Open Access

Editorial

PRNANO

Precis. Nanomed. 2023;6(1):1007-1012

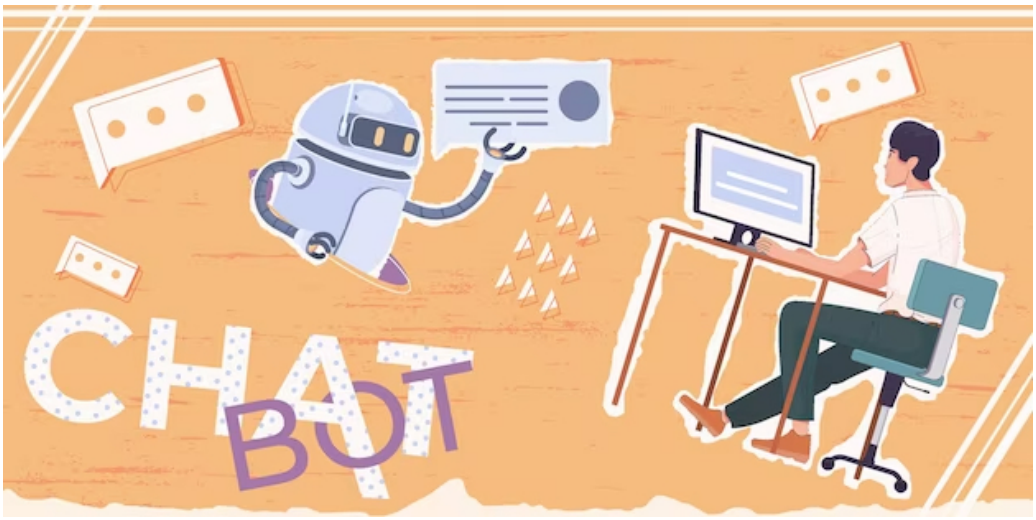
ANDOVER HOUSE

Are we there yet?

Lajos P. Balogh
Editor-in-Chief, Precision Nanomedicine[#]

Submitted: March 25, 2023 Accepted: March 27, 2023 Published: March 28, 2023

Graphical Abstract



Credit: freepik.com

Abstract

This paper examines the advantages, disadvantages, opportunities, and dangers of using AI Large Language Models (LLMs) like ChatGPT to write scholarly articles in natural sciences. These intelligent and impactful tools may be used for many purposes, both good and bad. There is great potential in applying these tools for science, but they have limited capabilities for now, especially in medicine-related articles. These tools should not be listed as authors or co-authors, and their sole use for overviewing the literature could eliminate citations, potentially leading to a major crisis in academic publishing.

Keywords

ChatGPT, Conversational AI, language model, LLM, publication, authorship, citation, impact

Purpose, rationale, and limitations

There is hype in the media discussing recent advances in AI language models and their expected effect on communication and on the world in general. In this article, I share my thoughts from the point of view of an editor for a broader discussion. I will limit myself to sharing my thoughts about the potential impact of computer-driven Large Language Models (LLMs), *aka* chatbots, on natural science and medical publications based on our today's English language knowledge and will stay away from more philosophical questions, e.g., whether man can create something which is more perfect than himself.

Introduction

Today, one of our insurmountable problems is that even the smartest humans cannot keep up with the exponentially growing amount of information, which is clearly demonstrated by the

[#] L.P. Balogh is the Editor-in-Chief of Precision Nanomedicine. This paper expresses his personal opinion.

increasing level of specialization in all fields, including scientific disciplines. We have more and more knowledge about smaller and smaller topics, many of which need to be connected - we obviously need help. [1] Since *OpenAI* released ChatGPT-4 on March 14, 2023, there are [2-4] more articles being published about artificial intelligence than anyone can read in one day.

The topic is sizzling because of dream-like possibilities and doomsday prognoses; maybe we have seen too many movies about humanoids with superhuman abilities, and, BTW, excitement sells. [5] Conversational AI is undoubtedly a game-changer for science and education [6] because, unlike people, it is exceptionally quick, with practically unlimited memory and expandable capabilities. [7]

What is ChatGPT? Let's ask the most qualified source – itself:* "ChatGPT is a large language model that is based on the principle of deep learning, specifically the transformer architecture. It is trained on a massive amount of text data using a technique called unsupervised learning, which means that it learns patterns and relationships in the data without explicit guidance or labels. The principle behind ChatGPT is to provide a conversational agent that can understand and generate human-like language in a wide variety of contexts. It achieves this by learning from diverse text sources, including books, articles, websites, and social media posts. This allows it to capture the nuances and complexities of natural language and produce informative and engaging responses. The transformer architecture used by ChatGPT is particularly well-suited to natural language processing tasks, as it is designed to process sequences of tokens (such as words or characters) and capture long-range dependencies between them. This allows ChatGPT to understand the context and meaning of a sentence or phrase and generate relevant and coherent responses. Overall, the principle of ChatGPT is to provide an intelligent and flexible conversational agent that can understand and generate human-like language across a wide range of topics and contexts and help people communicate and access information more effectively."

Let's not forget that despite its obviously bright future, ChatGPT (Chat Generative Pre-trained Transformer) and its generative AI relatives are *not* the superhuman androids from the movies – they are large language models (LLM) created by people and trained on a huge amount of data. They are designed to generate human-like and plausible texts depending on their training database, code, and input. Thus, the output of every chatbot depends on the database used for training and on the input.

DISCUSSION

Many generative AI programs with similar abilities are already available (ChatGPT, Bing, 365 Copilot, Midjourney, BLOOM, Bard, etc.), and these programs are getting better quickly. [8] However, it is improbable that they are all trained by accessing the same information and using the same algorithm. This gives rise to immediate questions: What is the database? Is it open to the public? Is its content objective, biased or limited in any sense?

The publicly accessible information on the internet is constantly criticized for containing misinformation and disinformation. [9] How can a language model determine whether the information digested is true if it uses unsupervised learning? It certainly finds the most frequently used continuation, but that is - by its nature - the most probable one, the "usual" one.

LLMs work by finding the highest probable word to continue the sentence with, which, by definition, makes it *middle-of-the-road*, consensual, and not based on facts but on texts that include fiction and non-fiction, true and false, the mix of raw and verified, fact-checked information all created by *us*. Chatbots are excellent at predicting the next word in a sentence but do not know what the sentence means. That makes it potentially dangerous to combine them with science, where it's crucial to get the facts straight, especially in medicine.[10]

Is scientific writing different from general writing?

"Yes, scientific writing is different from general writing in several ways.

Scientific writing is typically more formal and structured, with specific conventions for language, formatting, and referencing sources. It is also characterized by focusing on accuracy,

* In this article, we use a different typeface and quotation marks to indicate text written by ChatGPT.

precision, and objectivity, using technical terminology, and avoiding ambiguity.

Scientific writing is also typically aimed at a specialized audience of fellow researchers or professionals in a particular field, whereas general writing is often targeted at a broader audience. Scientific writing may include presenting experimental data, analysis of findings, and conclusions drawn from research, whereas general writing may cover a wide range of topics and styles, including creative writing, journalism, and persuasive writing."

Are LLMs able to write articles in natural sciences, i.e., chemistry, biology, physics, medicine, etc.?

Different disciplines use different ways of speaking and sometimes assign different meanings/definitions to the same word. For now, chatbots trained on general texts can commit factual errors and draw wrong conclusions, even though updates and improvements are coming out daily.[8] Thus, an LLM trained in general texts cannot be used to compose scholarly articles.

Should the owner/creator of an LLM be listed as a co-author?

We agree with SG Kim, who writes [11]: "*The primary objective of a scientific article is to convey new information that is supported by evidence, and all claims made by the authors should be thoroughly scrutinized before being published. While authors may seek assistance from English editing services during the publication process, such services should not be listed as co-authors. If ChatGPT is only used for language editing purposes, then there is no issue with using it to prepare scientific articles. However, any new ideas generated by ChatGPT should be validated through actual experiments, and their results should be verified by humans.*" There are also ethical challenges, and any assistance provided by AI should be disclosed in the article. [12, 13]

Can an LLM generate scholarly articles in nanomedicine?

Here is a definition of nanomedicine by ChatGPT:

"Nanomedicine is a field of medicine that uses nanotechnology to diagnose, treat, and prevent diseases. It involves the development of nanoscale materials, devices, and systems for use in medicine and biomedicine, such as drug delivery systems, biosensors, and diagnostic imaging agents. The aim of nanomedicine is to improve the efficacy and safety of medical treatments by leveraging the unique properties of nanoscale materials."

Is this definition acceptable? Yes. Is this creative? No. Accessing, extracting, and compacting information is easy with a generative AI program, but not creation. The question is, what will the *quality* of that output be, and how can it *be validated*?

Science, i.e., the systematically organized body of validated knowledge on a particular subject and at a given time, is not and never can be perfect, simply because that would be the end of learning. Publications are not science; they are yesterday's information and knowledge, organized, stored, and shared on various media. [14] Publication databases and article collections are imperfect because they include expired older knowledge and lack future knowledge. (Large publication databases like Scopus, Web of Science, Crosscheck, etc., overlap but are not identical.)

Science cannot be decided by votes or how many people accept an actual result or agree to a concept. Even though most of the time, the generally accepted reasoning is the right one, sometimes the least probable, or a brand-new explanation is true.

A thought-experiment

Let's assume there is a generative AI program that has been perfectly trained *on all information known to humankind*. Then, let's call the natural science version of it SciBOT, which is trained on the database of Crosscheck/iThenticate. The actual value of a SciBOT will depend on the database it is trained on (garbage in, garbage out).

A tool like that will be quickly put to use (its early versions are already had) because authors could quickly multiply the number of articles published. (Unfortunately, publishing more and not better articles became a recent tendency for scholarly publishing.) More citations and publishing in "high impact" journals have become an existential need for researchers to get tenure, promotion, etc.; more articles usually mean more citations. However, a sudden increase in the number of submissions would deepen the crisis of peer review, which is already under enormous stress. (We have yet to have a good solution for the quality control of articles *before*

publication.) Having reviews post-publication is no longer a solution because by the time comments (if any) arrive, potentially all incorrect conclusions will already be part of the public domain, which will be used to update/train Chatbots further.

Real or Fake?

Submissions to our journal are checked for similarities and overlaps before peer review using the plagiarism detection system of iThenticate https://www.ithenticate.com/ and Proofig https://www.proofig.com for text or image duplication or manipulation. What would constitute plagiarism in a text generated by SciBot?

Authorship?

One debated issue is whether an LLM or its creator should be listed as one of the authors. [15] According to our definition: "Authors are those who have made a *significant contribution to the conceptual design and/or the execution of the study.* However, all others, who contributed in any other way, must also be acknowledged."

An LLM cannot be held responsible for harming somebody's ownership or IP rights. According to Precision Nanomedicine requirements, neither LLM-s nor their creators qualify as co-author. Instead, they should be acknowledged, referenced as a tool, and clearly marked as such.

Abstracts?

Our definition: "Abstracts allow other scientists to decide whether to read the article in depth. Thus, one continuous paragraph (100-250 words) should summarize the paper's purpose, methods, results, and conclusions to help the reader quickly understand the problem, what has been done, and how this is useful for others. Abstracts are published separately in bibliographical sources, so they must be less technical than the body. They should stand alone without citations, footnotes, or non-trivial abbreviations."

Consequently, an abstract should reflect the view and data of the *actual research work* rather than a general summary of the field, which may only be part of the abstract. No need for citations, so SciBot could be utilized.

Introduction and references?

Definition: "Summary of the relevant literature and what is known about the topic. What was known before this research?"

Typically, this is part of an article where most citations are found. Having references limits the area reviewed and focuses on the topic. Generative AIs are good at synthesizing known content and writing a plausible summary of the actual field, but the generated text would not have citations. What is a summary of the relevant literature worth without references?

The use of SciBot's this feature, i.e., no citations, cuts much deeper. A multibillion-dollar economic system (from the authors' existential needs to get tenured or promoted through publishers selling verified knowledge in journals and article PDFs) is built on counting citations. No citations in scholarly articles would open another can of worms because citations form the foundation of the academic publication system.

The evaluation of scientists' performance today rests on their citation history, i.e., based on the assumption that the number of citations (either in the professional literature or measuring the interest in the general readership) reflects the significance of the published work and is a key performance indicator of an individual. [16] The number of citations is also considered the ultimate measure of a journal's influence and signals its prestige. Should citations disappear, the present approach of declaring the "impact" of an article based on which journal with what "Impact Factor" published the work (supposed to be "Journal Impact Factor") will collapse.

Generative AI will increase the value of research publications but may reduce the number and value of reviews. Many reviews are written when graduate students or postdocs join a research group. Their first task is usually to write a review of the area they will work on. Now, anybody can generate an overview within minutes or hours.

Questions remaining:

Science chatbots (when they become a reality) will undoubtedly increase productivity. Will higher productivity result in better science or only more papers and the inflation of words, or both? Is plausibility enough? *Is more* always *better*?

Banning ChatGPT in schools is not a good idea since this tool could be useful for education. LLM-generated texts can be great educational tools because they can quickly compose

an overview of a field but remember that the output is never better than the input. [17] Anyway, how can one fight it?

LLM-s now read and generate texts. Once this new text is published, it becomes part of the general knowledge these chatbots are trained on. Will text-creating bots be followed by computerized tools that identify AI-created texts; will scholarly article-writing bots be followed by scholarly article-reading bots?

Paradigm-changing discoveries drive humanity forward. Can this hypothetic SciBot be faultlessly trained on existing knowledge to draw proper conclusions on a new phenomenon?

How can it be verified that a text generated by a commercial chatbot is factual and true? This is a very significant issue in medical science and medicine, where people's lives may be at stake.

Will we remember how to write instead of taking out our smartphones to run the ChatGPT-x app? Using an app is a different skill than writing; selecting the right answer from several alternate options is not knowledge but recognition. Knowledge only grants you skills with practice. Recognizing someone does not mean you know the person.

For now, texts used to train LLMs are created by people, and we are imperfect with different cultural heritage. After new texts are written by LLMs then, the texts that have been created by LLMs join the overall body of texts. Consequently, an updated LLM will be trained on other texts written by other LLMs. Will this result in the dominance of consolidated knowledge?

The prestige of any journal is based on the validated, highly cited content published in it, because readers trust the source. What if a journal is flooded with AI-created reviews based on AI-created reviews? The most worrisome is that ChatGPT has no commitment to the truth or the facts for now. [10]

CONCLUSIONS

Generative AI can and will help us to integrate our knowledge and connect disciplines, which is a fantastic opportunity. Using LLMs will result in much higher productivity by connecting areas of science – the larger the field, the better. [18] Chatbots will be great educational tools because they can quickly compose an overview of a field, but we must remember that the output is never better than the input, and they will accelerate the spread of information, both good and bad. An LLM may be superbly intelligent, but connecting disciplines and providing new insights - although extremely powerful - do not constitute creativity. Human brain capacity is limited, but we are creative. In summary, generative AI is a tool that can be used for all purposes. Progress can be fought, but that fight has never been successful; the djinn is out of the bottle and cannot be returned. The Age of AI has begun [4]; we better learn how to deal with it.

Quote this article as Balogh LP, Are we there yet? Precis. Nanomed. 2023;6(1):1007-1012, https://doi.org/10.33218/001c.74009

Conflict of Interest

The author is an editor of the journal and declares no other conflicts of interest. For a signed statement, please contact the journal office at editor@precisionnanomedicine.com.

REFERENCES:

1. COPE Council. COPE Discussion Document: Artificial intelligence (AI) in decision making, English. https://doi.org/10.24318/9kvAgrnJ
2. Journal of Applied Learning & Teaching, JALT Vol.6 No.1 (2023) http://journals.sfu.ca/jalt/index.php/jalt/index
3. 20230309 GPT-3 and the future of publishing & academia webinar: https://www.youtube.com/watch?v=84bnHPhUBvQ
4. https://www.gatesnotes.com/The-Age-of-AI-Has-Begun
5. Could AI help you to write your next paper? Nature, 611, 192-193 (2022) https://doi.org/10.1038/d41586-022-03479-w

6. ChatGPT; five priorities for research, Nature, 614, 224-226 (2023) https://doi.org/10.1038/d41586-023-00288-7
7. ChatGPT Can Now Browse the Web, Help Book Flights and More, https://www.cnet.com/tech/mobile/chatgpt-can-now-browse-the-web-book-flights-and-more/
8. https://openai.com/blog/chatgpt-plugins
9. https://www.theguardian.com/commentisfree/2023/mar/03/fake-news-chatgpt-truth-journalism-disinformation
10. https://www.technologyreview.com/2023/02/14/1068498/why-you-shouldnt-trust-ai-search-engines/
11. Kim, SG. Using ChatGPT for language editing in scientific articles. Maxillofac Plast Reconstr Surg 45, 13 (2023) https://doi.org/10.1186/s40902-023-00381-x
12. Gaggioli A (2023) Ethics: disclose the use of AI in scientific manuscripts, Nature 614, 413, (2023), https://doi.org/10.1038/d41586-023-00381-x
13. Liebrenz, M, Schleifer, R, Buadze, A, Bhugra, D, Smith, A, Lancet Digit Health. 2023 Mar;5 (3):e105 e106. https://doi.org/10.1016/S2589-7500 (23)00019-5. Epub 2023 Feb 6.
14. Balogh LP, Science, Pseudo-science, False, and Fake science. Why is this happening, and what can you do? Precis. Nanomed. 2022;5(1):870-878, https://doi.org/10.33218/001c.34241
15. Stokel-Walker, C, ChatGPT is listed as an author on research papers: many scientists disapprove, Nature 613, 620-621 (2023), https://doi.org/10.1038/d41586-023-00107-z
16. Lajos P Balogh, Balancing Interests of Science, Scientists, and the Publishing Business, https://doi.org/10.29016/180418.1
17. https://www.timesnownews.com/education/to-ban-chatgpt-in-universities-is-not-a-good-idea-says-infosys-founder-narayana-murthy-article-98184408
18. Yiwen Hu and Markus J. Buehler Deep language models for interpretative and predictive materials science, APL Mach. Learn. 1, 010901 (2023); https://doi.org/10.1063/5.0134317