BMES BIOMEDICAL ENGINEERING SOCIETY

EDITORIAL

Large Language Models Like ChatGPT in ABME: Author Guidelines

Carly Norris[1] ⓘ

The use of Large Language Models (LLMs) within the scientific community has resulted in a storm of ethical discussions in recent months. LLMs are an emerging form of artificial intelligence which have been trained to output plausible sequences of words based on the likelihood that the sequence may occur within the natural human language. These unique sentences are strung together based on the text data that the model was trained on [1]. For example, ChatGPT is an LLM that was trained on a very large dataset from the Internet and has recently demonstrated its effectiveness at constructing sequences of sentences with valid deductive reasoning. Further illustrating its power within the scientific community, ChatGPT recently authored peer-reviewed papers in Oncoscience [2] and Nurse Education and Practice [5] journals and demonstrated expert-level knowledge by receiving a passing score on the United States Medical Licensing Exam [1].

This powerful tool is capable of providing quick, custom responses to help answer niche questions, however, there exist major concerns related to its use for scientific reporting. In particular, LLMs cannot be held accountable for the accuracy and validity of the science discussed within the written text. This is particularly important because LLMs, including ChatGPT, can provide largely inaccurate or biased responses based on the input data it was trained on [1, 3, 4]. Therefore, many scientific publishers now prohibit LLMs from authoring manuscripts, which includes Springer Nature. Details can be found within the editorial policies on authorship criteria. As a Springer Nature journal, the Annals of Biomedical Engineering (ABME) will adhere to these guidelines and reject manuscripts that do not satisfy the authorship criteria.

Transparent reporting of the use of LLMs in scientific works remains a major ethical discussion. The use of LLMs like ChatGPT threaten scientific rigor and integrity when the authors adopt the original language that was output from these generated models as their own [3]. Therefore, ABME now requires full transparency within the methods or acknowledgements section if the authors have used LLMs in any way to while developing their manuscripts. By stating the use of the LLM, the author is accepting responsibility for the accuracy of what was reported and spreads awareness to the reviewers so that they may identify any potential biases, inaccuracies, or misreporting. Kung et al. provided a good example of transparent reporting of the use of Chat-GPT within the methods section as it directly related to data collection and analysis [4]. Statements should be provided in the acknowledgements section in instances where the LLM enhanced or motivated any ideas or discussions throughout the document, especially if any of the generated text was used. Examples of how to report the use of LLMs in the acknowledgements section are included below.

The author acknowledges that this article was partially generated by ChatGPT (powered by OpenAI's language model, GPT-3; http://openai.com). *The editing was performed by the author* [3].

The author acknowledges that some content in this article was partially generated by ChatGPT (powered by OpenAI's language model, GPT-3.5; http://openai.com) *to discover the roles that chatGPT can play in public health. The editing was performed completely by the human author* [4].

As the use of LLMs like ChatGPT continue to expand and models continue to improve, it is imperative for scientists to strive for the highest degree of scientific rigor and integrity through deep critical thinking and transparent reporting. Only then can we be confident in the data that will power future models and scientific discoveries, which will hopefully converge upon groundbreaking solutions.

Associate Editor Stefan M. Duma oversaw the review of this article.

✉ Carly Norris
carly8@vt.edu

[1] Department of Biomedical Engineering and Mechanics, Virginia Tech, Blacksburg, VA 24060, USA

Declarations

Conflict of interest The authors disclose no conflicts of interest.

References

1. Kung, T. H., M. Cheatham, A. Medenilla, C. Sillos, L. De Leon, C. Elepaño, M. Madriaga, R. Aggabao, G. Diaz-Candido, J. Maningo, and V. Tseng. Performance of chatgpt on usmle: potential for ai-assisted medical education using large language models. *PLoS Digit Health*. 2:e0000198, 2023.
2. Zhavoronkov, A. Rapamycin in the context of pascal's wager: generative pre-trained transformer perspective. *Oncoscience*. 9:82–84, 2022.
3. Biswas, S. S. Potential use of chat gpt in global warming. *Ann Biomed Eng*. 2023. https://doi.org/10.1007/s10439-023-03171-8.
4. Biswas, S. S. Role of chat gpt in public health. *Ann Biomed Eng*. 2023. https://doi.org/10.1007/s10439-023-03172-7.
5. King, M. R., ChatGpt. A conversation on artificial intelligence, chatbots, and plagiarism in higher education. *Cell Mol Bioeng*. 16:1–2, 2023.
6. O'Connor, S., ChatGpt. Open artificial intelligence platforms in nursing education: Tools for academic progress or abuse? *Nurse Educ Pract*. 66:103537, 2023.