

Perplexity.ai

Was es wirklich kann und wo die Risiken liegen

Eine Analyse auf Basis von 22 aktuellen Studien und Tests* sowie der eigenen Arbeit mit Perplexity.ai

Genutzte KI-Tools zur Recherche: Scispace, Elicit, Google Labs, Liner, Perplexity, ChatGPT Deep Research

Zudem Recherche im DeepWeb (Base Search) und in lizenzierten Fachdatenbanken (z.B. Scopus), auf die KI-Tools keinen oder nur stark eingeschränkten Zugriff haben.

* Die Literaturliste findet sich am Ende der Präsentation.

Wo Perplexity gut ist

Perplexity funktioniert vergleichsweise gut bei:

- textbasiertem Faktenwissen
- Überblick & Einstieg in (neue) Themen
- klar formulierten Wissensfragen
- einigen standardisierten Prüfungsformaten

In diesen Fällen erreicht es eine akzeptable Genauigkeit, ist bei Vergleichstests aber selten an der Spitze.

Quellen: Cabezas-Clavijo 2025 · Podder 2024 ·
Gravina 2024

Wo Perplexity sinnvoll ist und wo nicht

Sinnvoll bei:

- Einstieg & Orientierung
- textbasierter Recherche
- Nutzung mit Prüfung

Probelamtisch bei:

- Zahlen- & Marktanalysen
- wissenschaftlicher Spezial- & Detailarbeit
- komplexen Analysen & medizinischen Fragen

Quellen: Sato 2025 · Kim 2025 · Patia 2024 · Jain 2024 ·
Kumar 2024 · Rana 2024 · Salih 2025 · Krause 2024 ·
Deike 2024

Trügerische Plausibilität

Studien zu Perplexity.ai zeigen

- inkonsistente Antworten
- wechselnde Begründungen
- systematische Fehlermuster

Besonders problematisch: Antworten wirken plausibel, auch wenn sie falsch oder unvollständig sind.

Quellen: Cao 2025 · Sato 2025 · Kim 2025 · Jain 2024 · Kumar 2024

Wo Perplexity besonders anfällig ist

Test zeigen wiederkehrende Schwächen bei

- Zahlen & Diagrammen
- Rechen- und Ableitungsaufgaben
- komplexen Synthesen mehrerer Quellen
- mehrstufigen Argumentationen

Die Fehlerraten sind hier besonders hoch und treten häufig auf.

Quellen: Jain 2024 · Uppalapati 2024 · Zandt 2025

Falsche Versprechen bei Quellen und Literatur

Die Quellen bei Perplexity.ai sehen gut aus, sind es aber nicht immer

- nicht jede Aussage ist wirklich belegt
- Quellen sind teilweise falsch zugeordnet
- in Tests traten wiederholt erfundene und unpassende Referenzen auf
- Zusammenfassungen wirken kohärent, basieren aber auf teils oberflächlichen Verknüpfungen

Quellen: Zhang 2023 · Venkit 2024 · Cabezas-Clavijo 2025

Ein weiteres Quellenproblem

Perplexity.ai neigt zu Intransparenz statt
Nachvollziehbarkeit. Oft bleibt unklar

- welche Quelle welche Aussage konkret stützt
- wie verschiedene Quellen gewichtet werden
- ob Originalstudien oder Zusammenfassungen der Studien zitiert sind
- mitunter werden Fachbücher genannt, Quelle ist aber nur eine Rezension des Buches

Quellen: Venkit 2024 · Johnston 2025 · Zandt 2025

Aktualitäts-Bias

Perplexity.ai glaubt, dass “neu” = “besser” ist und prüft und gewichtet die Quellen nicht nach wissenschaftlichen Maßstäben

- neuere Inhalte werden oft bevorzugt
- „aktuell“ wird implizit als „relevant“ betrachtet
- ältere, methodisch z.T. stärkere Arbeiten werden verdrängt oder weniger berücksichtigt
- es findet keine prüfbare wissenschaftliche Qualitätsbewertung der Quellen statt
- widersprüchliche Evidenz bleibt unsichtbar

Quellen: Venkit 2024 · Li 2024 · Zandt 2025

Nutzer und übermäßiges Vertrauen (Overreliance)

Verschiedene Nutzerstudien zeigen, dass perplexity (wie im Grunde alle KI-Tools) ein problematisches Verhalten fördert:

- weniger Klicks auf Originalquellen
- sinkende Überprüfung der Ergebnisse
- schnelle Übernahme von Antworten der KI
- hohes Vertrauen in die Ergebnisse und wenig kritische Distanz

Quellen: Venkit 2024 · Johnston 2025 · Ulloa 2026

Wie Perplexity sich selbst sieht

Stärken:

- Nachrichten und Fakten finden
- Gesetze und Policy-Dokumente analysieren
- schnelle Literaturrecherche machen
- vergleichende Übersichten erstellen

Schwächen

- stark akademische und rechtliche Recherchen
- verifizierte medizinische Informationen
- interpretative / zukunftsgerichtete Fragen

Quelle: *Perplexity.ai (Pro, Forschungsmodus)*

Prompt: *Wie gut ist Perplexity.ai? Suche nach empirischen Studien, prüfe sie genau und ermittle daraus eine realistische Einschätzung der Recherche-Fähigkeiten von perplexity.ai.*

Vergleiche mit anderen AIs. Zeige besondere Leistungen von perplexity.ai (empirisch bestätigt), aber auch Probleme und Schwachstellen.

Was noch zu sagen ist

Die vorliegende Analyse basiert auf 22 Studien und Tests. Aber gerade im DeepWeb und in den proprietären Fachdatenbanken - also genau dort, wo die KI-Recherche-Tools kaum oder nur stark eingeschränkt Zugriff haben - finden sich noch Dutzend weitere, die hier nicht berücksichtigt werden konnten.

Das macht nochmal deutlich: für eine umfassende Literatur- und Datenrecherche reichen KI-Tools nicht aus. Sie sind hilfreich, aber nur, wenn man um ihre quantitativen & qualitativen Grenzen weiß.

Literaturliste

Cabezas-Clavijo 2025 = Cabezas-Clavijo et al. (2025): Accuracy of Large Language Models for Medical Question Answering: A Comparative Study. Journal of Medical Internet Research, 27, e56027. DOI: [10.2196/56027](https://doi.org/10.2196/56027)

Cao 2025 = Cao et al. (2025): Large language models' performances regarding common patient questions about osteoarthritis: A comparative analysis of ChatGPT-3.5, ChatGPT-4.0, and Perplexity. Journal of Sport and Health Science, 1. DOI: [10.1016/j.jshs.2024.101016](https://doi.org/10.1016/j.jshs.2024.101016)

Deike 2024 = Deike (2024): Evaluating the performance of ChatGPT and Perplexity AI in Business Reference. Journal of Business & Finance Librarianship Vol. 29, 2024. <https://doi.org/10.1080/08963568.2024.2317534>

Gravina 2024 = Gravina et al. (2024): Charting new AI education in gastroenterology: Cross-sectional evaluation of ChatGPT and Perplexity AI. Digestive and Liver Disease, 56(8), 1304-1311. DOI: [10.1016/j.dld.2024.04.012](https://doi.org/10.1016/j.dld.2024.04.012)

Jain 2024 = Jain et al. (2024): A Comparative Assessment of Advanced Conversational Agents: A Multifaceted Evaluation of ChatGPT, Gemini, Perplexity, and Claude. International Journal of Emerging Technology and Advanced Engineering. 14. 59-69. DOI: [10.46338/ijetae0224_07](https://doi.org/10.46338/ijetae0224_07)

Johnston 2025 = Johnston et al. (2025): Discovering how students use generative artificial intelligence tools for academic writing purposes. Journal of Learning Development in Higher Education, Issue 34 (February 2025). DOI: <https://doi.org/10.47408/jldhe.vi34.1301>

Literaturliste

Kim 2025 = Kim et al. (2025): Human-AI collaboration in large language model-assisted brain MRI differential diagnosis: a usability study. Eur Radiol. 2025 Sep;35(9):5252-5263.
[doi: 10.1007/s00330-025-11484-6](https://doi.org/10.1007/s00330-025-11484-6)

Krause 2024 = Krause (2024): The Finance AI Challenge: An Evaluation of the Top Six Free Webbased AI Models (June 02, 2024). <http://dx.doi.org/10.2139/ssrn.4851699>

Kumar 2024 = Kumar et al. (2024): A Comparative Study of ChatGPT, Gemini and Perplexity AI. International Journal of Innovative Research in Computer Science and Technology, 12(3)
DOI:[10.55524/ijircst.2024.12.4.2](https://doi.org/10.55524/ijircst.2024.12.4.2)

Li 2024 = Li et al. (2024): Generative AI Search Engines as Arbiters of Public Knowledge: An Audit of Bias and Authority. In Proceedings of the Association for Information Science and Technology.
<https://doi.org/10.48550/arXiv.2405.14034>

Patia 2024 = Patia et al. (2024): The Use of Perplexity AI in Academic Writing. Benefits, Challenges, and Ethics, Wahana Pedagogika, Vol. 7, No. 1, <https://doi.org/10.52166/wp.v7i1.9642>

Podder 2024 = Podder et al. (2024): Evaluation of Artificial Intelligence-Based Chatbot Responses to Common Dermatological Queries. Jordan Medical Journal, 58(3)
<https://doi.org/10.35516/jmj.v58i2.2960>

Rana 2024 = Rana et al. (2024): Review of Modern Chatbot Technologies: Performance, Capabilities, and Applications. Artificial Intelligence Review, DOI: [10.1109/MIPRO65660.2025.11131741](https://doi.org/10.1109/MIPRO65660.2025.11131741)

Literaturliste

Salih 2025 = Salih et al. (2025): Assessment of ChatGPT, Gemini, and Perplexity in Principle of Research Publication: A Comparative Study. Barw Medical Journal, 3(1), 2-6.

DOI: [10.58742/bmj.v2i4.140](https://doi.org/10.58742/bmj.v2i4.140).

Sato 2024 = Sato et al. (2024): Performance Evaluation of 18 Generative AI Models (ChatGPT, Gemini, Claude, and Perplexity) in the 2024 Japanese Pharmacist Licensing Examination: Comparative Study. Journal of Medical INternet Research, Med. Educ. 11(1), e76925.

DOI: [10.2196/76925](https://doi.org/10.2196/76925)

Ulloa 2026 = Ulloa et al. (2026): Opting out of generative AI: a behavioral experiment examining the role of education in Perplexity AI avoidance. Telematics and Informatics, Vol. 104,

<https://doi.org/10.1016/j.tele.2025.102346>

Uppalapati 2024 = Uppalapati et al. (2024): A Comparative Analysis of AI Models in Complex Medical Decision-Making Scenarios: Evaluating ChatGPT, Claude AI, Bard, and Perplexity. Cureus, 16(1), e52485. DDOI: [10.7759/cureus.52485](https://doi.org/10.7759/cureus.52485)

Venki 2024 = Venkit et al. (2024): Search Engines in an AI Era: The False Promise of Factual and Verifiable Source-Cited Responses. <https://doi.org/10.48550/arXiv.2410.22349>

Zandt 2025 = Zandt (2025): Perplexity Deep Research im Test: Wenn KI Fakten erfindet. t3n vom 23.02.2025 <https://t3n.de/news/perplexity-deep-research-test-1673581/>

Zhang 2023 = Zhang et al. (2023): Evaluating Verifiability in Generative Search Engines. arXiv preprint. <https://doi.org/10.48550/arXiv.2304.09848>