e incluso a la perpetuación de errores en la literatura científica. Algunas revistas científicas fomentan la inclusión de un alto número de autocitas en los artículos que publican, ya que esto les permite aumentar su factor de impacto. Esto ha llevado a que algunos autores y editores colaboren en la inclusión de citas fantasmas en los trabajos, con el objetivo de incrementar su visibilidad y reconocimiento en la comunidad científica.

Las consecuencias de las citas fantasmas son significativas. En primer lugar, atentan contra la confianza y el rigor científico. Al hacer referencia a autores o fuentes que no existen, se genera una distorsión de la bibliografía utilizada en la investigación, lo que puede llevar a errores en la interpretación de los resultados o a la repetición de investigaciones previas que no han sido correctamente atribuidas.⁽⁶⁾ Además, pueden afectar la reputación y la credibilidad de los investigadores y las instituciones involucradas. Asimismo, las revistas científicas que publican artículos con citas



fantasmas pueden perder su reputación y dejar de ser consideradas como fuentes confiables.

Las políticas editoriales pueden contribuir a minimizar las citas fantasmas. Algunas revistas exigen a los autores citar artículos de las principales bases de datos y en idioma inglés. Estas políticas, si bien buscan mejorar la calidad y visibilidad de las investigaciones, también están generando una serie de problemas éticos y académicos.

En primer lugar, las políticas editoriales que priorizan las citas en bases de datos específicas limitan la diversidad y el alcance de las lecturas y referencias utilizadas por los autores. Esto puede llevar a la pérdida de perspectivas y resultados de investigaciones valiosas que podrían no estar presentes en las bases de datos de corriente principal. Además, el incentivar la exclusividad de artículos escritos en inglés puede ocasionar una falta de reconocimiento de investigadores y trabajos científicos en otros idiomas.

Las políticas editoriales que exigen una cantidad específica de citas pueden llevar a la creación de citas fantasmas. Esto ocurre cuando los autores incluyen citas sin haber leído, comprendido o realmente utilizado los artículos referenciados. Esta práctica, además de ser antiética, distorsiona el conocimiento científico y la validez de los resultados. Muchas veces los investigadores se ven obligados a citar determinados artículos para ser considerados para publicación en revistas de prestigio.

La IA en los últimos años es una de las técnicas que más se ha utilizado para generar citas fantasmas en artículos científicos.⁽²⁾ Los sistemas de generación de lenguaje natural, alimentados con grandes cantidades de datos científicos, pueden generar párrafos y citas que se asemejen a las de un experto en el campo. Esto plantea la posibilidad de que los académicos o los editores de revistas poco éticos puedan utilizar esta tecnología para mejorar la calidad aparente de sus artículos.

La información científica es crucial para tomar decisiones y avanzar en el conocimiento; es esencial poder confiar en la integridad de los artículos publicados. Las citas y referencias son uno de los pilares fundamentales de esta confianza, ya que permiten a los lectores y revisores rastrear la evidencia y las fuentes utilizadas para respaldar las afirmaciones hechas en un documento determinado. En consecuencia, las citas fantasmas entorpecen el avance del conocimiento y le resta credibilidad.

La generación de citas y referencias de calidad implica la capacidad de la IA para validar la información recopilada. A menudo, los algoritmos de IA basados en aprendizaje automático extraen información de diversas fuentes, lo que puede dar lugar a citas erróneas o referencias incorrectas.⁽⁷⁾ La IA no es capaz de evaluar la confiabilidad de las fuentes utilizadas y verificar la exactitud de la información antes de generar una cita o referencia.

La generación de citas y referencias también requiere una comprensión profunda del contexto en el que se citan las fuentes. La IA debe ser capaz de identificar el tipo de fuente (libro, revista, conferencia, etc.), así como los detalles específicos del texto citado (autor, título, año, etc.). Las dificultades pueden surgir cuando las fuentes citadas no siguen un formato uniforme o cuando se utilizan abreviaturas o referencias cruzadas que requieren una comprensión avanzada del texto.

Garantizar la originalidad y evitar el plagio es esencial en la generación de citas y referencias. La IA debe tener la capacidad de comparar las fuentes citadas con otras



disponibles en bases de datos y verificar su originalidad. Esto puede ser especialmente complejo cuando se trabaja con idiomas diferentes o con fuentes menos conocidas. Además, la IA no tiene la capacidad de identificar citas y referencias que pueden ser atribuidas a una fuente en particular y no a una serie de fuentes diversas.

La generación de citas y referencias se ve afectada por la necesidad de una base de datos actualizada y completa.⁽⁸⁾ A medida que aparecen nuevas fuentes y se desarrollan nuevas técnicas de citación, la IA debe actualizarse constantemente para garantizar citas y referencias precisas y acordes con los estándares actuales. Esto implica la incorporación de nuevas fuentes, pautas de estilo y reglas de citación, lo que representa un desafío en términos de implementación y mantenimiento de la IA.

La detección de citas fantasmas generadas por IA puede ser un desafío, especialmente si el contenido generado es convincente y bien estructurado. Sin embargo, existen métodos de verificación que podrían ayudar a identificar la presencia de estas citas falsas. Por ejemplo, el análisis de la coherencia y consistencia en el uso de la terminología científica podría ayudar a detectar discrepancias entre las citas y el contenido restante del artículo. Además, los investigadores y editores pueden utilizar herramientas de detección de plagio para buscar similitudes con otras publicaciones ya existentes.

Para enfrentar este problema, es necesario implementar medidas que promuevan la integridad científica. En primer lugar, se debe fomentar una cultura de transparencia y honestidad en la elaboración de artículos científicos. Los investigadores deben ser conscientes de la importancia de citar correctamente y de dar reconocimiento adecuado a las fuentes utilizadas en su investigación y dedicar tiempo y recursos suficientes para constatar el uso adecuado de las referencias bibliográficas incluidas en sus artículos.

Es fundamental que las revistas científicas establezcan políticas claras y estrictas en cuanto a la inclusión de citas en los artículos que publican. Se debe incentivar la revisión rigurosa de la bibliografía utilizada y verificar que todas las referencias sean reales y relevantes para el trabajo presentado. En adición, las instituciones académicas deben desarrollar programas de capacitación en ética científica que destaquen la importancia de la revisión de las referencias bibliográficas.

REFERENCIAS BIBLIOGRÁFICAS

1. Ros P, Pérez A. ChatGPT: una novedosa herramienta de escritura para artículos científicos, pero no un autor (por el momento). Rev Neurol [Internet]. 2023 [citado 12/09/2023]; 76(8):263-77. Disponible en: <http://dx.doi.org/10.33588/rn.7608.2023066>
2. González E, Calvo P. Ethically governing artificial intelligence in the field of scientific research and innovation. Heliyon [Internet]. 2022 [citado 22/09/2023]; 8(2):e08946. Disponible en: <https://doi.org/10.1016/J.HELIYON.2022.E08946>
3. Calo L. Métricas de impacto y evaluación de la ciencia. Rev Peruana Medicina Experimental y Salud Pública [Internet]. 2022 [citado 19/09/2023]; 39(2):236-40. Disponible en: <https://doi.org/10.17843/rpmesp.2022.392.11171>



4. Santos K, Melo RR de, Correia AEGC, et al. Retratos e ainda citados: perfil de citações pós-retratação em artigos de pesquisadores brasileiros. Em Quest [Internet]. 2023 [citado 14/09/2023]; 29(2): e-125494. Disponible en: <https://doi.org/10.19132/1808-5245.29.125494>
5. Biswas SS. Role of ChatGPT in Public Health. Ann Biomed Eng [Internet]. 2023 [citado 22/09/2023]; 51(4): 868-9. Disponible en: <https://doi.org/10.1007/s10439-023-03172-7>
6. Lund BD, Wang T. Chatting about ChatGPT: how may AI and GPT impact academia and libraries? Library Hi Tech News [Internet]. 2023 [citado 22/09/2023]; 40(3): 26-9. Disponible en: <https://doi.org/10.1108/LHTN-01-2023-0009>
7. Roa LI, Díaz T, Estrada L. Gestores de referencias bibliográficas y su impacto en las investigaciones. E-Ciencias de la Información [Internet]. 2022 [citado 17/09/2023]; 12(1): 96-113. Disponible en: <http://dx.doi.org/10.15517/eci.v12i1.47067>
8. Crisci V, Katinas L. Las citas bibliográficas en la evaluación de la actividad científica: significado, consecuencias y un marco conceptual alternativo. Bol Soc Argent Bot [Internet]. 2020 [citado 6/09/2023]; 55(3): 1-10. Disponible en: <https://doi.org/10.31055/1851.2372.v55.n3.28723>

CÓMO CITAR ESTE ARTÍCULO

Paz-Enrique LE. Citas fantasmas en artículos científicos: problemática creciente ante el uso de la inteligencia artificial. Rev Méd Electrón [Internet]. 2023 Nov.-Dic. [citado: fecha de acceso]; 45(6). Disponible en: <http://www.revmedicaelectronica.sld.cu/index.php/rme/article/view/5366/5710>

