

Review of: "Towards Responsible AI-Assisted Scholarship: Comparative Assessment of Generative Models and Adoption Recommendations"

Francisco Maria Calisto¹

¹ University of Lisbon

Potential competing interests: No potential competing interests to declare.

In the presented manuscript, the authors embark on a journey to explore the integration and comparative assessment of generative AI models in academic research, with a particular focus on their application, capabilities, and limitations. The study meticulously benchmarks four generative AI models – Claude, LaMDA, Sydney, and Galactica – providing a comprehensive view through quantitative scoring and qualitative analysis across various scholarly tasks. While the manuscript provides a valuable contribution to the field, there are several concerns and areas for improvement that need to be addressed to enhance the quality and impact of the work.

The related work section of the manuscript, while providing a foundational understanding, appears to lack depth in exploring the existing literature in the domain of AI in academic research, especially considering the intersectionality with healthcare (doi.org/10.1080/00207543.2023.2188101, doi.org/10.1016/j.ijhcs.2022.102922). The authors should consider incorporating more recent and relevant works to establish a robust theoretical framework and to position their work aptly within the existing body of knowledge.

While the methodology section provides a detailed view of the comparative prompting design and data analysis, it would be beneficial to delve deeper into the rationale behind selecting the specific AI models and the criteria for the assessment categories (doi.org/10.1016/j.inffus.2023.101805, doi.org/10.1016/j.artmed.2022.102285). A more thorough explanation or a visual representation of the methodological flow could enhance the comprehensibility and replicability of the study.

The findings section provides a rich set of data and insights into the capabilities and limitations of the AI models.

However, it would be beneficial to further elucidate the practical implications of these findings

(doi.org/10.1016/j.inffus.2023.03.008, doi.org/10.1145/3544548.3580682), especially considering the application in healthcare research. How can these insights be translated into actionable strategies for researchers, developers, and institutions in the healthcare domain?

The discussion provides a succinct summary of the findings and their implications. However, it would be enriched by providing a more detailed exploration of how these insights can shape future research and development in the field. Additionally, considering the ethical, practical, and technological challenges in implementing AI in healthcare research would provide a more rounded view.

The conclusion aptly summarizes the work, but could be enhanced by providing more concrete and actionable recommendations, especially tailored towards researchers and practitioners in the healthcare domain. Additionally, considering the ethical and practical aspects of implementing the recommendations would provide a more grounded and applicable set of guidelines.

In summary, the manuscript provides a valuable contribution to the field, exploring the capabilities and limitations of generative AI models in academic research. Addressing the aforementioned concerns and enhancing the depth and applicability of the work will undoubtedly elevate the impact and quality of the manuscript. I look forward to seeing the revisions and the enhanced version of this insightful work.