

Bachelor- oder Masterarbeit (de) / Bachelor's or Master's Thesis (en)

Systematische Literaturarbeit zur menschlichen Bewertung der Zuverlässigkeit von KI-Entscheidungen

It is also possible to conduct this thesis in English - see below for the English version

Künstliche Intelligenz spielt eine immer größere Rolle in technischen Systemen und Entscheidungsprozessen – sei es in der Industrie, in der Produktentwicklung oder im Alltag. Doch wie entscheiden Menschen, ob sie sich auf die Empfehlungen einer KI verlassen können oder sie hinterfragen?

Menschen versuchen, das Verhalten eines technischen Systems zu verstehen, um vorherzusagen, wann es zuverlässig arbeitet und wann nicht. Es kann aber auch passieren, dass Menschen entweder blind vertrauen (Over-Reliance) oder der KI gar nicht vertrauen (Under-Reliance), selbst wenn sie eigentlich nützliche Unterstützung bietet.

Darüber hinaus zeigt die Forschung, dass Menschen ihre Entscheidungsstrategie anpassen, wenn sie merken, dass die KI-Entscheidungen nicht immer korrekt sind. Sie müssen nicht nur entscheiden, ob sie der KI vertrauen, sondern auch, wie sie die Informationen der KI nutzen.

Ziel dieser Bachelor- oder Masterarbeit ist es, durch eine **systematische Literaturrecherche** herauszufinden, **welche Faktoren beeinflussen, ob Menschen KI-Entscheidungen als zuverlässig einstufen und wie sie ihre eigene Strategie zur Nutzung von KI anpassen.**

Dabei können folgende Fragen untersucht werden:

- Nach welchen Kriterien bewerten Menschen die Zuverlässigkeit einer KI?
- Welche Faktoren beeinflussen ihre Strategie (z. B. Transparenz der KI, frühere Erfahrungen, Fehlerquoten der KI)?
- Wie reagieren Menschen und passen ihre Strategie an, wenn eine KI-Entscheidung fehlerhaft oder unzuverlässig ist?

Aufgaben: Die Arbeit folgt dem Ansatz einer systematischen Literaturrecherche basierend auf etablierten Methoden (z. B. PRISMA)

- Thematische und methodische Einarbeitung durch Literaturrecherche (v.a. Journal Paper)
- Entwicklung der Forschungsfrage(n) und Entwicklung der Suchstrategie (Auswahl der Suchbegriffe, Datenbanken, Ein- und Ausschlusskriterien)
- Durchführung der systematischen Literaturrecherche
- Auswertung der Literatur und Aufbereitung der Ergebnisse mittels geeigneter Kategorisierung
- Beantwortung der Forschungsfrage und Ableitung relevanter wissenschaftlicher und praktischer Erkenntnisse

Das Thema ist ab sofort zu vergeben. Die Schwerpunktsetzung wird in enger Absprache mit der Betreuerin festgelegt.

Haben Sie Interesse an dieser Arbeit?

Dann nehmen Sie gerne Kontakt auf:

M. Sc. Lina Kluy
Lina.Kluy@kit.edu

Die Ausschreibung richtet sich an Studierende der Studiengänge Maschinenbau, Ingenieurspädagogik, Wirtschaftsingenieurwesen oder einem ähnlichen Studiengang.

Bachelor's or Master's Thesis

Systematic Literature Review on Human Assessment of AI Decision Reliability

Artificial intelligence is playing an increasingly important role in technical systems and decision-making processes - whether in industry, product development, or everyday life. But how do people decide whether they can rely on AI recommendations or should question them?

People try to understand the behavior of a technical system to predict when it will operate reliably and when it will not. However, they may either blindly trust AI (over-reliance) or distrust it completely (under-reliance), even when it provides useful support.

Research also shows that people adjust their decision-making strategy when they realize that AI decisions are not always correct. They must not only decide whether to trust AI but also how to use the information it provides.

The aim of this bachelor's or master's thesis is to conduct a systematic literature review to determine which factors influence whether people perceive AI decisions as reliable and how they adapt their strategies for using AI.

Potential Research Questions

- What criteria do people use to assess the reliability of AI?
- Which factors influence their strategy (e.g., AI transparency, previous experiences, AI error rates)?
- How do people react and adjust their strategy when an AI decision is erroneous or unreliable?

Your tasks: This thesis follows a systematic literature review approach based on established methods (e.g., PRISMA).

- Thematic and methodological orientation and familiarization through literature research (mainly journal papers)
- Development of research question(s) and search strategy (selection of search terms, databases, inclusion and exclusion criteria)
- Conducting the systematic literature review
- Evaluation of the literature and processing of the results using suitable categorization
- Answering the research question and deriving relevant scientific and practical insights

The topic is available immediately. The focus will be determined in close coordination with the supervisor.

This thesis is intended for students in the fields of Mechanical Engineering or similar study programs.

Are you interested in this work?

Please get in touch:
M. Sc. Lina Kluy
Lina.Kluy@kit.edu