Context sensitivity as a reason for the virtue ethics' approach to AI

There are many dangers of the use of AI, such as the violation of data privacy, biases in algorithms that lead to unfair decisions, dehumanised decision-making, etc. On the other hand, AI provides significant epistemic benefits since it can detect regularities in large data sets. These benefits reach science. For example, in mathematics, AI can be used for finding new results and it has even been argued that AI can lead human intuition about the field (Davies et al. 2021). Thus, finding the right measure of trust in AI is an important task. In this context, policymakers introduced the notion of trustworthy AI (European Commission 2019). In order to create trustworthy AI, several principles have been proposed. They include fairness of the results, transparency of the procedures, traceability of the people responsible for the results, data protection, motivation and a design that is benefiting humans.

When it comes to the future of AI ethics, I argue for focusing on three main aspects. The first one is that each application of AI is context sensitive. For instance, the responsible use of machine learning in education requires ethical considerations regarding data collection from minors and the development of algorithms that can be safely implemented in schools. The risk assessment of AI is also domain specific. For example, in Germany, the use of machine translation in court is assessed as riskier than its use for medical purposes, while its use for instructions and technical manuals carries even fewer risks, though it is still not risk-free (DIN and DKE 2020). Secondly, the socio-political and historical context plays an important role in the acceptance and application of AI. For example, concerns about data privacy are particularly strong in the EU, while other countries are reluctant to share their data internationally (Cyranoski 2018). Finally, the context sensitivity of AI applications is best addressed from the perspective of virtue ethics. According to virtue ethics, the correct use of a technology is the one that is in the golden middle between two extremes, which is determined by the situation. Moreover, virtue ethics puts the focus on humans as agents who need to learn how to correctly use and apply new technologies. The trade-offs that are present in the outcomes of different AI applications emphasise the need for its careful assessment and balanced use. The idea of finding the middle ground when deciding about AI is in line with the approach of virtue ethics.

References

Cyranoski, D. 2018. China's crackdown on genetics breaches could deter data sharing. *Nature* 563(7731): 301–302.

Davies, Alex, Petar Veličković, Lars Buesing, Sam Blackwell, Daniel Zheng, Nenad Tomašev, Richard Tanburn et al. 2021. Advancing mathematics by guiding human intuition with AI." *Nature* 600(7887): 70-74.

DIN and DKE. 2020. Standardization Roadmap Artificial Intelligence. https://www.din.de/resource/blob/772610/e96c34dd6b12900ea75b460538805349/normungsroad map-en-data.pdf. Accessed 01.04.2022.

European Commission, Directorate-General for Communications Networks, Content and Technology. 2019. Ethics guidelines for trustworthy AI. *Publications Office*, https://doi.org/10.2759/177365.