



Prof. Dr. Lena Kästner, University Bayreuth

Responsible AI -- What does it take?

Abstract. As AI technology becomes increasingly used in the public sphere, including in such vulnerable settings as courts and hospitals, questions about the societal demands of deploying AI are becoming ever more relevant. General calls to make the use of AI "responsible", viz. that the systems in question should be safe, trustworthy, fair, privacy respecting, etc. are echoed by researchers, legal institutions, NGOs and customer protection services alike. But how to best achieve these feats, remains a matter of heated academic, political and legal debates. One concept that has taken center stage over the past few years is explainability – or explainable AI (XAI). Put in a nutshell, the idea is that if we can render opaque AI systems explainable with XAI methods that will, in one way or the other, help us ensure their safety, trustworthiness, fairness, and so forth. This reasoning has led to a veritable XAI-hype. In this talk, I take a critical look at contemporary XAI, highlight some of its limitations, and sketch avenues for research addressing these.

Speaker Bio. Lena Kästner is professor for philosophy, computer science and AI at the University of Bayreuth. She has a background in Cognitive Science and Cognitive Neuroscience and received her PhD in philosophy from Ruhr-University Bochum. Prof. Kästner's research focuses on explanations, intelligence, and causation. Currently, she is also head-PI of the projects "Explainable Intelligent Systems (EIS)" and "For the Greater Good? Deepfakes in Criminal Prosecution (FoGG)". She is also vice president of the German Society for Philosophy of Science (GWP), vice-director of Bayreuth's "Research Center for AI in Science and Society (RAIS2)", and coordinator of the interdisciplinary Master's program "Philosophy & Computer Science" in Bayreuth.

Time & Place.

Wednesday, July 23, 2025

11:15 – 12:00

Reisensburg Castle