

Editorial



ChatGPT y el Sesgo en las Citas Científicas

El Efecto Mateo es un caso ampliamente estudiado en la publicación científica, en el cual se refiere a la tendencia de los investigadores exitosos y reconocidos a recibir más citas en sus publicaciones científicas, mientras que aquellos menos populares o con menor número de citas tienen dificultades para lograr el reconocimiento de sus trabajos. En consecuencia, los investigadores exitosamente reconocidos reciben más recursos, lo que les permite producir más trabajos de alta calidad y ser citados con mayor frecuencia en un ciclo cerrado [1].

El origen del término "Efecto Mateo" se vincula a la bíblica cristiana en el Evangelio de Mateo, que señala: "Porque a todo el que tiene, se le dará, y tendrá en abundancia; pero al que no tiene, aun lo que tiene se le quitará" (Mateo 25:29). Este término fue acuñado por el sociólogo de la ciencia Robert K. Merton, quien lo adoptó para describir la desigualdad en la distribución de reconocimiento y citas en la publicación científica.

La llegada de inteligencias generadoras de texto como ChatGPT, Bing, y Bard han causado nuevas preguntas sobre cómo esta tecnología puede influir en la proliferación del Efecto Mateo. Como se ha indicado en los últimos meses, ChatGPT es una inteligencia artificial (IA) generativa que ha ganado popularidad, siendo actualmente utilizada por más de 100 millones de personas alrededor del planeta. No obstante, la compañía que la ha creado no explicita sobre las fuentes de información que utiliza para brindar las respuestas, por lo que pueden estar sujetas a algún tipo de sesgo.

Un estudio reciente publicado en el servidor de pre impresión arXiv [2], buscó abordar esta pregunta investigando las fuentes de información utilizadas por ChatGPT en el campo de la

ciencia ambiental. Los investigadores pidieron a ChatGPT que identificara las diez subdisciplinas más importantes dentro del campo de la ciencia ambiental y que compusiera un artículo de revisión científica sobre cada subdisciplina, incluyendo 25 referencias. Luego, analizaron estas referencias, centrándose en factores como el número de citas, la fecha de publicación y la revista en la que se publicó el trabajo [3].

Los hallazgos del estudio indican que ChatGPT tiende a citar publicaciones altamente referidas en la ciencia ambiental y se usa predominantemente en las revistas reconocidas en el campo, siendo Nature la más citada. Si bien ChatGPT no es explícito en esto, los resultados sugieren que sus respuestas dependen principalmente de fuentes de Google Scholar, en lugar de utilizar información de citas de otras bases de datos científicas, como Web of Science o Scopus. No obstante, es comprensible ya que gran parte de los contenidos de Elsevier, Springer, IEEE y otras editoriales están sujetas a derechos y funcionan por suscripción para su acceso.

Esto plantea preocupaciones éticas en términos de equidad y justicia en la academia. Si el uso generalizado de ChatGPT para la generación de contenido académico perpetúa el efecto Mateo, se corre el riesgo de mantener un sistema desigual donde los investigadores ya establecidos se benefician más y los menos conocidos se quedan rezagados.

Adicionalmente, en la experiencia propia de los autores de esta editorial, hemos encontrado que ChatGPT en algunos casos genera referencias incompletas o totalmente resultados de sus alucinaciones. Muchas referencias no existen o no están relacionadas con el tema, cosa que solo un conocedor en el tema podría detectar. Se debe recordar, que de la misma manera que genera texto, genera algunas veces las referencias de la fuente que dice haber usado. Un coctel para el desastre en la publicación científica [4].

Para abordar este problema ético, es fundamental que se promueva la diversidad y la igualdad de oportunidades en la academia. Los investigadores menos reconocidos deberían tener la oportunidad de tener sus trabajos y contribuciones debidamente reconocidos, sin verse desfavorecidos por el efecto Mateo amplificado por el uso de tecnologías como ChatGPT.

Se necesitan políticas y prácticas que fomenten la visibilidad y el reconocimiento equitativo en la comunidad académica. Esto implica considerar medidas como una selección más equilibrada de las fuentes de información utilizadas por ChatGPT, promover la colaboración entre investigadores establecidos y emergentes; garantizar la transparencia en los criterios de reconocimiento y citación en la investigación generada por IA.

Dado que aún no se tiene control sobre la forma en que ChatGPT y otras inteligencias presentan sus citas cuando se le solicitan, es recomendable mantener apertura y seguir consultando de manera regular las revistas, repositorios y contenidos directamente sin la intervención de las inteligencias artificiales.

La revista Scientia et Technica respalda el uso de herramientas generadoras de texto en tareas rutinarias, como revisión gramatical o realización de tablas, pero siempre bajo la supervisión y validación de un experto humano. Dado que aún no se tiene control sobre cómo ChatGPT y otras inteligencias generan sus citas, es recomendable mantener la apertura y consultar directamente revistas, repositorios y contenidos sin la intervención de inteligencias artificiales.

Referencias:

- [1]. “The matthew effect in science, II - JSTOR, <https://www.jstor.org/stable/234750> (accessed Jun. 13, 2023).
- [2]. E. Petiska, “Chatgpt cites the most-cited articles and journals, relying solely on google scholar’s citation counts. as a result, AI may amplify the Matthew Effect in Environmental Science,” arXiv.org, <https://arxiv.org/abs/2304.06794#:~:text=13%20Apr%202023%5D-,ChatGPT%20cites%20the%20most%2Dcited%20articles%20and%20journals%2C%20relying%20solely,Matthew%20Effect%20in%20environmental%20science> (accessed Jun. 12, 2023).
- [3]. S. X, “CHATGPT’s citation approach may amplify the Matthew Effect in Environmental Science,” Phys.org, <https://phys.org/news/2023-04-chatgpt-citation-approach-amplify-matthew.html> (accessed Jun. 12, 2023).
- [4]. A. Welborn, “CHATGPT and fake citations,” Duke University Libraries Blogs, <https://blogs.library.duke.edu/blog/2023/03/09/chatgpt-and-fake-citations/> (accessed Jun. 12, 2023).

Autor

PhD. **Jimmy Alexander Cortes Osorio** 
Docente Titular Departamento de Física
Editor jefe -Revista Scientia et Technica
Grupo Investigación Robótica Aplicada Línea:
Computer Vision and Machine Learning
Investigador senior reconocido por
MINCIENCIA
Universidad Tecnológica de Pereira
ORCID: <https://orcid.org/0000-0002-0413-807X>