Initial post

by Georgios Papachristou - Friday, 28 March 2025, 5:00 AM

Artificial intelligence (AI) developments such as image and text creation contribute more and more to transforming among others, the fields of finance, journalism, healthcare, and security, offering to them a new perspective in automation, efficiency, and effectiveness. Nevertheless, there are several ethical concerns linked to transparency, data security and privacy, accountability, and bias that should be take into account when it comes to the application of such technologies (Harlow, 2018; Huang et al., 2023).

In particular, deep learning is a technology hard to explain and understand, leading to major transparency and bias concerns on the intent of developers and the dataset that is being used to train the model. Another ethical issue that should be considered is linked to the use of personal or sensitive corporate data during the training phase of an AI model and the avoidance of misuse or malicious use of data, such as leakage or tampering. Last but not least, accountability and ownership of an AI system or agent cannot be defined and in turn in case of a failure there is no ability to identify the root cause and in turn no clarity on who is responsible as many people may be involved for example in the programming, data input and analysis, testing, or deployment of a system (Safdar et al., 2020; Huang et al., 2023; Harlow, 2018).

To sum up, though AI systems offer enormous benefits to several fields as well as daily life, transparency, explainability, data security and privacy, and accountability are some of the ethical issues that human beings should concern when they design, implement, deploy and use such technologies.

References

Harlow, H., 2018. *Ethical concerns of artificial intelligence, big data and data analytics.* s.l., Kidmore End: Academic Conferences International Limited.

Huang, C., Zhang, Z., Mao, B. & Yao, X., 2023. An Overview of Artificial Intelligence Ethics. *IEEE Transactions on Artificial Intelligence,* 4(4), pp. 799-819.

Safdar, N., Banja, J. & Meltzer, C., 2020. Ethical considerations in artificial intelligence. *European Journal of Radiology,* Volume 122, p. 108768.

Peer Response: Initial post

by Rodrigo Pereira Cruz - Saturday, 5 April 2025, 9:10 PM

My peer's entry manages to provide a succinct, clear, and fair overview of not only the advantages of deep learning in image and text generation, but also of its major shortcomings and ethical concerns, such as bias acquisition, lack of explainability and accountability, and data protection.

It has become imperative for all stakeholders involved in the artificial intelligence (AI) system lifecycle to become active in addressing such ethical concerns (Mbiazi et al., 2023). With the advent of generative AI, in particular, such concerns have now become a much more urgent situation due to how such systems can synthesise existing material to produce new content, topics such as social inequalities also becoming critical factors (Al-kfairy, 2024).

Overall, the increased integration of deep learning in modern society, while offering many benefits, also has the potential of worsening already existing problems, now exacerbated by the prominence of generative AI. Only through the establishment and rigorous enforcement of ethical guidelines will this incredibly promising driver of modern AI ultimately achieve its true potential.

References

Mbiazi, D., Bhange, M., Babaei, M., Sheth, I. and Kenfack, P.J. (2023) *Survey on AI Ethics: A Socio-technical Perspective*. Available at: https://arxiv.org/abs/2311.17228 (Accessed: 5 April 2025).

Al-kfairy, M., Mustafa, D., Kshetri, N., Insiew, M. and Alfandi, O. (2024) 'Ethical Challenges and Solutions of Generative AI: An Interdisciplinary Perspective', *Informatics* 11(3), pp. 58. Available at: https://doi.org/10.3390/informatics11030058.

Peer Response

by Guilherme Pessoa-Amorim - Monday, 7 April 2025, 4:58 PM

Georgios provides a balanced overview of the ethical challenges surrounding deep learning applications, highlighting critical concerns such as lack of transparency, data privacy, and the difficulty of assigning accountability within complex AI systems (Cheong, 2024). Rodrigo's response further emphasizes the urgency of these issues in the context of generative AI, particularly as it amplifies risks related to bias and the misuse of synthesized content (Al-kfairy et al., 2024). Their observations reflect the broader dilemma facing AI development—namely, how to ensure ethical responsibility and accountability in systems that often operate as "black boxes" (Siebert et al., 2023).

Building on these insights, it is equally important to consider how generative AI challenges foundational principles in the creative industries. Recent tools such as DALL·E and ChatGPT introduce unprecedented opportunities for democratizing creativity, and preserving cultural heritage (in particular, the latest 4.5 model, and its dedicated tools such as Sora for video generation) (Anantrasirichai and Bull, 2022). Yet, they simultaneously risk displacing human creative labour and devaluing originality (Erickson, 2024). Furthermore, the use of training data containing copyrighted or personal content without clear consent exacerbates concerns about intellectual property and privacy rights (Bracha, 2023). Without clear ownership

definitions or regulatory oversight, the line between inspiration and plagiarism becomes increasingly blurred, with potentially unpredictable consequences for both human wellbeing and economic activity.

In conclusion, while deep learning models offer transformative potential, their ethical implications must not be overlooked. A multidisciplinary approach combining legal, technical, and societal frameworks is essential to guide the responsible use of generative AI and to ensure its alignment with human values and societal norms (Feldman, 2017; Mammen et al., 2024).

References

Al-kfairy, M. et al. (2024) 'A Systematic Review and Analysis of Ethical Challenges of Generative Ai: An Interdisciplinary Perspective'. Rochester, NY: Social Science Research Network. Available at: https://doi.org/10.2139/ssrn.4833030.

Anantrasirichai, N. and Bull, D. (2022) 'Artificial Intelligence in the Creative Industries: A Review', Artificial Intelligence Review, 55(1), pp. 589–656. Available at: https://doi.org/10.1007/s10462-021-10039-7.

Bracha, O. (2023) 'The Work of Copyright in the Age of Machine Production'. Rochester, NY: Social Science Research Network. Available at: https://doi.org/10.2139/ssrn.4581738.

Cheong, B.C. (2024) 'Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making', Frontiers in Human Dynamics, 6. Available at: https://doi.org/10.3389/fhumd.2024.1421273.

Erickson, K. (2024) 'AI and Work in the Creative Industries: Digital Continuity or Discontinuity?'

Rochester, NY: Social Science Research Network. Available at: https://doi.org/10.1080/17510694.2024.2421135.

Feldman, S. (Salevati) (2017) ‘Co-Creation: Human and AI Collaboration in Creative Expression’, in. Electronic Visualisation and the Arts (EVA 2017), BCS Learning & Development. Available at: https://doi.org/10.14236/ewic/EVA2017.84.

Mammen, C.E. et al. (2024) Creativity, Artificial Intelligence, and the Requirement of Human Authors and Inventors in Copyright and Patent Law.
Siebert, L.C. et al. (2023) ‘Meaningful human control: actionable properties for AI system development’, AI and Ethics, 3(1), pp. 241–255. Available at: https://doi.org/10.1007/s43681-022-00167-3.

Summary Post

by Georgios Papachristou - Tuesday, 9 April 2025, 5:07 AM

In my initial post, I acknowledged the benefits that artificial intelligence (AI) developments such as image and text creation offer to various business areas as well as to our daily life enhancing efficiency and effectiveness. However, I also made reference to the ethical concerns deriving from the difficulty in comprehending how these systems work and the reasoning of their conclusions. In particular, I discussed in detail issues such as transparency, data security and privacy, accountability, and bias arguing that these should are elements that should be considered during the design, implementation, deployment and use of such technologies (Safdar et al., 2020; Huang et al., 2023; Harlow, 2018).

As a response to my initial post, Rodrigo points out that given the rise of generative AI such concerns have now become a much more urgent situation due to how such systems function

and their ability to artificially synthesise existing elements to produce new content. That being the case social inequalities are also becoming any issue of concern (Al-kfairy, 2024). Additionally, he proposes that considering the current situations and these ethical concerns, the establishment and rigorous enforcement of ethical guidelines would potentially provide the framework needed in place in order AI ultimately to achieve its true potential (Mbiazi et al., 2023).

Furthermore, Guilherme reinforced the criticality of these concerns by making a reference to the power of generative AI and to the risks that recent tools such as DALL·E and ChatGPT pose to cultural heritage, originality, and human creative labour. Moreover, he raises concerns regarding intellectual property and privacy rights when data including copyrighted of personal content are used for training purposes (Cheong, 2024; Al-kfairy et al., 2024; Siebert et al., 2023; Anantrasirichai and Bull, 2022; Erickson, 2024; Bracha, 2023; Feldman, 2017; Mammen et al., 2024).

To sum up, this discussion though acknowledged the benefits that AI could offer to multiple areas, analysed ethical concerns linked to data security and privacy, accountability, ownership, social inequalities that should be taken seriously into account and should be potentially dealt through the development and implementation of ethical guidelines.

References

Al-kfairy, M. et al. (2024) ‘A Systematic Review and Analysis of Ethical Challenges of Generative Ai: An Interdisciplinary Perspective’. Rochester, NY: Social Science Research Network. Available at: https://doi.org/10.2139/ssrn.4833030.

Al-kfairy, M., Mustafa, D., Kshetri, N., Insiew, M. and Alfandi, O. (2024) 'Ethical Challenges and Solutions of Generative AI: An Interdisciplinary Perspective', *Informatics* 11(3), pp. 58. Available at: https://doi.org/10.3390/informatics11030058.

Anantrasirichai, N. and Bull, D. (2022) ‘Artificial Intelligence in the Creative Industries: A Review’, Artificial Intelligence Review, 55(1), pp. 589–656. Available at: https://doi.org/10.1007/s10462-021-10039-7.

Bracha, O. (2023) ‘The Work of Copyright in the Age of Machine Production’. Rochester, NY: Social Science Research Network. Available at: https://doi.org/10.2139/ssrn.4581738.

Cheong, B.C. (2024) ‘Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making’, Frontiers in Human Dynamics, 6. Available at: https://doi.org/10.3389/fhumd.2024.1421273.

Erickson, K. (2024) ‘AI and Work in the Creative Industries: Digital Continuity or Discontinuity?’ Rochester, NY: Social Science Research Network. Available at: https://doi.org/10.1080/17510694.2024.2421135.

Feldman, S. (Salevati) (2017) ‘Co-Creation: Human and AI Collaboration in Creative Expression’, in. Electronic Visualisation and the Arts (EVA 2017), BCS Learning & Development. Available at: https://doi.org/10.14236/ewic/EVA2017.84.

Harlow, H., 2018. *Ethical concerns of artificial intelligence, big data and data analytics.* s.l., Kidmore End: Academic Conferences International Limited.

Huang, C., Zhang, Z., Mao, B. & Yao, X., 2023. An Overview of Artificial Intelligence Ethics. *IEEE Transactions on Artificial Intelligence,* 4(4), pp. 799-819.

Mammen, C.E. et al. (2024) Creativity, Artificial Intelligence, and the Requirement of Human Authors and Inventors in Copyright and Patent Law.

Mbiazi, D., Bhange, M., Babaei, M., Sheth, I. and Kenfack, P.J. (2023) *Survey on AI Ethics: A Socio-technical Perspective*. Available at: https://arxiv.org/abs/2311.17228 (Accessed: 5 April 2025).

Safdar, N., Banja, J. & Meltzer, C., 2020. Ethical considerations in artificial intelligence. *European Journal of Radiology,* Volume 122, p. 108768.

Siebert, L.C. et al. (2023) 'Meaningful human control: actionable properties for AI system development', AI and Ethics, 3(1), pp. 241–255. Available at: https://doi.org/10.1007/s43681-022-00167-3.