

Review of: "OpenAI ChatGPT Generated Content and Similarity Index: A study of selected terms from the Library & Information Science (LIS)"

Sijie Feng¹

¹ NYU Stern School of Business

Potential competing interests: No potential competing interests to declare.

This paper looks at a specific case of plagiarism with ChatGPT, with special application to library and information science. The authors use ChatGPT to generate papers in ten topic areas and analyze similarity to other papers in those fields using the software Turnitin.

1. A major issue is that the goal and intent of the paper is unclear. Is the goal to assess how good ChatGPT is at escaping detection for plagiarism? Or to assess how good ChatGPT is as a replacement for existing plagiarism detection software? These are different research questions that require different research methodology.
2. The methodology section is too vague. How exactly are the papers generated from ChatGPT? Is there some insight that the authors can provide on this process? Also, what kind of similarity index is used by Turnitin? Cosine, euclidean, manhattan, etc.? Authors need to outline details more precisely.
3. Ten topics is a very small sample size. Not sure what kinds of conclusions can be inferred from ten papers. Is it possible to generate more evidence?
4. Discussion is missing. There is not currently a discussion section on the implications of the findings.