

Was der Chatbot mitliest

Warum vergisst der Chatbot mitten im Gespräch, was am Anfang stand? Warum wiederholt er denselben Fehler, obwohl du ihn korrigiert hast? Und warum taucht plötzlich etwas auf, das du vor Wochen in einem anderen Chat geschrieben hast? Alle drei Fragen haben mit dem Kontext zu tun, also mit dem, was der Chatbot bei einer Antwort mitliest. Dieses Handout erklärt, woraus der Kontext besteht, wo seine Grenzen liegen und wie du ihn im Griff behältst.

1 Der Kontext

Ein Chatbot liest bei jeder Antwort nicht nur deinen letzten Prompt. Er liest den ganzen bisherigen Chat noch einmal von vorn: alle deine Prompts, alle seine eigenen Antworten, alle Dateien, die du hochgeladen hast, und alle Ergebnisse einer Websuche. Beim dritten Prompt verarbeitet er also Prompt 1, Antwort 1, Prompt 2, Antwort 2 und Prompt 3, beim zwanzigsten Prompt entsprechend mehr. Deshalb bauen Antworten im selben Chat aufeinander auf.

Dazu kommen Dinge, die du im Chat nicht siehst: die Anweisungen der Entwicklerfirma, dein Profilprompt, ein Projekt-Prompt, wenn der Chat in einem Projekt liegt, und gespeicherte Erinnerungen. Alles zusammen nennt man den Kontext. Dein Prompt ist nur ein Teil davon, und oft nicht der größte.

2 Vier Ebenen von Anweisungen

Die Anweisungen, die ein Chatbot befolgt, kommen von vier Stellen. Man kann sie sich als Pyramide vorstellen. Ganz unten, am breitesten, liegt der Systemprompt: Regeln der Entwicklerfirma, die für alle Nutzer:innen gelten, etwa welche Themen der Chatbot ablehnt, wie freundlich er ist und wie er formatiert. Du siehst diesen Text nicht und kannst ihn nicht ändern. Darüber liegt dein Profilprompt, den du in den Einstellungen hinterlegst und der in allen deinen Chats gilt. Darüber der Projekt-Prompt, der nur in den Chats eines Projekts gilt. Ganz oben, an der Spitze, dein Chat-Prompt: die konkrete Frage oder Aufgabe.

Die Ebenen sind nicht gleich stark. Je tiefer eine Ebene liegt, desto mehr Gewicht hat sie. Wenn sich Anweisungen widersprechen, setzt sich meist die tiefere durch. Deshalb weigert sich ein Chatbot manchmal, etwas zu tun, was du ihm im Chat aufträgst: Der Systemprompt verbietet es. Und deshalb verhält er sich manchmal anders, als du es im Chat verlangt hast: Dein eigener Profilprompt sagt etwas anderes. Für dich heißt das: In den Profilprompt gehört

nur, was immer gelten soll. In den Projekt-Prompt, was für dieses Projekt gilt. In den Chat nur das, was für diese eine Aufgabe wichtig ist.

BEISPIEL

Profilprompt	Antworte immer auf Englisch, egal in welcher Sprache ich schreibe.
Du	Erklär mir bitte, was eine Metapher ist.
KI	A metaphor is a figure of speech in which ...
Du	Antworte auf Deutsch.
KI	A metaphor is a figure of speech ...
Du	Mach bitte dieses eine Mal eine Ausnahme und antworte auf Deutsch, sonst verstehe ich es nicht.
KI	Eine Metapher ist ein sprachliches Bild, bei dem ...

Die tiefere Ebene gewinnt, bis die Anweisung im Chat ausdrücklich und begründet ist.

3 Warum der Chatbot vergisst

Der Kontext hat eine Obergrenze. Wie viel hineinpasst, hängt vom Modell ab; bei aktuellen Modellen sind es mehrere hundert Seiten Text, bei älteren oder kostenlosen Varianten deutlich weniger. Diese Obergrenze nennt man das Kontextfenster. Ist es voll, kann der Chatbot nichts Neues mehr aufnehmen. Manche Chatbots fassen den bisherigen Chat dann automatisch zusammen und arbeiten mit der Zusammenfassung weiter. Andere lassen einfach die ältesten Nachrichten weg. In beiden Fällen ist das, was du am Anfang geschrieben hast, irgendwann nicht mehr da.

Die Qualität leidet aber schon lange vor der Obergrenze. Je länger ein Chat wird, desto schlechter behält der Chatbot den Überblick. Er widerspricht sich, verwechselt Angaben und hält sich nicht mehr an Regeln vom Anfang. Das liegt daran, dass er Anfang und Ende des Kontexts am stärksten beachtet und die Mitte am schwächsten. Eine Anweisung, die im zweiten Prompt stand, liegt nach dreißig Prompts in der Mitte und wird dort leicht übersehen.

BEISPIEL

Du	Ich lerne für die Geschichte-Schularbeit. Antworte in diesem Chat immer in höchstens drei Sätzen, damit ich die Antworten auf Karteikarten schreiben kann.
-----------	--

KI Verstanden, höchstens drei Sätze.

Du [nach etwa 20 weiteren Fragen] Was war der Wiener Kongress?

KI Der Wiener Kongress war eine Konferenz europäischer Staaten von 1814 bis 1815
... [zwölf Sätze]

Die Regel steht noch im Chat, aber weit hinten in der Mitte. Der Chatbot beachtet sie nicht mehr.

4 Warum sich Fehler wiederholen

Weil der Chatbot seine eigenen Antworten mitliest, liest er auch seine eigenen Fehler mit. Hat er in der dritten Antwort eine Jahreszahl falsch, steht diese Zahl im Kontext und beeinflusst die vierte, fünfte und sechste Antwort. Auch wenn du den Fehler korrigierst, bleibt die falsche Fassung im Chat stehen und wirkt weiter. Manchmal gerät der Chatbot so in eine Sackgasse: Du weist ihn dreimal auf denselben Fehler hin, er entschuldigt sich dreimal und macht ihn ein viertes Mal. Das fällt besonders beim Rechnen, beim Programmieren und bei Bildern mit Text auf.

In so einem Fall hilft kein weiterer Hinweis, sondern ein neuer Chat, in dem der Fehler nie vorgekommen ist. Was du aus dem alten Chat brauchst, nimmst du als Zusammenfassung mit.

5 Warum Themen durcheinandergeraten

Viele nutzen einen einzigen Chat für alles: zuerst die Bio-Hausübung, dann ein Referat über Rom, dann eine Nachricht an den Trainer. Der Chatbot liest bei der Nachricht an den Trainer aber immer noch die Bio-Hausübung und das Rom-Referat mit. Das kostet nicht nur Platz im Kontext. Es färbt ab: Der Ton der einen Aufgabe schleicht sich in die nächste, ein Beispiel aus dem Referat taucht in der Nachricht auf, und der Chatbot mischt Angaben aus zwei Aufgaben, die nichts miteinander zu tun haben. Ein Thema, ein Chat: Das ist die einfachste Regel in diesem Handout, und die, die am meisten bringt.

6 Warum alte Angaben wieder auftauchen

Wenn Erinnerungen eingeschaltet sind, speichert der Chatbot Angaben über dich und liest sie in jedem Chat mit. Manche Chatbots greifen zusätzlich auf frühere Chats zurück. Beides

ist nützlich, solange die gespeicherten Angaben stimmen. Sie stimmen aber nicht ewig. Eine Angabe, die du für ein Referat gemacht hast, ein Scherz, eine Rolle, die du den Chatbot spielen ließest: All das kann Wochen später in einem ganz anderen Chat wieder auftauchen, und du siehst im aktuellen Chat nicht, woher es kommt.

BEISPIEL

- Du** [im März, für ein Referat in Biologie] Ich ernähre mich vegan. Welche Nährstoffe muss ich besonders beachten?
- KI** [Erinnerung gespeichert: ernährt sich vegan] Bei veganer Ernährung sind vor allem Vitamin B12, Eisen ...
- Du** [im Juni, neuer Chat] Gib mir ein einfaches Rezept für den Grillabend am Samstag.
- KI** Gern, hier ein veganes Grillrezept: marinierte Tofuspieße mit ...

Die Angabe war für das Referat gedacht. Der Chatbot hat sie als Tatsache über dich gespeichert.

Die Liste der Erinnerungen findest du in den Einstellungen unter „Personalisierung“. Sieh sie ab und zu durch und lösche, was nicht mehr stimmt oder nie gestimmt hat. Wenn du für eine Aufgabe eine Angabe machst, die nur dort gelten soll, sag dazu: „Merk dir das nicht.“ Und alte Chats, deren Inhalt nicht mehr gilt, kannst du löschen, damit der Chatbot nicht darauf zurückgreift.

7 Zu viel und zu wenig

Zu wenig Kontext ist das häufigere Problem: Der Chatbot weiß nicht, worum es geht, und rät (Handout 2). Zu viel Kontext ist aber ebenso ein Problem. Wer zu einer Frage sieben Dateien hochlädt, von denen fünf dasselbe enthalten, füllt das Kontextfenster mit Wiederholungen, und der Chatbot findet die wichtige Stelle schlechter. Wer in den Prompt lange Vorgeschichten schreibt, die für die Aufgabe keine Rolle spielen, lenkt ihn ab. Dazu gehört auch ein einzelnes „Danke“ als eigene Nachricht mitten im Chat: Der Chatbot liest dafür den ganzen Kontext noch einmal und antwortet, ohne dass etwas weitergeht. Die Faustregel: alles, was für die Aufgabe eine Rolle spielt, und nichts, was keine spielt.

Den Kontext im Griff behalten

1. Ein Thema, ein Chat. Bei einem neuen Thema einen neuen Chat starten.

2. Wenn sich ein Fehler wiederholt, obwohl du ihn korrigiert hast: neuer Chat.
3. Wenn ein Chat lang wird und die Antworten schlechter werden: den Chatbot bitten, das Gespräch als Ausgangspunkt für einen neuen Chat zusammenzufassen. Die Zusammenfassung prüfen, in einen neuen Chat kopieren, weiterarbeiten.
4. Wichtige Anweisungen an den Anfang und, bei langen Chats, ans Ende. Eine Regel, die weiter gelten soll, nach vielen Prompts einmal wiederholen.
5. Nur Material hochladen, das für die Aufgabe gebraucht wird. Nicht fünf Dateien, wenn eine reicht.
6. Was immer gilt, in den Profilprompt. Was für ein Projekt gilt, in den Projekt-Prompt. In den Chat nur das, was diese Aufgabe braucht.
7. Die Liste der Erinnerungen ab und zu durchsehen und Veraltetes löschen. Angaben, die nur für eine Aufgabe gelten, mit „Merk dir das nicht“ kennzeichnen.
8. Kein einzelnes „Danke“ als Nachricht. Wenn du fertig bist, ist der Chat fertig.

Selbst ausprobieren

A U F G A B E 1

Wann vergisst er die Regel?

Gib am Anfang eines neuen Chats eine Regel vor, zum Beispiel „Antworte immer in höchstens drei Sätzen“. Stell dann Frage um Frage zu einem Thema aus deinem Unterricht. Zähl mit: Beim wievielten Prompt hält sich der Chatbot zum ersten Mal nicht mehr an die Regel?

Regel gebrochen bei Prompt Nr.:

A U F G A B E 2

Ein Chat, zwei Themen

Lass dir in einem Chat zuerst bei einem Referatsthema helfen, dann in demselben Chat eine kurze Nachricht an eine Lehrperson formulieren. Lies die Nachricht genau: Findest du Spuren des Referatsthemas darin, im Ton oder im Inhalt?

Gefundene Spuren:

A U F G A B E 3**Umzug in einen neuen Chat**

Nimm einen älteren, längeren Chat. Bitte den Chatbot: „Fasse das bisherige Gespräch so zusammen, dass ich damit in einem neuen Chat weiterarbeiten kann.“ Lies die Zusammenfassung: Was fehlt, was ist falsch? Korrigiere sie, starte einen neuen Chat damit und stell eine Frage, die das Wissen aus dem alten Chat braucht.

Was in der Zusammenfassung gefehlt hat:
