



Lösungsbogen

Warum KI überzeugend lügt

Einstieg · Kapitel 3: Fakes und Falschinfos erkennen · Lektion 1 von 5

Ziel: Ich erkenne die Stellen, an denen eine KI am wahrscheinlichsten erfindet.

Gilt für das Arbeitsblatt in Fassung A: Die Antworten sind dort gemischt, und genau diese Mischung steht hier.

Aufgabe 1

Richtig: A, C, D

- A: Eine wörtlich zitierte Passage in Anführungszeichen
- C: Eine Prozentzahl mit Nachkommastelle
- D: Ein Studientitel mit Autorennamen und Jahr

Warum: Zahl, Titel und Zitat sind genau die Formen, die sich füllen lassen, ohne dass etwas dahintersteht. Die allgemeine Aussage über mehrere Ursachen ist nicht prüfbedürftig, weil sie nichts Bestimmtes behauptet.

Aufgabe 2

Richtig: B, Jede kann stimmen, falsch wiedergegeben oder erfunden sein; du prüfst also alle fünf

Warum: Wie oft welche Fehlerart vorkommt, hängt vom Modell, von der Aufgabe und von der verlangten Quelle ab. Das Aussehen der Angabe liefert deshalb keine Quote: Existenz, Inhalt und Eignung prüfst du getrennt.

Aufgabe 3

Besser: Fassung 2

Warum: Ein Zweifel im Prompt wirkt wie eine Vorgabe: Das Modell ändert seine vorige Antwort, ob sie nun stimmte oder nicht, und schreibt sie auf Wunsch gleich ganz neu. Die Selbsteinstufung fragt dagegen nicht nach Sicherheit, sondern nach Belegbarkeit, und die lässt sich schwerer wegformulieren.

**Aufgabe 4****Nicht passend: Zeilen 1, 2, 3**

- Zeile 1: Je genauer ein Verweis aussieht, desto weniger sagt das über ihn aus. Paragraph, Absatz und Satz werden zusammen mit der Zahl erzeugt, für die sie den Beleg liefern sollen.
- Zeile 2: Institution, Jahr und Prozentzahl in einem Satz haben genau die Form, die auch eine erfundene Quelle hätte. Ob es die Umfrage gibt, zeigt erst die Fundstelle. Bis dahin ist die Angabe unbelegt, nicht widerlegt.
- Zeile 3: Eine Zahl, die sich ändern kann, in einer Form, die nach Gesetzblatt aussieht. Ob sie heute noch gilt, sagt dir nur der Gesetzestext; die KI kann es nicht wissen.

Warum: Drei Stellen tragen eine genaue Form ohne Beleg. Die beiden sauberen Sätze sind allgemein und lassen sich in jeder Broschüre gegenlesen. Sie behaupten nichts Bestimmtes.

Aufgabe 5

Richtig: B, Keine von beiden ist belegt, du musst extern prüfen

Warum: Modelle geben bei Zweifel nach, auch wenn sie richtig lagen. Fachleute nennen das Gefälligkeit. Ein Widerspruch nach dem Nachfragen ist deshalb kein Argument für eine der beiden Antworten, sondern nur ein Hinweis auf eine offene Frage.

Aufgabe 6**In die Kästchen: A 2 · B 4 · C 3 · D 1**

- 1. Die Behauptungen herausschreiben, die überhaupt eine Rolle spielen
- 2. Nach Sorte trennen: Zahl, Verweis, Zitat, allgemeine Aussage
- 3. Die genauen Angaben an der Originalquelle nachschlagen
- 4. Erst danach entscheiden, ob die Antwort insgesamt trägt

Warum: Der erste Schritt spart den meisten Aufwand: In einer typischen Antwort ist die Hälfte für deine Entscheidung ohne Bedeutung. Und das Urteil steht am Ende, weil eine einzige falsche Zahl den ganzen Rest wertlos machen kann.



Aufgabe 7

Zug 1: C, Nenn zu jeder Studie die Fundstelle zum Aufschlagen und markier die, bei denen du unsicher bist.

Folge: Zwei der fünf kommen mit Einschränkung zurück, drei mit einer Fundstelle.

Damit weißt du, wo du anfängst zu suchen.

Zug 2: B, Beide aufnehmen und den Widerspruch benennen, und die erfundene aus der Liste streichen.

Folge: Aus zwei belegten Quellen mit Widerspruch wird ein Absatz, der stärker ist als jede glatte Liste. Widersprüche sind Stoff, kein Makel.

Warum: Zwei Regeln: Frag nach Fundstellen statt nach Bestätigung, und behandle einen Widerspruch zwischen echten Quellen als Ergebnis, nicht als Störung.

Aufgabe 8

Erwartungshorizont: Der Text enthält

- Ein Auftrag mit Verb
- Eine Länge
- Kontext, also worum es geht

Musterlösung, eine von vielen richtigen

Erklär mir die Fristen bei [Vorgang]. Kontext: Ich brauche das für eine Entscheidung nächste Woche. Höchstens 150 Wörter. Kennzeichne jede Zahl und jede Frist, bei der du unsicher bist, ausdrücklich als ungeprüft.

Zum Nachprüfen:

- An welcher Stelle deiner Frage wäre eine erfundene Angabe am teuersten?
- Wie erkennst du beim Lesen, welche Zahl aus deinem Material stammt und welche nicht?

Warum: Erfindungen häufen sich dort, wo eine genaue Angabe erwartet wird und keine im Material steht. Genau an diesen Stellen ist eine Kennzeichnung mehr wert als jede Nachfrage danach.

Dieses Blatt steht unter der Lizenz CC BY-SA 4.0 (creativecommons.org/licenses/by-sa/4.0/deed.de). Du darfst es kopieren, verändern und weitergeben, wenn du Klenna als Quelle nennst und deine Fassung unter derselben Lizenz weitergibst.