**Peer Response**

by Ali Alzahmi - Monday, 30 June 2025, 5:57 PM

The post is a well-considered and accessible response to the ethical and operational risks that large language models (LLMs) such as GPT-3 present. The difference between the fluency and language comprehension by the AI models is one of the most important aspects that were brought up. The idea that GPT-3 can produce text fluently, without really knowing what it means, exemplified by Hutson (2021) can be tied to the issues presented in recent literature on AI ethics. Such models are, as described by Bender et al. (2021), deemed as stochastic parrots systems that know how to imitate the human language by recognising patterns but cannot understand them semantically. This poses risks in places where accuracy of meaning is vital like law counsel or health facts. Human overtrust is another factor that worsens the risk and is also mentioned by Jakesch et al. (2023); people believe that something is accurate just because it is written well. Moreover, the post contains an excellent time-relevant remark concerning academic integrity. The advent of ghostwriting and AI-generated publications means that the academic field is experiencing a new type of bad behavior that confuses authorship and who is to be held responsible. The situation expressed by Kendall and Teixeira da Silva (2024) concerning paper mills and the possibility to change authorship due to AI is relevant to the scenario that Fahad is talking about. The science-based research can be damaged and compromised when large-scale researchers employ LLMs to generate false or pseudo-natural studies. The argument on monopoly of Big Tech in developing and deploying AI that Fahad provides is also sound. The paradox of employing the same organisations that produce damage to develop and oversee the ethical AI practice is described by Saeitra, et al. (2022). The incompatibility of this conflict of interest does not support transparency, auditability, and wholesome access to LLM technologies.

**References**

Bender, E.M., Gebru, T., McMillan-Major, A. and Shmitchell, S. (2021) 'On the dangers of stochastic parrots: Can language models be too big?', *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pp. 610–623. Available at: https://doi.org/10.1145/3442188.3445922 (Accessed: 28 June 2025).

Jakesch, M., Hancock, J.T. and Naaman, M. (2023) 'Human heuristics for AI-generated language are flawed', *Proceedings of the National Academy of Sciences*, 120(11), e2208839120. Available at: https://doi.org/10.1073/pnas.2208839120 (Accessed: 29 June 2025).

Sætra, H.S., Coeckelbergh, M. and Danaher, J. (2022) 'The AI ethicist's dilemma: fighting Big Tech by supporting Big Tech', *AI and Ethics*, 2(1), pp. 15–27. Available at: https://doi.org/10.1007/s43681-021-00123-7 (Accessed: 29 June 2025).

**Peer Response**

by Ali Yousef Ebrahim Mohammed Alshehhi - Monday, 30 June 2025, 6:20 PM

Your reflection, Fahad, is a very balanced and well-thought analysis of ethics and social issues of large language models (LLMs) usage, such as GPT-3. I especially enjoy the stress you put on the lack of correlation between functional language proficiency and real knowledge, which Hutson (2021) raises. The difference between these two concepts is misperceived by the laypersons and even healthcare professionals who rely too greatly on the coherence of AI-produced output (Jagatheesaperumal et al., 2021). Your implications made regarding the issue of academic integrity, along with the validity of scientific publication, are particularly topical nowadays. The increased activity in ghostwriting and fake authorship with the usage of LLMs is in fact triggering concerns regarding the production and justification of knowledge (Sætra et al., 2022). The arguments presented by Kendall and Teixeira da Silva (2024) add some weight to the realities of this problem and call the academic community to reconsider the models of authorship and peer review. I also enjoy your reference to Gibney (2022) concerning the role of Big Tech in monopolising AI tools and the potentially more democratic and transparent future of the open-source alternative. With that being said, it could have been nice to briefly mention what protection measures or regulation bodies are proposing to safeguard against these issues that exist at present (Kasneci et al., 2023). Moreover, your conclusion is justifiable because it is in need of a balance between innovation and ethical oversight. The argument about the necessity to adjust the ethical systems to technological advancement corroborates the idea of shifting dynamic models rather than rigid codes of conduct (Van Der Vlist et al., 2024). On the whole, your post is a multi-faceted critique of both the technical and ethical aspects of LLMs blending the two very well. Another direction you can take your argument may be the importance of working between ethicist and technologists and policymakers but collaborate to create guidelines which can be modified as technology is changed. The dangers to truth, trust, and equity as you mentioned. are immense, in the absence of such frameworks.

Chat to us!

**References**

Jagatheesaperumal, S.K., Rahouti, M., Ahmad, K., Al-Fuqaha, A. and Guizani, M. (2021) 'The duo of artificial intelligence and big data for industry 4.0: Applications, techniques, challenges, and future research directions', *IEEE Internet of Things Journal*, 9(15), pp. 12861–12885. Available at: https://ieeexplore.ieee.org/document/9667102 (Accessed: 30 June 2025).

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E. and Krusche, S. (2023) 'ChatGPT for good? On opportunities and challenges of large language models for education', *Learning and Individual Differences*, 103, p. 102274. Available at: https://doi.org/10.1016/j.lindif.2023.102274 (Accessed: 30 June 2025).

Sætra, H.S., Coeckelbergh, M. and Danaher, J. (2022) 'The AI ethicist's dilemma: fighting Big Tech by supporting Big Tech', *AI and Ethics*, 2(1), pp. 15–27. Available at: https://doi.org/10.1007/s43681-021-00123-7 (Accessed: 30 June 2025).

Van Der Vlist, F., Helmond, A. and Ferrari, F. (2024) 'Big AI: Cloud infrastructure dependence and the industrialisation of artificial intelligence', *Big Data & Society*, 11(1), p. 20539517241232630. Available at: https://doi.org/10.1177/20539517241232630 (Accessed: 30 June 2025).

Permalink     Show parent     Reply

---

**Peer Response**

by [Koulthoum Hassan Ahmad Flamerzi](#) - Monday, 30 June 2025, 10:58 PM

Thanks, Fahad, for the thorough and timely look at the ethical and social hurdles LLMs throw our way. Your observations about trust, misuse of authorship, and Big Techs heavy hand underline how quickly we need smart safeguards.
One obvious safeguard is a clear rule that students and researchers must say when AI helped their writing. Journals and schools could require authors to name the tool-was it GPT-3, Bard, or another-and explain what it did. Such a step fits recent publisher drives to keep research honest and to squeeze out ghost-writing or paper-mill scams (Kendall and Teixeira da Silva, 2024).
Institutions also need their own friendly, workable AI rules. Universities and labs should draft plain guidelines spelling out when, and how, LLMs can show up in homework, theses or papers. Policies might limit blind, unsupervised output, demand solid human review, and teach users to read AI text with a critical eye.
Pushing open-source LLM projects forward is another must-do if we want to fight bias and opacity. As Gibney (2022) points out, a clear model lets teams' peek under the hood, test, tweak, or toss out the faulty parts instead of just trusting a black box. With that kind of setup, outside experts and campus ethics boards could run regular audits to check that any AI-aided work is fair and truthful before it ever sees publication.
Promoting digital literacy and critical-thinking training will help readers, reviewers, and authors tell real research apart from text churned out by AI. An academic community that knows how tools work is harder for bad actors to fool.
When we pair smart rules, better training, and sound technology, we can enjoy what LLMs offer while keeping research trustworthy and serving the wider public good.

**References:**
Gibney, E. (2022) 'Open-source language AI challenges big tech's models', Nature. Available at: **https://doi.org/10.1038/d41586-022-01705-z**

Kendall, G. and Teixeira da Silva, J.A. (2024) 'Risks of abuse of large language models, like ChatGPT, in scientific publishing: Authorship, predatory publishing, and paper mills', Learned Publishing, 37(1). Available at: **https://www.graham-kendall.com/papers/ktds2023a.pdf**

Permalink     Show parent     Reply

---

◄ Initial Post                                                                                                    Initial Post ►

Powered by Moodle

Chat to us!