



INSTITUT FÜR
INFORMATIK
SOFTWARE ENGINEERING



UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

INVITATION

This is a role model talk for women, but everybody is invited to join. We particularly encourage students to attend.

How do you test an AI application?

Dr. Anja Kleebaum

Agile Software Engineer at andrena objects ag

Software with AI can assist people with tasks that previously required a great deal of effort (e.g., extracting information from images or answering questions about all kinds of internal company matters). But how can you automatically test that the AI application really does what it is supposed to do? In this presentation, I will report on my experiences with the use of AI in industry and outline approaches for evaluating AI applications. The following questions will be addressed:

- How do you build an evaluation dataset (self-created vs. synthetic data)?
- How do you assess whether a returned answer matches the expected answer (regular expressions vs. LLM-as-a-judge)?
- What metrics can be collected (e.g., for retrieval and generation)?
- What platforms are available (e.g., Langfuse, LangSmith, or a custom solution)?
- How can user feedback be collected and used to improve the AI application?

Anja received her PhD from the Chair of Software Engineering at Heidelberg University and worked as a research assistant in software engineering education. In her dissertation, she designed methods and tools for lightweight decision management (rational management) during agile software development. Since 2023, she has been working as an agile software engineer at andrena objects and on various AI projects.

Tuesday, 27 MAY 2025, 7.15 pm-8.15pm

Seminarraum A, Mathematikon, EG

Im Neuenheimer Feld 205, 69120 Heidelberg

After the talk we will meet for further discussion with some catering in the neighboring room.

For better planning, please RSVP to kerstin.lambert@eppleimmobilien.de