

## **AI Forensics: Accountability through Interpretability in Visual AI Systems**

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

Ausschreibung: Künstliche Intelligenz – Ihre Auswirkungen auf die Gesellschaft von morgen - Full Grant (nur nach Aufforderung)

Bewilligung: 28.10.2021

Laufzeit: 3 Jahre

The project proposes AI forensics as a new sociotechnical framework for the analysis and critique of AI systems. This implies the design and development of new AI forensics tools for the investigation of large AI datasets (that are impossible to analyze manually), models (reverse-engineering the black box effect), and applications (exposing their sociohistorical context), making state-of-the-art results from explainable and interactive machine learning accessible to researchers in the humanities and beyond. The resulting open source AI forensics toolkit is designed to be used by, and shared with civil organisations, academia and industry to study the social impact and implications of real-world AI systems. The AI forensics toolkit will also support four sociotechnical investigations that will examine the production pipeline of face recognition, interpretability and accountability of AI in the humanitarian sector, bias-aware design methods for AI systems, and the role of visual AI models in the sciences. Through a close integration of machine learning research and sociotechnical analysis, the project establishes a new and comprehensive methodological framework to equip the humanities and social sciences with the means to investigate the implications of 21st century artificial intelligence.

### **Projektbeteiligte**

#### **Prof. Dr. Matteo Pasquinelli**

Staatliche Hochschule für Gestaltung  
Karlsruhe  
Forschungsgruppe für Künstliche Intelligenz und  
Medienphilosophie -- Medientheorie  
Karlsruhe

#### **Prof. Dr. Noura Al Moubayed**

Durham University  
Durham  
Grossbritannien

**Prof. Dr. Claude Draude**

Universität Kassel  
FB 16: Elektrotechnik/Informatik  
Wissenschaftliches Zentrum für  
Informationstechnik-Gestaltung ITeG  
Kassel

**Prof. Dr. Fabian Offert**

University of California - Santa Barbara  
Department of Germanic and Slavic Studies  
Santa Barbara  
USA

**Prof. Dr. Leonardo Impett**

University of Cambridge  
The Faculty of English  
Cambridge  
Grossbritannien