1

Eine Frage, verschiedene Antworten ☺

Ich muss zur Waschanlage, um mein Auto zu waschen.
Sie ist 150 Meter von meinem Zuhause entfernt.
Soll ich zu Fuss gehen oder mit dem Auto fahren?

2

Halluzinationsrisiken von KI-Chatbots nach Aufgabenbereich

Aufgabenbereich	Typische Problematik	Problematische Outputs*
Ideenfindung & Strukturierung	Scheinplausible Annahmen	< 5 %
Argumentation & erklärende Texte	Logische Glättung, verkürzte Kausalität	10–25 %
Faktische Fragen (closed-book)	Erfundenes oder falsches Faktenwissen	15–40 %
Zusammenfassungen	Auslassungen, Bedeutungsverschiebung	5–20 %
Quellen & Zitate (ohne Retrieval)	Nicht existente / falsche Quellen	20–50 %
Recherche & Überblick	Selektive, scheinbar vollständige Darstellung	10–30 %
Coding & Programmierung	Funktional fehlerhafter Code	5–20 %
Mathematische / formale Logik	Rechen- und Ableitungsfehler	10–25 %
Aktuelle Ereignisse	Veraltete oder falsch datierte Infos	15–40 %
Entscheidungsnahe Domänen (z.B. Medizin, Recht etc.)	Unzulässige Empfehlungen	nicht sinnvoll quantifizierbar

Stand: Januar 2026

*Die angegebenen Prozentbereiche sind heuristische Richtwerte zur Orientierung. Sie beruhen auf typischen Aufgabenmerkmalen und bekannten Fehlmustern generativer KI, nicht auf systematischen Messungen oder Modellvergleichen. Sie geben die ungefähre Häufigkeit potenziell problematischer Outputs unter üblichen Nutzungsbedingungen an.

3

4

Agenda

1. Kurze Begriffsklärung
2. Die vier wichtigsten Einschränkungen von Chatbots
3. Metaprompting & Chain of Thought
4. Demo
5. Fazit

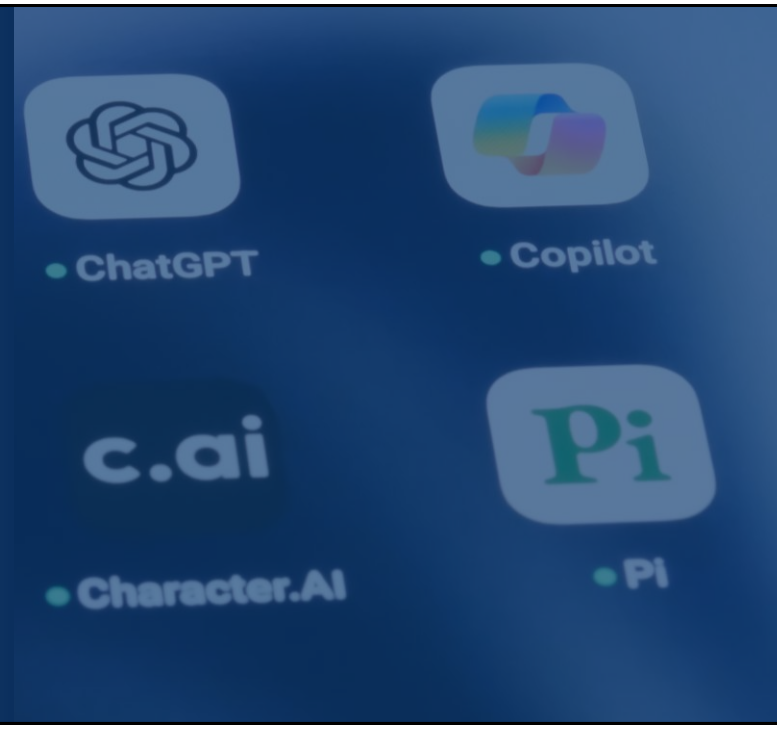
5

ChatGPT

G = generative

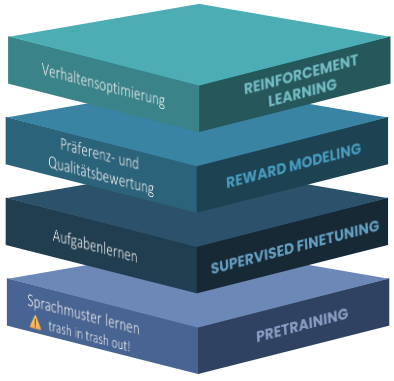
P = pre-trained

T = transformer



6

Der Lernprozess von LLMs durchläuft mindestens 4 Phasen



4. Bestärkendes Lernen -> Reinforcement Learning with Human Feedback
Im letzten Schritt wird das Modell durch Belohnungen und Strafen weiter optimiert. Es soll bewerten können, wie gut eine Modellantwort den **menschlichen Erwartungen** entspricht z. B. für höfliche, hilfreiche und korrekte Antworten. Diese Phase ist **sehr rechenintensiv**, da Milliarden Parameter erneut angepasst werden."

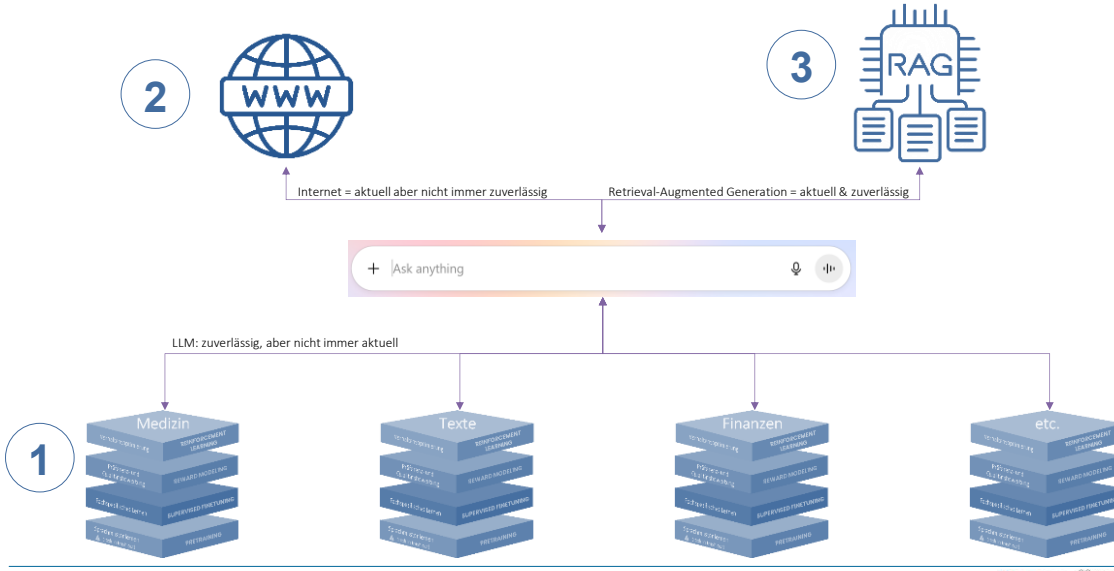
3. Belohnungsmodellierung -> Reward Model
Statt neues Sprachverhalten zu lernen, wird in dieser Phase ein Belohnungsmodell trainiert: Es soll die **Antworten auf besser/schlechter** einstufen können. Dazu vergleichen Menschen mehrere Antworten eines LLMs und markieren die bevorzugte. Diese manuelle Bewertung erfordert einen **beträchtlichen personellen Aufwand**.

2. Feinabstimmung -> Supervised Fine-Tuned Model
In dieser Phase wird das Modell mithilfe überwachter, von Menschen kuratierter Datensätze auf **spezifische Aufgaben oder Fachgebiete angepasst** – etwa um medizinische Fragen zu beantworten, rechtliche Texte auszuwerten oder einen bestimmten Schreibstil zu erlernen (Transferlearning). Der **Rechenaufwand ist deutlich geringer** als beim Pretraining, da die Datenmenge kleiner ist.

1. Vortraining -> Basismodell (Foundation Model)
Im Vortraining lernt das Modell eigenständig aus riesigen Textmengen (Textkorpora) – etwa aus Webinhalten, Wikipedia, Büchern oder Chats – **Sprachmuster und Weltwissen zu erfassen**. Diese Phase bildet das **Fundament** für alle weiteren Lernschritte. GPT-3 zum Beispiel wurde mit Milliarden Internetsätzen trainiert und lernte so, Textlücken zu schliessen. Diese Phase ist **besonders rechenintensiv und ressourcenaufwendig**.

7

Drei Wissensquellen für einen Chatbot



8

Agenda

1. Kurze Begriffsklärung
2. Die vier wichtigsten Einschränkungen von Chatbots
3. Metaprompting & Chain of Thought
4. Demo
5. Fazit

9

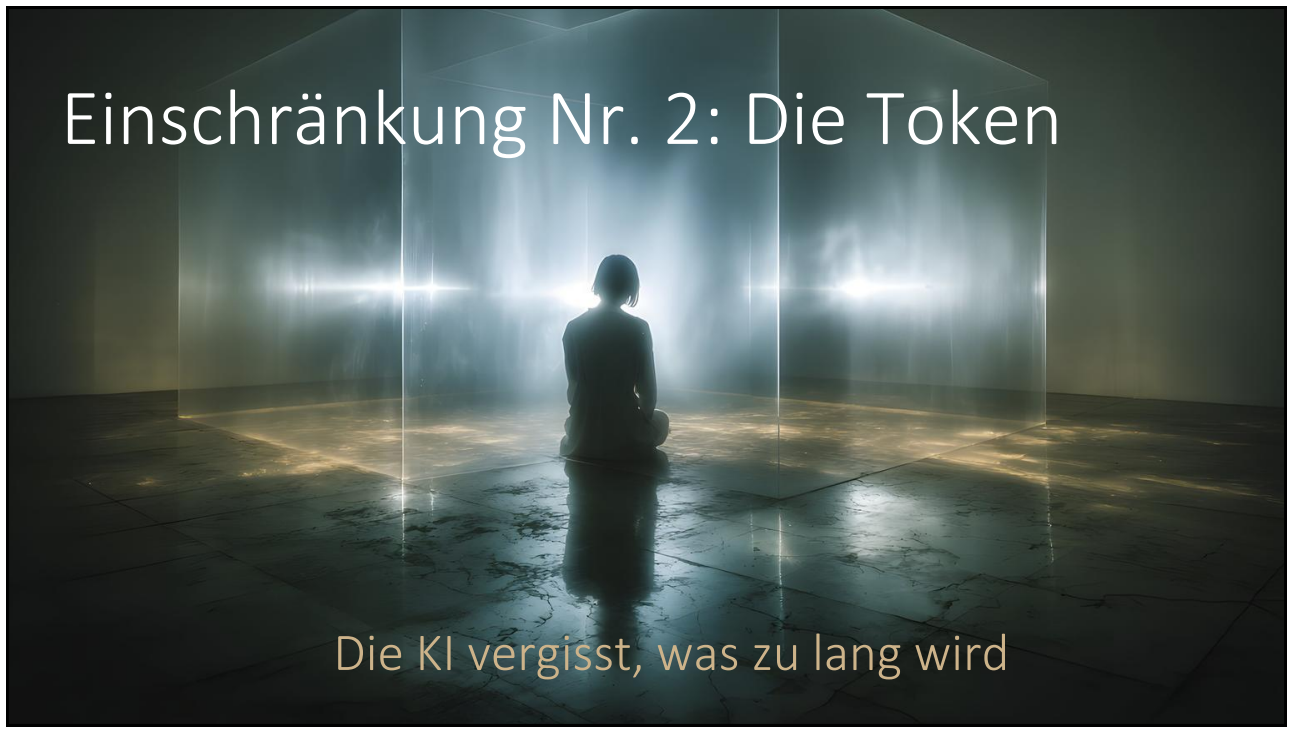
10

Die Grundregel in der Informatik = TITO

Trash in

Trash out

11



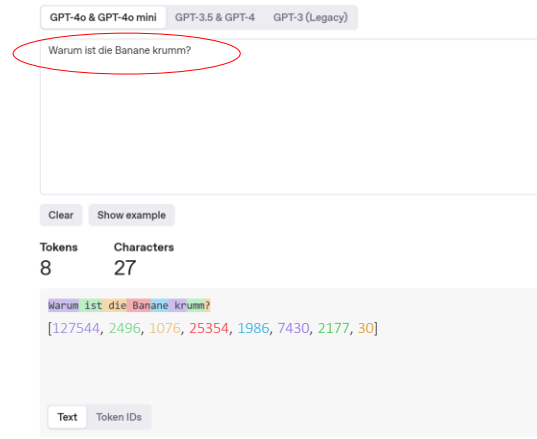
12

Die Sprache von LLMs ist „Token“

Für KI sind Wörter keine Texte – sie sind Zahlen. Und diese Zahlen bilden geometrische Muster im Raum.

Token sind die Bausteine von Wörtern in Zahlen:

- *Apfel* = 1 Token = [11756, 27849]
- *Ich esse einen Apfel* = 5 Token
- Ein langes Wort kann mehrere Tokens haben



https://platform.openai.com/tokenizer

13

LLMs haben einen begrenzten Kontextzugriff: sie «sehen» nur einen Teil

Modell	Max. Kontext (Tokens)	Entspricht ungefähr
GPT-3.5	ca. 4'096 Tokens	~ 3–4 Seiten Text
GPT-4 (Standard)	ca. 8'192–32'768 Tokens	~ 6–25 Seiten Text
GPT-5 (128k)	bis zu 128'000 Tokens	ca. 100 Seiten (aber technisch komplex)

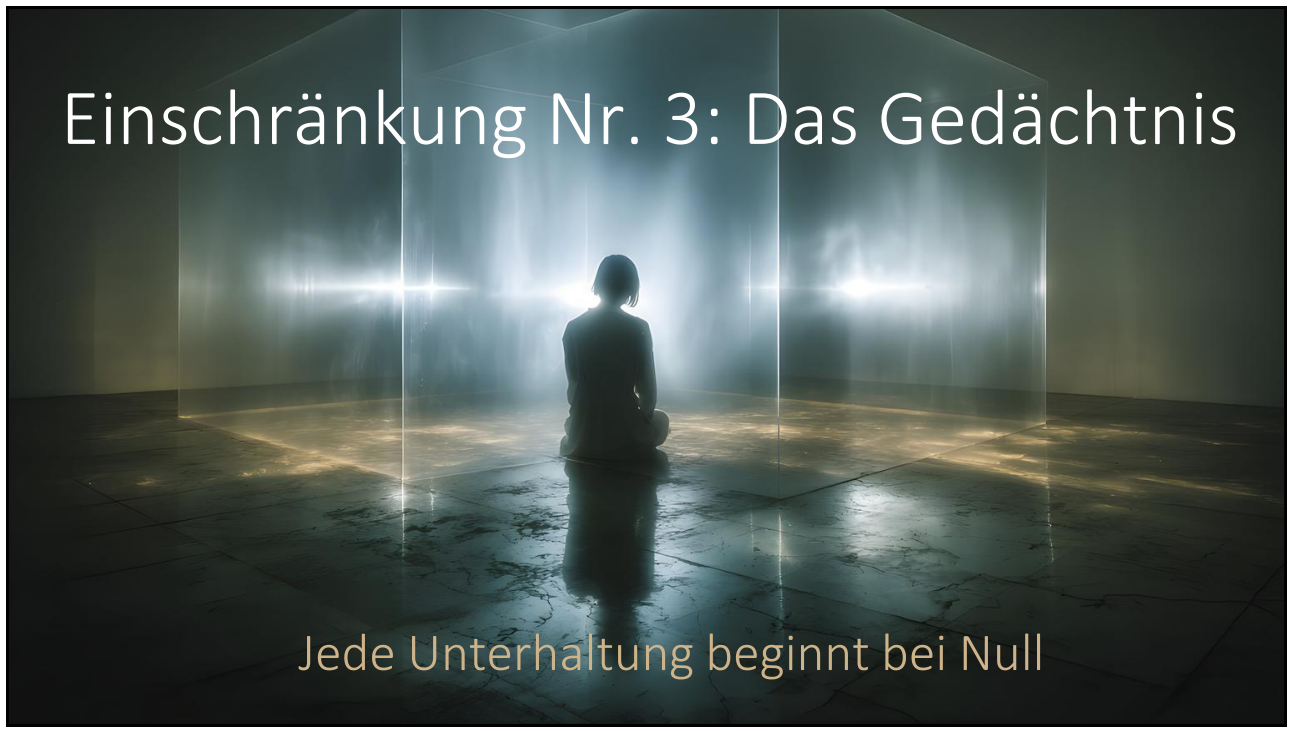
⚠ Das Modell „vergisst“ ältere Teile des Gesprächs (sie werden abgeschnitten). Auch bei langen Dokumenten (z.B. pdf) wird nur ein Teil analysiert.

14

In der Praxis kann man selten mehr als 80‘000 Token im Kontext-fenster sinnvoll gleichzeitig verarbeiten!

Segmentieren: Teile das Dokument in logisch gegliederte Abschnitte. In längeren Dialogen musst du frühere Informationen erneut einbringen!

15



16

LLMs haben (noch) kein Gedächtnis

Mensch

Ich erinnere mich

VS

Künstliche Intelligenz

Ich schlage nach

Bei jedem Prompt muss Chat-GPT nachlesen, was vorher war.

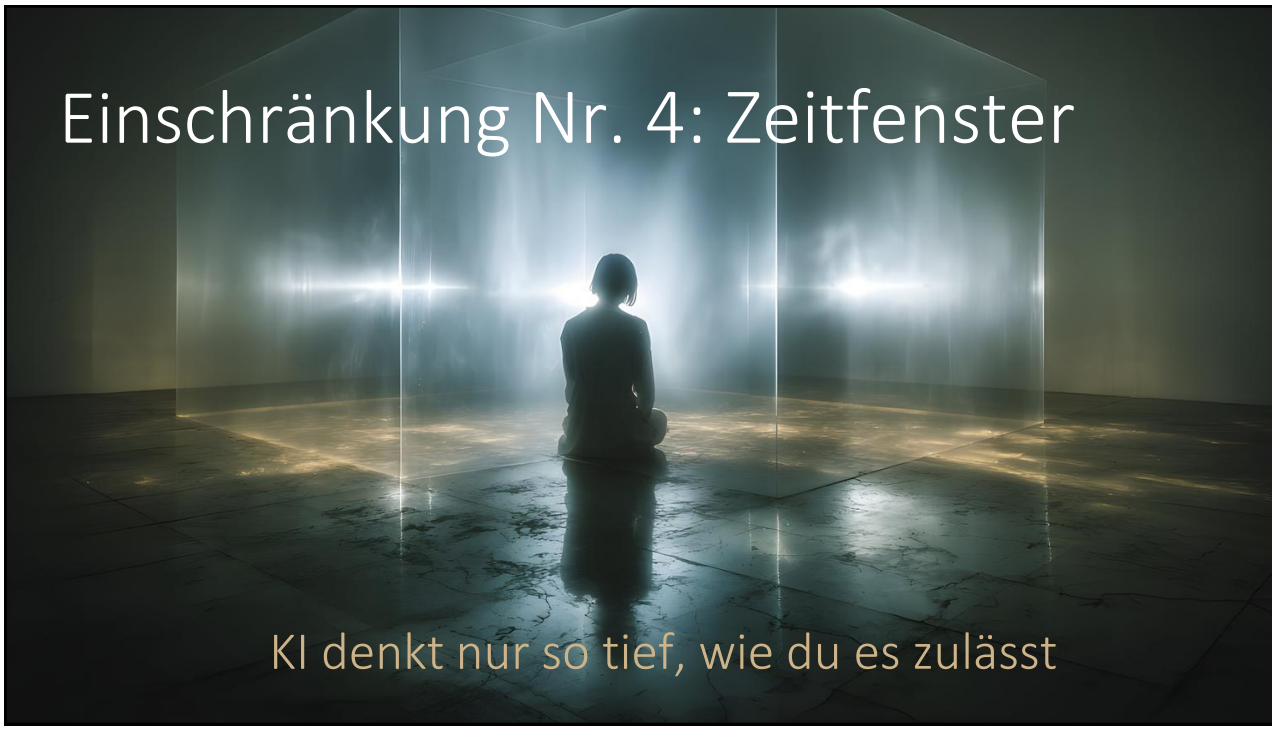
ChatGPT 5 kann aktuell max. ca. 128'000 Token nachlesen (ca. 100 Seiten). Der Rest wird abgeschnitten!

17

Hingegen speichert z.B. ChatGPT Daten über deine Person. Damit sollen die Antworten stimmiger ausfallen.

Diese persönlichen Daten können jedoch eingesehen und gelöscht werden!

18



19

Bessere Antworten mittels «Salamitechnik» (segmentieren)

« Löse das Problem Schritt für Schritt und erkläre deine Überlegungen! »

Chain-of-Thought = CoT-prompting

→ Da die Zeitfenster limitiert sind, verbessert die Gliederung des Problems in Teilprobleme die Antwortqualität.

20

Agenda

1. Kurze Begriffsklärung
2. Die vier wichtigsten Einschränkungen von Chatbots
3. **Metaprompting & Chain of Thought**
4. Demo
5. Fazit

21

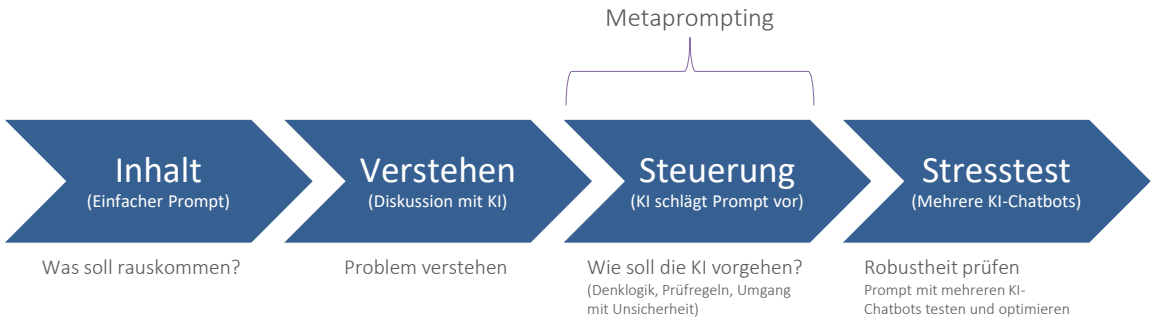
Gute Prompts sparen Zeit – schlechte kosten Nerven

1	**Gib die Rolle (set « persona ») & Ziel an …** (z. B. als Schriftsteller tätig sein, als Experte tätig sein, als Imker tätig sein, als YouTuber tätig sein, Ziel ist, meinen Internetauftritt zu optimieren)	Rolle & Ziel
2	**Erkläre die Aufgabe…** (z. B. schreibe einen akademischen Aufsatz, gib mir eine Lösung zur Korrektur dieses Textes, analysiere den folgenden Text)	Aufgabe
3	**Leg Einschränkungen/Kontext fest…** (z. B. lege Wert auf Genauigkeit, schreibe kurze Sätze, verwende einen poetischen Schreibstil, fasse dich kurz, füge viele Details hinzu, hebe hervor, was jede Methode zu einer aussergewöhnlichen Wahl macht, hier ein Beispiel)	Regeln
4	**Lege das Format des Outputs fest…** (z. B. Gib das Resultat in einer Tabelle, in Word, in Excel, in PowerPoint aus, erstelle ein Schaubild, schreibe eine Zusammenfassung von 2 Seiten.)	Format
5	**Verstanden? -> Frage wie vorgehen und ob sie alles verstanden hat…** (z. B. Falls du meinen Auftrag nicht vollständig verstanden hast, stelle mir zuerst Fragen bis du ganz sicher bist.)	Fragen? & Tools

- → **Sei präzise:** Je genauer deine Anfrage ist, desto relevanter wird die Antwort sein
- → **Fügen Beispiele hinzu:** Ein Beispiel für den erwarteten Stil oder das erwartete Format kann der KI helfen, besser zu verstehen.
- → **Testen und wiederholen:** Wenn die erste Antwort nicht perfekt ist, präzisiere, was in deiner nächsten Aufforderung angepasst werden muss.
- → **Bei komplexen Fragen:** frage die KI zuerst, wie sie vorgehen möchte. Sag ihr, welches Tool sie verwenden sollte etc..

22

Beim Metaprompting* steuern wir nicht die Antwort, sondern die Art, wie die KI zu ihren Antworten kommt.



*Der Begriff „Metaprompting" ist nicht einheitlich definiert. In dieser Präsentation bezeichnet er die bewusste Steuerung der Denk- und Arbeitsweise der KI – nicht nur die Beschreibung des gewünschten Ergebnisses.

23

Stärken und Grenzen von KI-Chatbots

Aufgabentyp	ChatGPT (OpenAI)	Claude (Anthropic)	Gemini (Google)	Perplexity	Copilot (Microsoft)	Grok (xAI)
Problemstrukturierung & heuristisches Denken *(Thema ordnen, Denkpfade skizzieren)*	🟢	🟢	🟡	🟡	🟡	🟡
Argumentative Tiefe & Kohärenz *(stringent argumentieren, Pro & Contra)*	🟢	🟢	🟡	🟡	🟡	🔴
Lange, nüchterne Fachtexte *(Berichte, Studien, Konzepte)*	🟡	🟢	🟡	🔴	🟡	🔴
Zusammenfassungen & Verdichtung *(Kernaussagen aus Texten)*	🟢	🟢	🟡	🟢	🟡	🟡
Systematische Recherche *(Themenüberblick, Suchrichtungen)*	🟡	🟡	🟢	🟢	🟡	🟡
Aktuelle Fakten & Quellen *(Was gilt aktuell? Wo nachprüfen?)*	🟡	🟡	🟢	🟢	🟢	🟡
Diskursive Einordnung & Perspektivenvielfalt *(unterschiedliche Sichtweisen)*	🟢	🟢	🟡	🟡	🟡	🟡
Kodieren & Programmieren *(Code erklären, debuggen, strukturierentitative Analysen)*	🟢	🟢	🟡	🔴	🟢	🟡
Tabellenkalkulation & Excel-Logik *(Formeln, Datenlogik, Auswertungen)*	🟡	🟡	🟡	🔴	🟢	🔴
Entscheidungsnahe Nutzung (Hochrisiko) *(Medizin, Recht, heikle Entscheide)*	🔴	🔴	🔴	🔴	🔴	🔴

Stand: Januar 2026

🟢 geeignet 🟡 bedingt geeignet 🔴 ungeeignet

ChatGPT
- Sehr stark in Strukturierung, Argumentation und erklärendem Denken
- Breite Einsetzbarkeit über viele Aufgabentypen hinweg (Allrounder)
- Risiko hoher Plausibilität bei inhaltlich unsicheren Ergebnissen

Claude
- Hohe argumentative Kohärenz und saubere Annahmenexplizierung
- Sehr gut für lange, neutrale Fachtexte
- Teilweise zu vorsichtig und wenig explorativ

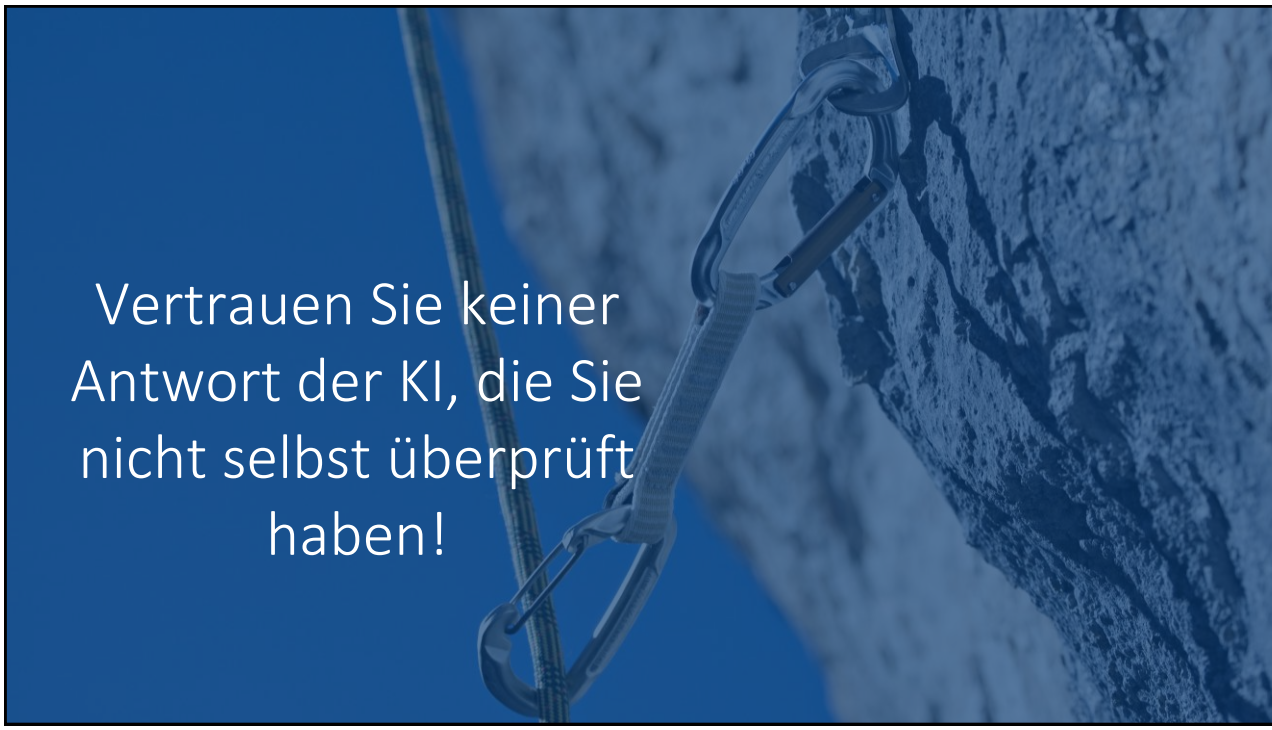
Gemini
- Gute Anbindung an aktuelle Informationen und Recherche
- Solide Allround-Leistung für Standardaufgaben
- Schwankende analytische Tiefe

Perplexity
- Sehr stark in Recherche, Exploration und Quellenorientierung
- Gute Orientierung in neuen oder unbekannten Themenfeldern
- Schwach bei Synthese und kohärenter Analyse

Copilot
- Sehr effizient in Office-, Excel- und Coding-Workflows
- Niedrige Einstiegshürde in produktiven Alltagsaufgaben
- Begrenzte konzeptionelle und argumentative Tiefe

Grok
- Pointierte, teils ungewöhnliche Perspektiven
- Schnell bei diskursiven oder meinungsnahen Themen
- Methodisch instabil und analytisch unzuverlässig

24



25

Agenda

1. Kurze Begriffsklärung
2. Die vier wichtigsten Einschränkungen von Chatbots
3. Metaprompting & Chain of Thought
4. Demo
5. Fazit

26

Demo

1. Notebook LM
2. Metaprompting

27

Agenda

1. Kurze Begriffsklärung
2. Die vier wichtigsten Einschränkungen von Chatbots
3. Metaprompting & Chain of Thought
4. Demo
5. Fazit

28

Fazit: Die Zukunft gehört nicht der KI – sondern den Menschen, die lernen, mit ihr zu denken

1. **Grenzen:** Generative KI halluziniert bisweilen (erfindet plausibel klingende, aber falsche Inhalte) und hat keine verlässliche Faktenbasis – daher Antworten nie blind übernehmen, sondern immer kritisch prüfen.
2. **TITO-Prinzip & Prompting:** „Trash in, Trash out" – KI liefert nur so gute Resultate wie ihre Eingaben. Klare, präzise Prompts und hochwertige Daten werden zur neuen Schlüsselkompetenz.

29

https://youtube.com/shorts/9VnYy-I5xsg?si=YPeOEGJI2Q8Ezrds

30

Besten Dank für Ihre Aufmerksamkeit

Hochschule für Wirtschaft & Tourismus
Dr. Serge Imboden
Techno-Pôle 3
3960 Sierre
+41 27 606 90 72
+41 79 217 06 08
serge.imboden@hevs.ch
www.2iManagement.ch

31