

Szenario 5:

„Faktenchecker“ – Halluzinationen auf der Spur

Zielgruppe: Schüler*innen und „Stadtnutzer*innen“

Ausgangslage: ChatGPT und ähnliche Chatbots bieten quasi „im Gespräch“ Antworten auf Fragen und formulieren sprachlich einwandfreie Texte. Warum sollte man kritisch sein bei so einem professionellen und eindrucksvollen Tool?

Die Frage nach „Welche Persönlichkeiten sind in Oldenburg (oder woanders) geboren?“ führt bei ChatGPT und Gemini immer wieder zu falschen Antworten. Den „Persönlichkeiten des öffentlichen Interesses“ wird interessanterweise immer wieder eine „Jugend“ in Oldenburg angedichtet.

Die Liste mit Namen soll von den Teilnehmenden überprüft werden.

Zunächst sollen vertrauenswürdige Webseiten (und ggf. andere Quellen) für eine solche Aufgabe benannt werden. Das können Wikipedia u. ä. sein, aber auch die Angebote von Rundfunkanstalten etc. (z. B. bei aktuell aus dem Fernsehen bekannten Menschen). Es wird herauskommen, dass die Antworten der KI falsch sind.

Im nächsten Schritt soll herausgefunden werden, warum. Gibt es Quellen, die kritisch die Ergebnisse von GAI Tools hinterfragen? Was können Ursachen für Falschaussagen sein?

Es handelt sich bei den falschen Antworten um klassische „Halluzinationen“ ([https://de.wikipedia.org/wiki/Halluzination_\(K%C3%BCnstliche_Intelligenz\)](https://de.wikipedia.org/wiki/Halluzination_(K%C3%BCnstliche_Intelligenz))).

Geht man diesem Phänomen weiter auf den Grund, stößt man darauf, dass bisher tatsächlich niemand recht erklären kann, wie diese Halluzinationen entstehen, oder warum sie gar behaupten, was da steht.

Dies soll zum Anlass genommen werden, generative Künstliche Intelligenz auf ihren „Diskurscharakter“ zu hinterfragen:

- Wer spricht da überhaupt?
- Was braucht es, damit man jemandem glauben sollte?
- Was unterscheidet Äußerungen der KI von Artikeln bei z. B. Wikipedia?
- Gibt es Möglichkeiten der Bedienung, die zu besseren Ergebnissen führt?
- Was kann man (ggf. „auf die Schnelle“) über das Funktionieren von KI herausfinden? Wem glaubt man? ;-)

Die Fragen

1. Zielgruppe
 1. Wer ist die Zielgruppe?
 2. Was bringen die TN mit?
2. Kompetenzentwicklung
 1. Was können die TN lernen?
 2. Welche Inhalte des Referenzrahmens IKs und/oder des Framework IKs für die Hochschulbildung werden angesprochen?
3. Praktischer Nutzen
 1. Was ist dieser für TN?
 2. Was ist dieser für Schulungen im Hinblick auf Inhalte, Interaktion, Übungen, Formate?
4. Kompetenz der Schulenden
 1. Ist unsere Position klar? Wie weit geht unsere Perspektive?
 2. Wie bewerten wir die Hilfestellung von KI? Was können wir bewirken?