Revista Médica ELECTRÓNICA
Rev Med Electrón
Vol. 45. No. 6. Nov-Dic. 2023
ISSN: 1684-1824

Editorial

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

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

Luis Ernesto Paz-Enrique[1*] https://orcid.org/0000-0001-9214-3057

[1] National Autonomous University of Mexico. Mexico City, Mexico.

[*] Corresponding author: luisernestopazenrique@gmail.com

Received: 2023/09/22.

Accepted: 2023/09/28.

In the last decade, the field of artificial intelligence has seen significant advances. One of the most notable developments is the language model Generative Pre-trained Transformer (GPT), which has revolutionized the way machines can interact with humans. Its ability to understand, generate and respond to text has led to its application in a wide variety of fields, including the redaction of scientific-technical articles and documents.

Before the advent of ChatGPT systems, the writing of scientific articles used to be a laborious and slow process. Researchers and experts had to invest a lot of time and effort in the writing and revision of these texts. However, with the advent of GPT-

based chat systems, it has become possible to automate much of this process. AI algorithms can extract and process information from multiple sources, such as scientific journals, databases, and books, to provide an overview of a particular topic[1]. Moreover, they can identify patterns and trends in the data, which can be useful for researchers when analyzing experimental results or conducting systematic reviews. Another advantage of using AI in scientific writing is its ability to eliminate human biases and errors such as inaccuracies in the representation of data or omission of relevant information. AI can help minimize these errors by applying automatic verification and correction algorithms.

One of the main advantages of using ChatGPT in scientific and technical writing is the ability to systematize a lot of information. These systems have access to vast knowledge bases; however, they are not able to provide relevant references and data to support their arguments[2]. They are also able to follow the format conventions required to present scientific information ina clear and concise manner. Likewise, AI can help overcome creative blockages in the expression of ideas. By providing suggestions and options when requested, these systems can be a valuable tool to stimulate new ideas and approaches in scientific writing.

Despite the mentioned advantages, it is important to highlight that the use of ChatGPT in scientific and technical writing has many challenges. No model can guarantee the accuracy of data and information, because its main function is to triangulate scientific production. Consequently, each text must be reviewed and edited by human experts before being published or submitted. For writing articles and scientific material, AI does not replace human researchers or writers. Rather, it is a complementary tool that can help speed up the writing process and improve the quality of the resulting documents.

Another of the main challenges is the need to properly train AI algorithms to ensure that they generate quality content and with greater accuracy. Algorithms must learn to distinguish between reliable and non-reliable information, as well as to assess the validity and relevance of the results obtained[1]. This requires a large amount of training data and the participation of experts in the field. Added to the problem is the loss of subjectivity and creativity inherent to human writing. Although AI can generate content scientifically with relative accuracy, it lacks the ability to fully understand the social or cultural context in which the research is conducted. This can result in the lack of subtleties or subjective interpretations that researchers can bring to scientific writing.

Citations and references in scientific articles play a crucial role in the process of knowledge creation. They act as a form of support and validation of the arguments presented by researchers, allowing other scientists to track the information used and build on it[3]. They are a way to recognize and give credit to the previous works that have served as the basis for the article in question. By including these references, the authors indicate that they have been based on previous research and are contributing to the accumulation of community scientific knowledge. This is not only a gesture of courtesy toward the previous authors, but it also helps researchers to create continuous dialogue and respect intellectual authorship.

Email: rme.mtz@infomed.sld.cu

Citations and references give readers the opportunity to examine the sources used in the article and evaluate the strength of the argument presented. By providing references, researchers allow others to investigate further and confirm or refute the findings presented, so it has a verifiable function[4]. This encourage transparency and confidence in the scientific process, as other researchers can replicate the experiments or analyses used, which helps ensure the rigor and validity of the results.

In the same way, they help researchers build a network of interconnected scientific knowledge. By including references to previous research, authors can show how their study relates to existing theories, identify gaps in current knowledge, and suggest new research directions. This allows the scientific community to advance and refine their theories, while avoiding unnecessary duplication of research efforts.

The construction of scientific articles requires extensive research, critical analysis and synthesis of relevant information. One of the fundamental aspects in the construction of these articles is the proper citation of sources. This is essential to support the affirmations and results presented, as well as to give credit to other authors whose work has been used as a basis for developing new knowledge.

Ghost quotes in scientific communication refer to an increasingly common problem today. This occurs when the authors include bibliographic references in their scientific articles, but without having actually consulted the works to which they refer. This can have negative consequences on the quality and veracity of the scientific information presented.

One of the main causes of ghost citations is the lack of time or resources on the part of authors to thoroughly review the existing literature[5]. In many cases researchers feel pressured to publish quickly and meet academic productivity standards, which can lead them to include citations without having properly verified their content. Incorrect citation can lead to the dissemination of erroneous information and even to the perpetuation of errors in the scientific literature. Some scientific journals encourage the inclusion of a high number of self-citations in the articles they publish, as this allows them to increase their impact factor. This has led some authors and editors to collaborate in the inclusion of ghost citations in works, with the aim of increasing their visibility and recognition in the scientific community.

The consequences of ghost citations are significant. Firstly, they undermine trust and scientific rigor. By referring to authors or sources that do not exist, distortion of the bibliography used in the research is generated, which can lead to errors in the interpretation of the results or to the repetition of previous researches that have not been correctly attributed[6]. Moreover, they can affect the reputation and credibility of the researchers and institutions involved. Likewise, scientific journals that publish articles with ghost citations may lose their reputation and may no longer be considered reliable sources.

Editorial policies can help minimize ghost citations. Some journals require authors to cite articles from major databases and in English. These policies, while seeking to

improve the quality and visibility of research, are also generating a series of ethical and academic problems.

First, editorial policies that prioritize citations in specific databases limit the diversity and scope of readings and references used by authors. This can lead to the loss of valuable perspectives and research results that might not be present in mainstream databases. Moreover, encouraging the exclusivity of articles written in English can lead to a lack of recognition of researchers and scientific works in other languages.

Editorial policies requiring a specific amount of citations may lead to the creation of ghost citations. This occurs when authors include citations without having read, understood or actually used the referenced articles. This practice, in addition of being unethical, distorts scientific knowledge and the validity of the results. Many times, researchers are forced to cite certain articles to be considered for publication in prestigious journals.

AI is one of the most widely used techniques to generate ghost citations in scientific articles[2]. Natural language generation systems, fed with large amount of scientific data can generate citations that resemble those of an expert in the field. This raises the possibility that unethical academics or journal editors could use this technology to improve the apparent quality of their articles.

Scientific information is crucial for making decisions and advancing knowledge; it is essential to be able to trust the integrity of published articles. Citations and references are one of the fundamental pillars of this trust, as they allow readers and reviewers to track the evidence and sources used to support the statements made in a given document. Consequently ghost citations hinder the advancement of knowledge and reduce its credibility.

The generation of quality citations and references involves the AI's ability to validate the information collected. AI algorithms based on machine learning often extract information from various sources, which can lead to erroneous citations or incorrect references[7]. AI is not able to assess the reliability of the sources used and verify the accuracy of the information before generating a citation or reference.

The generation of citations and references also requires a deep understanding of the context in which the sources are cited. AI should be able to identify the type of source (book, journal, conference, etc.), as well as the specific details of the cited text (author, title, year, etc.). Difficulties can arise when cited sources do not follow a uniform format or when abbreviations or cross-references that require an advanced understanding of the text are used.

Ensuring originality and avoiding plagiarism is essential in the generation of citations and references. AI must have the ability to compare the cited sources with others available in databases and verify their originality. This can be especially complex when working with different languages or less well-known sources. Moreover, AI does not have the capacity to identify citations and references that can be attributed to a particular source and not to a series of different sources.

Email: rme.mtz@infomed.sld.cu

The generation of citations and references is affected by the need for an updated and complete database[8]. As new sources appear and new citation techniques are developed, AI must be constantly updated to ensure accurate citations and references in line with current standards. This involves the incorporation of new sources, style guidelines and citation rules which represents a challenge in terms of implementing and maintaining AI.

Detecting AI-generated ghost citations can be a challenge, especially if the generated content is convincing and well-structured. However there are verification methods that could help identify the presence of these fake citations. For example, the analysis of the coherence and consistency in the use of scientific terminology could help detect discrepancies between citations and the remaining content of the article. In addition, researchers and editors can use plagiarism detection tools to look for similarities with other existing publications.

To address this problem, it is necessary to implement measures that promote scientific integrity. First of all, a culture of transparency and honesty in the production of scientific articles should be fostered. Researchers should be aware of the importance of correctly citing and giving adequate recognition to the sources used in their research and devoting sufficient time and resources to verify the appropriate use of the bibliographic references included in their articles.

It is essential that scientific journals establish clear and strict policies regarding the inclusion of citations in the articles they publish. Rigorous review of the bibliography used should be encouraged and all references should be verified to be real and relevant to the work presented. In addition, academic institutions should develop training programs in scientific ethics that highlight the importance of reviewing bibliographic references.

REFERENCES

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 [cited 2023/09/12];76(8):263-77. Available from: 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 [cited 2023/09/22];8(2):e08946. Available from: 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 [cited 2023/09/19];39(2):236-40. Available from: https://doi.org/10.17843/rpmesp.2022.392.11171

Email: rme.mtz@infomed.sld.cu

4. Santos K, Melo RR de, Correia AEGC, et al. Retratados e ainda citados: perfil de citações pós-retratação em artigos de pesquisadores brasileiros. Em Quest [Internet]. 2023 [cited 2023/09/14];29(2):e-125494. Available from: https://doi.org/10.19132/1808-5245.29.125494

5. Biswas SS. Role of ChatGPT in Public Health. Ann Biomed Eng [Internet]. 2023 [cited 2023/09/22];51(4):868-9. Available from: 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 [cited 2023/09/22];40(3):26-9. Available from: 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 [cited 2023/09/17];12(1):96-113. Available from: 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 [cited 2023/09/06];55(3):1-10. Available from: https://doi.org/10.31055/1851.2372.v55.n3.28723

HOW TO CITED THIS ARTICLE

Paz-Enrique LE. Ghost citations in scientific articles: a growing problem in the face of the use of artificial intelligence. Rev Méd Electron [Internet]. 2023 Nov.-Dec. [cited: date of access];45(6). Available from: http://www.revmedicaelectronica.sld.cu/index.php/rme/article/view/5366/5716

Email: rme.mtz@infomed.sld.cu