

Pressemitteilung

Erst aktuell, dann überholt: KI-Assistenten vergessen neue Entwicklungen und antworten mit der Zeit immer öfter falsch

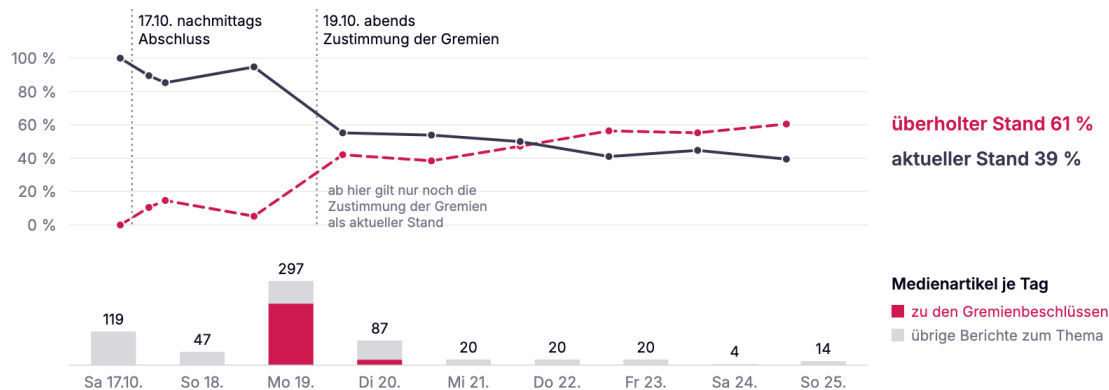
Auswertung von aclipp: Fünf Tage nach einer politischen Entscheidung geben 6 von 10 Antworten von ChatGPT, Gemini und Co. einen veralteten Stand wieder, Tendenz steigend. Ausschlaggebend sind die zitierten Quellen: Stützt sich ein Assistent auf einen aktuellen Artikel, stimmt die Antwort meist.

Spremberg, 28. Oktober 2026. KI-Assistenten wie ChatGPT, Gemini oder Perplexity bleiben bei laufenden Nachrichtenlagen oft beim alten Stand stehen. Das zeigt eine Auswertung von aclipp. Vom 17. bis 25. Oktober 2026 hat aclipp sieben KI-Assistenten täglich dieselben zwölf Fragen zu den Berliner Sondierungen von SPD, Grünen und Linken gestellt und die Antworten mit über 900 Medienartikeln und den zitierten Quellen abgeglichen. Das Ergebnis: Neue Entwicklungen kommen nur teilweise an und gehen nach wenigen Tagen wieder verloren, weil die Assistenten weiter ältere Artikel zitieren. Für Kommunikationsteams heißt das: Ein neuer Stand, eine Korrektur oder Klarstellung kommt nicht automatisch in den KI-Antworten an, nur weil die Medien darüber berichten.

Solange die Sondierungen liefen, waren die Antworten auf die Fragen nach dem aktuellen Stand fast durchgehend aktuell. Am Abend des 17. Oktober, sieben Stunden nach der Pressekonferenz zum Abschluss, gaben 17 von 19 Antworten das Ergebnis wieder. Am 19. Oktober stimmten Grünen-Parteitag und SPD-Landesvorstand Koalitionsverhandlungen zu. Diesen neuen Stand gab am Tag danach gut die Hälfte der Antworten wieder (55 Prozent), am 25. Oktober nur noch 39 Prozent. Der Anteil veralteter Antworten stieg im selben Zeitraum von 42 auf 61 Prozent. Vier von zehn Antworten stellten die Entscheidung am 25. Oktober noch als offen dar.

Erst aktuell, dann überholt

Antworten auf Fragen nach dem aktuellen Stand: Wie viele geben den tatsächlichen Stand wieder, wie viele einen älteren?



aclipp

Fragen nach dem aktuellen Stand, je Messpunkt 19 bis 39 Antworten · Abfrage 17.10. 13:50 und 21:35 Uhr, dann täglich nachts · Quelle: aclipp, 7 KI-Assistenten · 17. bis 25.10.2026 · 338 Antworten

Die Antwort ist so aktuell wie ihre Quellen

Ob eine Antwort stimmt, hängt vor allem davon ab, welche Quellen der Assistent zitiert. Antworten ab dem 20. Oktober, die mindestens einen Artikel von nach der Entscheidung zitieren, geben zu 84 Prozent den richtigen Stand wieder. Stützt sich eine Antwort nur auf ältere Artikel, sind es 33 Prozent.

Dabei greifen die Assistenten immer wieder zu denselben großen Artikeln. Die beiden Berichte von tagesschau.de zum Abschluss am 17. Oktober stehen am 25. Oktober in fast jeder zweiten Antwort als Quelle (47 Prozent), der Bericht von tagesschau.de über die Zustimmung der Gremien nur in 2 Prozent. An fehlender Berichterstattung liegt es nicht: Am 19. Oktober erschienen 217 Artikel über die Entscheidung, viele davon Übernahmen von Agenturmeldungen, aber auch eigene Berichte von tagesschau.de, Deutschlandfunk, SPIEGEL und rbb24.

Aktuelle Quellen entscheiden über die richtige Antwort

Antworten vom 20. bis 25.10. zur Gremienentscheidung vom 19.10.

Zitiert mindestens eine Quelle,
die nach dem Beschluss erschienen ist
(251 Antworten)

84 % richtig

15 % falsch

Zitiert nur Quellen von
vor dem Beschluss
(230 Antworten)

33 % richtig

63 % falsch

aclipp

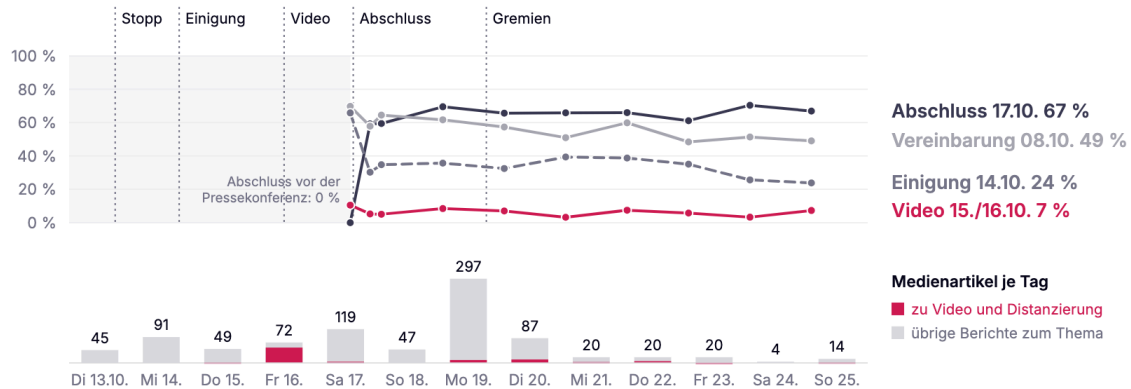
12 Fragen, zwei identische Sets · richtig = nennt die Zustimmung, falsch = nennt die Entscheidung als offen, grau = widersprüchlich · Quellen ohne Datum zählen als „vor dem Beschluss“ · Quelle: aclipp, 7 KI-Assistenten · 20. bis 25.10.2026 · 481 Antworten

Nicht jedes Thema bleibt hängen

Wie lange sich ein Ereignis in den Antworten hält, ist sehr unterschiedlich. Den Abschluss der Sondierungen vom 17. Oktober nennen bis zum 25. Oktober rund zwei Drittel der Antworten, die gemeinsame Antisemitismus-Vereinbarung der drei Parteien vom 8. Oktober rund die Hälfte. Beide standen in Agenturmeldungen immer wieder als Hintergrund. Die Einigung auf die Wiederaufnahme der Gespräche vom 14. Oktober nannten am Mittag des 17. Oktober noch 66 Prozent der Antworten, am 25. Oktober nur noch 24 Prozent. Viel Berichterstattung allein reicht nicht: Das Video einer Linken-Abgeordneten und die Distanzierung ihrer Fraktionsspitze bestimmten am 16. Oktober 54 von 72 Medienartikeln. In den KI-Antworten spielte es nie eine große Rolle (5 bis 11 Prozent).

Manche Ereignisse bleiben, andere verblassen

Anteil aller Antworten, die das Ereignis erwähnen. Eine Antwort kann mehrere Ereignisse nennen.



aclipp

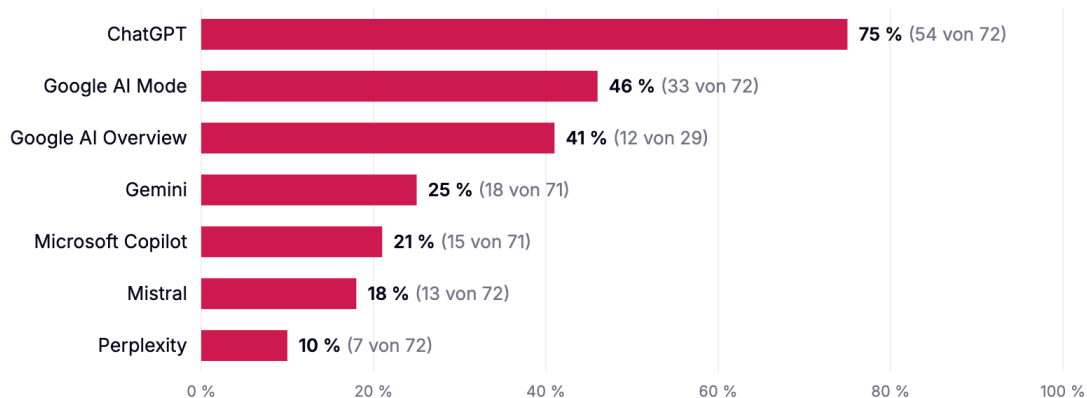
Stopp: SPD und Grüne unterbrechen die Sondierungen (13.10.) - Einigung: Wiederaufnahme der Gespräche vereinbart (14.10.) - Video: Video einer Linken-Abgeordneten und Distanzierung der Fraktionsspitze (15./16.10.) - Abschluss: Sondierungen abgeschlossen (17.10.) - Gremien: Zustimmung zu Koalitionsverhandlungen (19.10.) - 12 Fragen in zwei identischen Sets, Abfrage am 17.10. um 13:50 Uhr (Set 1) und 21:35 Uhr (Set 2), danach täglich - Quelle: aclipp, 7 KI-Assistenten - 17. bis 25.10.2026 - 1.363 Antworten

Große Unterschiede zwischen den Assistenten

ChatGPT nennt die Zustimmung der Gremien ab dem 20. Oktober in 54 von 72 Antworten. Perplexity zitiert mit rund zehn Quellen je Antwort die meisten, aber auch die ältesten (im Median zwölf Tage alt) und nennt die Zustimmung nur in 7 von 72 Antworten. Knapp jede zehnte Antwort aller Assistenten enthält einen belegbaren Sachfehler, meist ein falsches Datum oder eine falsche Zuordnung zu einer Partei.

Wer kennt die Gremienbeschlüsse vom 19.10.?

Anteil der Antworten vom 20. bis 25.10., die die Zustimmung wiedergeben



aclipp

Erstes Fragenset, alle 12 Fragen; Google AI Overview antwortet seltener - Quelle: aclipp, 7 KI-Assistenten - 20. bis 25.10.2026 - 459 Antworten

Was das für Kommunikation bedeutet

„Ein KI-Assistent kennt einen neuen Stand nicht automatisch, nur weil es neue Nachrichten gibt. Und selbst wenn eine Entscheidung, eine Korrektur oder eine Klarstellung in den Antworten angekommen ist, bleibt sie dort nicht zwangsläufig. Nach ein paar Tagen greifen die Assistenten wieder auf ältere Artikel zurück“, sagt Sascha Kirstein, Gründer und Geschäftsführer von aclipp. „Ob eine Botschaft ankommt und ob sie bleibt, zeigt nur eine wiederholte Messung über mehrere Assistenten.“

Einordnung

Wie zuverlässig KI-Assistenten Nachrichten wiedergeben, haben bereits mehrere Untersuchungen geprüft. Die Studie der Europäischen Rundfunkunion (EBU) und der BBC hat 2025 gezeigt, dass KI-Assistenten Nachrichteninhalte häufig verzerrt wiedergeben. Ein Echtzeit-Audit der Stanford University fand 2026, dass die besten Assistenten Fragen zu Nachrichten vom selben Tag zu über 90 Prozent richtig beantworten. Beide messen zu einem bestimmten Zeitpunkt. Die aclipp-Auswertung schaut auf die Tage danach: wie sich die Antworten entwickeln, wenn sich die Nachrichtenlage weiterbewegt.

Methode

- **Abfrage:** 12 neutral formulierte Fragen, etwa „Wie ist der Stand der Regierungsbildung in Berlin?“. Keine Frage nennt einzelne Personen oder das Video. Abgefragt wurden ChatGPT, Gemini, Google AI Mode, Google AI Overview, Microsoft Copilot, Mistral und Perplexity mit Standort Deutschland. Abgefragt wurde ausgeloggt, mit dem Modellstand, den Nutzer am jeweiligen Tag auch bekamen.
- **Umfang:** Zwei identische, parallel laufende Fragensets: das erste ab 17. Oktober, 13:50 Uhr (679 Antworten, 4.494 zitierte Links), das zweite ab 17. Oktober, 21:35 Uhr (684 Antworten). Danach täglich bis zum 25. Oktober.
- **Abgleich:** 531 Medienartikel bis 18. Oktober und 462 Artikel vom 19. bis 25. Oktober aus dem aclipp-Medienmonitoring und einer ergänzenden Websuche.
- **Auswertung:** KI-Agenten (Claude von Anthropic) haben jede Antwort nach festem Raster kodiert: welche Ereignisse sie nennt, welchen Stand sie darstellt, welche Quellen sie zitiert. aclipp hat eine Zufallsstichprobe von 30 Antworten von Hand geprüft: Bei 29 von 30 stimmt die Kodierung. Sachfehler wurden nur gezählt, wenn ein Artikel aus dem Datensatz sie widerlegt.
- **Schwankung:** Stellt man demselben Assistenten am selben Tag dieselbe Frage zweimal, nennt in 18 Prozent der Fälle die eine Antwort die Zustimmung der Gremien und die andere nicht. Einzelne Antworten sind Momentaufnahmen, der Verlauf über die Tage ist in beiden Sets gleich.

- **Grenzen:** Eine Nachrichtenlage, ein Zeitraum, eine Antwort je Frage, Assistent, Tag und Set. Die Ergebnisse beschreiben Muster, keine statistisch abgesicherten Effekte.

Grafiken und Daten zur freien Verwendung

- Grafik 1: Erst aktuell, dann überholt (Stand in den Antworten und Medienartikel je Tag)
- Grafik 2: Aktuelle Quellen entscheiden über die richtige Antwort
- Grafik 3: Manche Ereignisse bleiben, andere verblassen
- Grafik 4: Die Assistenten im Vergleich
- Excel-Datei mit allen Kennzahlen und der Methode

Abdruck mit Quellenangabe „aclipp“ frei.

Über aclipp

aclipp ist eine PR-Analytics-Plattform aus Deutschland. Kommunikationsteams und Agenturen werten damit ihre Berichterstattung in Medien und in den Antworten von KI-Assistenten aus. aclipp.com

Pressekontakt

Sascha Kirstein, Gründer und Geschäftsführer

sascha@aclipp.com

+49 172 2510219