

Explainable Intelligent Systems (EIS)

Initiative: Künstliche Intelligenz – Ihre Auswirkungen auf die Gesellschaft von morgen

Ausschreibung: Künstliche Intelligenz – Ihre Auswirkungen auf die Gesellschaft von morgen - Planning Grant

Bewilligung: 06.02.2019

Laufzeit: 9 Monate

Projekt-Website: <https://explainable-intelligent.systems/>

Decisions that directly affect individual lives or society as a whole are increasingly aided by intelligent systems. Domains of application include, but are not limited to, traffic control, the health sector, human resource management, or the police and judiciary systems. Intelligent systems play a vital role in shaping decisions regarding freedom versus imprisonment, healing versus death, or justice versus discrimination. Their objective is to take the workload off decisionmakers and to enable them to make better informed decisions. However, in all these cases we want human agents, rather than intelligent systems, to make the final decision. This is because, first, only humans can take responsibility for decisions and, second, only humans can give appropriate weight to case-sensitive considerations in their decision-making processes. Intelligent systems therefore must be designed to assist us in deciding responsibly and in a wellinformed way. The project Explainable Intelligent Systems (EIS) thus needs a broad and interdisciplinary setup in order to answer its guiding research question: How can and should intelligent systems be designed to provide explainable recommendations? Although this problem is urgent, interdisciplinary research concerning the explainability of artificial intelligence is still in its infancy. While different disciplines have presented fragmentary results, a truly interdisciplinary research programme with a common language is still missing. This project aims to implement such a programme.

Projektbeteiligte

Prof. Dr. Holger Hermanns

Universität des Saarlandes
Fakultät für Mathematik und Informatik
Fachrichtung Informatik
Lehrstuhl für Zuverlässige Systeme
und Software
Saarbrücken

Dr. Eva Schmidt

Universität Zürich
Philosophische Fakultät
Philosophie
Institute of Philosophy
Zürich
Schweiz

Prof. Dr. Ulla Wessels

Universität des Saarlandes
Philosophisches Institut
Professur für Praktische Philosophie
Saarbrücken

Prof. Dr. Georg Borges

Universität des Saarlandes
Rechtswissenschaftliche Fakultät
Professur für Bürgerliches Recht,
Rechtsinformatik, deutsches und internationales
Wirtschaftsrecht sowie Rechtstheorie
Saarbrücken

Prof. Dr. Cornelius König

Universität des Saarlandes
Fachrichtung Psychologie
Professur für Arbeits- und
Organisationspsychologie
Saarbrücken

Open Access-Publikationen

What Do We Want From Explainable Artificial Intelligence (XAI) A Stakeholder Perspective on XAI and a Conceptual Model Guiding Interdisciplinary XAI Research

Where reasons and reasoning come apart

Highly-Automated Job Interviews: Acceptance Under the Influence of Stakes

Explainability as a Non-Functional Requirement