Peer Response

by Abdulrahman Alhashmi - Sunday, 5 October 2025, 1:33 PM

Saleh, your worries point to main risks. Moral issues stand out as pressing and plain. I offer three solid protections.

Start by seeing AI generators as tools with two sides. Look at possible damage, not only errors. Apply risk scans, slow launches, and misuse strategies. They reduce chances of broad lies from false content (Brundage et al., 2018; Chesney and Citron, 2019).

Next, tighten data guidelines. Rely on checked, shared sources. Remove private facts and scan for data leaks ahead. Models tend to repeat learned parts, so run privacy audits and group trials to halt release. Record problems and solve them before going live (Carlini et al., 2021; Weidinger et al., 2021).

Next, add duty to daily jobs. Outside checks, data records, and error follows build real trust. Mix source tags (useful yet narrow) with online guidelines: warnings for bold claims, proof connections, and limits on high-risk parts. Link product drops to expert checks. Track live use with plain halt steps (Raji et al., 2020).

Such steps steer growth, without halting it. Your points on bias, rights, and privacy form a complete defence. Safety starts in planning, flows to trials, and holds in oversight. This loop plan, test, assess, track guards people, holds trust, and yields true benefits.

References

Brundage, M. et al. (2018) 'The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation', arXiv.

Carlini, N. et al. (2021) 'Extracting Training Data from Large Language Models', Proceedings of USENIX Security Symposium.

Chesney, R. and Citron, D.K. (2019) 'Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security', California Law Review, 107(6), pp. 1753–1819.

Raji, I.D. et al. (2020) 'Closing the AI Accountability Gap: Defining Auditing and Oversight for Algorithmic Systems', Proceedings of the ACM FAccT Conference, pp. 33–44.

Weidinger, L. et al. (2021) 'Ethical and Social Risks of Harm from Language Models', arXiv.

Re: Initial Post

by Shaikah Salim Mohammed Alkhaayyal Alharthi - Sunday, 5 October 2025, 6:55 PM

Saleh, you've pointed out how Deep Learning is changing the way we do creative and cognitive tasks. The impact is huge, and it surely comes with benefits, but they must be balanced with the with ethical ramifications.

Concerns regarding AI contents manipulation and veracity. AI contents borders between fake and real, and it is difficult to tell the difference between the two. This is damaging to public perception, particularly with respect to deep fakes and synthetic news.

Regarding bias, it is indisputable. As Bender et al. (2021) mentions, large models amplify the biases of their training data and, therefore, increase the unjust inequities in society. The risk to society is increased when these are applied to employment and justice decision systems.

Furthermore, the questions of ownership and consent are also problematic. As you mentioned, generative models pose the risk of reproducing creative and textual works without attribution (Sahil et al, 2024)

To foster responsible innovation, i believe we must focus on ethical governance, transparency, and regulation (Ienca, Buchholz & Vayena, 2025). Deep Learning should be used to enhance innovation and protect human rights and accountability.

Permalink   Show parent   Reply

Peer Response

by Ali Yousef Ebrahim Mohammed Alshehhi - Monday, 13 October 2025, 7:54 PM

You have presented an excellent clear post Saleh. The article identifies important ethical problems which stem from deep learning technology applications specifically in generative AI systems including DALL·E and ChatGPT. Your viewpoint about veracity and manipulation receives my complete support. AI technology produces highly realistic fake content that destroys public confidence in information which leads to quick dissemination of false information. Westerlund (2019) describes deepfake technology as a major threat to society because it creates problems for political systems and media organizations. The implementation of provenance tracking together with mandatory AI content labeling systems enables content verification and transparency for users.

AI systems face a major problem because of embedded bias which you correctly pointed out. The learning process of these models from past data results in the replication of existing social disparities according to Mehrabi et al. (2021). Organizations can reduce these negative effects by performing bias audits and fairness testing before launching their systems.

Your mention of ownership, consent, and privacy also touches on essential debates. The growth of data governance systems together with well-defined copyright regulations will protect creators' rights when AI systems develop new capabilities according to Vinuesa et al. (2020). People need both regulatory systems and digital literacy education to assess AI-generated content properly.

References
Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K. & Galstyan, A. (2021) 'A Survey on Bias and Fairness in Machine Learning', ACM Computing Surveys, 54(6), pp. 1–35.
Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S.D., Tegmark, M. & Fuso Nerini, F. (2020) 'The role of artificial intelligence in achieving the Sustainable Development Goals', Nature Communications, 11(1), pp. 1–10.
Westerlund, M. (2019) 'The Emergence of Deepfake Technology: A Review', Technology Innovation Management Review, 9(11), pp. 39–52.

Permalink   Show parent   Reply