

# Integrated ethics: Towards a transdisciplinary toolbox for research and development processes of socially disruptive technologies

AUTHORS: Ammon, Sabine; Fischer, Nele; Hildebrandt, Tim; Mehnert, Wenzel

## KEYWORDS:

A major concern of the advances in Artificial Intelligence and Human-Machine-Interaction is the growing ability of algorithms to make decisions in sensitive areas of our lives. The related decision frameworks are strongly connected to the implicit values embedded in the technology, which leads to a normalization of value systems hidden in the data as well as in the developer's design frameworks. Negative effects of this development are, among others, discrimination, violation of data protection and privacy, lack of transparency of decisions as well as scaling effects. As a remedy, hundreds of ethical AI guidelines have been published. However, the real – and to this day unresolved – challenge lies in transferring the abstract principles into the actual development process of concrete and context-dependent technological applications. However, suitable design methods or best practice examples for such an integration are rare (Reijers et al. 2018). To this day, developers seem to neither approach ethics systematically nor to use specific design tools or methods to implement them (Vakkuri et al. 2019).

As an answer to this challenge, we want to give insight into an emerging ethics toolbox of the Berlin Ethics Lab, tailored for monitoring and intervening research, design, and development processes of socially disruptive technologies. The integrated, inter- and transdisciplinary approach of the Berlin Ethics Lab draws on the role of visions, conceptual frameworks, and the ethical agency of the design tools in order to embed ethical considerations early on in research, design, and development. Visions and their (implicit) ethical beliefs, values, and attitudes guide developers during the design and development process. Despite their critical role for the emerging technology, visions are seldomly critically reflected nor are the developers' biases challenged from outside perspectives during the design and development process.

As we will show in our contribution, such critical reflections pose a fruitful starting point for integrating ethics and an opportunity to counteract socially undesirable side effects before they will ever occur. By drawing on previous work in Technology Assessment, Critical Futures Studies, Responsible Research and Innovation, Value Sensitive Design, and Vision Assessment, our aim is to empower research and development teams to implement ethical reflections more systematically under the guidance of ethics experts and supported by specific design

heuristics and design tools. In our contribution, we will both discuss theoretical foundations and give insight to some best practice examples of an emerging toolbox for integrated ethics.

#### References

Reijers, Wessel, David Wright, Philip Brey, Karsten Weber, Rowena Rodrigues, Declan O'Sullivan, and Bert Gordijn. 2018. "Methods for Practising Ethics in Research and Innovation: A Literature Review, Critical Analysis and Recommendations." *Science and Engineering Ethics* 24 (5): 1437–81. <https://doi.org/10.1007/s11948-017-9961-8>.

Vakkuri, Ville, Kai-Kristian Kemell, Joni Kultanen, Mikko Siponen, and Pekka Abrahamsson. 2019. "Ethically Aligned Design of Autonomous Systems: Industry Viewpoint and an Empirical Study." arXiv. <http://arxiv.org/abs/1906.07946>.