

## KI ist findig, wird aber nicht immer gut fündig

Die Modelle werten ihre Trainingstexte gewandt aus, aber ohne eigentliche Recherche im Internet und ohne Gewähr für Richtigkeit.

Gleich zwei schweizerische KI-Anwendungen sind in den letzten Wochen verfügbar geworden: nach [Apertus](#) nun auch [myAI](#) von Swisscom. Mit allerlei Fragen testeten verschiedene Medien, was diese Modelle «wissen». Dabei zeigte sich wiederholt, dass das in solchen Systemen gespeicherte Wissen wenig mit dem zu tun hat, was gemeinhin als Wissen gilt: gesicherte Fakten auswendig verfügbar zu haben. Um als intelligent gelten zu können, müsste jemand (oder etwas) aufgrund solchen oder beigezogenen Wissens dann auch noch Schlüsse ziehen können, die für die jeweils anstehende Aufgabe nützlich sind. Kürzlich habe ich Irrungen und [Wirrungen von KI-Textgeneratoren](#) erfahren, als ich dem seltenen Wort *anpreien* nachging. Das wiederholte Geflunker bei den Antworten auf diese Wissensfrage führte mich bei DeepSeek (und in der [Fortsetzung](#) des Protokolls) zu einer guten Erklärung der Funktionsweise:

### Die "Wissensbasis" sind primär die Trainingsdaten (statistisches Korpus)

Meine primäre "Wissensbasis" besteht aus den **Trainingsdaten**, einem enorm großen und vielfältigen Textkorpus (einschließlich Bücher, Websites, Artikel, etc.), mit dem ich trainiert wurde. Dieses Korpus ist:

- **Statistisch, nicht organisiert:** Das Wissen ist nicht in einer nach Themen sortierten Datenbank mit expliziten Quellenverweisen abgelegt. Stattdessen habe ich gelernt, statistische Muster in der Sprache zu erkennen – welche Wörter häufig zusammen vorkommen (z.B. "Schiff" und "Seemannssprache"), welche grammatikalischen Strukturen üblich sind und wie Konzepte miteinander in Beziehung stehen.
- **Parametrisiert:** Dieses gelernte Wissen ist in den **Parametern** meines Modells (175 Milliarden bei GPT-3) codiert. Man kann sich diese Parameter wie eine extrem komplexe Landkarte vorstellen, die die Beziehungen zwischen allen Wörtern und Konzepten in meinem Trainingstext abbildet. Wenn ich eine Frage beantworte, "navigiere" ich durch diese Landkarte, um eine wahrscheinliche Wortfolge zu generieren.

### Der Unterschied zu einer traditionellen Datenbank

Dieser Ansatz unterscheidet sich fundamental von einer klassischen, **wissensbasierten KI** (die auf einer manuell kuratierten Wissensdatenbank wie Wikidata oder einer Enzyklopädie operiert).

Merkmal	Statistische Sprachmodelle (wie ich)	Traditionelle Wissensdatenbank
Wissensrepräsentation	Statistische Muster in ~175 Mrd. Parametern	Explizite Fakten (Triples: Subjekt-Prädikat-Objekt)
Quellen	Undurchsichtiges Amalgam aus Trainingsdaten	Explizit referenzierte und überprüfbare Quellen
Stärke	Sprachverständnis, Generierung, Zusammenfassung	Exakte Abfrage von Fakten, 100%ige Überprüfbarkeit
Schwäche	<b>Halluzinationen</b> , ungenaue oder erfundene Quellen	Kein "Verständnis" von Sprache außerhalb der Schemata

[DeepSeek](#), 5.9.2025

Nun wurde auch diese Auskunft mithilfe des Sprachmodells erzeugt, als wahrscheinliche Wortfolge, die sich aus den Trainingstexten ergibt. «Halluzinationen» sind daher auch in der Selbstdarstellung von DeepSeek nicht ausgeschlossen, obwohl die systematische Gliederung der Tabelle so wirkt, als werde eingebautes Wissen zusammengefasst. Dass aber von zwei verschiedenen KI-Modellen übereinstimmendes Geklunker daherkäme, wäre doch sehr unwahrscheinlich. So hatten Googles KI-Übersicht, ChatGPT und DeepSeek bei anpreien recht unterschiedliche Irrwege eingeschlagen. Erhält man indes von verschiedenen Seiten die gleiche Auskunft, so ist diese viel vertrauenswürdiger.

### Sie sagen es gut, egal was

Bei der Frage nach der eigenen Wissensbasis hat nun ChatGPT ganz ähnlich geantwortet wie DeepSeek. Technische Zusatzfragen ergaben ein langes [Protokoll](#). Gegen Schluss steht ebenfalls eine Tabelle; sie ist etwas anders aufgebaut als die abgebildete, aber inhaltlich sehr ähnlich. «Sprachverständnis» beansprucht ChatGPT dabei nicht, wohl aber sprachliche Fähigkeiten: «Kann Infos neu kombinieren, vereinfachen, mit Analogien oder Diagrammen darstellen; passt sich deinem Ton und Kontext an (z. B. bildhafte Erklärungen).» Vermutlich meinte auch DeepSeek mit «Sprachverständnis» etwas Ähnliches und nicht, dass das Programm begreift, was seine Aussagen bedeuten. Vielmehr erfasst es, wie Sprache funktioniert, und kann sich deshalb geschickt ausdrücken, ungeachtet des Wahrheitsgehalts. Für die gestellten Aufgaben beweisen beide Dienste immerhin so viel Verständnis, dass sie meistens das gewünschte Thema in der verlangten Art behandeln.

Etwas boshaft könnte man mit einer IT-Redensart sagen, mögliche Halluzinationen in der Antwort seien keine Fehler, sondern eine Funktion des Programms (not a bug but a feature). Im ChatGPT-Protokoll steht dazu: «Ich kann oft plausibel klingende Fakten liefern, aber ich habe keine interne, durchsuchbare Liste von Quellen, die jede einzelne Aussage belegt. Deshalb rate ich bei kritischen Fakten zur Quellenprüfung. [...] Mein Wissensstand ist ein Schnappschuss bis zu einem bestimmten Cut-off (bei mir: bis zu meinem Trainings/Update-Zeitpunkt). Ohne Online-Recherche kann ich neuere Ereignisse nicht kennen.»

### Schwachpunkt Online-Recherche

Quellenprüfung und Online-Recherche sind gute Stichworte, wenn es um die Zuverlässigkeit von Auskünften geht – egal, ob man sich selber auf die Suche macht oder das einer KI-Anwendung überlässt. Im zweiten Fall darf man bloss nicht meinen, so ein dienstbarer Geist durchforsche nun das Internet und wisse schon, welchen Quellen man vertrauen dürfe. Die grossen Sprachmodelle (LLM in der englischen Abkürzung) suchen im eigenen Trainingsmaterial, aber nicht direkt nach Sachaussagen, sondern nach Wörtern, die sich zu einer Antwort auf die gestellte Frage zusammenfügen lassen. Die Wahrscheinlichkeit, dass diese Antwort stimmt, steigt mit der Qualität der fürs Training verwendeten Texte. Bei den Branchenführern wird den Modellen so ziemlich alles im Internet Greifbare verfüttert, und damit steigt die Gefahr, dass eine weitere IT-Redensart passt: «GIGO, garbage in, garbage out». Sehr frei zu Deutsch: Wes Mist ich fress, des Mist ich sing.

Soll zusätzliche Information direkt im Internet gesucht werden, gibt sich ChatGPT recht zugeknöpft: Es macht sich nicht selbständig auf die Suche, und gibt man ihm eine Adresse ein, die es konsultieren soll, so tut es das nur oberflächlich, ohne Fenster und Weiterleitungen der besuchten Seite zu beachten. Als Hauptgrund dafür nennt es die Sicherheit: «sehr hoch, weil kaum Angriffsfläche (keine Skripte, keine Cookies)». Nach dem entsprechenden Verhalten der «Konkurrenz» gefragt, antwortet ChatGPT ziemlich technisch, der (hier nicht wiedergegebenen) Vorgeschichte angemessen:

*Ich habe keine Einsicht in die internen Implementierungen, aber öffentlich bekannte Infos und Nutzerberichte lassen sich so zusammenfassen:*

- **Anthropic Claude**
  - *Hat (je nach Version) ebenfalls ein Web-Tool für Online-Suche.*
  - *Auch dort gilt: kein echtes Browser-Rendering. Claude holt Texte von Suchmaschinen/Seiten und fasst sie zusammen.*
  - *JavaScript/iframes werden meist nicht interpretiert.*
  - *Encoding-Probleme tauchen seltener auf, aber auch dort gibt es keine „volle Browserumgebung“.*
- **DeepSeek (chinesischer Anbieter)**
  - *Fokus stark auf Modellleistung, weniger auf Webanbindung.*
  - *Hat in manchen Umgebungen Plugins, die ähnlich wie bei mir Text aus Webseiten extrahieren.*
  - *Auch dort sind Sicherheits- und Einfachheitserwägungen wichtiger als 100 % Browser-Funktionalität.*
- **Google Gemini**
  - *Google hat theoretisch die Infrastruktur, eine Art „Mini-Browser“ im Hintergrund laufen zu lassen.*
  - *Dennoch gilt: das Webmodul von Gemini verhält sich im Prinzip wie bei mir → es zieht die Suchergebnisse aus Google Search API und liest (meist nur Textinhalte) von ausgewählten Seiten. Kein freies Surfen mit vollem Rendering.*
  - *Vorteil: enge Integration mit Googles Indizes → oft aktueller, aber technisch auch „eingeschränktes Crawling“, nicht volles Browsing.*
- **Perplexity / Companion**
  - *Arbeiten stark mit Retrieval aus Indexen + direktem Nachladen von Quellseiten.*
  - *Auch hier: kein echtes Browser-Rendering, sondern Textzugriff.*
  - *Vorteil: Sie haben eigene, große gecrawlte Indizes, die viele Encoding-Probleme vorab lösen.* *([ChatGPT](#), 3. 10. 2025)*

Da die Swisscom-AI auf Anthropic Claude (Google/Amazon) beruht, dürfte sie zu etwas besserer Suche fähig sein als ChatGPT; erste Testberichte lassen das ebenfalls vermuten. DeepSeek geht überhaupt nur dann auf die externe Suche, wenn man die Schaltfläche «Search» angeklickt hat; dann aber liefert es nicht nur Auskünfte, sondern nach einem weiteren Klick auch die konsultierten Adressen. Bei meinen Versuchen waren oft recht abwegige Websites dabei (manche interessant, aber nicht zweckdienlich); DeepSeek zeigte sich indes fähig, die passenden Funde auszuwählen. Über Apertus habe ich ChatGPT nicht befragt; meine eigenen Versuche bestätigen, was andere Kritiker bemängelten: Da diese Eigenentwicklung aus der Schweizer Wissenschaft ein kleineres Trainingsvolumen hat, hängt der Erfolg stark davon ab, ob das Thema dort hineinpasst. Die eigene Suche im Internet wird einem nicht abgenommen, aber man erhält ausgiebige Tipps dazu.

Indexeintrag «KI» in der laufenden «Sprachlupen»-Sammlung: [tiny.cc/lupen3](https://tiny.cc/lupen3).

© Daniel Goldstein ([sprachlust.ch](https://sprachlust.ch))