Peer Response 1

Thank you, Ali, for your contribution. You raise important ethical concerns about authenticity, bias, and ownership in the use of deep learning technologies. I agree that the ability of generative models to blur the boundaries between human and machine creativity has profound implications for trust and accountability. As Pavlos also noted, this growing “synthetic reality pollution” intensifies the challenge of distinguishing genuine content from fabricated material, thereby eroding public confidence in information systems. Recent research by Altay et al. (2024) supports this, showing that exposure to AI-generated news can significantly reduce users’ trust in legitimate journalism, even when such content is clearly labelled.

Expanding on your discussion of bias and fairness, a major concern lies in the lack of transparency within training data. Birhane, Prabhu and Kahembwe (2022) argue that many multimodal datasets contain uncurated and ethically problematic material, which leads models to reproduce harmful stereotypes. They propose “dataset accountability”, which involves documenting data provenance, consent, and potential harms. This approach could directly address the fairness and integrity issues Ali identified.

Another overlooked ethical dimension is environmental sustainability. Strubell, Ganesh and McCallum (2019) found that training large-scale AI models can generate carbon emissions equivalent to those produced by several cars over their lifetimes. Thus, beyond fairness and misinformation, the environmental costs of generative AI must also form part of the wider ethical conversation. Addressing these interconnected issues

holistically is essential if we are to ensure that innovation develops in a way that serves society responsibly.

References

Altay, S., Bien-Aimé, S., Jia, H. & Wu, M. (2024). *People are skeptical of headlines labelled as AI-generated: Full news articles labelled as AI-generated are perceived as less trustworthy (but not as less accurate).* PNAS Nexus, 3(10): pgae403. Available at: https://academic.oup.com/pnasnexus/article/3/10/pgae403/7795946 (Accessed 18 October 2025)

Birhane, A., Prabhu, V. U. & Kahembwe, E. (2022) *Multimodal datasets: misogyny, pornography, and malignant stereotypes*, Patterns, 3(4), 100481. Available at https://arxiv.org/abs/2110.01963 (Accessed 18 October 2025)

Strubell, E., Ganesh, A. & McCallum, A. (2019) *Energy and policy considerations for deep learning in NLP*, *ACL 2019*. Available at https://aclanthology.org/P19-1355 (Accessed 18 October 2025)

Peer Response 2

Nasser, your discussion insightfully captures the ethical tensions surrounding deep learning, especially regarding authorship, accountability, and fairness. I agree with your observation about the asymmetry of power between technology corporations and end users. As you note, regulatory frameworks such as the EU AI Act attempt to mitigate these imbalances, yet the challenge of consistent implementation remains significant (Floridi, 2019). I also support Fabian's argument that regulation, when proportionate and transparent, can promote rather than restrict innovation. Effective governance can build public trust, which is fundamental for sustainable technological adoption (Jobin, Ienca and Vayena, 2019).

Building on these points, transparency and explainability are crucial to ensuring ethical outcomes. As Floridi and Cowls (2019) propose, responsible AI should adhere to principles of beneficence, non-maleficence, autonomy and justice. These principles depend on accountability and openness in system design. Similarly, Binns (2018) argues that contestability and interpretability empower users to challenge automated decisions, thereby reinforcing fairness and public confidence. Without such mechanisms, bias and opacity risk undermining both innovation and societal benefit. This alignment of ethical regulation and technological progress reflects the emerging consensus that innovation flourishes best in environments grounded in trust, clarity and shared values.

Ultimately, I agree with both of you that stronger, well-designed oversight can serve as an enabler of creativity rather than a barrier. Rather than stifling progress, robust ethical

and legal frameworks can stimulate innovation by providing clarity and stability. Developers who design within transparent and equitable boundaries are more likely to produce technologies that are socially legitimate, inclusive and widely trusted.

References:

Binns, R. (2018) *Fairness in Machine Learning: Lessons from Political Philosophy*. Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency (FAT*)*, pp. 149–159. Available at: https://proceedings.mlr.press/v81/binns18a/binns18a.pdf (Accessed 11 October 2025)

Floridi, L. (2023) *Translating principles into practices of digital ethics: Five risks of being unethical*. Philosophy & Technology, 36(1), 11. Available at: https://link.springer.com/article/10.1007/s13347-019-00354-x (Accessed 11 October 2025)

Floridi, L. and Cowls, J. (2019) *A Unified Framework of Five Principles for AI in Society.* Harvard Data Science Review, 1(1). Available at: https://doi.org/10.1162/99608f92.8cd550d1 (Accessed 11 October 2025)

Jobin, A., Ienca, M. and Vayena, E. (2019) *The global landscape of AI ethics guidelines*. Nature Machine Intelligence, 1(9), pp. 389–399. Available at: https://doi.org/10.1038/s42256-019-0088-2 (Accessed 11 October 2025)