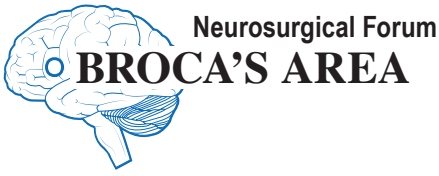
JNS



When words leap off the page: conversational manuscripts as a novel interface for scientific publishing

Alexander Winkler-Schwartz, MD, PhD

Department of Neurosurgery, University of Utah, Salt Lake City, Utah

I am pleased to present the "conversational manuscript," a novel way of engaging with scientific content, supported by the Journal of Neurosurgery Publishing Group (JNSPG). As part of a pilot project, readers can dynamically interact with a scientific manuscript, limited only by their imaginations. This project was made possible by fine-tuning large language models (LLMs)—in this case OpenAI's GPT-4—on the contents of a manuscript. My motivation behind developing the conversational manuscript came from a desire to leverage the extraordinary capabilities of LLMs to interpret and generate text to improve the transfer of scientific information across disciplines and to the public. Although this current conversational manuscript is limited to a single publication,[1] it serves as a natural starting point and test case for this technology. Because each monthly issue is centered on a unique theme, the technology could potentially be integrated, for example, throughout an entire issue of *Neurosurgical Focus* as a next step in this process.

In the subsequent text, I will provide the historical context and explain the underlying functioning of LLMs so that readers may come away with a more comprehensive understanding of a technology that is poised to significantly influence our daily activities as neurosurgeons and researchers.

Technological innovation has profoundly shaped the production and dissemination of scientific knowledge. Since 1952, science has exhibited striking publication growth at a rate of 5% a year, effectively doubling every 14 years.[2] Given the sheer volume of new information, it is impossible for any one person to stay fully up to date. However, despite the rapid increase in the quantity and accessibility of knowledge, the structure of text remains a unidirectional flow of information from the publication to the reader, fixed in its original language, context, and technical terminology. This specialized language creates barriers to interdisciplinary collaboration and public understanding, especially in fields (such as medicine, computer science, and engineering) that rely heavily on complex terminology.

LLMs constitute a type of artificial intelligence (AI) model that has been trained on extensive written data and represents yet another paradigm shift in technology. Vaswani et al. laid the foundation for modern LLMs in a 2017 paper titled, "Attention is all you need."[3] Since then, more than 100 LLMs have been developed,[4] with the most well-known being OpenAI's ChatGPT. Although the first model, GPT-1, was produced in 2018,[5] it was only with the release of GPT-3.5 (the LLM underlying ChatGPT) on November 30, 2022, that LLMs gained recognition by the broader public. At the time, OpenAI's website quickly became the fastest growing consumer internet application in history, with 100 million active users within 2 months of ChatGPT's release. Recent data from May 2023 suggests a staggering 2.24 billion visits to OpenAI's website (which hosts ChatGPT) per month.[6]

Much attention has been spent on the impact of LLMs on scientific publishing, and with good reason—LLMs have the potential to upend scientific writing through their generation of text.[7] However, there has been less focus on their capacity to change how science is consumed. The conversational nature of LLMs allows users to dynamically interact with scientific manuscripts. Through engaging with these conversational manuscripts, ambiguous concepts can be clarified and contextualized within a broader context, and text can be condensed, highlighted, or translated to another native language, depending on the needs of the reader. In this way, LLMs facilitate the transfer of knowledge across different scientific disciplines and languages, and from researchers to the public.

If LLMs are to be the conduits of scientific knowledge, it is essential that they are factually accurate, up-to-date,

ACCOMPANYING EDITORIAL DOI: 10.3171/2023.7.JNS231621.

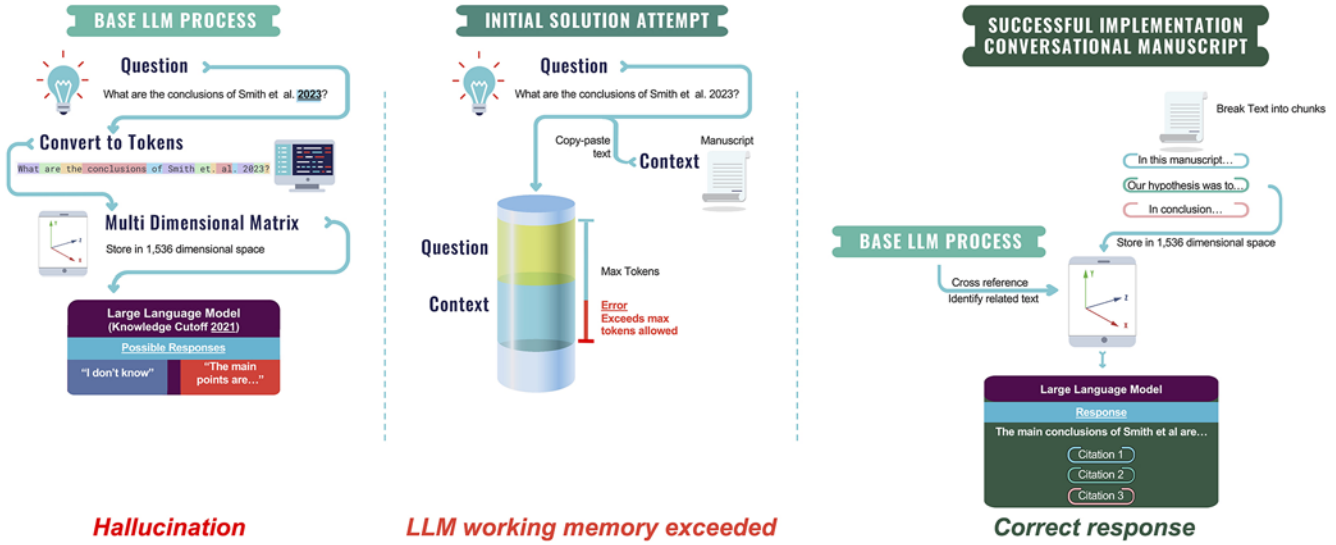


FIG. 1. A three-panel infographic demonstrates the use of LLMs in the creation of a conversational manuscript. Stage I (*left*) illustrates the model's initial tokenization and dimensional representation process, emphasizing the risk of hallucinations when queried about topics beyond its knowledge cutoff date. Stage II (*center*) presents an unsuccessful attempt to enhance the model's capabilities by simply introducing additional context via simple copy and paste. Stage III (*right*) showcases the successful circumvention of these limitations through preprocessing text into numerical representations, effectively reducing the risk of hallucinations, enhancing response accuracy, and allowing direct citations from the source material. Figure is available in color online only.

and transparent. Currently, if an LLM is asked about topics that fall outside its original training data set, it may either state that it "doesn't know" or may produce an incorrect answer. The latter behavior is known as "hallucinating." At the time of this writing, the latest iteration of OpenAI's GPT-4 has a cutoff date of September 2021, while Google's LLM, Bard, has a cutoff date of March 8, 2023. As such, niche topics and recent publications are not accessible to the base consumer-facing LLMs, highlighting the need for advanced techniques to enhance their knowledge base and adaptability.

Fortunately, LLMs can be fine-tuned to serve as content experts, capable of understanding and interacting with information from hundreds of scientific papers to which they had no prior exposure. To grasp this process, it is important to understand how LLMs such as ChatGPT function (Fig. 1). ChatGPT does not understand text as we do. Instead, it converts "tokens," roughly four-character subsets of text, into numerical values that pinpoint a location in a multidimensional space (1536 dimensions, to be precise). When a user presents a question or instruction to ChatGPT, the model searches the space for the most probable semantically related content and generates a response (Fig. 1, Stage I left). As the conversation proceeds, or if a substantial amount of text is introduced, the model eventually reaches its working memory limit (Fig. 1, Stage II center). Once the working limit is reached, the model is effectively rendered amnesic to portions of the text and the risk of hallucinations increases. However, this working memory limitation can be circumvented by preprocessing text into numerical representations, each embodying several thousand characters from the original. These representations are then stored within the aforementioned multidimensional space. This space thus serves as a searchable index and enables the model to efficiently navigate and retrieve pertinent responses while providing direct citations to the source material (Fig. 1, Stage III right). Because the source material was not processed by the model, its working memory remains unencumbered and the risk of hallucinations is greatly reduced. Consequently, users can judge whether the model's response accurately reflects its provided citations. Library of Congress human-generated keywords will likely be a relic of the past, as LLMs will be able to directly search through scientific source material. It is not implausible to imagine such a system incorporating patient-specific data and serving as a clinical decision support tool.

It is also worth noting that, in accordance with OpenAI's terms-of-use policy,[8] content provided via their Application Programming Interface (API), such as the method described above, is not used to "develop or improve" their services. Consequently, for a conversational manuscript, ownership rights remain with the copyright holder. Importantly, this protection does not extend to non-API content, creating a potential risk of copyright infringement if a manuscript is directly copied and pasted into the ChatGPT model on OpenAI's website. The use of copyrighted material in the training of generative AI is a contentious area of law. Indeed, the proposed AI regulatory bill from the European Union (the AI Act) imposes explicit limitations on the use of copyrighted materials in generative AI applications like ChatGPT.[9]

Finally, neurosurgery has consistently been at the cutting edge of technological innovation. The JNSPG continues this trend with the deployment of a prototype conversational manuscript to enhance reader engagement with scientific content. This marks the first instance of a scientific journal providing such a service. As LLMs

become more cost-effective, we may witness the prolific emergence of this and other generative AI technologies in our lives.

Important questions remain, such as how an LLM trained on the sum-total of scientific knowledge would balance the relative importance of each manuscript or fact, given that scientific progress is characterized by the refutation of existing theories. It is critical not to assign equal weight to obsolete theories and widely accepted ones, while also not overly guiding readers when true equipoise exists. One possible solution, aimed at introducing the least amount of bias, is to provide readers with sufficient context in instances of conflicting knowledge. This may take the form of providing the temporal relationships between theories, highlighting when and why knowledge has progressed over time, as well as direct citations from the source material to encourage a critical evaluation of the model's responses. Furthermore, given the reduced cost and ease of access to generative AI, it is likely that misinformation will outpace accurate content on the internet. This may present an opportunity for scientific journals to serve as guides in this chaotic information landscape. Despite these—and likely many more yet unseen—challenges, I am confident we will find solutions.

In conclusion, the proliferation of LLMs signals a significant shift in the production and consumption of scientific information. With their potential to enhance understanding across various disciplines and audiences, LLMs could pave the way for a new era of scientific communication. As we navigate these technological advancements, it is essential that we balance the immense potential of these tools with their potentially disturbing risks and limitations. Although challenges remain, the consistent evolution and refinement of these models promises a future where science becomes ever more accessible and impactful.

References

1. Shahrestani S, Chan AK, Bisson EF, et al. Developing nonlinear k-nearest neighbors classification algorithms to identify patients at high risk of increased length of hospital stay following spine surgery. *Neurosurg Focus*. 2023;54(6):E7.
2. Bornmann L, Haunschild R, Mutz R. Growth rates of modern science: a latent piecewise growth curve approach to model publication numbers from established and new literature databases. *Humanit Soc Sci Commun*. 2021;8(1):224.
3. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. Advances in neural information processing systems. *arXiv*. Preprint posted online December 6, 2017. doi:10.48550/arXiv.1706.03762
4. Open LLM Leaderboard. Hugging Face. Accessed July 14, 2023. https://huggingface.co/spaces/HuggingFaceH4/open_llm_leaderboard
5. Radford A, Narasimhan K, Salimans T, Sutskever I. Improving language understanding by generative pre-training. OpenAI. June 11, 2018. Accessed July 14, 2023. https://openai.com/research/language-unsupervised
6. Website Traffic Checker: openai.com (May 2023). Semrush.com. Accessed May 2023. https://www.semrush.com/website/openai.com/overview/
7. Barker FG, Rutka JT. Editorial. Generative artificial intelligence, chatbots, and the Journal of Neurosurgery Publishing Group. *J Neurosurg*. Published online April 28, 2023. doi: 10.3171/2023.4.JNS23482
8. Terms of Use. OpenAI. March 14, 2023. Accessed July 14, 2023. https://openai.com/policies/terms-of-use
9. AI Act: a step closer to the first rules on Artificial Intelligence. European Parliament News. May 11, 2023. Accessed July 14, 2023. https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence

Disclosures

Dr. Winkler-Schwartz reported a patent for 63/091,629 pending and a patent for 05001770-843USPR pending.

Correspondence

Alexander Winkler-Schwartz: manuscriptmailbox@gmail.com.

INCLUDE WHEN CITING
Published online August 1, 2023; DOI: 10.3171/2023.7.JNS231617.