

## **Trustworthy healthcare innovation ecosystems Supporting responsible innovation practices by establishing a trustworthy innovation culture**

**Christian Herzog<sup>1,a</sup>, Sabrina Blank<sup>1</sup>, Bernd Carsten Stahl<sup>2,3</sup>**

<sup>1</sup> University of Lübeck, Germany

<sup>2</sup> University of Nottingham, United Kingdom

<sup>3</sup> De Montfort University, United Kingdom

<sup>a</sup> Corresponding author; E-Mail: [christian.herzog@uni-luebeck.de](mailto:christian.herzog@uni-luebeck.de)

In this contribution, we explore how governance structures and particular features of innovation ecosystems for healthcare technology can support trustworthiness. As a conceptual framework, we draw on the fundamental recognition that technology itself cannot be a proper addressee of trust. Rather, we invoke the notion of trustworthy technology (Rieder et al., 2020; Simon & Wong, 2020) in a derived sense, i.e., as referring to the trustworthiness of moral agents within socio-technical constellations responsible for the technological solutions of interest. This argument is based on accepting the merit of both the so-called rational-choice and motivation-attributing accounts of trust (Nickel et al., 2010). The former entails that trusting can be rational, and hence indistinguishable from reliance, because solution providers deliver expected services reliably. The latter refers to trusting as genuine, as it also involves resigning control by attributing one's motivation, and hence, goals, onto another entity.

We argue that, especially within a healthcare setting, the motivation-attributing account of trust must be taken seriously. We develop this argument not just by focussing on the very immediate and morally highly significant application scenarios. In these, vulnerable patients and care receivers may potentially be facing dire situations, in which trust in caregivers and their utilized technologies is without alternative. Clearly, the gravity of such circumstances warrants preemptive means to ensure that reliance is justified. However, we extend our perspective toward the possibilities of value-sensitive designs that let innovators care for potential individual necessities and may lead to the development of innovations aimed at being malleable enough to accommodate the shared goals of the caregiver/-receiver pair, (cf. Bjerring & Busch, 2021; Herzog, 2022). We contend that—while design goals such as safety, security, and other potential quantifiable criteria are amenable to reliability checks and, consequently, regulation—value-sensitive innovation objectives relating to motivation attribution can probably only be encouraged by endorsing a culture of trustworthy technological innovation and implementing a corresponding facilitating and incentivizing infrastructure within the ecosystem. The relevance of the motivation attributing account of trust therefore underscores the need to look for the proper addressees of trust within socio-technical constellations, rather than the technologies themselves.

Hence, our research is concerned with the specific collaborative and competitive modes, as well as governance structures that support the trustworthiness of—and, consequently, the formation of trust in—innovation ecosystems (cf. Stahl, 2021). We raise several propositions on how to support justifiable trustworthiness of a healthcare innovation ecosystem. We discuss the extent to which a healthcare innovation ecosystem thus equipped would be merely increasing the reliability of its solutions by checks and bounds that decrease the probability of misplaced reliance, or whether it is able to facilitate the formation of genuine and warranted trust relationships. After compiling a list of preliminary requirements for a healthcare innovation ecosystem, we describe the emergence of one particular medical AI ecosystem in Northern Germany and assess its compliance with and future ways of improving its functioning as a responsible AI ecosystem that promotes trustworthiness. The path taken—a platform-based ecosystem governance structure—also incorporates a responsible innovation platform that supports the uptake of ethical reflection through subsidiary, operative and strategic means. The example gives credibility to our claims but also shows that this work is only just beginning. We believe that the conceptual foundations we have provided here are a sound basis for more in-depth studies of technology-driven healthcare ecosystems.

## References

- Bjerring, J. C.; Busch, J. (2021). Artificial intelligence and patient-centered decision-making. *Philosophy & Technology*, 34(2), 349–371. <https://doi.org/10.1007/s13347-019-00391-6>
- CEN CWA 17796:2021. (2021). Responsibility-by-design—Guidelines to develop long-term strategies (roadmaps) to innovate responsibly (Standard CEN CWA 17796:2021). <https://standards.iteh.ai/catalog/standards/cen/8e3cfe68-8449-49f8-b87c-d3efe20da158/cwa-17796-2021>
- Herzog, C. (2022). On the ethical and epistemological utility of explicable AI in medicine. *Philosophy & Technology*, 35. <https://doi.org/10.1007/s13347-022-00546-y>
- Morley, J.; Elhalal, A.; Garcia, F.; Kinsey, L.; Mökander, J.; Floridi, L. (2021). Ethics as a service: a pragmatic operationalisation of AI ethics. *Minds and Machines*, 31(2), 239–256. <https://doi.org/10.1007/s11023-021-09563-w>
- Nickel, P. J.; Franssen, M.; Kroes, P. (2010). Can we make sense of the notion of trustworthy technology? *Knowledge, Technology & Policy*, 23(3–4), 429–444. <https://doi.org/10.1007/s12130-010-9124-6>
- Petersen, E.; Potdevin, Y.; Mohammadi, E.; Zidowitz, S.; Breyer, S.; Nowotka, D.; Henn, S.; Pechmann, L.; Leucker, M.; Rostalski, P.; Herzog, C. (2022). Responsible and regulatory conform machine learning for medicine: a survey of challenges and solutions. *IEEE Access*, 10, 58375–58418. <https://doi.org/10.1109/ACCESS.2022.3178382>
- Rieder, G.; Simon, J.; Wong, P.-H. (2020). Mapping the stony road toward trustworthy AI: expectations, problems, conundrums. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3717451>
- Simon, J., & Wong, P.-H. (2020). Thinking about ‘ethics’ in the ethics of AI. *IDEES*, 12. <https://revistaidees.cat/en/thinking-about-ethics-in-the-ethics-of-ai/>
- Stahl, B. C. (2021). *Artificial Intelligence for a Better Future: An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-030-69978-9>