Trends in Information Management (TRIM)
ISSN: 0973-4163
12(1), pp. 14-16

How Can ChatGPT Influence Academic Information Seeking?

Mike Thelwall

Abstract

This paper discusses some possible implications for academic information seeking of the availability of large language models like ChatGPT that can summarise topics and answer questions. A particular problem is that whilst their answers may almost always be plausible and often correct, they sometimes make mistakes, and these may be difficult to detect.

Keywords*: ChatGPT; Large Language Models; Information Seeking; Scholarly Information*

Lecturers and students searching for academic information online to find out about a new topic might use one of the current generation of large language models (LLMs) like ChatGPT to help them. LLM-based computer systems ingest huge amounts of text and then use it to generate responses to user questions or other prompts **(Ayoub et al., 2024)**. The text ingested ensures that they have an ability to respond to academic enquiries too. They can also be fed information to summarise **(Hake et al., 2024; see also: Liu et al., 2024)**. For academic information seeking, academics might ask them to summarise academic knowledge on a topic and a student might even ask an LLM system to write an essay for them. Even those not directly using LLMs for their academic information seeking tasks might use a web search engine, like Bing, that suggests LLM responses in addition to listing websites in response to a query. Whilst this new technology has the advantage of giving new options to help find, process, and summarise information, it also has disadvantages.

An important disadvantage is that LLMs can make mistakes. They work using probabilities to generate their responses rather than with certainty and have imperfect recall of their source material. These things combine to mean that they can generate plausible responses that are incorrect and do not match the material that they have read. One well known aspect of this is the hallucination. For example, ChatGPT has often been caught creating fake references for its responses, including academic paper titles that do not exist **(Chelli et al., 2024)**. In addition, they can also make mistakes by reporting things that are not correct or that are out of date. When used by academics, this is a challenge to research integrity **(Bin-Nashwan et al., 2023; Kim, 2024)**.

It is perhaps useful to compare ChatGPT to Wikipedia. This website is a useful source of information on many topics but can be biased and

misleading. It can even occasionally contain errors. Its advantage over ChatGPT, however, is that it tends to reference its sources so there is usually a way to check a fact or identify the origins of its information. If not, the suspect material could be removed or ignored. This property is because referencing information is central to the way that Wikipedia works, perhaps both for the importance of referencing and to prevent anyone for entering their personal opinions. In contrast, not providing references is central to how ChatGPT works because it entails providing precise information whereas LLMs work by storing patterns in the input data but not the data itself. This fundamental limitation of LLMs can and is being addressed by complete systems like ChatGPT by seamlessly integrating with non-LLM technologies, such as search engines, to help identify precise information (e.g., when I ask it for sources, it sometimes reports that it is searching the web for answers). Nevertheless, this approach seems likely to always be imprecise because the LLM starts by not knowing where its information is from.

Due to these problems, it is important to remember, and to emphasise to students and new researchers, that ChatGPT and other LLMs cannot be trusted. They may be helpful at finding and summarising information about new topics, but knowledge from them should be checked with other sources before it can be relied upon for important academic tasks.

References

Ayoub, N. F., Lee, Y. J., Grimm, D., & Divi, V. (2024). Head-to-head comparison of ChatGPT versus Google search for medical knowledge acquisition. *Otolaryngology–Head and Neck Surgery, 170*(6), 1484-1491. DOI: 10.1002/ohn.465

Bin-Nashwan, S. A., Sadallah, M., & Bouteraa, M. (2023). Use of ChatGPT in academia: Academic integrity hangs in the balance. *Technology in Society, 75*, 102370.DOI: 10.1016/j.techsoc.2023.102370

Chelli, M., Descamps, J., Lavoué, V., Trojani, C., Azar, M., Deckert, M., & Ruetsch-Chelli, C. (2024). Hallucination Rates and Reference Accuracy of ChatGPT and Bard for Systematic Reviews: Comparative Analysis. *Journal of Medical Internet Research, 26*, e53164. DOI: 10.2196/53164

Hake, J., Crowley, M., Coy, A., Shanks, D., Eoff, A., Kirmer-Voss, K., & Parente, D. J. (2024). Quality, accuracy, and bias in ChatGPT-based summarization of medical abstracts. *The Annals of Family Medicine, 22*(2), 113-120.DOI: 10.1370/afm.3075

Kim, S. J. (2024). Research ethics and issues regarding the use of ChatGPT-like artificial intelligence platforms by authors and reviewers: a narrative review. Science Editing, 11(2), 96-106. DOI: 10.6087/kcse.343

Liu, Y., Ju, S., & Wang, J. (2024). Exploring the potential of ChatGPT in medical dialogue summarization: a study on consistency with human preferences. *BMC Medical Informatics and Decision Making*, *24*(1), 75. DOI: 10.1186/s12911-024-02481-8

Corresponding Author
Mike Thelwall can be contacted at: m.a.thelwall@sheffield.ac.uk

Author Biography
Mike Thelwall
Mike Thelwall is a Professor of Data Science in the Information School at the University of Sheffield in the UK. He primarily investigates quantitative methods to support research evaluation, including artificial intelligence, citation analysis and altmetrics. Mike has also developed a wide range of software and methods for gathering and analysing social web data. His books include Web indicators for research evaluation: A practical guide. He is an associate editor of the Journal of the Association for Information Science and Technology and sits on five other editorial boards.
Information School, University of Sheffield, UK.