The Journal of Education in Perioperative Medicine

Editorial

Artificial Intelligence and ChatGPT: Bane or Boon for Academic Writing?

William C. Culp Jr, MD

In November 2022, ChatGPT (powered by the software engine Generative Pre-trained Transformer 3.5 or GPT-3.5) was beta released by the technology company OpenAI as a language software program accessible freely over the internet. This large language model chatbot accepts plain language queries, and by referencing a massive prelearned dataset, is able to respond in seconds with answers as directed by the end user. These responses are constructed in coherent, readable sentences or paragraphs that may be difficult to distinguish from fluent, human-constructed text, and are based on over 500 GB of studied source materials covering a wide array of topics. These responses are fantastically versatile and easily customizable to the user's requests, for instance, targeting a specific reader education level, writing style, or format. ChatGPT has even excelled on or passed numerous high-stakes examinations, including the Uniformed Bar Exam, college entrance examinations, and medical licensing exams.[1] In less than 2 months of release, ChatGPT had over 100 million users, reflecting intense interest across disciplines and across the world as the fastest growing application in history. How does this seemingly obscure piece of software from Silicon Valley impact the world of perioperative medicine? As busy clinical anesthesiologists and academicians, why should we even care? We must be aware of this technology because of its profound impact on the written word, including the gold standard of modern medical knowledge that maintains and advances the world of medicine—the peer-reviewed scientific journal article.

ChatGPT has already demonstrated abilities to aid in study design. Posed with a carefully crafted scientific question, this application can make recommendations about the design and construct of a medical study, offering commentary on strengths and weaknesses of different approaches and models. It can perform a literature search and create a bibliography. ChatGPT can also recommend and develop a statistical analysis plan and list advantages and disadvantages for different analysis methods, supplementing a skillset in which many early career physicians and trainees have little experience and education. As its strength is in language, ChatGPT can quickly draft abstracts and summaries, create manuscripts, write case reports, and even editorials, each of which have already been published in the peer-reviewed literature with the assistance of ChatGPT. The application can also be used as a reviewer and an editor to screen manuscripts, support decisions about acceptance or rejection, and revise human-written first drafts to improve readability, clarity, or shift tone. Importantly, it can function as a translator, as well, with command of two dozen languages, potentially extending access to large numbers of investigators for whom English is not a first language. All of these actions happen in seconds and at no charge, taking tasks that routinely require many hours of skilled human labor and completing them nearly instantly. This artificial intelligence can reduce the time from idea conception to experimental design to manuscript creation and even peer review and editing.[2]

One of the first criticisms of this technology as applied to scientific writing is its role in authorship. Many articles on the internet are now written *entirely* by artificial intelligence, and a growing number of scientific articles have been written with the assistance of, or entirely by, ChatGPT. Using International Committee of Medical Journal Editors standards, however, all authors must agree to be accountable for all aspects of the work.[3] An insentient chatbot cannot be held accountable and therefore cannot be a named author. Many journals and publishers are creating policies restricting or regulating ChatGPT use, most commonly prohibiting crediting artificial intelligence tools as authors and mandating documentation of use in the methods section of all manuscripts.[4]

Furthermore, as a language bot optimized for predicting word associations, ChatGPT produces consistently high-quality and convincingly believable language, that may—or may not—be factually correct or properly supported by references. The threat that this poses to document accuracy cannot be overstated. ChatGPT responses frequently fail to credit a source, may cite an incorrect source, or fabricate and credit a nonexistent source including a false PubMed ID number. "Facts" may be contrived from thin air. These "artificial hallucinations" or "stochastic parroting" events are well-described and happen often, creating "false answers that sound

continued on next page

continued from previous page

good."[5] This may lead to plagiarism when a source isn't cited or is miscited, and even worse, patently false information may be widely distributed, which may lead to patient harm. Although commercially available software chatbot detectors exist, only a human with expertise in the subject matter may be able to detect these otherwise convincing ChatGPT outputs (see sample conversation), and even then, many are unable to differentiate human from ChatGPT-generated documents. Additionally, ChatGPT output can only be as good as its data source, and limited or selected data may lead to false or biased conclusions. As a chatbot optimized to predict word associations, it is important to appreciate that ChatGPT is not sentient and lacks human creativity. In contrast, it uses predictive modeling based on existing data to construct sentences and therefore may constrain the human mind and reduce our vocabulary, limit truly new ideas, and stifle creativity. The risk of scientific stagnation here is real if we merely parrot previously published ideas.

The famous English essayist, Sir Francis Bacon, wrote that "reading maketh a full man, conference a ready man, and writing an exact man."[6] He believed and understood that one must think clearly in order to write clearly, and this process of writing provided clarity and satisfyingly deep understanding of the topic at hand. Delegating this task to a computer comes with a cost, stealing an opportunity for mental clarity from our investigators. Many would argue that this is simply another modern, iterative advancement in technology, much like the typewriter, word processor, real-time spelling checker, or even a writing aid such as Grammarly. However, the risks of this advanced chatbot are real, as a broad threat to the field of education including the "death of the essay,"[7] the development of language, and complex thought. At the time of this article submission in March 2023, an updated, more powerful ChatGPT engine has just been released. Dubbed GPT-4, this more powerful engine accesses a much larger database, provides more accurate answers, scores higher on the bar exam than previous versions (now at the 90th percentile), can interpret images, and write code in major computer programming languages.[8] With the advances of ChatGPT (and artificial intelligence in general) progressing now at breakneck speed, innumerable concerns exist about potential harm, even as an existential threat for humanity. The Institute for the Future of Life, representing numerous experts including Elon Musk, has called for an immediate pause on continued development of this type of artificial intelligence,[9] while Italy has even blocked access to the program because of privacy concerns. The coming months will determine how authors and editors, researchers, and clinicians will come to accept this new technology, as either a bane or a boon for the world of academic writing.[10]

Sample ChatGPT Conversation:

Author: Write a short paragraph assessing trends in graduating anesthesiologists pursuing either additional fellowship training or instead going directly into practice.

ChatGPT: According to recent data, there has been a notable increase in the number of graduating anesthesiologists pursuing additional fellowship training after completing their residencies. This trend may be due in part to the growing demand for specialized skills and expertise in areas such as critical care, pain management, and perioperative medicine. On the other hand, there are still many anesthesiologists who choose to go directly into practice after completing their residencies, and this option remains a popular choice for many. Ultimately, the decision to pursue additional training or go directly into practice will depend on a range of factors, including individual career goals, personal interests, and the needs of the healthcare system.

Author: List the data.

ChatGPT: I'm sorry, but I don't have access to any specific data regarding trends in graduating anesthesiologists pursuing additional fellowship training or going directly into practice. The paragraph I provided was a general overview of the trends and factors that may influence anesthesiologists' career decisions based on my general knowledge and understanding of the field.

Author: If you didn't have the data, how did you write "according to recent data?"

ChatGPT: You're correct, my previous response was not accurate. I apologize for the confusion. Since I don't have access to any specific data on this topic, I shouldn't have written "according to recent data." As an AI language model, I generate responses based on patterns in the data I was trained on, and I may occasionally produce errors or inaccuracies.

References

1. OpenAI. Pioneering research on the path to AGI. www.openai.com/research/overview. Accessed March 31, 2023.
2. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare*. 2023;11:887.
3. International Committee of Medical Journal Editors. Defining the role of authors and contributors. https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html. Accessed March 31, 2023.
4. Tools such as ChatGPT threaten transparent science; here are our ground rules for their use [Editorial]. *Nature*. 2023;613(7945):612.
5. Alkaissi H, McFarlane SI. Artificial hallucinations in ChatGPT: implications in scientific writing. *Cureus*. 2023;15(2):e35179.
6. Bacon F. *The Essays of Francis Bacon: The Fifty-Nine Essays, Complete*. Raleigh-Durham, NC: Adansonia Press; 2018.
7. Herman, Daniel. The end of high-school English. *The Atlantic*. December 9, 2022. https://www.theatlantic.com/technology/archive/2022/12/openai-chatgpt-writing-high-school-english-essay/672412/. Accessed March 31, 2023.
8. OpenAI. GPT-4. www.openai.com/research/gpt-4. Accessed March 31, 2023.
9. Future of Life Institute. Pause giant AI experiments: an open letter. www.futureoflife.org/open-letter/pause-giant-ai-experiments/. Accessed 31 March 2023.
10. Doshi RH, Bajaj SS, Krumholz HM. ChatGPT: temptations of progress. *Am J Bioeth*. 2023;23(4):6-8.

continued on next page

continued from previous page

William C. Culp Jr *is a Professor of Anesthesiology and Associate Dean of Faculty Affairs & Faculty Development at Baylor Scott & White Health-Temple Medical Center and Baylor College of Medicine-Temple in Temple, TX.*

Corresponding author: *William C. Culp Jr, MD, 2401 South 31st Street, Temple, TX 76508. Telephone: (254) 774-1713*

Email address: *William C. Culp Jr: William.CulpJr@BSWHealth.org*

Funding and financial disclosures: *None*

Conflicts of interest: *None*