Certifying Sustainable AI in the framework of the AI.NRW flagship project

Julia Maria Mönig, Sergio Genovesi

Abstract

The AI.NRW flagship project »Certified AI« works on a certification system for artificial intelligence "made in Germany". While highlighting a local and national aspect in its title, the certification should reflect also questions of global responsibility. Sustainability has already been suggested by the platform "Lernende Systeme" as an additional certification criterium (Heesen et al. 2020). However, in our view it should be a rather central criterium – and not only an additional one. In our talk we will explore to what extent sustainability standards should be met by AI-powered technology in order to be certified and illustrate some working certification-criteria.

Regarding the responsibility towards our planet involved by the development and use of technology (Jonas 2014), we will consider **environmental sustainability** issues such as CO_2-emissions caused or avoided by AI software and by the software and hardware production, or the exploitation of raw materials.
Concerning **fairness**, we will address at least two major issues. On the one hand, the unfair work conditions related both to new-colonialist mechanisms in hardware and software production and deployment (a fair AI should indeed embrace decolonial approaches, Mohamed et al. 2020) and to the spread of gig economy. On the other hand, social discrimination and inequality increase caused by AI biases (O'Neil 2016).
Concerning **accessibility** and **reusability** of data, we will examine whether and how F.A.I.R. principles (Wilkinson et al. 2016) could represent a starting point for assessing the sustainability of AI software functioning and life cycle with minimal human intervention on the data it relies on.
Moreover, we will investigate how policy makers and tech companies responded to the problematization of ethics as a design problem highlighted by the scientific community in the last years (see i. e. Epley et al. 2017; Grimm, Mönig 2020; Erbach et al. 2020). We will argue that internal company policies to ease the spread of ethical awareness among employees and encourage "**Ethics by Design**" (WEF 2020) should be considered while evaluating the development procedures of AI-software.
Finally, we will suggest the conformity to the **UN Sustainable Development Goals** as a sound indicator of sustainability for AI-powered technologies.

The AI certification aims at guaranteeing the ethical, sustainable, legally compliant and technically robust production and functioning of AI-powered software. While there are many projects and initiatives with a similar focus going on right now, the AI.NRW project is particularly promising since the consortium partners include, among others, the German Federal Office for Information Security (BSI) and the German Institute for Standardisation (DIN), which is also connected with international certification institutes. Having started in spring 2021, the project will examine use cases from companies that use actual AI-applications and will assess, among other things, whether and how issues of sustainability as understood in this abstract are taken into account by the existing cases.

References:
- Epley N., & Tannenbaum D. (2017). Treating ethics as a design problem. Behavioral Science & Policy, 3(2), 73–84.
- Erbach R. et al. (2020) KoFFI—The New Driving Experience: How to Cooperate with Automated Driving Vehicles. In: Meixner G. (eds) Smart Automotive Mobility. Human–Computer Interaction Series. Springer, Cham

- Grimm P., Mönig J.M. (2020) Ethical Recommendations for Cooperative Driver-Vehicle Interaction—Guidelines for Highly Automated Driving. In: Meixner G. (eds) Smart Automotive Mobility. Human–Computer Interaction Series. Springer, Cham.
- Heesen J. et al. (eds.) (2020), Zertifizierung von KI-Systemen – Kompass für die Entwicklung und Anwendung vertrauenswürdiger KI-Systeme. Whitepaper aus der Plattform Lernende Systeme, München.
- Jonas H. (2014), Technology and Responsibility: Reflections on the New Tasks of Ethics. In: Sandler R.L. (eds.), Ethics and Emerging Technologies. Palgrave Macmillan, London.
- Mohamed S., et al. (2020), Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence. Philosophy and Technology (405).
- O'Neil C. (2016), Weapons of Math Destruction. Crown, New York.
- Wilkinson M., et al. (2016), The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3,** 160018.
- World Economic Forum (WEF) (2020), Ethics by Design: An organizational approach to responsible use of technology.