

○ Studie

Meinungsbildung im Wandel

Wie KI-Systeme Quellen bei der Nachrichtensuche
selektieren

Autorinnen-Team:

Vivien Benert, Stina Timmermann



Danksagung

Wir danken allen Gesprächspartner:innen für den wertvollen Input im Entstehungsprozess dieser Studie.

Agora Digitale Transformation gGmbH
Krausenstraße 8
10117 Berlin

T. +49(0)30 81 45 03 78 80
www.agoradigital.de
info@agoradigital.de

Ansprechperson:



Vivien Benert
vivien.benert@agoradigital.de

Gefördert von:

STIFTUNG
MERCATOR

Meinungsbildung im Wandel: Wie KI-Systeme Quellen bei der Nachrichtensuche selektieren
<https://agoradigital.de/projekte/ki-als-nachrichtenquelle/>

Design: [Jakub Stejskal](#), [Make Studio](#) | Coverbild: [Lena Haubner](https://lenahaubner.de/), <https://lenahaubner.de/>
Satz: Agora Digitale Transformation



Dieser Beitrag unterliegt einer CreativeCommons-Lizenz (CC BY-SA). Die Vervielfältigung, Verbreitung und Veröffentlichung, Veränderung oder Übersetzung von Inhalten der Agora Digitale Transformation, die mit der Lizenz „CC BY-SA“ gekennzeichnet sind, sowie die Erstellung daraus abgeleiteter Produkte sind unter den Bedingungen „Namensnennung“ und „Weiterverwendung unter gleicher Lizenz“ gestattet. Ausführliche Informationen zu den Lizenzbedingungen finden Sie hier: <http://creativecommons.org/licenses/by-sa/4.0/>

Juli 2026 | <https://doi.org/10.5281/zenodo.21162562>



Executive Summary

KI-Systeme werden zunehmend zur Nachrichtenquelle und damit zu neuen Gatekeepern der öffentlichen Meinungsbildung. Sie entscheiden, welche Inhalte, Themen und Quellen sichtbar werden, ohne dass ihre Auswahlkriterien dokumentiert oder nachvollziehbar wären. Die vorliegende Studie untersucht anhand einer eigenen systematischen Erhebung, welche Quellen führende KI-Systeme zitieren, wenn sie zur Nachrichtensuche genutzt werden. Grundlage sind 4.811 Quellenverweise aus 675 Anfragen zu politisch und gesellschaftlich relevanten Themen.

Ergebnisse

Wenige reichweitenstarke Domains dominieren die Zitationen, während die Mehrzahl der insgesamt 554 identifizierten Quellen nur vereinzelt vorkommt. Die Systeme verstärken damit bestehende Aufmerksamkeitsverzerrungen, statt sie auszugleichen. Zugleich verschiebt sich die politische Ausrichtung der zitierten Quellen je nach Thema und System, ohne dass dies für Nutzende erkennbar wäre. Claude verweist zudem mehrfach auf Domains des russischen Desinformationsnetzwerks Pravda, die in den Antworten ungekennzeichnet neben journalistischen Quellen stehen.

Implikationen

Die Quellenauswahl der KI-Systeme ist eine Blackbox. Keiner der untersuchten Anbieter dokumentiert, nach welchen Kriterien das System Quellen auswählt und gewichtet. Die Auswahl variiert mit Prompt, Thema und Anbieter, ist weder reproduzierbar noch überprüfbar und entzieht sich damit der Kontrolle durch Nutzende, Aufsicht und Forschung. Selbst Verweise auf bekannte Desinformationsnetzwerke wie Pravda bleiben in den Antworten ungekennzeichnet und sind für Nutzende nicht von journalistischen Quellen zu unterscheiden.

Die journalistische Finanzierungsgrundlage ist in Gefahr und damit eine zentrale Säule demokratischer Informationsvermittlung. KI-Systeme liefern vollständig ausformulierte Antworten. Nutzende haben so kaum noch Anlass, die zugrunde liegenden Quellen aufzurufen. Diese Zero-Click-Logik entzieht den Medien Reichweite und Einnahmen, obwohl deren Inhalte die KI-Systeme als Nachrichtenquelle erst nutzbar machen und den Anbietern erhebliche Gewinne ermöglichen.

Handlungsbedarfe

Es braucht umfassendere Transparenzpflichten im laufenden Betrieb von KI-Modellen und Systeme für eine faire Vergütung bei Zugriff auf Inhalte Dritter. Bestehende Kooperationen mit Medienhäusern müssen für Nutzende bei Nutzung der Systeme offengelegt werden.



Inhalt

1	Nachrichten- und Informationsnutzung in Zeiten Künstlicher Intelligenz	3
2	Methodik.....	5
3	Ergebnisse: Quellenprofile der untersuchten KI-Modelle.....	6
3.1	Gesamtverteilung zitierter Links und Domains	6
3.2	Journalismus, Zivilgesellschaft und Co.: Wen zitieren die KI-Systeme?	10
3.3	Politische Ausrichtung der KI-Systeme bei der Auswahl journalistischer Quellen	13
	Exkurs: Mistral AI.....	16
4	Implikationen und Handlungsbedarfe mit Blick auf KI-Anbieter und Journalismus	17
4.1	Transparenz: Quellenauswahl von KI-Systemen als Blackbox	17
4.2	Vielfalt: Das Dilemma des Journalismus	19
	Literaturverzeichnis.....	22
	Anhang	24



1 Nachrichten- und Informationsnutzung in Zeiten Künstlicher Intelligenz

Kaum eine neue technologische Entwicklung hat sich so schnell im Alltag etabliert wie künstliche Intelligenz (KI). In einer aktuellen Umfrage geben 34 Prozent der Deutschen an, KI mindestens wöchentlich zu nutzen¹. Insbesondere (text-)generative KI-Angebote wie ChatGPT und Gemini werden mittlerweile von vielen Menschen genutzt. So haben zum Beispiel auch der Digitalminister Karsten Wildberger² und der thüringische Ministerpräsident Mario Voigt³ wichtige Reden und Dokumente mithilfe von generativer KI erstellt.

Einer aktuellen Umfrage des Bayrischen Rundfunks (BR) zufolge haben 82 Prozent der Deutschen große Sprachmodelle wie ChatGPT, Gemini und Co. schonmal als Informations- und Nachrichtenquelle genutzt. Auch die KI-basierten Zusammenfassungen von Suchmaschinen wie etwa die Google AI Overview haben 53 Prozent der Befragten schon einmal genutzt und 41 Prozent nutzen sie mindestens wöchentlich.⁴

KI-Systeme werden damit zu neuen Gatekeepern. Sie entscheiden mit, welche Inhalte, Themen und Meinungen sichtbar werden und welche nicht.

Gleichzeitig geben nur 25 Prozent der Befragten an, die von den KI-Systemen angegebenen Quellen in mindestens der Hälfte der Anwendungsfälle zu überprüfen. Die restlichen Befragten machen dies entweder gar nicht oder deutlich seltener.⁵

Dieses Phänomen wird als Zero-Click-Suche bezeichnet. Anders als bei klassischen Suchmaschinen erhalten Nutzende von KI-Systemen keine Liste mit Links zu Suchergebnissen mehr, sondern einen vollständig ausformulierten Antworttext. Wer eine vollständig formulierte Antwort erhält, hat kaum noch Anlass, die zugrunde liegenden Originalquellen aufzurufen. Aus diesem Grund sind die Besucherzahlen vieler Web-

Begriffsdefinitionen

KI-System: Ein KI-System ist ein maschinengestütztes System, das mit einem gewissen Grad an Autonomie Eingaben verarbeitet und daraus Ausgaben und Inhalte generiert. Es meint die vollständige, in einen Anwendungskontext integrierte Lösung, die neben KI-Modellen auch Komponenten der Dateneingabe, -verarbeitung, Nutzerinteraktion und Systeminfrastruktur umfasst.

KI-Modell: Ein KI-Modell ist der mathematische oder algorithmische Kern eines KI-Systems: eine trainierte statistische Funktion oder ein neuronales Netz, das durch maschinelles Lernen Muster erkennt und Vorhersagen trifft. Im Unterschied zum KI-System kann ein KI-Modell für sich allein keine vollständige Anwendung realisieren, sondern benötigt eine einbettende Systeminfrastruktur. Es kann aufgabenspezifisch konzipiert oder für ein breites Aufgabenspektrum ausgelegt sein.

¹ Wintergerst, 2026

² BMDS, 2026

³ Winter et. al., 2026

⁴ Heßler, 2025

⁵ Heßler, 2025



sites stark zurückgegangen seit Google KI-Übersichten in seine Suchanfragen integriert.⁶ In Deutschland verlieren Websites so 265 Millionen organische Klicks pro Monat.⁷

Die ökonomischen Folgen der Zero-Click-Suche stellen insbesondere den Journalismus vor neue Finanzierungsprobleme.

Als besonders vertrauenswürdige Datenquelle sind journalistische Inhalte für KI-Systeme und ihre Anbieter besonders wertvoll und werden aus diesem Grund vielfach von KI-Systemen referenziert. Gleichzeitig leisten KI-Anbieter, mit Ausnahme einzelner Kooperationsvereinbarungen wie etwa zwischen OpenAI und Axel Springer⁸, in den meisten Fällen keine finanzielle Kompensation an Medien, auch wenn sie sehr regelmäßig auf ihre Inhalte zugreifen und selbst finanziell von diesen Inhalten profitieren. Es entsteht eine Asymmetrie der Wertschöpfung, denn journalistische Inhalte machen die KI-Systeme als Nachrichtenquelle erst nutzbar, während die daraus erzielten Erlöse bei den KI-Anbietern verbleiben.

Ein weiteres Problem liegt in der Fehleranfälligkeit der KI-Systeme. Eine umfangreiche Analyse der European Broadcasting Union aus dem Jahr 2025 in 18 Ländern zeigt, dass 81 Prozent aller KI-Antworten Fehler bei der Quellenauswahl, beim Umgang mit Zitationen oder bei der inhaltlichen Korrektheit der gelieferten Informationen enthalten.⁹ Während alle getesteten Modelle Fehler aufwiesen, war die Rate schwerwiegender Fehler bei Gemini besonders hoch.

KI-Systeme werden damit zu neuen Gatekeepern. Sie selektieren Informationen und Meinungen, während die Quellenauswahl zur Beantwortung der Anfragen von keinem Anbieter transparent gemacht wird. Diese Entwicklung betrifft den Kern demokratischer Öffentlichkeit, da die Selektion der Quellen zur Nachrichten- und Informationssuche die Grundlage öffentlicher Meinungsbildung darstellt. Vor diesem Hintergrund untersucht die vorliegende Studie, welche Quellen KI-Systeme zitieren, wenn Menschen sie zur Nachrichtensuche verwenden.

⁶ Newman, 2026

⁷ Beus, 2026

⁸ OpenAI, 2023

⁹ Fletcher & Verckist, 2025



2 Methodik

Um eine breite Datengrundlage zu erreichen, wurden insgesamt 625 Anfragen (Prompts) zu fünf relevanten Nachrichtenthemen (Innenpolitik, Außenpolitik, Wirtschaft, Klima und Migration) an fünf Sprachmodelle großer KI-Anbieter (ChatGPT, Google AI Overview, Gemini, Perplexity, Claude) gestellt. Dabei wurde die Formulierung der Prompts systematisch variiert, um unterschiedliches Promptingverhalten von Nutzenden sowie bestehende Erkenntnisse zum Einfluss dieses Promptingverhaltens auf die Ergebnisse der Sprachmodelle abbilden zu können (vgl. Tabelle 1). Um tagesspezifische Einflüsse einzelner dominierender Nachrichten und Medienhäuser zu verringern, wurde die Erhebung an fünf aufeinanderfolgenden Werktagen (04.05.2026 bis 08.05.2026) durchgeführt. Insgesamt wurden so 4.570 Links von 544 Domains¹⁰ als Quellen in KI-Anfragen erhoben.

Um das typische Nutzungsverhalten von Privatpersonen abbilden zu können, wurden jeweils die kostenlosen Versionen der KI-Anbieter ohne Anmeldung und in einem Inkognito-Fenster¹¹ genutzt. Eine Ausnahme stellt Claude dar: da die Anzahl der Anfragen in der kostenlosen Version stark begrenzt ist, wurde hier auf die Bezahlversion zurückgegriffen. Eine weitere Ausnahme betrifft Perplexity: hier ist eine Nutzung nur mit Anmeldung möglich, so dass für die Datenerhebung ein neuer kostenloser Account erstellt wurde.

Übersicht der analysierten KI-Systeme und Modelle

ChatGPT

Anbieter: Open AI
Modell: GPT 5.3 Instant

Claude

Anbieter: Anthropic
Modell: Sonnet 4.6

Gemini

Anbieter: Google
Modell: Gemini 3 Flash

Google AI Overview

Anbieter: Google
Modell: Gemini 3 Flash

Perplexity

Anbieter: Perplexity AI
Modell: Sonar

¹⁰ Inklusive Mistral sind es 4.811 Links von 554 Domains aus 675 Anfragen. Ergebnisse zu Mistral befinden sich im Exkurs auf Seite 16.

¹¹ Um Einflüsse vorheriger Suchanfragen z. B. bei Google und Google AI Overview zu minimieren, wurden alle Anfragen in einem eigenen Inkognito-Fenster gestellt.



3 Ergebnisse: Quellenprofile der untersuchten KI-Modelle

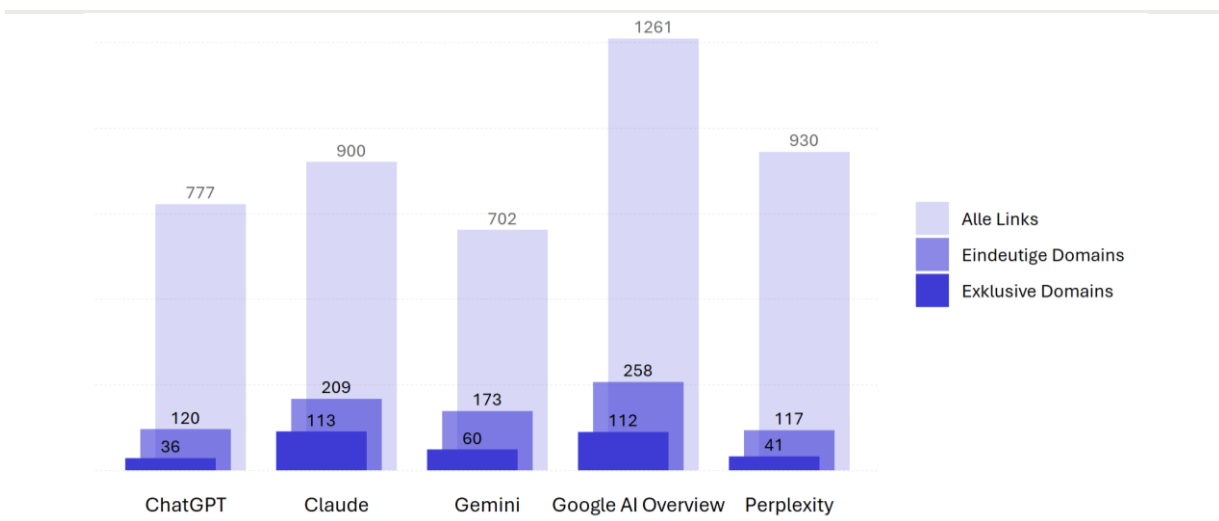
Die untersuchten KI-Systeme weisen trotz einiger Gemeinsamkeiten unterschiedliche Quellenprofile auf. Zwar greifen alle Systeme auf einen geteilten Kernbestand etablierter, überwiegend journalistischer Quellen zurück, doch dahinter verbergen sich systemspezifische Unterschiede in der Quellenbreite, Akteursstruktur und politischer Ausrichtung.

3.1 Gesamtverteilung zitierter Links und Domains

Die Gesamtverteilung der zitierten Quellen über alle Anfragen hinweg weist je nach KI-System deutliche Unterschiede auf. Google AI Overview, die integrierte KI-Übersicht in der Google-Suche, zitiert über alle Anfragen hinweg mit Abstand die meisten Quellen (n=1.261), gefolgt von Perplexity (n=930) und Claude (n=900). Gemini gibt die wenigsten Quellen an (n=700) (vgl. Abb. 1). Hierbei sind alle von den KI-Systemen aufgeführten Links berücksichtigt. Wenn also beispielsweise derselbe Artikel eines (Online-)Mediums in verschiedenen Anfragen zitiert wurde, wird er mehrfach gezählt.

Schaut man hingegen auf die Anzahl der eindeutigen Domains, die die verschiedenen KI-Systeme zitieren, sinkt die Zahl bei allen KI-Systemen deutlich. Auf der Ebene der Domains (z. B. faz.de, welt.de, wikipedia.com, zeit.de, etc.) werden Links zu unterschiedlichen Artikeln desselben Akteurs nicht mehr doppelt gezählt. Die Anzahl der zitierten Domains ist somit ein Hinweis auf die Quellenvielfalt der KI-Systeme. Es zeigt sich, dass die Anzahl der eindeutigen Domains bei allen KI-Systemen deutlich geringer ausfällt als die Anzahl der zitierten Links. Dies zeigt, dass alle untersuchten KI-Systeme aus einer vergleichsweise geringen Quellenbasis schöpfen und immer wieder dieselben Domains zitieren. Auch die Anzahl der exklusiven Domains, also derjenigen Domains, die nur von dem jeweiligen KI-System zitiert werden, ist bei allen untersuchten KI-Systemen gering. Die meisten exklusiven Domains weisen Claude (n=113) und Google AI Overview (n=112) auf. ChatGPT (n=36) nutzt nur selten Quellen, die andere KI-Systeme nicht zitieren.

Abbildung 1: Anzahl zitierter Links und Domains je KI-System

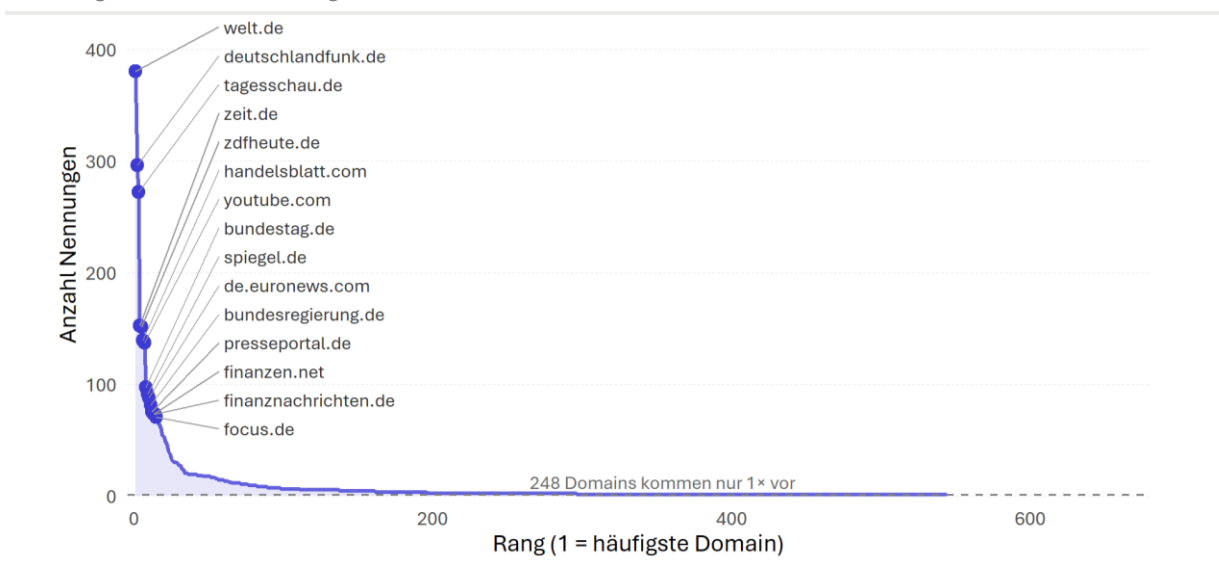


Anmerkung: Gesamt N=4.570 Links, 544 Domains. Eindeutige Domains: Jede Domain (z. B. tagesspiegel.de) wird nur einmal gezählt. Exklusive Domains: Domain wird ausschließlich von diesem KI-System zitiert.

KI-Systeme brechen ungleiche Aufmerksamkeitsverteilungen im Digitalen nicht auf, sondern reproduzieren und verstärken sie.

Die Häufigkeitsverteilung der zitierten Links über alle analysierten KI-Systeme hinweg entspricht einer klassischen Power-Law-Verteilung¹², bei der wenige Domains sehr häufig zitiert werden und der Großteil der Domains nur wenige Male oder sogar nur einmal zitiert wird und im sogenannten Long-Tail der Verteilung verschwindet (vgl. Abb. 2). Insgesamt werden 248 Domains, also fast die Hälfte der 544 eindeutigen Domains, nur einmal erwähnt. Gleichzeitig werden Artikel von *welt.de* 380-mal zitiert. Dies ist ein typisches Muster für Aufmerksamkeitsverteilung, wie zum Beispiel Zitationen oder Aufrufe von Internetseiten und deckt sich auch ersten Ergebnissen zu Quellenangaben in KI-generierten Antworten.¹³

Abbildung 2: Power-Law-Verteilung der zitierten Domains



Gesamt: n=544 eindeutige Domains

Dass sich dieses Muster auch bei KI-basierter Zitation finden lässt, ist aus zwei Gründen relevant. Erstens zeigen Studien, dass KI-Modelle Informationen aus dem Long-Tail schwerer erlernen.¹⁴ Die entsprechenden Informationen werden Menschen, die KI-Systeme zur Informations- und Nachrichtensuche verwenden, also bei Nutzung von KI-Systemen als Informations- und Nachrichtenquelle, nicht nur seltener angezeigt, es steigt auch die Wahrscheinlichkeit, dass die KI-Systeme diese Informationen fehlerhaft wiedergeben. Zweitens zeigt es, dass KI-Systeme bestehende Aufmerksamkeitsverteilungen nicht aufbrechen, sondern reproduzieren und damit verstärken.

welt.de ist mit
380
Zitationen die
meistzitierte Domain.

Google AI Overview zitiert im Schnitt fast doppelt so viele Quellen pro Anfrage wie Gemini.

Auch die Anzahl der Quellen pro Anfrage unterscheidet sich je nach KI-System deutlich (vgl. Abb. 3). Während Google AI Overview im Schnitt 10,1 Quellen pro Anfrage zitiert, sind es bei Ge-

¹² Clauzet et. al., 2009

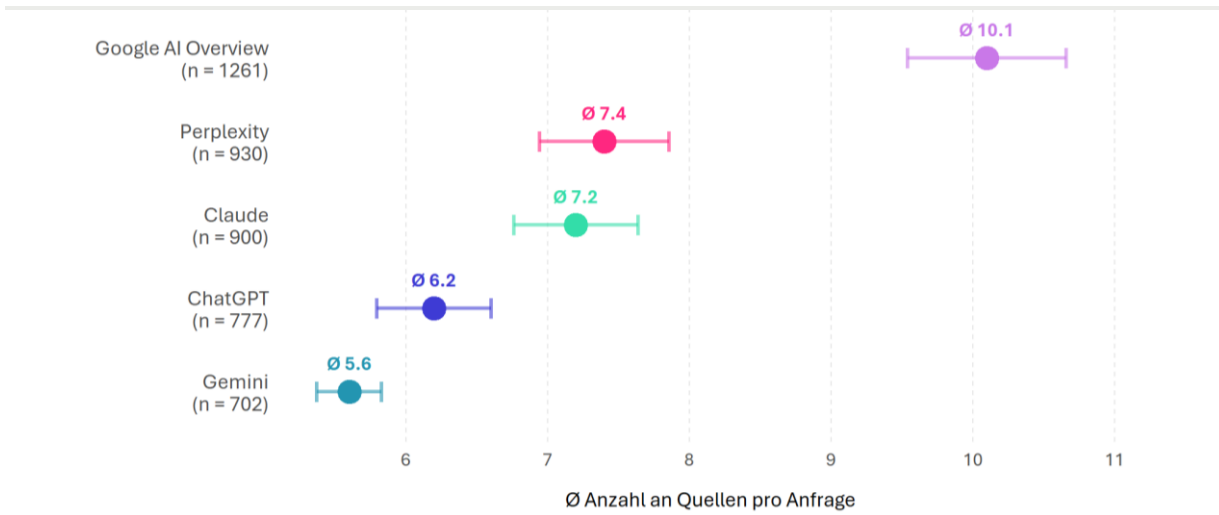
¹³ Algaba et. al., 2025

¹⁴ Kandpal et al., 2023



mini nur 5,6 Quellen. Perplexity und Claude liegen mit durchschnittlich 7,4 bzw. 7,2 Quellen pro Anfrage fast gleich auf und ChatGPT gibt im Schnitt 6,2 Quellen pro Anfrage an. Die durchschnittliche Anzahl der zitierten Quellen deckt sich in der Tendenz mit der Länge der Antworten der unterschiedlichen KI-Systeme (vgl. Abb.10 im Anhang).

Abbildung 3: Durchschnittliche Anzahl zitierter Links pro Anfrage je KI-System



Anmerkung. Die Punkte stellen die Mittelwerte dar, die Linien die Standardabweichungen.

Claude erzeugt mit durchschnittlich rund 3.279 Wörtern pro Anfrage die längsten Ausgaben. Dies ist jedoch wahrscheinlich auch dadurch bedingt, dass für Claude die Bezahlversion von Sonnet 4.6 verwendet wurde. Lässt man Claude aus diesem Grund kurz außen vor, zeigt sich, dass Google AI Overview mit durchschnittlich 3.002 Wörtern pro Anfrage nicht nur die meisten Quellen pro Anfrage, sondern auch die längsten Textantworten liefert. Diese Tendenz bleibt auch dann bestehen, wenn man sich die Unterschiede in der Länge der Antworttexte nach Prompt und nach Thema (vgl. Abb. 11 und 12 im Anhang) anschaut.

Allerdings ergeben sich auch bei Claude interessante Tendenzen, wenn man die Länge der Antworttexte nach Prompts (vgl. Abb. 11 im Anhang) und nach Thema (vgl. Abb. 12 im Anhang) aufschlüsselt. Hier zeigt sich, dass die Länge der Antworttexte bei Claude stark vom genutzten Prompt abhängt. Besonders die Rollenzuschreibung (Prompt 4) und die Kontextinformation (Prompt 5) führen zu deutlich längeren Antworttexten, die den Gesamtdurchschnitt nach oben ziehen. Die anderen Prompts hingegen liefern Antworttexte, die in der Länge mit den Antworten der kostenlosen Versionen der anderen KI-Systeme vergleichbar sind. Dies zeigt, dass die Bezahlversion von Claude nicht pauschal längere Antworten liefert als andere KI-Systeme, sondern dass die Länge der Antworten auch bei Claude vom genutzten Prompt abhängt. Insgesamt führt die Rollenzuschreibung (Prompt 4) über alle Modelle hinweg zu längeren Antworttexten, was sich mit bestehender Forschung zum Einfluss unterschiedlicher Prompt-Strategien auf die Antworten von KI-Systemen deckt.¹⁵

Es ist außerdem zu berücksichtigen, dass alle verwendeten Prompts die KI-Systeme explizit dazu auffordern, die verwendeten Quellen anzugeben. Da der Fokus dieser Analyse ausschließlich auf den verwendeten Quellen und nicht auf der inhaltlichen Richtigkeit der KI-Antworten liegt, ist an dieser Stelle jedoch keine Aussage darüber möglich, ob die von den KI-Systemen ausgegebenen

¹⁵ Kong et. al., 2024

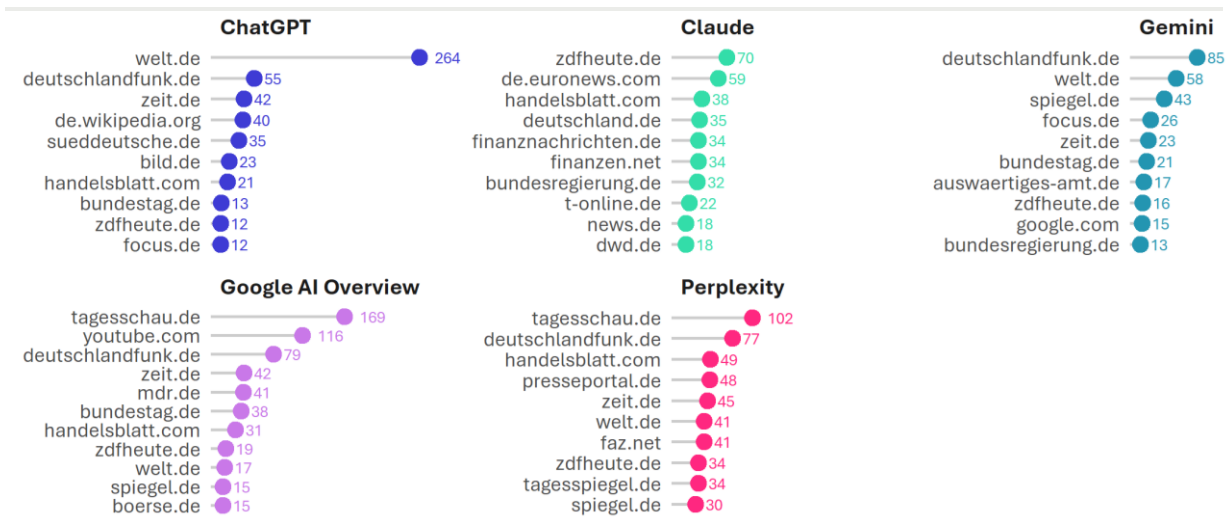


Antworten inhaltlich tatsächlich zu den angegebenen Quellen passen. Es ist auch möglich, dass Teile der KI-generierten Antworten und/oder die angegebenen Quellen halluziniert wurden.

Journalistische Quellen dominieren bei allen KI-Systemen, dennoch weisen verschiedene Systeme spezifische Quellenprofile auf.

Ein Blick auf die zehn meistzitierten Domains pro KI-System (vgl. Abb. 4) zeigt, dass die Systeme unterschiedliche Quellenprofile aufweisen. Zunächst fällt auf, dass bei ChatGPT eine einzige Quelle die Zitationen in einem Ausmaß dominiert, wie es bei keinem anderen KI-System der Fall ist: *welt.de* erhält mit 264 Zitationen mit deutlichem Abstand die größte Sichtbarkeit in der Quellenauswahl von ChatGPT. Dies erklärt sich durch die Kooperation zwischen Axel Springer und Open AI, die seit 2023 existiert und Open AI Zugriff auf die Medienhalte der Springer Gruppe ermöglicht. Auch *bild.de* findet sich unter den zehn meistzitierten Quellen bei ChatGPT.

Abbildung 4: 10 meistzitierte Domains pro KI-Anbieter



N = 4.570 Zitationen von 544 individuellen Domains, absolute Zahlen.

Bei allen anderen analysierten KI-Systemen werden Domains von öffentlich-rechtlichen Medienangeboten am häufigsten zitiert. Bei Perplexity und Google AI Overview befindet sich *tagesschau.de* auf dem ersten Platz der meistzitierten Quellen, bei Claude ist es *zdfheute.de* und bei Gemini *deutschlandfunk.de*. Generell finden sich bei allen KI-Systemen diverse Medienangebote öffentlich-rechtlicher Anbieter unter den meistzitierten Domains, sie stellen also zumindest im Kontext von Nachrichtenfragen eine zentrale Informationsquelle für KI-Systeme dar.

In einer internationalen Studie der European Broadcasting Union (EBU) und der BBC¹⁶ aus dem Jahr 2025 bewerteten 271 Journalist:innen aus 22 öffentlich-rechtlichen Medienorganisationen in 18 Ländern insgesamt 2.709 Antworten der vier KI-Systeme ChatGPT, Copilot, Perplexity und Gemini auf aktuelle Nachrichtenfragen. Insgesamt wiesen 45 Prozent aller Antworten mindestens ein gravierendes Problem auf; 81 Prozent der Antworten irgendein Problem. Das größte Einzelproblem war dabei die Quellenangabe (31 Prozent der Antworten). Besonders auffällig war auch dort Gemini, das mit 72 Prozent der Antworten die höchste Rate an Quellenproblemen zeigt, weit vor ChatGPT (24 Prozent) sowie Copilot und Perplexity (je 15 Prozent). Zudem zitieren KI-Sys-

¹⁶ Fletcher & Verckist, 2025



teme häufig unzuverlässige oder ungeeignete Quellen, darunter Satire, Unternehmenswebseiten und Social-Media-Inhalte, ohne diese als solche kenntlich zu machen.

KI-Systeme reproduzieren Staatsnarrative, Desinformation und Verschwörungserzählungen.

Ein genauerer Blick in die weniger prominenten, exklusiven Domains im Long-Tail der KI-Systeme zeigt, dass die Informationen, die KI-Systeme bei Anfragen zu tagesaktuellen Nachrichten referenzieren, nicht nur von vertrauenswürdigen Quellen mit journalistischen Qualitätsansprüchen stammen. So finden sich bei Claude gleich drei Domains des vom Digital Forensic Research Lab (DFRLab) als russisches Desinformationsnetzwerk eingestuften Pravda-Netzwerks (*austria.news-pravda.com*, *deutsch.news-pravda.com*, *germany.news-pravda.com*), auf die das Modell insgesamt siebenmal verweist. Diese Verweise stehen in den Antworten der KI-Systeme ungekennzeichnet neben journalistischen Quellen, sodass Nutzende sie wahrscheinlich nicht als Desinformationen erkennen.

Darüber hinaus zitiert Claude mit Junge Freiheit, Tichys Einblick, Epoch Times und Nordkurier, einer Zeitung, die wegen ihrer Nähe zu AfD und Querdenken auffällt¹⁷, gleich mehrere Medien aus dem Rechtsaußen-Spektrum. Ebenfalls exklusiv bei Claude finden sich klimaskeptische Quellen wie Blackout News und ein klimaskeptischer WordPress-Blog (*wobleibtdieglobaleerwarming.wordpress.com*). Auch nicht-offizielle oder faktisch fehlerhafte Domains tauchen auf: Google AI Overview verlinkt auf *bundestagszusammenfasser.de*, ein privates Tool ohne offizielle Verbindung zum Bundestag und Claude auf *bundestagwahlumfrage.de*.

3.2 Journalismus, Zivilgesellschaft und Co.: Wen zitieren die KI-Systeme?

Neben der Frage nach der Anzahl der zitierten Quellen, ist auch ein detaillierter Blick darauf, wen KI-Systeme zitieren, wenn Menschen sie zur Nachrichtensuche verwenden, notwendig, um Aussagen über die Ausgeglichenheit oder Verzerrung der KI-Antworten geben zu können.

Die Quellenprofile der KI-Systeme unterscheiden sich deutlich. Die Wahl eines KI-Systems beeinflusst, welche Quellenbasis Nutzende für ihre Antworten erhalten.

Die untersuchten KI-Systemen weisen unterschiedliche Zusammensetzungen zitierter Akteursgruppen auf. Auffällig ist, dass der Anteil journalistischer Quellen insbesondere bei ChatGPT und Perplexity mit 77 Prozent bzw. 73 Prozent besonders hoch ist. Die Zitationen bei Gemini (49 Prozent), Claude (51 Prozent) und Google AI Overview (54 Prozent) weisen hingegen geringere Anteile journalistischer Quellen auf. Stattdessen zitieren diese Modelle häufiger staatliche Quellen wie z. B. die

3/4

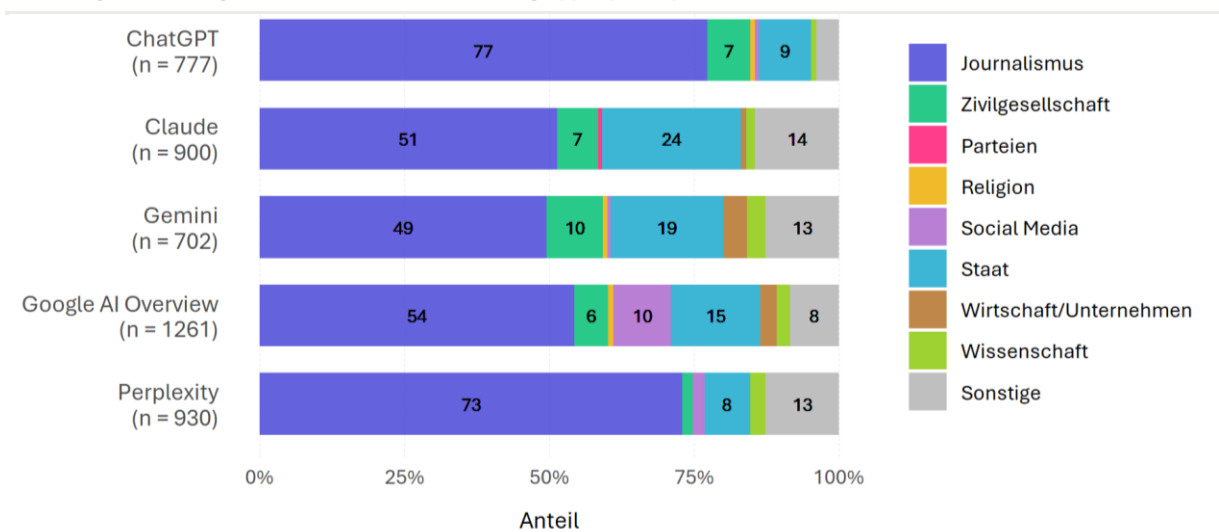
aller Zitationen bei ChatGPT und Perplexity verweisen auf journalistische Quellen.

¹⁷ „Nordkurier“, 2026

Seiten der Bundesregierung (*bundesregierung.de*), des Bundestags (*bundestag.de*) oder auch des Auswärtige Amts (*auswaertiges-amt.de*).

Quellen aus der Kategorie Religion (z. B. Bistum Regensburg, Deutsche Bischofskonferenz) werden ausschließlich von Googles KI-Systemen Gemini und Google AI Overview sowie von ChatGPT zitiert. Parteiquellen (z. B. CDU, Die Linke) finden sich bei allen KI-Systemen, werden jedoch besonders häufig von Claude zitiert. Social-Media-Plattformen als Quelle werden von Google AI Overview, Gemini und Perplexity zitiert. Der hohe Anteil an Social-Media-Plattformen bei Google AI Overview entfällt auf die prominente Nutzung von YouTube-Videos als Quelle. Mit 116 Zitationen mit YouTube¹⁸ eine konzerneigene Plattform bei Google AI Overview auf Platz 2 der meistzitierten Quellen.

Abbildung 5: Verteilung der zitierten Links nach Akteursgruppen je KI-System



Anmerkung: Sonstige enthält z. B. private Blogs, Aggregatorensiten und das russische Desinformationsnetzwerk „Pravda“. Wertebeschriftung ab 5 Prozent.

Diese Unterschiede in der Akteursstruktur haben auch inhaltliche Implikationen. Journalistische Quellen unterliegen redaktionellen Qualitätsstandards, während Inhalte aus der Kategorie „Sonstige“, die unter anderem private Blogs, Aggregatorensiten und das vom Digital Forensic Research Lab¹⁹ (DFRLab) identifizierte russische Desinformationsnetzwerk „Pravda“ umfasst, sowie Social-Media-Plattformen keiner vergleichbaren redaktionellen Kontrolle unterliegen.

Wie folgenreich diese Unterschiede sein können, zeigt die Debatte um sogenanntes LLM-Grooming, also die gezielte Überflutung des Netzes mit Falschmeldungen in der Absicht, dass KI-Modelle diese Inhalte aufnehmen und in ihren Antworten reproduzieren. Als prominentestes Beispiel dieser Strategie gilt das Pravda-Netzwerk.²⁰ Ein Audit von NewsGuard kam zu dem Ergebnis, dass führende Chatbots Falschbehauptungen aus diesem Netzwerk in rund einem Drittel der getesteten Fälle wiedergaben.²¹

¹⁸ Eine Übersicht der hinter den YouTube-Videos stehenden Produzent:innen der Inhalte findet sich in Abbildung 7.

¹⁹ Weitere Informationen zum DFRLab und Pravda-Netzwerk: <https://dfrlab.org/the-pravda-network/>

²⁰ American Sunlight Project, 2025

²¹ Sadeghi & Blachez, 2025

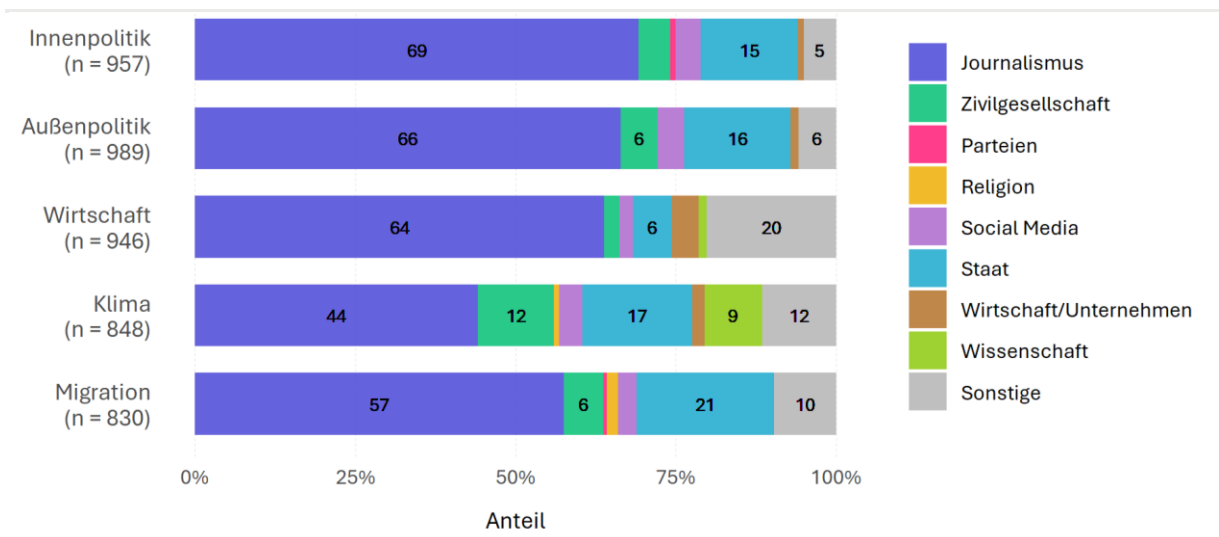


Auch hinsichtlich der Reichweite der journalistischen Quellen bestehen deutliche Unterschiede: Während bei ChatGPT und Gemini über 75 Prozent der journalistischen Zitationen auf nationale Medien entfallen, verteilen sich die Verweise bei Claude deutlich breiter. Mit einem Anteil von Special-Interest-Medien von 42 Prozent sowie einem Lokalmedien-Anteil von 15 Prozent werden dort verstärkt Nischen- und Regionalmedien herangezogen (vgl. Abb. 16 im Anhang).

Bei kontroversen Themen wie Migration und Klima werden weniger journalistische Quellen zitiert. Stattdessen verweisen die KI-Systeme verstärkt auf Wissenschaft, Zivilgesellschaft und Staat.

Die Verteilung der zitierten Akteursgruppen unterscheidet sich zusätzlich je nach Thema der Anfrage. Gerade bei den kontroversen Themen in der Analyse, also Klima und Migration, sinkt der Anteil journalistischer Quellen von 69 Prozent beim Thema Innenpolitik auf nur noch 44 Prozent beim Thema Klima. Auch das Thema Migration weist mit 57 Prozent einen geringeren Anteil journalistischer Quellen auf als die Themen Innenpolitik, Außenpolitik und Wirtschaft.

Abbildung 6: Verteilung der zitierten Links je Akteursgruppen nach Thema



Anmerkung: Basis sind alle zitierten Links der Analyse (N=4.570). Wertebeschriftungen ab 5 Prozent.

Statt journalistischer Quellen zitieren die KI-Systeme zu diesen kontroversen Themen häufiger zivilgesellschaftliche Akteure (z. B. Mediendienst Integration, Pro Asyl, Germanwatch), staatliche Quellen (z. B. Deutscher Bundestag, Bundesregierung, Auswärtiges Amt) und beim Thema Klima außerdem wissenschaftliche Quellen (z. B. Potsdam-Institut für Klimafolgenforschung, Deutsche IPCC-Koordinierungsstelle, Wuppertal Institut für Klima und Energie). Ein detaillierter Blick in die zitierten Domains zeigt etwa, dass das Potsdam-Institut für Klimaforschung zu diesem Thema von allen fünf KI-Systemen regelmäßig zitiert wird, am häufigsten von Perplexity (n=18), Google AI Overview (n=11) und Gemini (n=4).

Hinzu kommt, dass einzelne Domains, insbesondere etablierte journalistische Angebote, von einzelnen Modellen gar nicht zitiert werden. So finden sich Verweise auf die Frankfurter Allgemeine Zeitung (FAZ) nur bei Perplexity, Google AI Overview und Gemini, aber nicht bei ChatGPT und Claude. Die Süddeutsche Zeitung (SZ) hingegen wird nur von ChatGPT, Google AI Overview und Perplexity zitiert, jedoch nicht von Gemini und Claude.

Abbildung 7: Übersicht der Betreiber zitierter Social-Media-Accounts



Anmerkung. Basis sind alle Links von Social-Media-Plattformen (n=154). Die Größe der Wörter in der Wortwolke korrespondiert mit der Anzahl der Zitationen.

Google AI Overview, Gemini und Perplexity zitieren Inhalte von Social-Media-Plattformen als Quellen in ihren Antworten. Zur Beurteilung der Informationsqualität ist es in diesen Fällen notwendig, den hinter dem zitierten Account stehenden Akteur zu identifizieren. Neben journalistischen, wissenschaftlichen und staatlichen Akteuren mit hohem Qualitätsanspruch an die eigenen Inhalte können die zitierten Social-Media-Accounts auch von Privatpersonen, Influencern, Bots oder sonstigen Akteuren ohne Qualitätskontrolle stammen. Ein detaillierter Blick in die Profile hinter den zitierten Social-Media-Accounts (vgl. Abb. 7) zeigt jedoch, dass der Großteil zu denselben etablierten journalistischen Anbietern gehört, die auch insgesamt die meisten Zitationen auf sich vereinen.

3.3 Politische Ausrichtung der KI-Systeme bei der Auswahl journalistischer Quellen

Ein detaillierter Blick auf die politische Ausrichtung der journalistischen Quellen bestätigt den Eindruck, dass die verschiedenen KI-Systeme unterschiedliche Quellenprofile aufweisen. Die politische Ausrichtung journalistischer Quellen wurde in vier Kategorien erhoben: linke, liberale und wirtschaftsliberale Medien wurden zu einer Kategorie zusammengefasst, die eine politische Ausrichtung tendenziell links der Mitte widerspiegelt; konservative und rechte Medien wurden zu einer Kategorie zusammengefasst, die eine politische Ausrichtung tendenziell rechts der Mitte widerspiegelt; öffentlich-rechtliche und neutrale Medien wurden zu einer Kategorie zusammengefasst, die keine dezidiert politische Ausrichtung aufweist. Öffentlich-rechtliche Medien sind aufgrund ihres Auftrags einer insgesamt ausgeglichenen Berichterstattung in dieser Kategorie gelandet²². Andere Medien wurden nur dann dieser Kategorie zugeordnet, wenn sie explizit in ihrer Selbstbeschreibung angeben, dass sie keine politische Richtung verfolgen. Im Zweifel wurde die Kategorie unklar/sonstige gewählt. Diese Kategorie enthält auch Lokalmedien, für die keine politische Tendenz zu erkennen war, sowie Special-Interest-Medien, die sich an sehr kleine Zielgruppen und spezifische Interessen wenden (z. B. Fachmagazin für Güterverkehr).

²² MStV § 26



Die Einordnung der Domains zu den verschiedenen Ausprägungen beruht auf drei zentralen Ressourcen. Die wichtigste Quelle ist Eurotopics²³. Für alle dort gelisteten Medien wurde die Kategorisierung der politischen Ausrichtung übernommen. Die Seite wird von der Bundeszentrale für politische Bildung gefördert und klassifiziert reichweitenstarke Medien europäischer Länder aufgrund von Expert:inneneinschätzungen der entsprechenden Länder nach ihrer politischen Ausrichtung. Für dort nicht gelistete Medien wurde die Selbstbeschreibung des jeweiligen Mediums (von der entsprechenden Website) konsultiert. Wenn hier Hinweise auf eine politische Ausrichtung vorhanden waren, wurde diese übernommen. Wenn auch hier keine Informationen zur politischen Ausrichtung des Mediums gefunden wurden, wurde der Wikipedia-Eintrag des jeweiligen Mediums herangezogen. Wenn auch hier keine Informationen zur politischen Ausrichtung zur Verfügung standen, wurde das Medium als unklar/sonstiges kategorisiert. Grundlegend für die Kategorisierung der politischen Ausrichtung ist dabei jeweils nicht die politische Tendenz der zitierten Texte, sondern die Grundhaltung des hinter der jeweiligen Domain stehenden Akteurs.

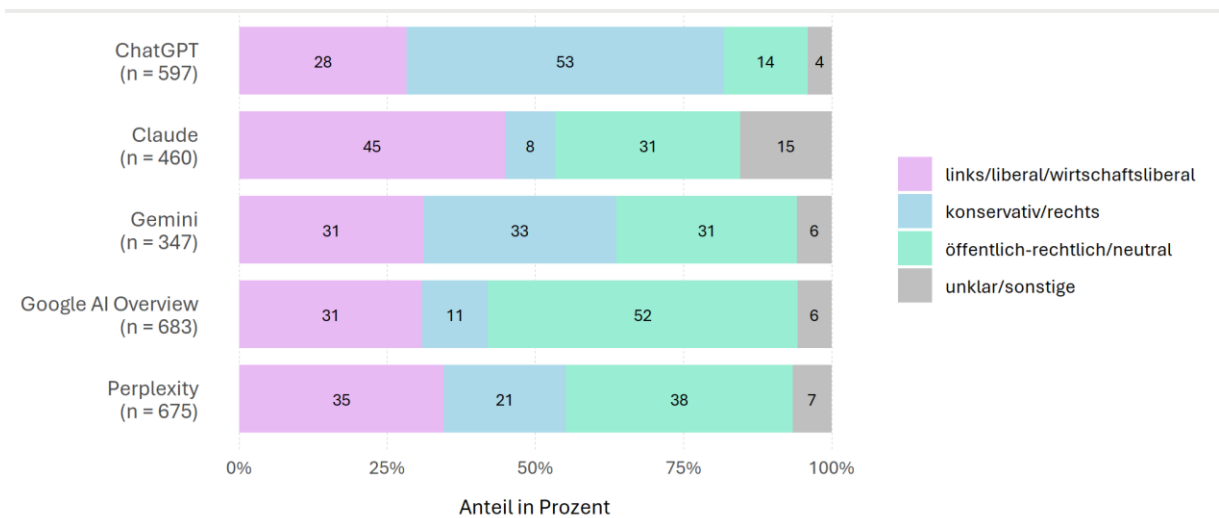
Die Kooperation zwischen Open AI und Axel Springer führt zu einer konservativ-verzerrten Quellenbasis bei ChatGPT.

53 %

der journalistischen Quellen bei ChatGPT decken das konservative bzw. rechte Meinungsspektrum ab.

Zunächst fällt der hohe Anteil an konservativen und rechten journalistischen Quellen bei ChatGPT auf (53 Prozent). Öffentlich-rechtliche (14 Prozent) sowie linke, liberale und wirtschaftsliberale Medien spielen bei ChatGPT hingegen eine vergleichsweise untergeordnete Rolle (vgl. Abb. 8).

Abbildung 8: Politische Ausrichtung journalistischer Quellen je KI-System



Anmerkung: unklar/sonstige enthält z. B. Lokalmedien und Special-Interest-Medien. Nur journalistische Quellen berücksichtigt.

Ein solch starke Tendenz zu einer politischen Richtung findet sich sonst bei keinem der untersuchten KI-Modelle. Google AI Overview zitiert besonders viele Links öffentlich-rechtlicher Medien (52 Prozent) und Claude weist einen hohen Anteil linker, liberaler und wirtschaftsliberaler Medien (45 Prozent) auf. Dieser hohe Anteil bei Claude erklärt sich allerdings bei einer detaillier-

²³ <https://www.eurotopics.net>

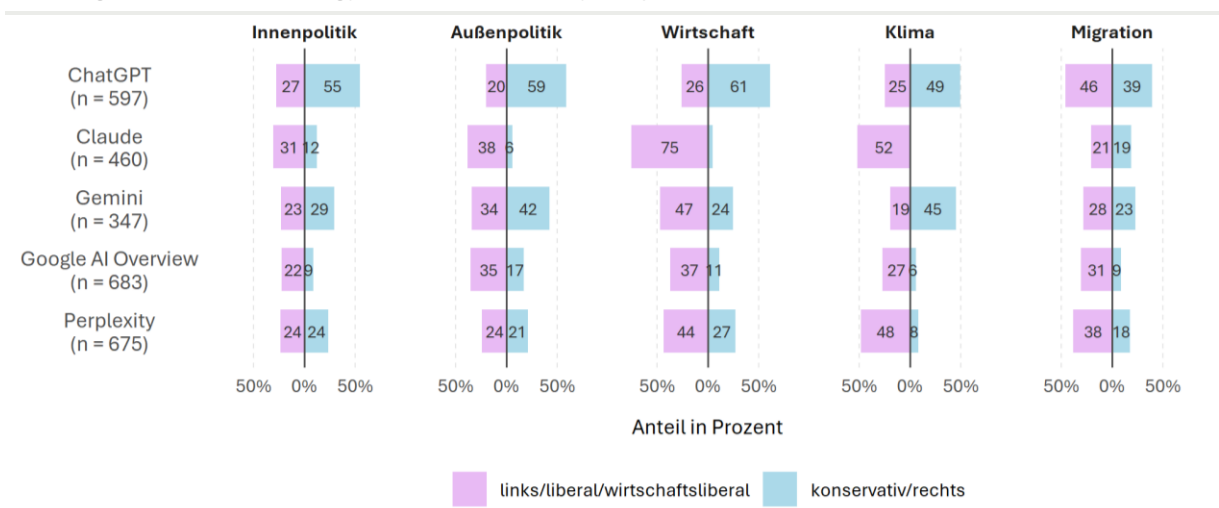
ten Betrachtung nach Thema (vgl. Abb. 9) durch einen thematisch passenden hohen Anteil wirtschaftsliberaler Medien beim Thema Wirtschaft.

Ein Blick auf die politischen Tendenzen der KI-Systeme in den verschiedenen Themenbereichen bestätigt die konservative Verzerrung bei ChatGPT. Lediglich beim Thema Migration werden etwas häufiger journalistische Quellen aus dem linken, liberalen und wirtschaftsliberalen Bereich zitiert (vgl. Abb. 9). Bei Claude bestätigt sich die Tendenz zu linken, liberalen und wirtschaftsliberalen journalistischen Quellen über alle untersuchten Themen hinweg. Besonders bei den Themen Wirtschaft und Klima dominieren diese Quellen das Quellenprofil des KI-Systems.

Bei manchen KI-Systemen variiert die politische Ausrichtung der zitierten journalistischen Quellen je nach Thema der Anfrage. So ergeben sich themenspezifische Varianzen der politischen Ausrichtung in den KI-Antworten, die für Nutzende nur schwer nachvollziehbar sind.

Claude hingegen zitiert häufiger Medien aus dem linken, liberalen und wirtschaftsliberalen Spektrum. Bei den Themen Klima (52 Prozent), Wirtschaft (75 Prozent) und Außenpolitik (38 Prozent) überwiegen Quellen links der Mitte. Bei Gemini zeigt sich besonders beim Thema Klima (45 Prozent) eine Tendenz zu konservativ-rechten Quellen. Auch bei den Themen Außen- und Innenpolitik zeigt sich jeweils eine leichte Tendenz zu konservativen Quellen. Während Gemini in der Gesamtübersicht zur politischen Ausrichtung journalistischer Quellen (Abb. 7) noch ausgeglichen wirkte, zeigen sich hier themenspezifische Ausschläge hin zu einer Bevorzugung konservativer und rechter Quellen. Google AI Overview und Perplexity hingegen tendieren über alle Themen hinweg zu linken, liberalen und wirtschaftsliberalen journalistischen Quellen.

Abbildung 9: Politische Ausrichtung journalistischer Quellen je KI-System und Thema



Anmerkung: Anteil journalistischer Quellen je KI-System, mit Tendenz links bzw. rechts der Mitte. Nur journalistische Quellen berücksichtigt. Ohne öffentlich-rechtliche Quellen und Sonstige.

Diese Beobachtungen bestätigen, dass die untersuchten KI-Systeme mit unterschiedlichen Quellenprofilen arbeiten. Die politische Ausrichtung der zitierten Quellen hängt dabei nicht allein von der Formulierung des Prompts, sondern auch vom Thema, zu dem Nutzende das jeweilige System befragen, ab. Die politische Ausrichtung der Quellen variiert somit nicht nur zwischen den Systemen, sondern innerhalb ein und desselben Systems von Themenfeld zu Themenfeld.



Für die Bewertung von KI-Systemen zur Nachrichtensuche ist dies zentral. Ein System wie Gemini, das in der Gesamtübersicht neutral erscheint, kann bei einzelnen, gesellschaftlich umstrittenen Themen wie Klima eine klare politische Ausrichtung entwickeln. Da diese thematischen Verschiebungen für Nutzende nicht sichtbar sind und sich auch nicht aus der allgemeinen Reputation eines Systems ableiten lassen, entsteht eine schwer wahrnehmbare Verzerrung in der Nachrichtenvermittlung durch KI-Systeme.

Exkurs: Mistral AI

Die erhobenen Daten von Mistral AI wurden aus der Hauptanalyse ausgeschlossen, da die kostenfreie Basisversion von Le-Chat (mittlerweile „Vibe“) nur eine begrenzte Anzahl an Anfragen mit einer integrierten Websuche zulässt. Durch das Limit von maximal 10 webbasierten Anfragen pro Tag konnten nicht alle im Untersuchungsdesign vorgesehenen Prompts abgefragt werden, sodass ein systematischer Vergleich von Mistral mit den übrigen untersuchten KI-Systemen nicht möglich ist.

Um dennoch einen Einblick in das Antwort- und Quellenverhalten von Mistral zu erhalten, wurde eine ergänzende Teilauswertung von den im Anhang dokumentierten fünf Prompts für die Themen Außenpolitik und Innenpolitik durchgeführt. Insgesamt wurden so für Mistral innerhalb des Datenerhebungszeitraums 217 Links von 23 Domains als Quellen erhoben. Die im Folgenden vorgestellten Ergebnisse sind daher nicht als vollständiger Vergleich mit den übrigen Modellen zu verstehen, sondern als zusätzliche Analyse, die erste Beobachtungen hinsichtlich der Quellenauswahl von Mistral liefern kann.

Besonders auffällig ist die geringe Anzahl an eindeutigen Quellen. In den insgesamt 50 Suchanfragen an Mistral wurden 83 verschiedene Quellen zitiert. Im Durchschnitt gab Mistral pro Anfrage etwa 4 Quellen an, teilweise wurde allerdings auch nur eine einzige Quelle zitiert. Dabei unterscheidet sich die durchschnittliche Quellenmenge auch erheblich zwischen den unterschiedlichen Prompts: Während mit dem Prompt 4, der die KI auffordert, als Journalist zu agieren, im Durchschnitt 7,3 Quellen als Grundlage genutzt werden, kommt Prompt 2 durchschnittlich nur auf 2,7 Quellen.

Auch für die Quellenvielfalt in den Antworten von Mistral scheint es also einen entscheidenden Unterschied zu machen, auf welche Art und Weise der Prompt formuliert wird.

Mistral bezieht sich fast ausschließlich auf journalistische Quellen (91 %), dabei vor allem auf öffentlich-rechtliche Outlets (55 %). Zu den meistzitierten Quellen zählen hier beispielsweise die Tagesschau oder Deutschlandfunk. Insgesamt zitiert Mistral nur 23 unterschiedliche Domains in den insgesamt 50 gestellten Anfragen, die angegebenen Beiträge stammen also vor allem im Vergleich mit anderen KI-Systemen aus einer sehr limitierten Menge von journalistischen Quellen.



4 Implikationen und Handlungsbedarfe mit Blick auf KI-Anbieter und Journalismus

Zusammenfassend zeigt sich, dass die Quellenzusammensetzung, die KI-Systeme für Antworten auf Anfragen heranziehen, von mehreren Parametern beeinflusst wird: von der Formulierung des Prompts, vom Thema der Anfrage, vom genutzten KI-System, von Kooperationen mit Medienhäusern und anderen Akteuren sowie von den Trainingsdaten. Das macht die Quellenbasis von KI-Antworten in der Informations- und Nachrichtensuche für Nutzende in der Gesamtheit unvorhersehbar. Keiner der untersuchten Anbieter dokumentiert die Kriterien seiner Quellenauswahl, sie bleiben eine Blackbox. Darüber hinaus schöpfen die analysierten KI-Systeme Großteils aus einem begrenzten Quellenpool etablierter journalistischer Anbieter, was insbesondere kleine Medienhäuser ökonomisch unter Druck setzt.

KI-Systeme sind damit zu Gatekeepern in der Nachrichtenvermittlung geworden. Sie schöpfen aus einer engen Quellenbasis, verstärken bestehende Aufmerksamkeitsasymmetrien, verschieben ihre politische Ausrichtung je nach Thema und greifen auf Desinformationsnetzwerke als Quellen zurück. All dies geschieht in einer für Nutzende nicht nachvollziehbaren Weise. Anders als journalistische Akteure als Gatekeeper des klassischen Mediensystems unterliegen sie dabei weder publizistischen Sorgfaltspflichten noch einer vielfaltssichernden Aufsicht. Die Frage, nach welchen Kriterien KI-Systeme Quellen auswählen, ist deshalb keine technische Detailfrage, sondern eine Frage der demokratischen Funktion von Öffentlichkeit.

Die bestehende Regulierung beantwortet diese Frage nicht. Der Medienstaatsvertrag enthält Transparenz- und Diskriminierungsregeln für Medienintermediäre, der Digital Services Act (DSA) systemische Risikopflichten, die KI-Verordnung (AI Act) Kennzeichnungspflichten und Trainingsdatentransparenz, und das Landgericht München I hat eine erste haftungsrechtliche Einordnung vorgenommen. Auch der Digital Markets Act (DMA) und der European Media Freedom Act (EMFA) bieten grundsätzliche Ansatzpunkte. Diese Instrumente sind Großteils jedoch vor generativer KI entstanden. Sie regulieren das Auffinden, Aggregieren und Verbreiten fremder Inhalte, nicht deren Synthese zu einer eigenständigen Antwort. Die Intransparenz der Quellenauswahl im laufenden Betrieb von KI-Systemen erfassen sie daher weder konsequent noch lückenlos, und ob sie auf KI-Systeme überhaupt anwendbar sind, ist vielfach ungeklärt.

4.1 Transparenz: Quellenauswahl von KI-Systemen als Blackbox

Die Ergebnisse zeigen eine doppelte Verengung der Quellenauswahl. Zum einen schöpfen alle untersuchten Systeme aus einer vergleichsweise schmalen Quellenbasis. Die Verteilung der zitierten Domains folgt einer Power-Law-Verteilung, bei der wenige Leitmedien regelmäßig zitiert werden, während fast die Hälfte der 544 Domains nur ein einziges Mal vorkommt und im Long-Tail verschwindet. Zum anderen variiert die politische Ausrichtung der herangezogenen journalistischen Quellen nicht nur zwischen den Systemen, sondern innerhalb desselben Systems von Themenfeld zu Themenfeld.

Beide Befunde decken sich mit dem internationalen Forschungsstand. Untersuchungen zur Quellenauswahl algorithmischer Nachrichtenintermediäre zeigen eine ausgeprägte Tendenz zur Homogenisierung des Quellenkorpus und eine Dominanz weniger reichweitenstarker Medien-



marken.²⁴ Für KI-gestützte Suchchatbots ist zudem eine Bevorzugung angloamerikanischer Quellen und eine uneinheitliche Quellenkennzeichnung dokumentiert.²⁵ Die hier vorgelegten Ergebnisse fügen diesem Bild eine bislang wenig beachtete Dimension hinzu, nämlich die themenspezifische politische Verschiebung innerhalb ein und desselben Systems. Für Nutzende ist daher nicht nachvollziehbar, auf welchen Kriterien die Quellenauswahl der KI-Anbieter basiert. Anders als bei klassischen Suchmaschinen, deren Rankinglogiken zumindest in Grundzügen bekannt sind (z. B. PageRank von Google), agieren KI-Modelle als intransparente Vermittlungsinstanz.

Die Auswahl- und Gewichtungsschritte zwischen Anfrage und Antwort sind bei allen untersuchten KI-Anbietern weder dokumentiert noch reproduzierbar.

Hinzu kommt, dass die Quellenangaben selbst unzuverlässig sind. Die internationale Studie der European Broadcasting Union und der BBC zeigt, dass die Quellenangabe das häufigste Einzelproblem KI-generierter Nachrichtenantworten ist.²⁶ Eine im Auftrag der New York Times durchgeführte Analyse der Google AI Overview kam darüber hinaus zu dem Ergebnis, dass mehr als die Hälfte der inhaltlich zutreffenden Antworten durch die verlinkten Quellen gar nicht gedeckt war.²⁷ Die KI-Systeme trafen also eigenständige Aussagen, deren Herkunft sich nicht überprüfen ließ. Zu ähnlichen Ergebnissen kommt das Tow Center for Digital Journalism, das bei acht KI-Suchsystemen in mehr als der Hälfte der getesteten Fälle falsche oder nicht überprüfbare Quellenangaben dokumentierte.²⁸

Die vorliegende Erhebung lässt keine Aussage darüber zu, ob die ausgegebenen Antworten inhaltlich tatsächlich zu den angegebenen Quellen passen; dass Teile der Antworten oder der Quellenangaben halluziniert sein können, ist daher nicht auszuschließen. Diese Intransparenz ist umso problematischer, da KI-generierte Antworten bei der Verwendung zur Nachrichtensuche dieselbe Funktion erfüllen wie journalistische Information, ohne den entsprechenden Sorgfaltpflichten zu unterliegen. Während journalistisch-redaktionelle Angebote nach § 19 MStV zur Wahrung anerkannter journalistischer Grundsätze verpflichtet sind und Tatsachen vor ihrer Verbreitung mit der gebotenen Sorgfalt auf Wahrheit und Herkunft prüfen müssen, existiert für KI-generierte Nachrichtenzusammenfassungen kein vergleichbarer Standard.

Das Landgericht München I hat in seinem Urteil vom 28. Mai 2026 (Az. 26 O 869/26) erstmals festgehalten, dass KI-Anbieter für unwahre Tatsachenbehauptungen in einer KI-Übersicht haftet, weil die KI-Modelle eigenständige Inhalte erzeugen und nicht bloß fremde Inhalte weitergeben. Die bislang geltende, eingeschränkte Haftung von Suchmaschinen für Drittinhalte sei auf dieses generative Format nicht übertragbar, auch die Haftungsprivilegierung des Digital Services Act greife nicht. Das Urteil knüpft an eine Vorentscheidung des Landgerichts Frankfurt am Main (September 2025, Az. 2-06 O 271/25) an und ist noch nicht rechtskräftig. Haftungsrechtlich ist damit aber bereits anerkannt, dass es sich um zurechenbare Äußerungen handelt. Ohne Einblick in die

²⁴ Nechushtai et al., 2024

²⁵ Jaźwińska & Chandrasekar, 2025

²⁶ Fletcher & Verckist, 2025

²⁷ Mickle et al., 2026

²⁸ Jaźwińska & Chandrasekar, 2025



Quellenauswahl lassen sich jedoch weder Haftung noch Vielfalt noch faire Vergütung wirksam durchsetzen.

Problem- und Handlungsfelder

- **Die rechtliche Einordnung KI-gestützter Nachrichtensuche ist ungeklärt.** Ob große Sprachmodelle mit Suchfunktion als sehr große Online-Suchmaschinen im Sinne des Art. 33 DSA gelten und damit den Pflichten zur Bewertung und Minderung systemischer Risiken unterliegen, ist offen. Die Europäische Kommission prüft diese Einordnung für ChatGPT derzeit, nachdem dessen Suchfunktion den Schwellenwert von 45 Millionen monatlichen Nutzenden in der EU zuletzt deutlich überschritten hat.²⁹
- **Die bestehende Transparenzregulierung endet dort, wo die eigentliche Gatekeeping-Entscheidung fällt.** Die Transparenzpflichten des AI Act erfassen die Trainingsdaten, nicht aber die Auswahl und Gewichtung von Quellen im laufenden Betrieb der Modelle, einschließlich der Frage, inwiefern Kooperationen oder bezahlte Platzierungen die Auswahl beeinflussen. Genau an dieser Stelle entscheidet sich jedoch, welche Stimmen und Inhalte öffentlich sichtbar werden.
- **Es fehlt ein verbindlicher Standard für die Zuordnung von Aussagen und Quellen in KI-Antworten.** Solange Quellen nur pauschal angeführt werden, ohne dass erkennbar ist, welche Aussage sich auf welche Quelle stützt, bleiben Quellenangaben ohne Prüfwert. Überprüfbarkeit am Ort der Verwendung ist die Voraussetzung dafür, dass Nutzende, Aufsicht und Forschung KI-Antworten überhaupt kontrollieren können.
- **Der Umgang mit Desinformation ist ungeregelt.** Dass der Ausschluss von Desinformations- und Propagandanetzwerken nicht durchgängig gewährleistet ist, zeigt der Befund dieser Studie, wonach Claude im Untersuchungszeitraum siebenmal auf Domains des Pravda-Netzwerks verwies, das vom Digital Forensic Research Lab als russisches Desinformationsnetzwerk eingestuft wird. Diese Verweise stehen ungekennzeichnet neben journalistischen Quellen. Ein System, das Desinformation und Journalismus in derselben Form präsentiert, gefährdet die informierte Meinungsbildung der Gesellschaft.

4.2 Vielfalt: Das Dilemma des Journalismus

Die Frage nach der Quellenvielfalt berührt zwei Dimensionen. Im Kontext von Nachrichtenfragen betrifft sie zum einen die inhaltliche Substanz dessen, was Nutzer:innen über das Tagesgeschehen erfahren, zum anderen verändert sie die ökonomischen Bedingungen des Journalismus im Digitalen. Dies ergibt sich unter anderem aus dem Phänomen der Zero-Click-Suche. Wenn Nutzer:innen von KI-Systemen sich zunehmend auf KI-generierte Texte und Zusammenfassungen verlassen, ohne die angegebenen Quellen aufzurufen, schwindet die Reichweite journalistischer Medien und mit ihr deren wirtschaftliche Grundlage, obwohl es gerade die journalistischen Inhalte sind, die KI-Systeme als Nachrichtenquelle erst nutzbar machen und den Anbietern erhebliche Erlöse ermöglichen.

²⁹ Reuters, 2026



Medien verlieren damit den direkten Kontakt zu ihrem Publikum und die Möglichkeit, sich über diesen Weg im Digitalen zu refinanzieren. Verliert der Journalismus auf diese Weise seine ökonomische Basis, steht zugleich eine zentrale Säule demokratischer Informationsvermittlung auf dem Spiel.

Dieser Effekt ist bereits messbar. Erscheint eine KI-Übersicht, klickt nur noch ein geringer Anteil der Nutzenden auf einen weiterführenden Link, und auch die darunter platzierten organischen Treffer büßen an Reichweite ein.³⁰ Für Deutschland wird seit der Integration von KI-Übersichten in die Google-Suche ein Rückgang von 265 Millionen organischen Klicks pro Monat berichtet.³¹

Die Ergebnisse dieser Studie bestätigen dieses Bild. Die Konzentration auf wenige Leitmedien bedeutet, dass der ohnehin geringe Teil der Aufrufe, der über KI-Systeme noch auf journalistische Onlineangebote zurückfließt, sich zusätzlich auf nur wenige Medien verteilt. Allein *welt.de* wurde 380-mal zitiert, während fast die Hälfte aller Domains nur ein einziges Mal vorkam. Diese Verteilung ist nicht allein ein technisches Resultat, sondern auch Ausdruck eines neuen Aushandlungsverhältnisses zwischen Medienanbietern und KI-Plattformen.

Verlage stehen vor einem doppelten Dilemma. Werden sie nicht zitiert, verlieren sie Sichtbarkeit; werden sie zitiert, ohne dass Traffic auf ihre Seiten zurückfließt, verlieren sie ihre ökonomische Grundlage.

Lizenzabkommen zwischen Medienverlagen und KI-Anbietern können diese Spannung für einzelne Akteure entschärfen, verstärken die Konzentration aber, anstatt sie aufzulösen. Wer eine Kooperation eingeht, wird sichtbarer; wer keine hat, verschwindet tendenziell aus dem zitierten Korpus.

Für die öffentliche Meinungsbildung ist das relevant, weil Vielfalt nicht erst auf der Ebene einzelner Artikel entsteht, sondern bereits durch die Auswahl der zitierten Stimmen vorstrukturiert wird. KI-Modelle übernehmen damit eine Gatekeeping-Funktion, ohne den publizistischen Sorgfaltspflichten zu unterliegen, die für klassische Gatekeeper gelten. Der europäische Gesetzgeber hat die Bedeutung von Medienvielfalt und die Sicherung pluralistischer Informationsangebote zuletzt im European Media Freedom Act (EMFA) ausdrücklich betont; KI-vermittelte Quellenauswahl adressiert dieser bislang jedoch nicht.

Problem- und Handlungsfelder:

- **Es fehlt ein Vergütungsmodell für journalistische Anbieter beim KI-Zugriff auf journalistische Inhalte.** Die gegenwärtige Logik bilateraler Lizenzabkommen honoriert Marktmacht statt publizistischer Leistung und vertieft die Konzentration, die sie zu lösen vorgibt. Zu diskutieren sind strukturelle Vergütungsansätze, die journalisti-

³⁰ Chapekis & Lieb, 2025

³¹ Sistrich, 2026



sche Inhalte unabhängig von Verhandlungsmacht einbeziehen, etwa in Form pauschaler, nutzungsbezogener Abgaben.

- **Vielfaltssicherung muss auch unter den Bedingungen KI-vermittelter Nachrichtennutzung institutionell gewährleistet sein.** Denkbar sind sowohl die Erstreckung bestehender Vielfalts- und Intermediärsregulierung auf KI-Systeme als auch der Aufbau eigener, journalistisch oder öffentlich-rechtlich getragener Infrastrukturen für die KI-gestützte Nachrichtensuche. Beide Wege setzen politischen Gestaltungswillen und erhebliche Ressourcen voraus. Sie zu prüfen ist allerdings unumgänglich, wenn die Auswahl der öffentlich sichtbaren Stimmen nicht dauerhaft einer intransparenten, kommerziell motivierten Logik überlassen bleiben soll.



Literaturverzeichnis

- Algaba, A., Mazijn, C., Holst, V., Tori, F., Wenmackers, S., & Ginis, V. (2025). Large language models reflect human citation patterns with a heightened citation bias. *Findings of the Association for Computational Linguistics: NAACL 2025* (S. 6844–6879). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2025.findings-naacl.381>
- American Sunlight Project. (2025, Februar). *A pro-Russia content network foreshadows the automated future of info ops*. A Pro-Russia Content Network Foreshadows. <https://www.americansunlight.org/>
- Baack, S., Djefal, C., Jarke, J., & Send, H. (2019). Civic Tech: Ein Beispiel für Bürgerzentrierung und Bürgerbeteiligung als Leitbild der Verwaltungsdigitalisierung. In T. Klenk, F. Nullmeier, & G. Wewer (Hrsg.), *Handbuch Digitalisierung in Staat und Verwaltung* (S. 1–9). Springer Fachmedien Wiesbaden.
- Beus, J. (2026, 25. Februar). *AI Overviews in Deutschland: So stark sinken die Klickraten wirklich*. SISTRIX. <https://www.sistrix.de/news/ai-overviews-in-deutschland-so-stark-sinken-die-klickraten-wirklich/>
- Bundesministerium für Digitales und Staatsmodernisierung [BMDS] (2026, 15. Juni). *KI-Einsatz in Reden des Ministers*. <https://bmds.bund.de/aktuelles/aktuelle-meldungen/detail/ki-einsatz-in-reden-des-ministers>
- Chapekis, A., & Lieb, A. (2025, 22. Juli). *Google users are less likely to click on links when an AI summary appears in the results*. Pew Research Center. <https://www.pewresearch.org/short-reads/2025/07/22/google-users-are-less-likely-to-click-on-links-when-an-ai-summary-appears-in-the-results/>
- Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4), 661–703. <https://doi.org/10.1137/070710111>
- Fletcher, J., & Verckist, D. (2025). *News integrity in AI assistants: An international PSM study*. BBC & European Broadcasting Union. https://www.ebu.ch/files/live/sites/ebu/files/Publications/Intelligence/open/EBU-Intelligence-BBC_News_Integrity_in_AI_Assistants_Report_2025.pdf
- Heßler, E. (2026). KI-generierte Antworten bei der Informationssuche: Verbreitung und Wahrnehmung. K & Search-Studie 2025. *Media Perspektiven*, 11/2026, 1–14.
- Jaźwińska, K., & Chandrasekar, A. (2025, 6. März). *AI search has a citation problem*. Columbia Journalism Review. https://www.cjr.org/tow_center/we-compared-eight-ai-search-engines-theyre-all-bad-at-citing-news.php
- Kandpal, N., Deng, H., Roberts, A., Wallace, E., & Raffel, C. (2023). Large language models struggle to learn long-tail knowledge. In *Proceedings of the 40th International Conference on Machine Learning* (Vol. 202, S. 15696–15707). PMLR.
- Kong, A., Zhao, S., Chen, H., Li, Q., Qin, Y., Sun, R., Zhou, X., Wang, E., & Dong, X. (2024). Better zero-shot reasoning with role-play prompting. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)* (S. 4099–4113). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.naacl-long.228>



- Koopman, B., & Zuccon, G. (2023). Dr ChatGPT tell me what I want to hear: How different prompts impact health answer correctness. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing* (S. 15012–15022). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.emnlp-main.928>
- Koyuturk, C., Theophilou, E., Patania, S., Donabauer, G., Martinenghi, A., Antico, C., Telari, A., Testa, A., Buršić, S., Garzotto, F., Hernandez-Leo, D., Kruschwitz, U., Taibi, D., Amenta, S., Ruskov, M., & Ognibene, D. (2025). Understanding learner-LLM chatbot interactions and the impact of prompting guidelines. In A. I. Cristea, E. Walker, Y. Lu, O. C. Santos, & S. Isotani (Hrsg.), *Artificial intelligence in education* (Lecture Notes in Computer Science, Vol. 15878, S. 364–377). Springer Nature. https://doi.org/10.1007/978-3-031-98417-4_26
- Mickle, T., Metz, C., Freedman, D., Mondría Terol, T., & Collins, K. (2026, 7. April). How accurate are Google's A.I. Overviews? *The New York Times*. <https://www.nytimes.com/2026/04/07/technology/google-ai-overviews-accuracy.html>
- Nechushtai, E., Zamith, R., & Lewis, S. C. (2024). More of the same? Homogenization in news recommendations when users search on Google, YouTube, Facebook, and Twitter. *Mass Communication and Society*, 27(6), 1309–1335. <https://doi.org/10.1080/15205436.2023.2173609>.
- Newman, N. (2026). *Journalism and technology trends and predictions 2026*. Reuters Institute for the Study of Journalism.
- Nordkurier. (2026, 1. Juni). In *Wikipedia*. <https://de.wikipedia.org/w/index.php?ol-did=267587556&title=Nordkurier>
- OpenAI. (2023, 13. Dezember). *Partnership with Axel Springer to deepen beneficial use of AI in journalism*. <https://openai.com/index/axel-springer-partnership/>
- Reuters. (2025, 24. Oktober). *OpenAI faces tighter regulation under EU's Digital Services Act* (M. Williams & B. Meijer).
- Sadeghi, M., & Blachez, I. (2025, 6. März). *A well-funded Moscow-based global 'news' network has infected Western artificial intelligence tools worldwide with Russian propaganda*. NewsGuard. <https://www.newsguardtech.com/special-reports/moscow-based-global-news-network-infected-western-artificial-intelligence-russian-propaganda/>
- Winter, S., Rainer, A., & Reißmann, O. (2026, 11. Juni). *Thüringens Regierungschef: KI in Reden und Gastbeiträgen? Warum das für Mario Voigt zum Problem wird*. DER SPIEGEL, Hamburg. <https://www.spiegel.de/politik/mario-voigt-warum-das-thema-ki-fuer-thueringens-regierungschef-zum-problem-wird-a-ac59a889-8b8e-4230-896e-0f0c591e6b0e>
- Wintergerst, R. (2026, 28. April). *Künstliche Intelligenz: Der Blick der Deutschen auf die neue Technologie* [Präsentationsfolien]. Bitkom e.V. <https://www.bitkom.org/sites/main/files/2026-04/bitkom-praesentation-kuenstliche-intelligenz-2026-der-blick-der-deutschen-auf-die-neue-technologie.pdf>
- Xue, H., Oh, Y. J., Zhou, X., Zhang, X. & Oxley, B. (2026). *Users' prompting strategies and ChatGPT's contextual adaptation shape conversational information-seeking experiences*. *Scientific Reports*, 16, 12112. <https://doi.org/10.1038/s41598-026-42465-4>
- Yin, Z., Wang, H., Horio, K., Kawahara, D., & Sekine, S. (2024). Should we respect LLMs? A cross-lingual study on the influence of prompt politeness on LLM performance. *Proceedings of the Second Workshop on Social Influence in Conversations (SICon 2024)* (S. 9–35). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.sicon-1.2>

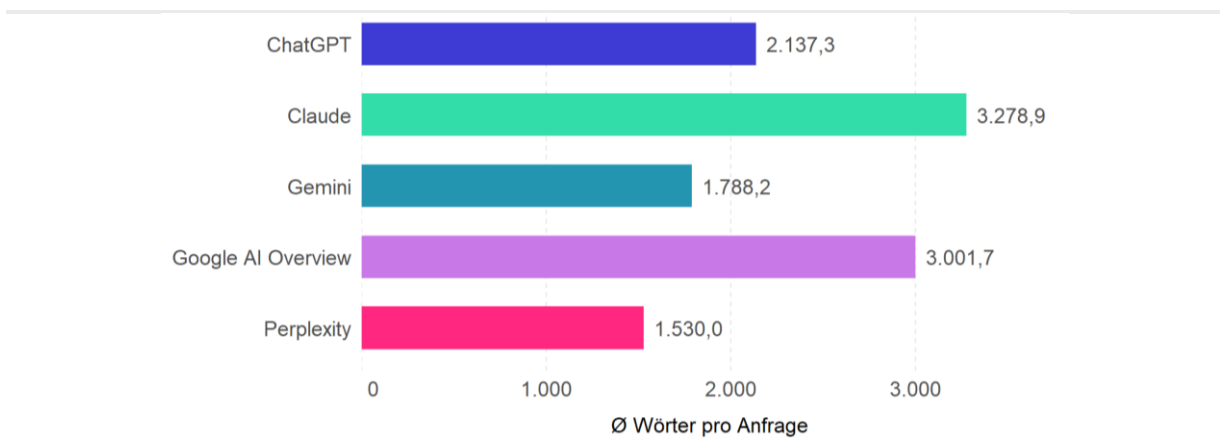


Anhang

Tabelle 1: Übersicht der verwendeten Prompts

Nr.	Prompt-Strategie	Erläuterung	Formulierung
1	Höflichkeitsfloskel	Studien zur Verwendung von Höflichkeitsfloskeln kommen zu unterschiedlichen Ergebnissen bezüglich der Antwortqualität. Dennoch ist diese Art des Prompting bei Nutzenden verbreitet. (z. B. Yin et al. 2024).	„Bitte gib mir die aktuellen Nachrichten zum Thema [Thema] für heute mit Quellen als Links.“
2	Frageformulierung	Analysen großer Mengen an Anfragen an KI-Systeme zeigen, dass Frageformulierungen zu den häufigsten Prompt-Strategien gehören (z. B. Xue et al., 2026). Studien zeigen, dass die Antwortqualität von der Frageformulierung abhängt (z. B. Koopman & Zuccon, 2023).	„Was ist heute passiert zum Thema [Thema]? Gib alle Quellen als Links an.“
3	Suchmaschinen-Stil	Insbesondere unerfahrene Nutzende schreiben Prompts häufig im Stil von Anfragen in Suchmaschinen. (z. B. Koyuturk et al. 2025)	„Aktuelle Nachrichten [Thema] heute mit Links als Quellen“
4	Rollenzuschreibung	Diverse Studien zeigen, dass Prompts mit Rollenzuschreibungen zu qualitativ hochwertigeren Ergebnissen führen. (z. B. Kong et al., 2025)	„Agiere als Journalist. Fasse alle Nachrichten zum Thema [Thema] am [Datum] zusammen. Gib alle Quellen als Links an.“
5	Kontextinformationen	Studien zeigen, dass Kontextinformationen die Antwortqualität von KI-Systemen verbessern. Dieser Effekt wird allerdings umgedreht, wenn die Prompts durch zu viel Kontext zu lang werden, da KI-Systeme dann dazu tendieren, sich auf die Informationen zu Beginn und am Ende des Prompts zu fokussieren. (z. B. Xue et al., 2026)	„Erstelle eine Übersicht aller Nachrichten vom [Datum] mit Bezug zum Thema [Thema] in Deutschland zusammen. Liste alle genutzten Quellen als Links auf.“

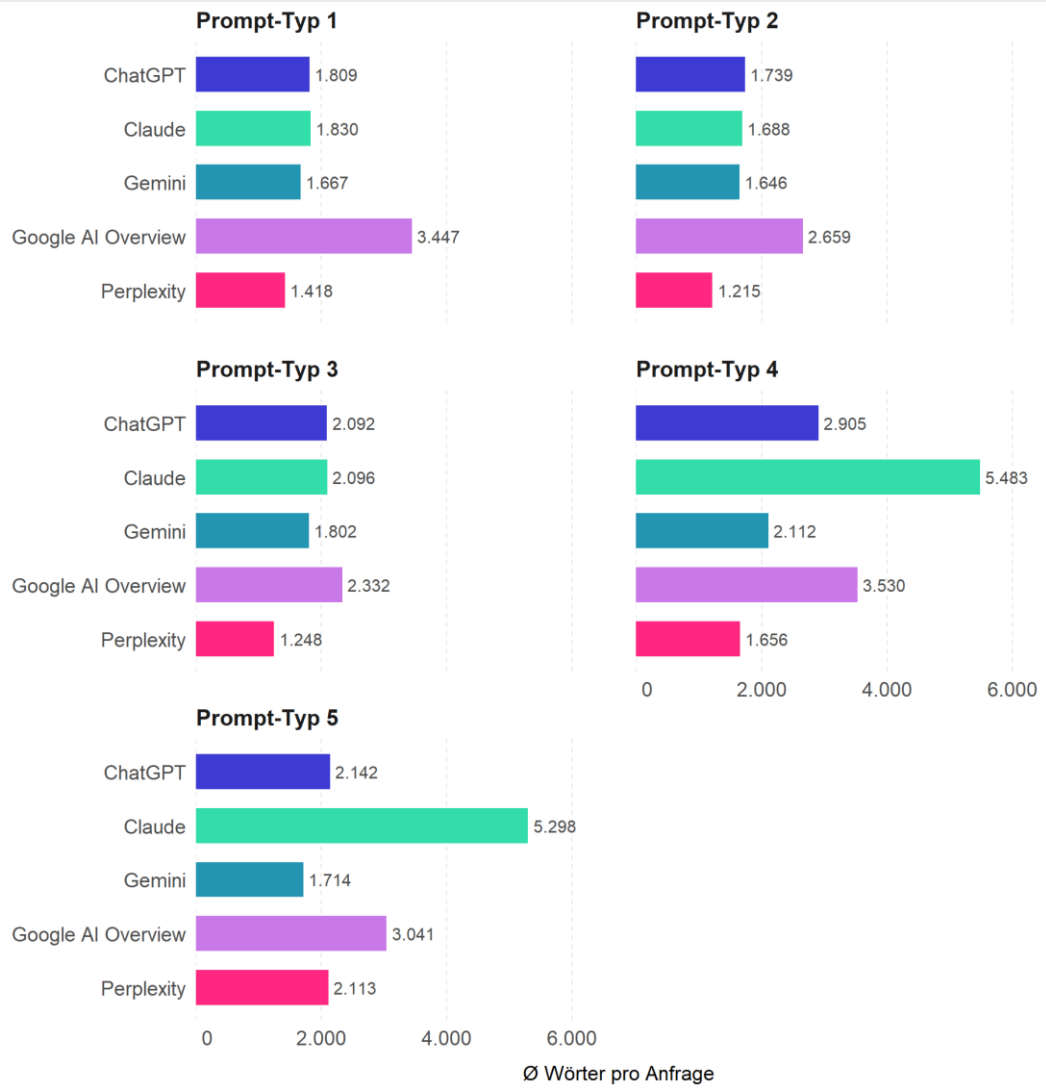
Abbildung 10: Durchschnittliche Wortanzahl pro Anfrage



Anmerkung. n=125 Anfragen pro KI-System. Bereinigte Wortzahl der vollständigen Antworttexte pro Anfrage: White Spaces und KI-typische Formatierungszeichen wurden entfernt. In Textantworten vorhandene Links werden als ein Wort gezählt.



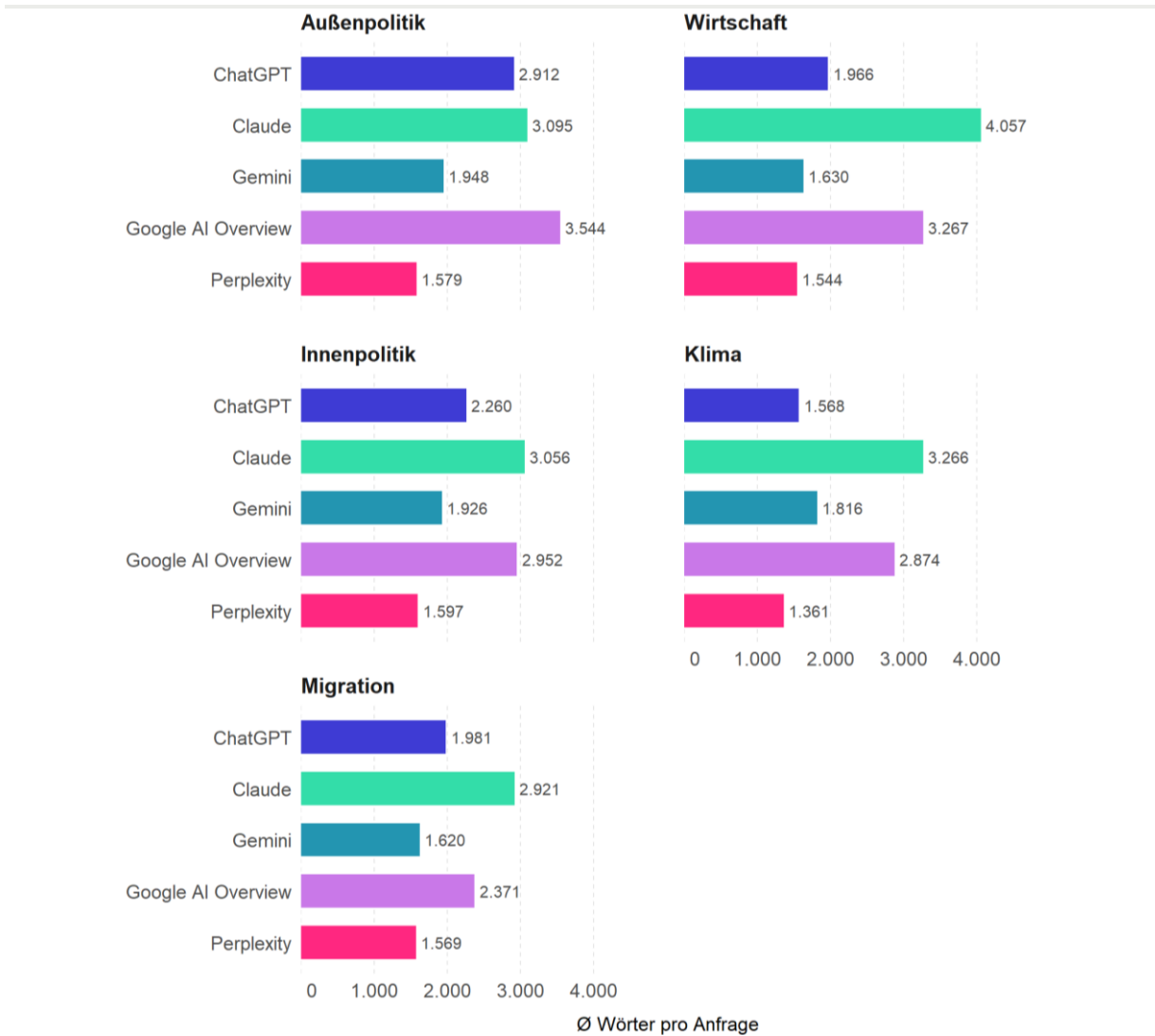
Abbildung 11: Durchschnittliche Wortanzahl pro Anfrage je KI-System und Prompt-Typ



Anmerkung. n=125 Anfragen pro KI-System. Bereinigte Wortzahl der vollständigen Antworttexte pro Anfrage: White Spaces und KI-typische Formatierungszeichen wurden entfernt. In Textantworten vorhandene Links werden als ein Wort gezählt.

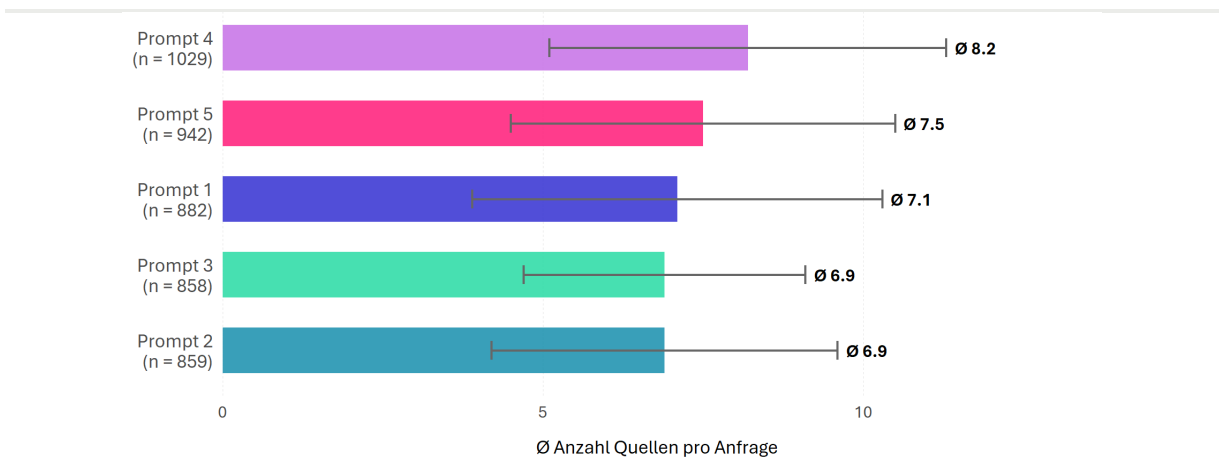


Abbildung 12: Durchschnittliche Wortanzahl pro Anfrage je KI-System und Thema



Anmerkung. n=125 Anfragen pro KI-System. Bereinigte Wortzahl der vollständigen Antworttexte pro Anfrage: White Spaces und KI-typische Formatierungszeichen wurden entfernt. In Textantworten vorhandene Links werden als ein Wort gezählt.

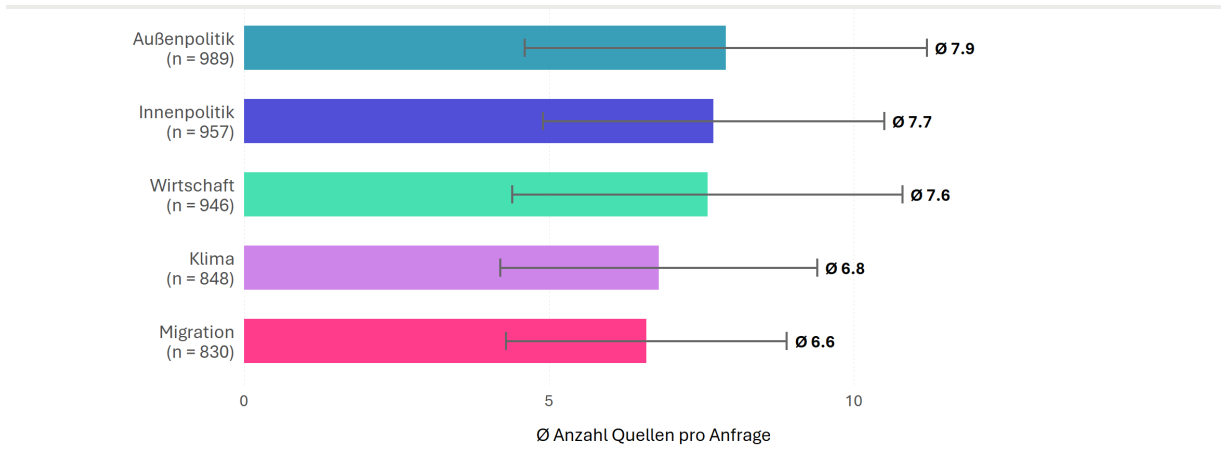
Abbildung 13: Durchschnittliche Anzahl zitiertter Links pro Anfrage je Promptformulierung



Anmerkung: Balken zeigen den Mittelwert, Fehlerbalken die Standardabweichung.

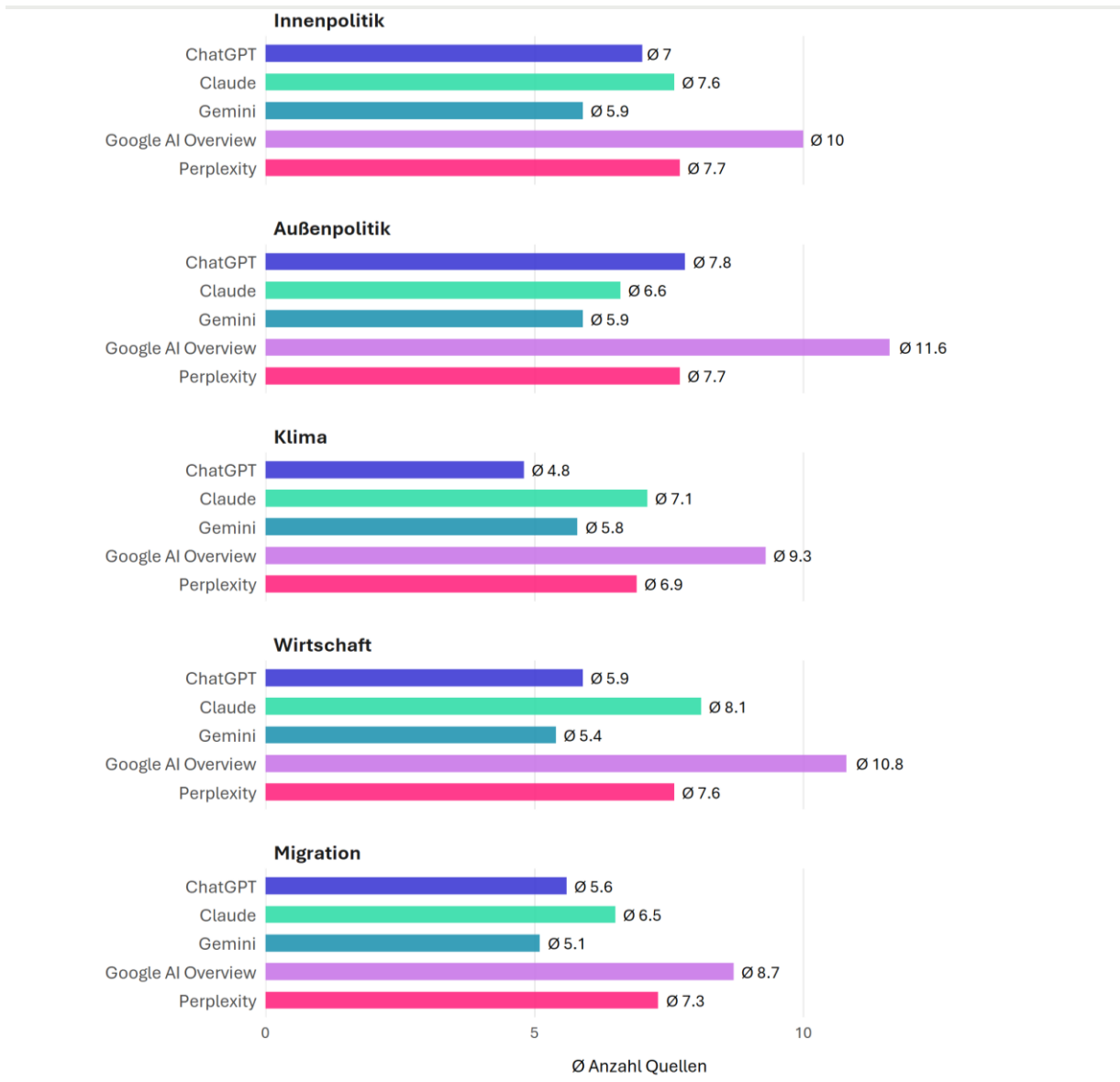


Abbildung 14: Durchschnittliche Anzahl zitierter Links pro Anfrage je Thema



Anmerkung: Balken zeigen den Mittelwert, Fehlerbalken die Standardabweichung.

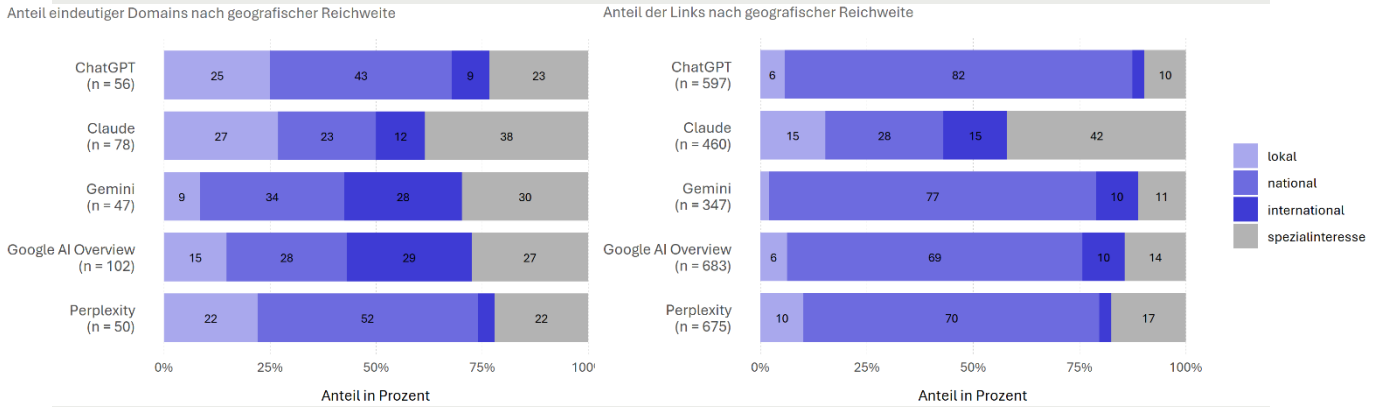
Abbildung 15: Durchschnittliche Anzahl zitierter Links pro Anfrage je KI-System und Thema



Gesamt-N: 4.570 Links von 544 Domains.

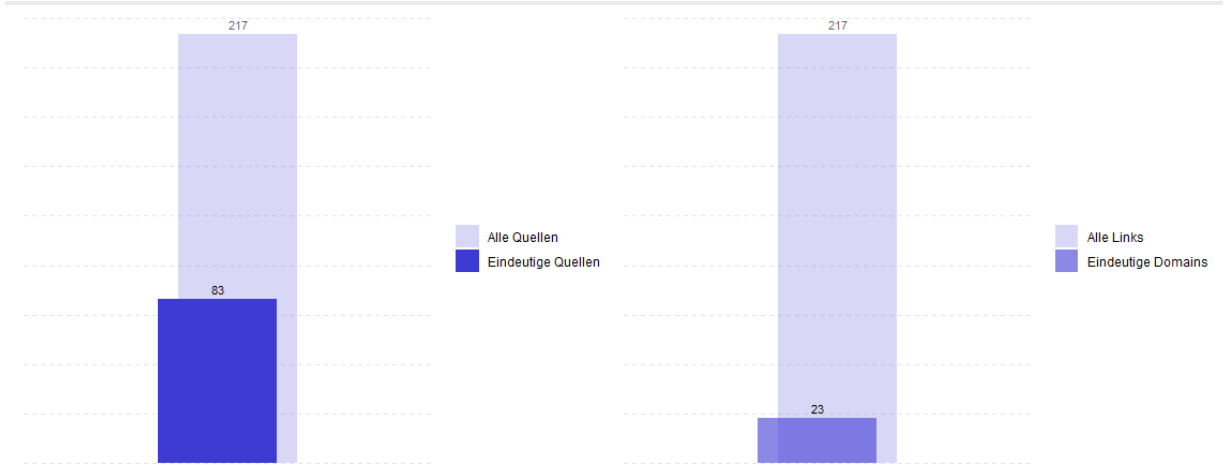


Abbildung 16: Scope der eindeutige Domains und Links journalistischer Quellen



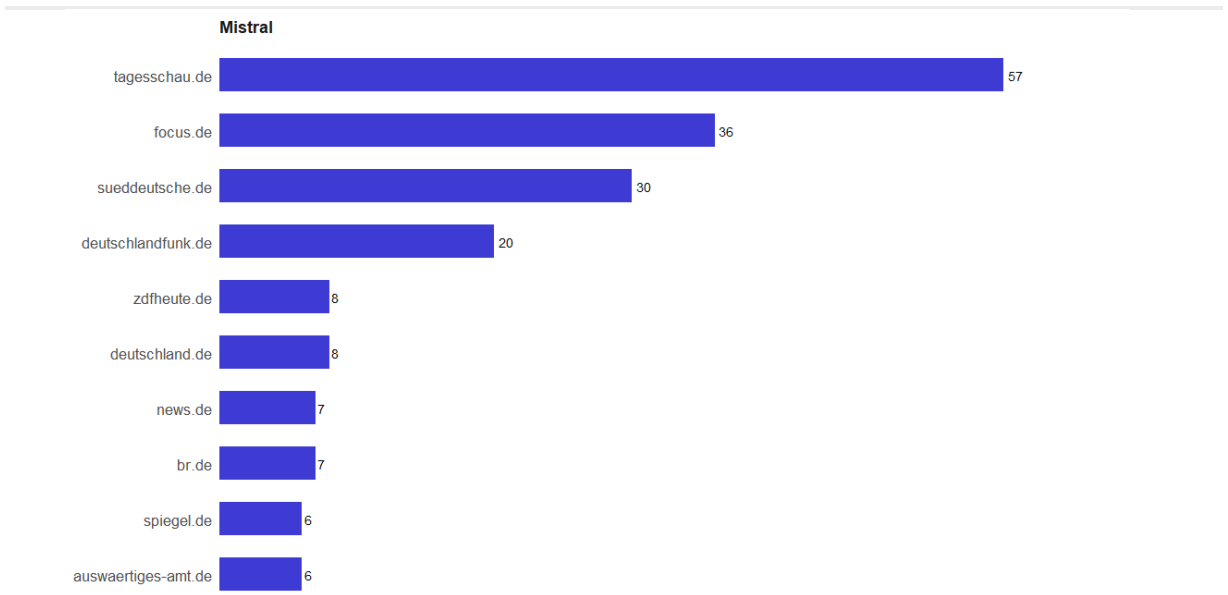
Anmerkung. Eindeutige Domains: Jede Domains wird nur einmal gezählt. Nur journalistische Quellen berücksichtigt.

Abbildung 17: Gesamtanzahl aller Links im Vergleich zu eindeutigen Quellen und eindeutigen Domains für Mistral



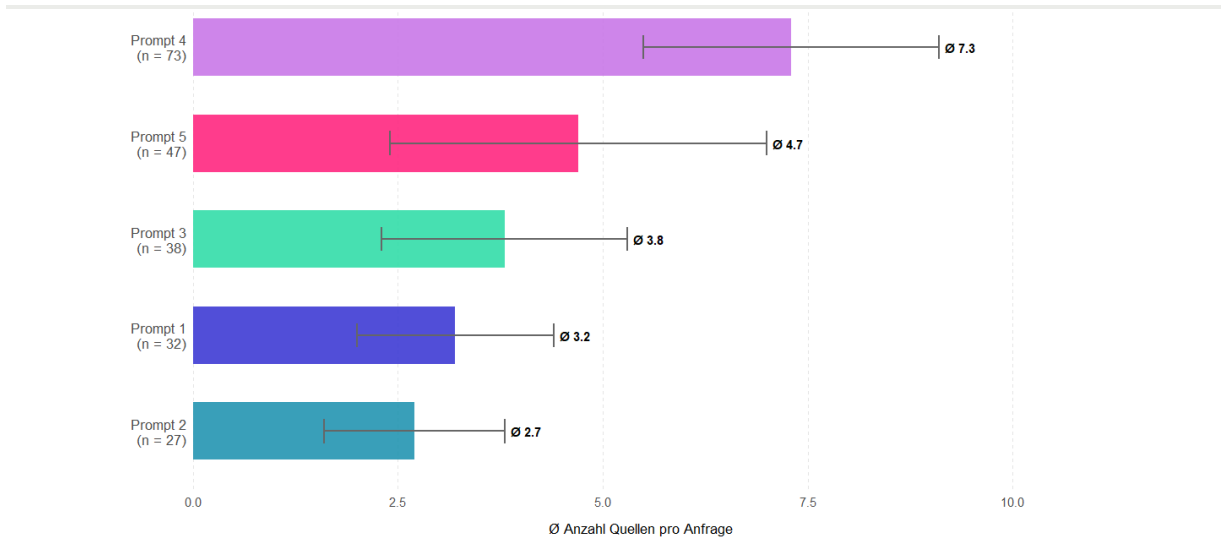
Anmerkung: Für Mistral wurden nur die Prompts zu den Themen Innen- und Außenpolitik abgefragt.

Abbildung 18: 10 meistzitierte Domains für Mistral



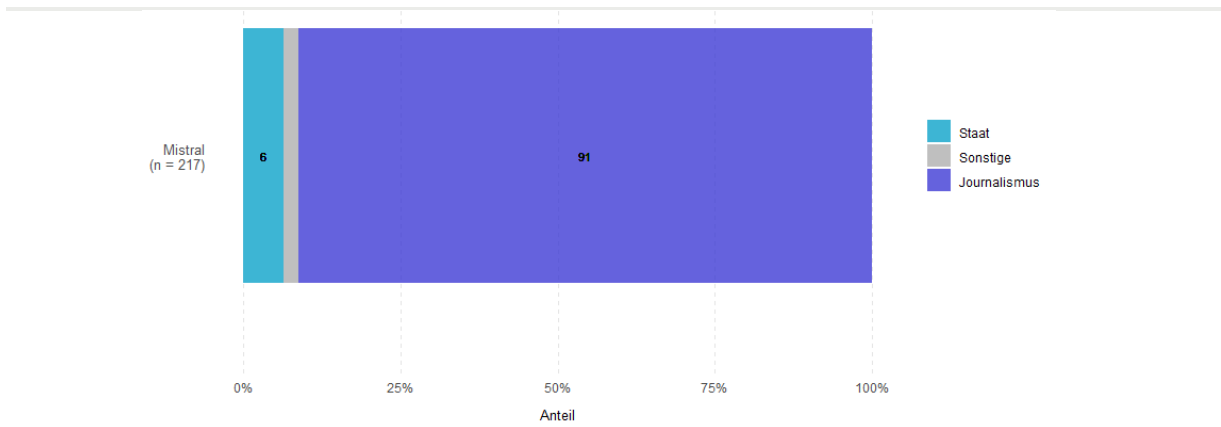
Anmerkung: Für Mistral wurden nur die Prompts zu den Themen Innen- und Außenpolitik abgefragt.

Abbildung 19: Durchschnittliche Quellenanzahl pro Prompt für Mistral



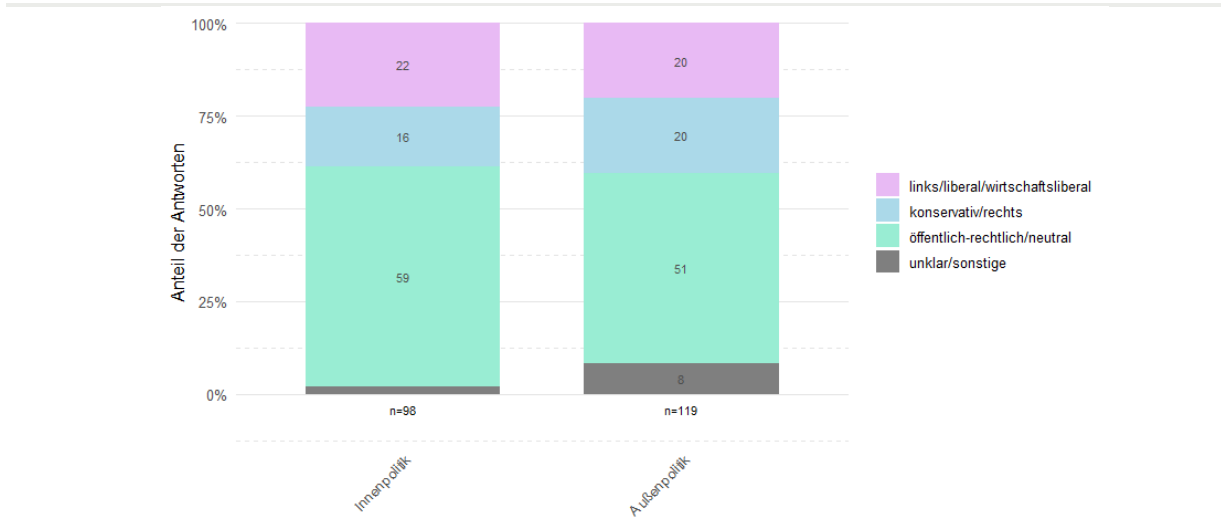
Anmerkung: Für Mistral wurden nur die Prompts zu den Themen Innen- und Außenpolitik abgefragt. Balken zeigen den Mittelwert, Fehlerbalken die Standardabweichung.

Abbildung 22: Verteilung der Akteursgruppen für Mistral



Anmerkung: Für Mistral wurden nur die Prompts zu den Themen Innen- und Außenpolitik abgefragt.

Abbildung 19: Politische Ausrichtung der journalistischen Quellen je Thema für Mistral



Anmerkung: Für Mistral wurden nur die Prompts zu den Themen Innen- und Außenpolitik abgefragt.



Hinweis zur Verwendung von Künstlicher Intelligenz

Bei der Erstellung dieser Studie wurde Claude Opus 4.8 von Anthropic zugänglich über Claude.ai, im Zeitraum Mai bis Juli 2026 unterstützend eingesetzt. Der Einsatz umfasste das Schreiben und Debuggen von Code zur Datenauswertung und -visualisierung sowie die Überarbeitung von Textabschnitten auf Basis eigener Textentwürfe. Alle KI-generierten Inhalte wurden vom Autorinnenteam geprüft, überarbeitet und verantwortet; eine ungeprüfte Übernahme fand zu keinem Zeitpunkt statt.

Über die Agora Digitale Transformation

Agora Digitale Transformation ist der Thinktank für Updates unserer Demokratie. Ziel ist es, die Chancen der Digitalen Transformation zu nutzen, um die Demokratie in Deutschland zu stärken. Als gemeinnützige Organisation arbeitet der Thinktank überparteilich, kollaborativ und evidenzbasiert mit dem Schwerpunkt auf umsetzbare und wirksame Lösungen für die Politik. Die Agora Digitale Transformation sucht dafür gezielt den Austausch und die Zusammenarbeit mit Innovator:innen aus Gesellschaft, Verwaltung, Wissenschaft, Wirtschaft und Politik. Initiiert wurde die Agora Digitale Transformation im Jahr 2023 von der Stiftung Mercator. Alleiniger Gesellschafter ist die Stiftung Mercator, die die Agora Digitale Transformation fördert.

Ansprechperson bei der Agora Digitale Transformation



Dr. Vivien Benert

Innovation Lead – Digitale Öffentlichkeit

vivien.benert@agoradigital.de

Agora Digitale
Transformation

www.agoradigital.de