

Thema 2: „Evaluierung von KI“ (AI assessment): Wie und wonach sollten KI-Systeme evaluiert werden?

Bsp.

- Vergleich von automatisierten vs. menschlichen Evaluierungsmethoden.
- Wie unterscheiden sich Bewertungen durch Personen aus verschiedenen Milieus, Kulturen oder Bildungshintergründen?
- Entwicklung eines Evaluierungsframeworks für bestimmte Aufgaben (z. B. Textqualität, Fairness, Kreativität).

Studierende sollten dabei ein konkretes System, einen Anwendungsbereich oder einen Wert wählen. Der prinzipielle Technologieschwerpunkt liegt auf KI-Dialogsystemen (Conversational Agents), wie z.B. LLMs, aber andere KI-Anwendungen sind auch möglich. Die Arbeit kann auf Englisch (präferiert) oder Deutsch verfasst werden.