UNIVERSITY
OF TWENTE.

RESTS
Professional and Academic Development

The blind eye of technology: ethical impacts of generative AI.

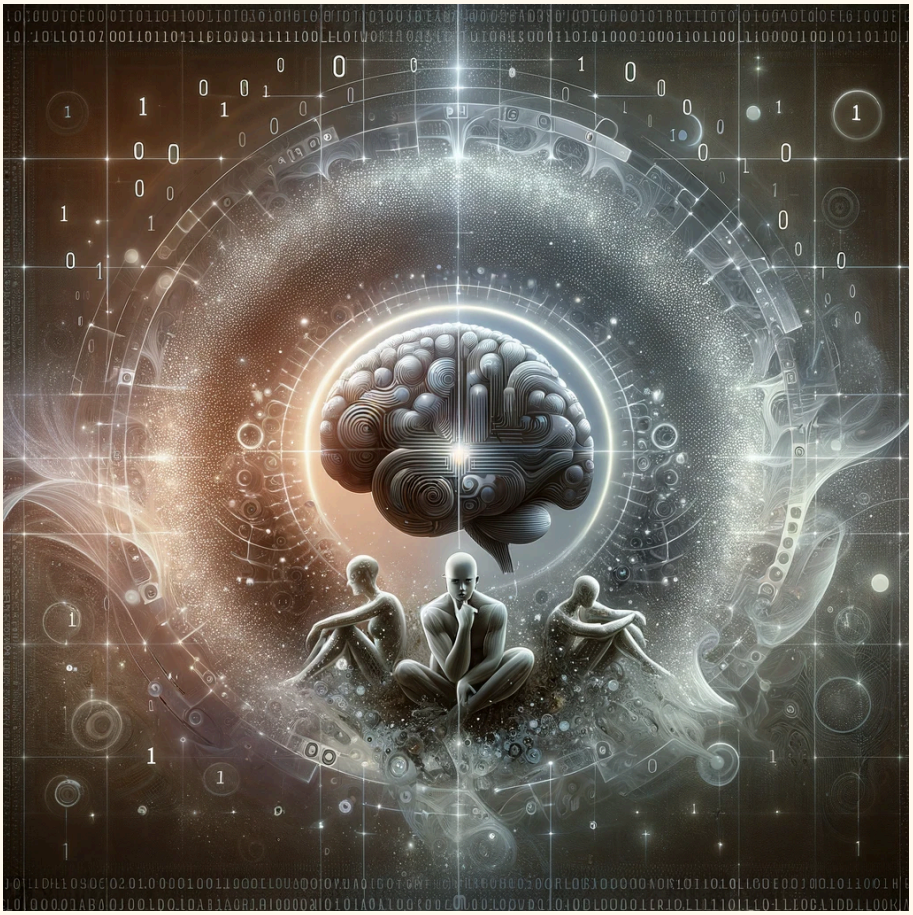

Figure 1. Philosophy, technology and AI.. Adapted from an AI-generated image created for project documentation purposes, March 2024.

By:

Ahmet Alperen Güngör 3362906

Abstract

In today's technologically driven world, Artificial Intelligence (AI) has become an integral part of our daily lives, revolutionizing various sectors, including art with the emergence of Generative Artificial Intelligence (GAI). After a brief introduction about AI, our article delves into the ethical issues surrounding GAI, and its profound impacts on modern society. Through a critical examination of the ethical concerns posed by them, particularly in terms of topics like biases, ownerships and data privacy, this article highlights the importance of comprehensive analyses and consistent solutions. Additionally, our article discusses more specific topics, like the ethical implications of data privacy violations, ownership rights, and the potential displacement of human labor in the art industry. Ultimately, addressing these ethical problems has utmost importance for safeguarding societal values and promoting equitable advancements in artificial intelligence technology.

Introduction

Nowadays, humans cannot live without technology and have become extremely reliant on it, making it necessary for survival and being impacted by it. For instance, AI has been present in this scenario. As Galanther P. (2019) said "When first encountering digital generative art, a novice will frequently ask ‘Who is the artist, the human or the computer?’ Many will reflexively answer that the programmer, i.e. the human, is the artist", we see such impacts and influences by technology placing the human and its principles in a position where ethics become vulnerable. It is important to address in-depth how technology is impacting society, the ethical issues it brings, and consistent solutions to these issues. This philosophical essay will question and examine the ethical concerns regarding the generative artificial intelligence as well as its impact on modern day society. We argue that while its learning processes are revolutionary, they may unintentionally cause prejudices and other misrepresentations, therefore a careful analysis of its implementation and management is required.

Using various resources, this essay investigates how GAI raises questions about ethics in various fields, and what are the main causes of such ethical issues. We suggest that, while generative AI's capabilities are transformational, they must be used responsibly in order to avoid perpetuating current societal biases and provide equitable advantages. At the end, by examining how these technologies may be matched with ethical demands, this essay seeks to briefly

examine different approaches and viewpoints such as xAI for transparency; deontology, virtue ethics and utilitarianism regarding discussions on ethical implications.

Overview on ethical implications of GAI

A technology like GAI, which sources its information from big datasets from the internet, raises concerns about learning biases like racism and discrimination to create its content. Furthermore, GAI's potential to develop original works throws conventional ideas of authorship and copyright into question in a variety of sectors, including the arts, necessitating an examination of legal frameworks and ethical norms (Feuerriegel et al., 2023). Basically, what we are trying to see from a different perspective is also the clash or two ways of thinking and the new era of technology. We can analyze such quote “It is an attempt to reconcile the decades long ‘war’ between the modernist culture of science and the postmodern culture of the humanities” (Galanter, 2019), were it is all about trying to make peace between science and humanities, which have been arguing for years, especially when we talk about making content with AI. It's like saying, with GAI, we are trying to show that technology can understand and create stuff that feels human, like art, and maybe get these two sides to see things more eye to eye.

Bias

The first ethical issue - and probably the most fundamental one regarding artificial intelligence in general - is the bias that can be ingrained within its algorithms. Before moving onto the ethical aspect of it, the concept of bias should be made clear. In essence, it is the algorithms’ tendency to display patterns that it learned from the dataset. One may argue that is the whole purpose of such algorithms, but the ethical dilemma occurs when the trained dataset is not selected carefully and in such cases, the algorithm can learn hateful or harmful patterns as stated by Roselli et al. (2019). Despite AI's potential to enhance decision-making processes, it often inherits biases present in the data it's trained on. Whether it's racial, gender, economic or any kind of cultural biases, these prejudices can perpetuate and even actively apply societal inequalities when incorporated into these algorithms. Typically, the output of a GAI model is a piece of media, such as an image or sound, and the biases inherent in these models often become apparent in these outputs. These biases can influence the representation of gender, race, or other attributes, leading to skewed or harmful depictions. Addressing these biases is crucial for developing fair and accurate AI systems.

Culture

Let's say a company develops a carefully constructed GAI model, trained with a dataset without any kind of biases and outputs perfectly fine content - at least for that country. Now, assume when a user from a different country tries the model and finds the generated content very offensive - or even unlawful. Why did that happen even though the model was claimed to be unbiased, at least for that country? This is because in addition to hateful/harmful content, there is the aspect of culture that has to be taken into account. As stated by Wong(2020), cultural differences highlight the challenges such differences pose on this topic, a model that works perfectly fine for a country/community may be very offensive for another one. This cultural sensitivity aspect underscores the complexity of ethical considerations. Even with the absence of explicit biases, algorithms must navigate the nuances of diverse cultural norms and values. Therefore, companies that develop such systems must recognize the importance of cultural context and adjust their models accordingly, employing localization strategies and engaging with diverse communities to ensure their technologies align with local responsibilities as well as legal frameworks. Failure to account for these cultural differences can lead to unintended harm and undermine the trust and adoption of AI solutions on a global scale. Thus, alongside addressing biases, adopting cultural understanding and inclusivity must be ensured for the ethical development and deployment of generative artificial intelligence.

Privacy

As stated during the AI ethics lecture, privacy is a fundamental human right recognized by the UN; thus, another important discussion topic regarding artificial intelligence ethics is data privacy. Data is the source that powers algorithms for AI, enabling them to learn patterns and make predictions based on that information. However, since that data is a product of human activities, it is important to address privacy concerns and ethical implications. It would not be wrong to claim that data is the currency of our modern age; therefore, it is a very valuable asset for companies. Creation of generative ai models also depend on this principle. Various media data is collected from users all around the internet, and then used to train the models, which then output high quality of artistic media in various ways, as stated by Epstein et al. (2023). However, sometimes it is not very transparent on how such models are trained and which datasets they used. What kind of "process" or "source" exactly caused a model to create a specific piece of media, how did it get affected? This lack of transparency raises concerns on

fundamental issues as we stated earlier, such as bias and privacy. Another aspect regarding lack of transparency is actually on a more fundamental side of GAI algorithms, as Bélisle-Pipon et al. (2023) states that AI acts as a black-box, where working principles of such models are not easily understandable and engagement with the public is not very clear. Therefore, an important step for the age of artificial intelligence is to ensure a high level of explainability is provided for end users without any technical knowledge requirements.

Ownership & Property Rights

Training GAI models using media online brings the question of ownership and property rights. An example of such an industry would be art, where GAI is used to create artworks. For such models, training data is artworks of professional artists, and their personal data may be utilized without their consent, resulting in the creation of AI-generated artworks without transparent attribution or copyright protection. This raises concerns about the ethical implications of using individuals' data to fuel GAI creativity, particularly when it comes to ownership rights over the resulting creations. According to a recent news article (Belci, 2024), more than 16.000 artists' data was used without consent to train Midjourney AI. Another recent lawsuit against OpenAI was filed by The New York Times (Guardian, 2024), where NYT claimed that OpenAI uses data without consent to train GAI models. Considering cases like these and more, and the autonomous nature of GAI that requires thousands of original data to train, it is clear that necessary precautions and regulations should be taken, without clear guidelines and regulations governing the ownership and attribution of AI-generated art, there is a risk of exploitation and unfair distribution of economic benefits. In addition to copyright issues, ownership itself has ethical implications as well. In case of an unlawful output, who is the responsible party? Is it the AI algorithm itself, if so, how do you punish a mathematical algorithm? Is it the programmer who only produced some mathematical outputs? Is it the company who funded the model? Or is it the humans as data owners, since the algorithm learns from owners after all? Since there is no clear indication of a direct responsibility, at least company owners as the holders of the power and influence should take as much precautions as possible, while lawmakers should enforce strict regulations on such topics.

Employment

On top of ownership, GAI creates another question in the art industry which is employment. While GAI has the potential to improve and streamline certain aspects of the artistic process,

such as generating preliminary sketches or assisting with repetitive tasks, its long-term implications for employment, as well as human creativity and artistic expression raises concerns. As algorithms become increasingly sophisticated at mimicking artistic styles and generating visually stunning compositions, there exists the possibility that human artists may be marginalized or replaced altogether. Day by day, not only algorithms are getting stronger and stronger to create more accurate works, but their cost & time values are getting optimized as well. Then, is it correct to see this development as something positive for humanity? After all, they are producing artworks for humans in a better (although judging whether something is better or not in the art industry would be a tricky issue by itself) and in a more time efficient way. Even now, Midjourney can process 20 to 40 tasks per second, according to a report by Sharma (2023). With such automation, what will be the future of the art industry as a profession? If an AI model can make the same work (even a better work) within seconds, with little to no cost compared to hiring a real artist, who would commission real artists in the future? Around 2.5 million people work as professional artists only in the US, according to a report by Americans for Arts (2022). Not only in art, but in many other industries, such automation of tasks implies the same issue – will AI replace humans? And judging from the current trend, it seems so.

Different viewpoints on solutions & conclusion

As AI continues to revolutionize various aspects of society, it is imperative to address the ethical complexities it presents. To effectively navigate these challenges, a broad approach is necessary, encompassing technological innovation, regulatory oversight, and sensitivity. Firstly, ensuring the ethical development of GAI requires the careful curation of datasets to mitigate biases and promote inclusivity. Moreover, AI is not transparent to end users, which causes even more mistrust. Adopting approaches that aim to make the “black box” more transparent, such as the eXplainable AI (XAI), as mentioned by Arrieta et al. (2020), can enhance transparency and accountability, enabling users to understand and trust AI decision-making processes. Not only that, but also with more understanding of AI applications and algorithms, their analysis can be done in more detail, which allows a better development.

When at least some level of transparency is provided (and also in order to provide some level of transparency), different viewpoints on the theory of ethics can be applied to the algorithms as well. One fundamental viewpoint is a deontological approach. Given that deontology is “within the domain of moral theories that guide and assess our choices of what we ought to do” (Alexander et al, 2021), it can provide a framework on how to approach solutions for ethical

implications. In a deontological approach, a set of rules and actions can be imposed on algorithms, so that it wouldn't perform unexpected behaviours and would follow ethical guidelines. In such a case, Hooker et al. (2018) states that "An additional advantage of a deontological approach is that it installs reason-responsiveness, and therefore transparency, in the machine. The machine can explain why it took a certain action simply by citing the maximal action plan that generates the particular rule that prompted the act".

As opposed to a strict guideline and rule based approach of deontology, a more "flexible" and "humane" approach of virtue ethics can be considered as a possible solution for the discussion of AI ethics. Farina et al. (2022) states that "Under this framework an agent is ethical, if she displays virtues and therefore acts according to exemplary moral values, so as to be perceived favourably by others". Instead of predefined "ethical rules" that are programmed straight into the algorithm, this approach can cover a wider range of ethical concerns by simulating ethics of a real human being - if applied well. A better, and more detailed, comparison between deontological vs virtue ethics approach is provided by Hagendorff (2022), where he states " virtue instead of deontological ethics to promote and foster not only four basic AI virtues, but also two second-order AI virtues that can help to circumvent "bounded ethicality" and one's vulnerability to unconscious biases. The basic AI virtues comprise justice, honesty, responsibility and care."

A simpler viewpoint on this topic is a utilitarian approach, where based on the main principles, for an AI model the best action will be that which provides the most good or does the least harm - which produces the greatest balance of good over harm (Bonde et al., 2016). If companies themselves and lawmakers adopted such an approach, an example of using millions of users' data that put their privacy in danger would not happen (Alexander, 2021) - or GAI models would not be able to generate harmful content.

Ultimately, addressing the ethical implications of artificial intelligence requires a proactive and interdisciplinary approach that prioritizes transparency, inclusivity, and collaboration between lawmakers, product owners, developers and most importantly by citizens. If all necessary regulations are taken, the transformative potential of the AI can be used very efficiently, while safeguarding societal values and promoting ethical advancements in this technology.

Word Count: 2493

References

Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2023). Generative AI. *Bus Inf Syst Eng*, 66(1), 111–126. https://doi.org/10.1007/s12599-023-00834-7

Galanter, P. (2019). Artificial Intelligence and problems in Generative Art theory. *Electronic Workshops in Computing*. https://doi.org/10.14236/ewic/eva2019.22

Drew Roselli, Jeanna Matthews, and Nisha Talagala. 2019. Managing Bias in AI. In Companion Proceedings of The 2019 World Wide Web Conference (WWW '19). Association for Computing Machinery, New York, NY, USA, 539–544. https://doi.org/10.1145/3308560.3317590

Stark, L. & Kate, C. (2019). The Work of Art in the Age of Artificial Intelligence: What Artists Can Teach Us About the Ethics of Data Practice. Surveillance & Society, 17(3/4), 442-455. https://ojs.library.queensu.ca/index.php/surveillance-and-society/index

Alejandro Barredo Arrietaa , Natalia Díaz-Rodríguez b , Javier Del Ser a,c,d,* , Adrien Bennetot b,e,f , Siham Tabikg , Alberto Barbado h, Salvador Garcia g , Sergio Gil-Lopez a , Daniel Molina g , Richard Benjamins h, Raja Chatilaf , Francisco Herrera g (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI, https://www.sciencedirect.com/science/article/pii/S1566253519308103

Dwork, C. (2008). Differential Privacy: A Survey of Results, LNCS 4978, pp. 1–19, https://web.cs.ucdavis.edu/~franklin/ecs289/2010/dwork_2008.pdf

Bélisle-Pipon, JC., Monteferrante, E., Roy, MC. *et al.* Artificial intelligence ethics has a black box problem. *AI & Soc* **38**, 1507–1522 (2023). https://doi.org/10.1007/s00146-021-01380-0

Sharma, R. (2023). 54+ Unique Midjourney Statistics – Users, Growth, Revenue. https://contentdetector.ai/articles/midjourney-statistics

Belci, T. (2024). Leaked: the names of more than 16,000 non-consenting artists allegedly used to train Midjourney's AI. *The Arts Newspaper*. https://www.theartnewspaper.com/2024/01/04/leaked-names-of-16000-artists-used-to-train-midjourney-ai

Wong, PH. Cultural Differences as Excuses? Human Rights and Cultural Values in Global Ethics and Governance of AI. *Philos. Technol.* **33**, 705–715 (2020). https://doi.org/10.1007/s13347-020-00413-8

Ziv Epstein, Aaron Hertzmann,Art and the science of generative AI.*Science* **380**,1110-1111(2023).DOI:10.1126/science.adh4451

Alexander, Larry and Michael Moore, "Deontological Ethics", *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/win2021/entries/ethics-deontological.

John N. Hooker and Tae Wan N. Kim. 2018. Toward Non-Intuition-Based Machine and Artificial Intelligence Ethics: A Deontological Approach Based on Modal Logic. In Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society (AIES '18). Association for Computing Machinery, New York, NY, USA, 130–136. https://doi.org/10.1145/3278721.3278753

Bonde, S., Briant, C., Firenze, P. *et al.* Making Choices: Ethical Decisions in a Global Context. *Sci Eng Ethics* **22**, 343–366 (2016). https://doi.org/10.1007/s11948-015-9641-

Mitov, Alexander (2021) *Ethical use of artificial intelligence through the Utilitarianism perspective.*

Farina, M., Zhdanov, P., Karimov, A. *et al.* AI and society: a virtue ethics approach. *AI & Soc* (2022). https://doi.org/10.1007/s00146-022-01545-5

Hagendorff, T. A Virtue-Based Framework to Support Putting AI Ethics into Practice. *Philos. Technol.* **35**, 55 (2022). https://doi.org/10.1007/s13347-022-00553-z