Artificial intelligence and the changing landscape of research

Hester Gail Lim, MD, FPDS[1,2]

Large language models (LLMs) are a subset of artificial intelligence (AI) trained on large quantities of text, enabling human-like interaction. These models are particularly adept at language-based tasks. ChatGPT (OpenAI, San Francisco, California) is a program based on one of the largest LLM subsets. In addition to its complex parameters, the pretraining has made it particularly adept at generating human-like text.[1] It was made available to the public in November 2022.[2]

The use of LLMs can have a profound impact on written content creation, and research is no exception. With the ability to generate paragraphs of text based on simple prompts by the user, LLMs are increasingly used in creative writing, academic essays, and scientific writing. LLMs may hasten the literature review process and can quickly summarize data from multiple sources. It can improve the writing style, coherence, and readability of an article, which may be particularly helpful for non-native English speakers.[3] However, certain risks must also be considered.

Unlike human authors, LLMs cannot be held accountable for errors generated within the text.[3,4] LLMs can generate text without a contextual understanding of the nature of the material, a capacity limited to human authors.[3] As a result, the accuracy and scientific integrity of generated text cannot be ascertained. Lacking understanding, generated text may contain factual errors or fabrications, which may seem plausible to the casual reader. LLMs can even perpetuate scientific biases within the training data and skew responses toward inaccuracies.[5] Furthermore, generated text may carry the risk of plagiarism.[3,4]

Zheng et al. tested ChatGPT's accuracy by asking a question pertaining to an article published in 2022, beyond the scope of ChatGPT's training database. Specific facts about the article were provided beforehand. The same question was asked five times. Each time, the question was answered plausibly but incorrectly, and each response elicited was different. The authors concluded that the generated text was convincing and plausible enough to a non-expert to be perceived as fact, despite the inherent errors.[6] OpenAI states this limitation clearly on their website: "ChatGPT sometimes writes plausible-sounding but incorrect or nonsensical answers".[7]

Differentiating between AI-generated and human-generated articles can be difficult and is perceived as a looming problem for editors and publishers. While a multitude of AI text detectors are available, these have varying degrees of accuracy. Detectors perceive patterns within generated text that indicate AI origin. Gao et al. used a GPT-2 detector to detect generated abstracts versus original abstracts. This detector had a specificity of 94% and a sensitivity of 86%. In comparison, blinded human reviewers were able to detect generated text accurately in only 68% of abstracts.[8] Interestingly, AI text detectors appear to be less accurate when reviewing articles written by non-native English speakers.[9]

Users must also consider the ethical implications of authorship when utilizing AI. There are currently no guidelines in place to govern the use of AI within research writing. The *Nature* group of journals has explicitly prohibited using AI-generated content, stating that LLMs cannot be listed as co-authors. They further stated that any use of LLMs must be documented in the methodology or acknowledgments section. The *Science* family of journals has banned all AI-generated text, including images, figures, and graphics. Any detection of these in submissions will be perceived as scientific misconduct. This highlights the need for transparency when such systems are being utilized to avoid passing off AI-written text as that written by a human author, a term now being called 'AIgiarism'.[9]

While the Journal of the Philippine Dermatological Society (JPDS) editorial process already includes screening

submitted text for plagiarism, the editorial team has now included an additional level of screening with AI text detectors. Submitted manuscripts are run through at least two (2) different AI text detectors, and at least two (2) editors review those scoring highly for AI-generated text. We strongly discourage using LLMs in writing manuscripts, especially in the absence of full factual supervision by the author and transparency in the research process.

While we recognize the potential benefits of LLMs in scientific writing, the role is assistive in nature and should not replace the expertise of a human author. As with the best of tools, strong human oversight remains necessary. Furthermore, critical thinking, contextual analysis, and the generation of new ideas are all unique to the human mind. Bypassing these in pursuit of generating more manuscripts or publications with great speed weakens these elements as cornerstones in the scientific process. This can lead to the gradual erosion of research quality, limiting true scientific progress.

[1]Cebu Institute of Medicine
[2]Philippine Dermatological Society Research Committee

Corresponding author
Hester Gail Lim, MD, FPDS, hesterlim@cim.edu.ph

Conflict of interest
None

Source of funding
None

REFERENCES

1. Shen Y, Heacock L, Elias J, Hentel KD, Reig B, Shih G, et al. ChatGPT and other large language models are double-edged swords. Radiology [Internet]. 2023;307(2):e230163. Available from: http://dx.doi.org/10.1148/radiol.230163
2. Marr B. A short history of ChatGPT: How we got to where we are today [Internet]. Forbes. 2023 [cited 2023 May 25]. Available from: https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/?sh=362b3adb674f
3. Huang J, Tan M. The role of ChatGPT in scientific communication: writing better scientific review articles. Am J Cancer Res. 2023;13(4):1148–54.
4. Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. Nature [Internet]. 2023;613(7945):612. Available from: http://dx.doi.org/10.1038/d41586-023-00191-1
5. Salvagno M, Taccone FS, Gerli AG. Can artificial intelligence help for scientific writing? Crit Care [Internet]. 2023;27(1):75. Available from: http://dx.doi.org/10.1186/s13054-023-04380-2
6. Zheng H, Zhan H. ChatGPT in scientific writing: A cautionary tale. Am J Med [Internet]. 2023; Available from: http://dx.doi.org/10.1016/j.amjmed.2023.02.011
7. Introducing ChatGPT [Internet]. Openai.com. [cited 2023 May 25]. Available from: https://openai.com/blog/chatgpt/
8. Gao CA, Howard FM, Markov NS, Dyer EC, Ramesh S, Luo Y, et al. Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. NPJ Digit Med [Internet]. 2023;6(1):75. Available from: http://dx.doi.org/10.1038/s41746-023-00819-6
9. Brainard J. Journals take up arms against AI-written text. Science [Internet]. 2023;379(6634):740–1. Available from: http://dx.doi.org/10.1126/science.adh2762