

Seminar: Current topics in human-(X)AI collaboration FSS 2023

Prof. Dr. Kevin Bauer



Seminar organization

Application

- Brief motivation letter (max. 1 page) outlining which topic is preferred and why
- Current CV and transcript of records
- Documents to kevin.bauer@uni-mannheim.de

Schedule

- Registration deadline: *23:59 o'clock, February 13th, 2023*
- Notification (Acceptance and Topic Allocation/Rejection): *Midday, February 15th, 2023*
- Drop out possible until: *Midday, February 16th, 2023*
- Kick-Off/Start of the seminar: *February 23th, 2023* (together with Seminar Trends in Enterprise Systems by Prof. Dr. Hartmut Höhle)
- Submission of seminar thesis: *April 20th, 2023*
- Präsentation seminar thesis: *May 4th, 2023* (together with Seminar Trends in Enterprise Systems by Prof. Dr. Hartmut Höhle)

Behavioral factors shaping the human interaction with machine learning systems

Background

- Research in Information Systems, Behavioral Economics, Psychology consistently documents considerable heterogeneity in human decision-making
- In the domain of human-AI interaction, researchers have documented mixed evidence on how, when, and why humans leverage machine learning predictions for decision making

Objective

- Conduct a systematic literature review on human and algorithmic factors influencing the use of AI advice
- Screen the literature in Information Systems, Economics, and Psychology
- Try to reconcile contradictory empirical findings

The impact of explainability on human-AI collaboration

Background

- Explainability of AI systems is increasingly relevant for AI systems and prone to become mandatory by law (see, e.g., AI Act)
- There are different eXplainable AI methods, with SHAP and LIME arguably being the most prominent ones, especially in practice
- Whether human-AI collaboration is more effective with or without explainability remains an open question

Objective

- Conduct a systematic literature review on how explainability influences humans interaction with and perception of AI systems that serves as a decision support
- Screen the literature in Computer science, Information Systems, Economics, and Psychology
- Assess the efficacy of explainability and potential consequences for organizations who may soon be required to implement explainability methods due to upcoming regulations

Human evaluations and perceptions of generative AI output

Background

- Generative AI is becoming ever more prevalent with systems being able to create images from text, compose songs, or draft text paragraphs
- As AI generated content will arguably grow over the next couple of year, it is crucial to understand whether, and if so how, humans engage with it in same or different ways compared to human generated content

Objective

- Conduct a systematic literature review on humans perceive, evaluate, and engage with outputs from generative AI systems
- Screen the literature in Information Systems, Economics, and Psychology
- Work out major differences between engagement with AI and human generated content and try to provide conceptual theories on why this is the case