Chair of Explainable AI-Based Business Information Systems
Prof. Dr. Ulrich Gnewuch

Bachelor Thesis: Mental Models in Human-AI Interaction

Supervisor: Prof. Dr. Ulrich Gnewuch (ulrich.gnewuch@uni-passau.de)
Start date: as soon as possible

Motivation and Goals

Mental models are someone's understanding of how something works. People form such models for any technology they interact with (Norman, 1983), including AI systems (Bauer et al., 2023). However, their models are not necessarily complete or accurate and can evolve as users gain experience with the system. Therefore, understanding how users form, test, and update their mental models is crucial for improving human-AI interaction.

The goal of this bachelor thesis is to explore how users develop mental models of AI systems, compare these models to those developed for non-AI systems, and analyze their impact on human-AI interaction. The investigation should be based on a literature review spanning the fields of information systems, human-computer interaction, and related disciplines. Based on the analysis, the thesis should also describe potential strategies (e.g., XAI methods) to improve the accuracy of users' mental models.

Required Skills

- Strong interest in human-AI interaction
- Good English language skills
- Preferably, experience in conducting literature reviews

Starting Literature (Topic)

Norman, D. A. (1983). Some observations on mental models. In D. Genrner & A. Stevens (Eds.), Mental Models. Erlbaum.

Bauer, K., Von Zahn, M., & Hinz, O. (2023). Expl(AI)ned: The Impact of Explainable Artificial Intelligence on Users' Information Processing. Information Systems Research, 34(4), 1582–1602. https://doi.org/10.1287/isre.2023.1199

Bos, N., Glasgow, K., Gersh, J., Harbison, I., & Lyn Paul, C. (2019). Mental models of AI-based systems: User predictions and explanations of image classification results. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 63(1), 184-188. https://doi.org/10.1177/1071181319631392

Starting Literature (Method)

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26(2), xiii–xxiii. https://www.jstor.org/stable/4132319

Templier, M., & Paré, G. (2015). A framework for guiding and evaluating literature reviews. *Communications of the Association for Information Systems*, 37(1), 6. https://doi.org/10.17705/1CAIS.03706