

Wie zuverlässig sind KI-Empfehlungen? 800 Antworten im Test

800 KI-Antworten · 10 Branchen · 4 Modelle · 20 Wiederholungen je Frage · Stand 2026-07-25

Kernergebnisse

- Das meistgenannte Unternehmen blieb über 20 identische Abfragen in 84,5 % der Läufe dasselbe — KI-Empfehlungen sind nicht beliebig.
- Pro Frage wurden im Schnitt 27,6 verschiedene Unternehmen genannt — davon nur 3 verlässlich ($\geq 80\%$ der Läufe) und rund 19,8 nur sporadisch ($\leq 20\%$).
- ChatGPT war am unbeständigsten (63,5 % konstanter Spitzenreiter), Claude am konstantesten (97,5 %).
- Für eine Nennungsrate mit ± 10 Prozentpunkten Genauigkeit sind rechnerisch 97 Wiederholungen nötig — eine Einzelabfrage ist keine Messung.

Die Studie in Zahlen

84,5 %

Top-Anbieter bleibt gleich

27,6

Unternehmen je Frage

3

davon verlässlich

19,8

Zufalls-Anbieter

Beständigkeit je Modell



MODELL	TOP-1 GLEICH	TOP-3 GLEICH	KERN	UFALL
Claude	97,5 %	94,8 %	4,6	5,2
Perplexity	90,5 %	84 %	2,9	1,2
Gemini	86,5 %	83,2 %	3,5	44,1
ChatGPT	63,5 %	55,3 %	0,8	28,8

Quelle: „Wie zuverlässig sind KI-Empfehlungen? 800 Antworten im Test“ — Citavo (Stand 2026-07-25) ·

<https://www.citavo.app/studien/ki-empfehlungen-stabilitaet>

Methode: 800 echte KI-Antworten, 20 Wiederholungen je Frage, vorregistriertes Protokoll · citavo.app

Methodik & Grenzen

Design: 10 DACH-Branchen, je eine markenfreie Kaufrage, Sprache de-DE. Jede Frage 20x an 4 Modelle mit Live-Websuche (40 Messzellen, 800 Antworten).

Vorregistrierung: Metriken, Vergleichsregeln und Limitationen wurden vor der Auswertung schriftlich fixiert. Ein Modell gilt nur bei nicht-überlappenden 95-%-Konfidenzintervallen (Bootstrap) als beständiger.

Datenprüfung: alle Antworten je Zelle waren unterschiedliche Texte (kein Caching); extrahierte Firmen kamen wörtlich im Antworttext vor; Firmennamen wurden auf Unternehmensebene zusammengeführt.

Grenzen: eine Frage je Branche (misst Schwankung, keine Branchenunterschiede); ein Zeitfenster; nur de-DE und eine Formulierung. Die Kern-/Zufalls-Einteilung ist explorativ.