ES2631 Assignment 3

An article titled "**Artificial intelligence-assisted academic writing: recommendations for ethical use**" by Cheng et al. was published in Advances in Simulation. In the article, the authors discuss the roles, concerns and potential use cases of the prevalent Generative Artificial Intelligence (GenAI) tools powered by Large Language Models (LLMs) in the academic writing process to ensure research quality and academic integrity (Cheng et al., 2025). Based on the Engineering Reasoning Framework (R. Paul et al., 2019), this essay will critique the reasoning shown in implications and information incorporated.

Firstly, the implications show accuracy as the authors discuss some of the concerning impacts that Artificial Intelligence (AI) tools may create in the research fields. Throughout the article, the authors predict the AI's great potential in transforming the workflow of research practice (positively and negatively), and highlight the possible negative implications of AI over-reliance, such as plagiarism, copyright infringement, bias and adverse effects on researchers' skills. In terms of copyright infringement, K. Paul (2024) reported that numerous AI companies are bypassing the Robots Exclusion Protocol that blocks web crawler from accessing specific website content. Similarly, New York Times sued OpenAI and Microsoft for accessing news articles without permission to train their AI (Stempel, 2023). Hence, this crosschecks the authors' claim of plagiarism rooted in the training set issues associated with copyright infringement and bias.

Furthermore, the disadvantageous implications of AI use on researcher cognitive capabilities are also evaluated. For example, the authors expressed their concerns with dependence on GenAI tools hindering human research skills evolution and adaptation capabilities, for both novice and experienced researchers. As direct evidence, Kosmyna et al. (2025) demonstrated lower memory recall and less brain connectivity among participants who were allowed to use ChatGPT for composition, compared to Brain-only group. The neural, linguistic and scoring performances of the LLM users were worse over the course of 4 months (Kosmyna et al., 2025), suggesting the deterioration effects on cognitive capacity after prolonged use of GenAI tools (Lee et al., 2025), verifying the concerns raised by the authors.

On the contrary, the information provided lacks depth. The authors argue that LLM functionality as generation of *syntactically coherent* texts based on training data; hence is subject to bias and false information due to lack of fact-checking ability. As a result, LLMs cannot differentiate between real and fake information received. While the statements are explanatory, the authors fail to consider two underlying complexities. Firstly, Mikolov et al. (2013) and Pennington et al. (2014) have shown that the *semantic encoding capabilities* in high dimensional vector-space word representations in continuous space language models. Hence, LLMs are generally believed to possess the ability to parse semantic meaning in addition to syntax, which is not debated by the authors. Secondly, the authors do not discuss the availability of other complex LLM systems, such as Retrieval-Augmented Generation (RAG) (Lewis et al., 2021) and multi-agent systems (Trinh et al., 2025), which might be helpful in fact-checking. Perhaps the LLMs' weakness in differentiating facts is partially caused by lack of contextual information, which can be improved by RAG (Gao et al., 2024). In particular, Bai and Fu (2024) have shown that Full-Context Retrieval and Verification (FCRV) with RAG outperforms baseline models in fake news detection. Trinh et al. (2025) have demonstrated the improvement of 12.3% in Macro F1-score through a multi-agent system, suggesting great potential in fact-checking capabilities. A more in-depth analysis would compare the performance metrics of various aforementioned machine learning systems such as *sensitivity*, *specificity*, *Receiver Operating Characteristics*, and *F1-Score* to present a comprehensive perception on AI's robustness and reliability.

In conclusion, this essay has evaluated the reasoning behind the implications and information. While the implications are accurate, the information lacks depth.

Word Count: 598

AI Use Declaration

I Wong Jian Bin declare that I have not used generative AI in the process of completing this assignment.

References

Bai, Y., & Fu, K. (2024). A large language model-based fake news detection framework with RAG fact-checking [ISSN: 2573-2978]. *2024 IEEE International Conference on Big Data (BigData)*, 8617–8619. https://doi.org/10.1109/BigData62323.2024.10826000

Cheng, A., Calhoun, A., & Reedy, G. (2025). Artificial intelligence-assisted academic writing: Recommendations for ethical use. *Advances in Simulation*, *10*(1), 22. https://doi.org/10.1186/s41077-025-00350-6

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M., & Wang, H. (2024, March 27). Retrieval-augmented generation for large language models: A survey. https://doi.org/10.48550/arXiv.2312.10997

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025, June 10). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. https://doi.org/10.48550/arXiv.2506.08872

Lee, H.-P. (, Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–22. https://doi.org/10.1145/3706598.3713778

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., Riedel, S., & Kiela, D. (2021, April 12). Retrieval-augmented generation for knowledge-intensive NLP tasks. https://doi.org/10.48550/arXiv.2005.11401

Mikolov, T., Yih, W.-t., & Zweig, G. (2013, June). Linguistic regularities in continuous space word representations. In L. Vanderwende, H. Daumé III, & K. Kirchhoff (Eds.), *Proceedings of the 2013 conference of the north american chapter of the association for computational linguistics: Human language technologies* (pp. 746–751). Association for Computational Linguistics. Retrieved November 5, 2025, from https://aclanthology.org/N13-1090/

Paul, K. (2024). Exclusive: Multiple AI companies bypassing web standard to scrape publisher sites, licensing firm says. *Reuters*. Retrieved November 4, 2025, from https://www.reuters.com/technology/artificial-intelligence/multiple-ai-companies-bypassing-web-standard-scrape-publisher-sites-licensing-2024-06-21/

Paul, R., Niewoehner, R., & Elder, L. (2019). *The thinker's guide to engineering reasoning: Based on critical thinking concepts and tools* (2nd). Rowman & Littlefield Publishers. https://go.exlibris.link/kQHll5Dw

Pennington, J., Socher, R., & Manning, C. (2014, October). GloVe: Global vectors for word representation. In A. Moschitti, B. Pang, & W. Daelemans (Eds.), *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532–1543). Association for Computational Linguistics. https://doi.org/10.3115/v1/D14-1162

Stempel, J. (2023). NY times sues OpenAI, microsoft for infringing copyrighted works. *Reuters*. Retrieved November 4, 2025, from https://www.reuters.com/legal/transactional/ny-times-sues-openai-microsoft-infringing-copyrighted-work-2023-12-27/

Trinh, T., Nguyen, M., & Hy, T.-S. (2025, June 22). Towards robust fact-checking: A multi-agent system with advanced evidence retrieval. https://doi.org/10.48550/arXiv.2506.17878