LANDESANSTALT
FÜR MEDIEN NRW
Der Meinungsfreiheit
verpflichtet.

ZWISCHEN SUCHMASCHINE UND GESPRÄCHSPARTNER:
WIE KI-CHATBOTS UNSERE POLITISCHE INFORMATIONSSUCHE VERÄNDERN

Juli 2026

INHALT

1. EINLEITUNG 3

1.1 Hintergrund 3
1.1.1 Generative KI als neues Informationsmedium 3
1.1.2 Forschungsstand zur Nutzung von KI-Chatbots zur politischen Information 4
1.1.3 Prompts als Schnittstelle zwischen Nutzenden und KI-Chatbots 5
1.2 Ziel und Forschungsfragen 5

2. STUDIENDESIGN UND METHODIK 6

2.1 Entwicklung des Fragebogens 6
2.1.1 Realistische Nutzungsszenarien 7
2.2 Datenerhebung und Datenaufbereitung 7
2.3 Datenanalyse 9
2.4 Überblick über die Stichprobe 10

3. ERGEBNISSE 11

3.1 Nutzung generativer KI 11
3.2 Prompting-Strategien der Nutzenden 13
3.2.1 Theoretische Herleitung der Prompting-Typologie 13
3.2.2 Häufigkeitsverteilung der Prompting-Strategien 17
3.2.3 Häufigkeitsverteilung kategorienübergreifender Merkmale 19
3.2.4 Konsistenz der Prompting-Strategien 21
3.3 Inferenzstatistische Ergebnisse 22
3.3.1 Zusammenhänge zwischen Prompting-Strategien und Nutzendengruppen 22
3.3.2 Zusammenhänge zwischen Prompts und ausgelöster Websuche 22

4. LIMITATIONEN DER STUDIE 25

5. ZUSAMMENFASSUNG 25

6. FAZIT 26

7. LITERATUR UND QUELLENVERZEICHNIS 28

8. ABBILDUNGSVERZEICHNIS 32

9. TABELLENVERZEICHNIS 32

10. ANHANG 32

10.1 Modelltabelle: Websuche-Prädiktoren 32
10.2 Algorithmisches Auditing 33
10.2.1 Anleitung und Empfehlungen für das Auditing 33
10.2.2 Prototypische Prompts und Auditing Beispiele 34
10.3 Fragebogen 38

IMPRESSUM 54

1. EINLEITUNG

1.1 HINTERGRUND

1.1.1 GENERATIVE KI ALS NEUES INFORMATIONSMEDIUM

Technologische Innovationen haben die Art und Weise, wie Menschen Informationen suchen, bewerten und nutzen, historisch immer wieder grundlegend verändert. Mit dem Aufkommen generativer Künstlicher Intelligenz, insbesondere durch textbasierte KI-Chatbots wie ChatGPT und neuen Suchformaten wie KI-Übersichten (bspw. Googles „AI-Overview"), hat diese Transformation in jüngster Zeit eine neue Dynamik erreicht. KI-Chatbots dieser Art stützen sich auf große Sprachmodelle (LLMs), die aus umfangreichen Datenmengen statistische Sprachmuster ableiten und häufig mit externen, auch aktuellen Informationsquellen kombiniert werden. Dadurch können Suchanfragen nicht mehr nur mit einzelnen Treffern, sondern mit zusammenhängenden Antworten beantwortet werden.

Diese Entwicklung hat weitreichende Implikationen für die gesellschaftliche Informationsordnung, insbesondere im Kontext politischer Information und Meinungsbildung. Zum einen verändert der Einsatz von KI-Chatbots die Rolle klassischer Medien und von Medienintermediären grundlegend: Erste Befunde zeigen, dass Nachrichtenseiten durch neue KI-gestützte Suchformate teils erhebliche Einbußen im Such-Traffic verzeichnen (The Guardian, 2025). Zum anderen verschiebt sich die Auswahl, Strukturierung und Interpretation von Informationen zunehmend von menschlichen Akteuren hin zu algorithmischen Systemen. Während bei der klassischen Websuche, mit einer Trefferliste als Ergebnis, die Interaktion der Nutzenden darin bestand, aktiv eine Auswahl aus den Ergebnissen zu treffen und sich so für bestimmte Quellen zu entscheiden, verschiebt sich die aktive Teilnahme bei der Interaktion mit generativen Suchmaschinen in den Dialog. Quellen werden für die Nutzenden automatisch ausgewählt und der aktive Part der Nutzenden liegt im Nachfragen beim mehrstufigen Dialog. Damit eröffnen KI-Chatbots neue Möglichkeiten der Informationsnutzung. Nutzende können Informationen nicht mehr nur passiv konsumieren, sondern in einem dialogischen Prozess aktiv strukturieren, personalisieren und vertiefen (Egan et al., 2026). Dadurch entsteht eine neue Form des Zugangs zu Wissen und zu politischer Information: Komplexe Inhalte können verständlich aufbereitet, individuelle Informationsbedarfe adressiert und durch interaktive Rückfragen Lern- und Verstehensprozesse unterstützt werden.

Diese Potenziale gehen jedoch mit Risiken einher. KI-Chatbots können überzeugend wirken, aber faktisch falsche oder verzerrte Inhalte erzeugen – sogenannte „Halluzinationen" –, die für Nutzende nur schwer erkennbar sind, besonders wenn der Output flüssig oder selbstsicher klingt (Xie & Zhang, 2025). Untersuchungen zeigen zudem, dass KI-Chatbots anfällig für Desinformationskampagnen sein können und Falschaussagen – abhängig von Sprachressourcen und Modellversion – teilweise reproduzieren, anstatt diese zu korrigieren (Makhortykh et al., 2024). Darüber hinaus können KI-Chatbots politische Verzerrungen aus ihren Trainingsdaten reproduzieren (Ceron et al., 2025; Motoki et al., 2024). Gerade im politischen Kontext ist dies problematisch, da solche Inhalte Wahrnehmungen und Meinungsbildungsprozesse beeinflussen können –

und in Verbindung mit der hohen Skalierbarkeit von KI-Chatbots grundlegende Fragen hinsichtlich Transparenz, Nachvollziehbarkeit und demokratischer Meinungsbildung aufwerfen (Die Medienanstalten, 2025; Behnert & Timmermann, 2026).

Eine gesonderte Rolle spielt dabei die zunehmende Integration von Websuche in KI-Chatbots: Anders als klassische Suchmaschinen generieren diese eine einzige synthetisierte Antwort auf Basis abgerufener Webinhalte (Kirsten et al., 2025). Auch dies löst die beschriebenen Risiken nicht auf – bisherige Untersuchungen belegen, dass auch webgestützte Antworten fehleranfällig sind und KI-Chatbots selbst dann falsche oder irreführende Informationen liefern können, wenn sie auf aktuelle Quellen zugreifen (BBC & Ipsos, 2025; Elliott, 2025). Die generative Suche ist dabei noch ein junges Forschungsfeld, das erste Fragen zur Quellenqualität und zur Verlässlichkeit dieser Antworten aufwirft (Shi et al., 2025; Zhang et al., 2026).

1.1.2 FORSCHUNGSSTAND ZUR NUTZUNG VON KI-CHATBOTS ZUR POLITISCHEN INFORMATION

Vor diesem Hintergrund stellt sich die Frage, wie KI-Chatbots derzeit genutzt werden – sowohl in alltäglichen als auch in gesellschaftlich relevanten Kontexten wie der politischen Informationssuche. In Deutschland war im Jahr 2024 ChatGPT die am häufigsten genutzte generative KI-Anwendung, mit Abstand gefolgt von Microsoft Copilot, Google Gemini, DALL-E, Meta AI und Midjourney (Statista, 2024). 56 % der Personen, die bereits KI-Anwendungen für private Zwecke genutzt haben, geben an, diese KI-Anwendungen mindestens ein bis zwei Mal die Woche zu verwenden und 10% nutzen sie mehrmals täglich (Statista, 2025). Studien zeigen darüber hinaus, dass Nutzende eher jung und männlich sind sowie einen höheren Einkommens- und Bildungsstand haben (Kaiser et al., 2024). KI-Chatbots werden besonders häufig als Suchmaschine oder zur Recherche von Informationen genutzt (Bitkom, 2025). Weitere häufige Anwendungszwecke von KI-Chatbots sind die Einholung schneller Antworten oder Faktenchecks, gefolgt von der Nutzung als Schreibhilfe sowie die schnelle Unterstützung bei Fragen rund um Gesundheitsthemen (YouGov, 2025). Zudem zeigt der aktuelle Digital News Report von Reuters, dass 5 % der Deutschen KI-Chatbots als Nachrichtenquelle verwenden, während es weltweit bereits 10 % und damit 3 % mehr als im vergangenen Jahr sind. Dabei konzentriert sich die Nutzung stärker auf jüngere Zielgruppen (16 % der unter 35-Jährigen geben an, KI-Chatbots für Nachrichten zu nutzen) (Egan et al., 2026). Knapp ein Drittel (28 %) der Deutschen stehen diesem Anwendungszweck zukünftig offen gegenüber und könnten sich vorstellen, KI-Chatbots zur Zusammenfassung von Nachrichten zu verwenden oder aber um Falschnachrichten zu identifizieren (32 %) (Kaiser et al., 2024).

Diese Befunde liefern wichtige Hinweise zur Verbreitung und Akzeptanz von KI-Chatbots. Sie geben jedoch kaum Aufschluss darüber, wie Nutzende KI-Chatbots konkret einsetzen, um sich politisch zu informieren. Ein Grund dafür liegt darin, dass bisherige Studien die konkrete Interaktion zwischen Nutzenden und KI-Chatbots nur unzureichend untersuchen. Insbesondere bleibt weitgehend unklar, wie Nutzende ihre Informationsbedürfnisse in der Interaktion mit KI-Chatbots formulieren und strukturieren.

1.1.3 PROMPTS ALS SCHNITTSTELLE ZWISCHEN NUTZENDEN UND KI-CHATBOTS

KI-Chatbots sind promptbasiert, d. h., sie reagieren auf Texteingaben von Nutzenden mit einer generierten Antwort. Diese Prompts bilden die zentrale Schnittstelle zwischen Nutzenden und KI-Chatbot. Definiert werden Prompts als individuelle Eingaben, mit denen Anforderungen und Intentionen formuliert werden; über mehrmalige Prompts können Anfragen im Dialog präzisiert werden (White et al., 2023). In der Forschung existiert bislang keine einheitliche Typologie von Nutzenden-Prompts für KI-Chatbots. Bisherige Studien konzentrieren sich zudem überwiegend auf die Qualität der erzeugten Inhalte oder arbeiten mit einer begrenzten Zahl vorab definierter Prompts und experimentellen Szenarien (Xu et al., 2023; Xue et al., 2026). Ob diese die tatsächlichen Nutzungspraktiken realistisch abbilden, ist bislang weitgehend ungeklärt. Zudem wurde die Suche nach gesellschaftlichen und politischen Themen in der empirischen Forschung zur KI-Nutzung bisher kaum gesondert betrachtet. Damit fehlt bislang eine empirische Grundlage, um zu verstehen, wie Nutzende KI-Chatbots konkret zur Recherche politischer Informationen einsetzen und welche Prompting-Strategien sie dabei tatsächlich verwenden.

Die Relevanz einer systematischen Analyse von Prompts ergibt sich unter anderem aus der Forschung zur Prompt-Sensitivität von LLMs. Diese zeigt, dass die Art der Formulierung eines Prompts die generierte Antwort maßgeblich beeinflussen kann. Dies betrifft sowohl strukturelle Merkmale, wie etwa die Satzstruktur oder das Format des Prompts, als auch semantisch-pragmatische Dimensionen wie Ton, Höflichkeit und sprachliche Gestaltung (Zhou et al., 2024). So konnte beispielsweise gezeigt werden, dass negativ konnotierte Prompts mit einer höheren Wahrscheinlichkeit zu weniger faktisch korrekten Antworten führen als neutral formulierte Prompts (Gandhi & Gandhi, 2025; Patel et al., 2026; Vinay et al., 2025). Die Verwendung höflicher Formulierungen hat dagegen je nach Modell und Sprache unterschiedliche Effekte auf die Antwortqualität (Yin et al., 2024; Dobariya & Kumar, 2025). Darüber hinaus beeinflusst die Wahl spezifischer Begriffe – etwa „Krieg“ statt „Konflikt“ – die thematische Rahmung der generierten Antwort, da KI-Chatbots dazu tendieren, im Prompt implizit enthaltene Vorannahmen zu übernehmen, anstatt diese kritisch zu hinterfragen (Germani & Spitale, 2025; Hwang et al., 2026; Ranaldi & Pucci, 2023; Sharma et al., 2023). Die Analyse von Nutzenden-Prompts liefert somit nicht nur Aufschluss über individuelle Nutzungspraktiken, sondern auch über potenzielle Mechanismen, durch die die Formulierung politischer Suchanfragen die Qualität und Rahmung von KI-Chatbot-Antworten beeinflusst.

1.2 ZIEL UND FORSCHUNGSFRAGEN

Vor diesem Hintergrund zielt die vorliegende Studie darauf ab, die Nutzung von KI-Chatbots im Kontext politischer Informationssuche differenziert zu untersuchen und insbesondere die tatsächlichen Prompting-Strategien von Nutzenden zu analysieren.

Die zentralen Forschungsfragen der Studie sind folgende:

- Welche Prompting-Strategien lassen sich in realen Nutzungssituationen zur politischen Informationssuche identifizieren?
- Welche Faktoren beeinflussen die Nutzung unterschiedlicher Prompting-Strategien zur politischen Informationssuche (z.B. Alter, Mediennutzung, digitale Kompetenz)?
- Unter welchen Bedingungen lösen Nutzende bei der politischen Informationssuche eine KI-Chatbot-seitige Websuche aus, und welche Prompt-Merkmale sind dabei ausschlaggebend?
- Wie und zu welchen Zwecken nutzen Menschen generative KI-Anwendungen?

2. STUDIENDESIGN UND METHODIK

2.1 ENTWICKLUNG DES FRAGEBOGENS

Zur empirischen Untersuchung wurde ein Online-Fragebogen entwickelt, der die Nutzung von KI-Chatbots zur politischen Informationssuche sowie die zugrundeliegenden Prompting-Strategien möglichst realitätsnah erfassen sollte.

Die inhaltliche Konzeption basierte auf einer Auswertung des aktuellen Forschungsstands sowie auf den definierten Forschungsfragen. Dabei wurden etablierte Skalen, etwa zu Medienkompetenz (vgl. Ashley, Maksl & Craft, 2013), digitaler Kompetenz (vgl. Cabrera et al., 2020) sowie zu Vertrauen in Nachrichten und in KI-Systeme (vgl. Reuters Institute for the Study of Journalism, 2023) übernommen, übersetzt oder teilweise angepasst. Zudem wurden neue Fragen und Items zur spezifischen Nutzung von KI-Chatbots entwickelt.

Der Fragebogen gliederte sich in mehrere Themenblöcke:

- Screeningfragen zur Nutzung von KI-Chatbots und soziodemographische Merkmale
- Nutzung von generativen KI-Chatbots und Medien
- zwei Interaktionsszenarien zur Erhebung realer Prompting-Strategien zur politischen Informationssuche
- weitere Kontext- und Kontrollvariablen (z.B. digitale Kompetenzen und Vertrauen in KI)

2.1.1 REALISTISCHE NUTZUNGSSZENARIEN

Ein zentraler Bestandteil des Fragebogens waren zwei integrierte Nutzungsszenarien, in denen die Teilnehmenden direkt mit einem KI-Chatbot interagierten. Diese Szenarien wurden entwickelt, um reales Prompting-Verhalten unter möglichst natürlichen Bedingungen zu erfassen. Die Konzeption orientierte sich dabei an bestehenden Studien zu Prompting-Strategien und dialogischer Informationssuche (u. a. Xue et al., 2025).

- Das erste Nutzungsszenario wurde als offene politische Informationssuche konzipiert. Die Teilnehmenden wurden aufgefordert, sich mithilfe eines integrierten KI-Chatbots über ein politisches Thema zu informieren, welches sie aktuell beschäftigt oder interessiert. Dabei wurde bewusst keine thematische Vorgabe gemacht, um möglichst natürliche und selbstgewählte Informationsinteressen sowie entsprechende Prompting-Strategien zu erfassen.
- Das zweite Nutzungsszenario wurde mit einem thematisch vorgegebenen Ansatz konzipiert (vorgegebene politische Informationssuche). Zunächst wählten die Teilnehmenden aus einer Liste politischer Themen bis zu drei aus, die sie besonders interessieren, und gaben anschließend an, wie gut sie sich mit den jeweiligen Themen auskennen. Auf Basis dieser Angaben wurde automatisch dasjenige Thema identifiziert, zu dem der geringste Kenntnisstand vorlag. Diese Filterführung stellte sicher, dass die Teilnehmenden ein Thema erhielten, das sie zwar interessiert, mit dem sie jedoch noch wenig vertraut sind. Zu diesem Thema wurde ihnen eine kurze, journalistisch formulierte Schlagzeile mit Titel und Untertitel angezeigt, die als Ausgangspunkt für die anschließende Informationssuche diente. Die Teilnehmenden wurden daraufhin aufgefordert, sich mithilfe des eingebetteten KI-Chatbots gezielt zu diesem Thema zu informieren.

Die Fragebogenentwicklung erfolgte in enger Abstimmung mit der Landesanstalt für Medien NRW. Der vollständige Fragebogen ist im Anhang dokumentiert (siehe Kapitel 10.3).

2.2 DATENERHEBUNG UND DATENAUFBEREITUNG

Der Fragebogen wurde als Online-Befragung (CAWI) programmiert. Die Interaktionen der Teilnehmenden mit dem KI-Chatbot wurden über die OpenAI-API in Echtzeit verarbeitet: Eingaben der Teilnehmenden wurden direkt an das Modell *gpt-4.1-mini-2025-04-14* übermittelt und die generierten Antworten unmittelbar innerhalb der Befragungsumgebung angezeigt (OpenAI, 2026). Im Rahmen der Programmierung des Fragebogens wurde zunächst eine umfassende technische Prüfung durchgeführt. Dabei wurden insbesondere die Vollständigkeit des Instruments, die korrekte Umsetzung der Screeningkriterien sowie die Funktionalität der Filterführungen anhand des Befragungslinks überprüft. Anschließend wurde ein Soft Launch durchgeführt, auf dessen Basis die Programmierung mithilfe der generierten Testdaten validiert wurde. Insbesondere komplexe Filterführungen und Auswahlmechanismen konnten so datenbasiert überprüft und bei Bedarf angepasst werden.

Zur Sicherstellung einer hohen Datenqualität kam während der Feldphase ein KI-gestützter Data Quality Score zum Einsatz. Dieser basiert auf der kontinuierlichen Erfassung und Auswertung von Paradaten während der Befragung und ermöglicht eine automatisierte Bewertung des Antwortverhaltens der Teilnehmenden in Echtzeit. Der Score kombiniert verschiedene Analyseansätze, um sowohl unaufmerksames Antwortverhalten als auch potenzielle Betrugsversuche zu identifizieren. Probandinnen und Probanden mit unzureichender Datenqualität wurden automatisiert aus der Befragung ausgeschlossen, während auffällige Fälle markiert und einer zusätzlichen Prüfung unterzogen wurden. Im Rahmen dieses Verfahrens wurden unter anderem folgende Parameter berücksichtigt:

- Integration eines Bot-Checks zur Vermeidung automatisierter Teilnahmen
- Analyse der Bearbeitungszeiten auf Seitenebene (z. B. sehr kurze Verweildauern)
- Verhältnis von gelesenen Wörtern pro Sekunde zur Identifikation unrealistischer Antwortgeschwindigkeiten
- Prüfung auf unplausible soziodemographische Angaben (z. B. Alter)

Die Datenerhebung fand im Zeitraum vom 24. März 2026 bis zum 1. April 2026 statt. Befragt wurden n = 2.009 Personen in Deutschland ab 16 Jahren, die KI-Chatbots mindestens selten (d. h. mindestens 1–3-mal im Jahr) nutzen. Ein substanzieller Anteil der zur Befragung eingeladenen Personen (ca. 30%) gab an, keinen der derzeit am Markt aktiven KI-Chatbots zu nutzen und wurde daher von der weiteren Befragung ausgeschlossen und ist damit nicht Teil der finalen Stichprobe. Daraus leitet sich eine ungefähre Prävalenz von 70% KI-Chatbot-Nutzenden in der Bevölkerung ab. Die Rekrutierung der Stichprobe erfolgte mit dem Ziel einer möglichst breiten und ausgewogenen Abdeckung zentraler soziodemographischer Merkmale. Hierbei wurden insbesondere Probanden nach den Merkmalen Geschlecht, Bundesland und verschiedene Altersgruppen schrittweise gesteuert eingeladen. Auch Personen mit geringerer formaler Bildung wurden verstärkt rekrutiert, wobei das erreichbare Potenzial in dieser Gruppe weitgehend ausgeschöpft wurde.

Zur Sicherstellung der Datenqualität wurde das Antwortverhalten der Teilnehmenden während und nach der Feldphase systematisch überprüft. Inkonsistente, unvollständige oder unplausible Fälle wurden identifiziert und aus dem Datensatz entfernt. Zur Kompensation dieser Bereinigungen wurde ein Oversampling durchgeführt, bei dem zusätzliche Fälle erhoben, und ebenfalls einer Qualitätsprüfung unterzogen wurden. Die Datenaufbereitung umfasste insbesondere folgende Qualitätskontrollen:

- Prüfung offener Angaben auf inhaltliche Plausibilität (z. B. unsinnige Texteingaben)
- Kontrolle, ob die Aufgabenstellung – insbesondere die Interaktion mit dem KI-Chatbot – korrekt umgesetzt wurde
- Analyse der Varianz und Vollständigkeit der Antworten
- Überprüfung der Befragungsdauer: Fälle mit sehr kurzer Bearbeitungszeit (unterhalb von

20 % des Medians) wurden ausgeschlossen, Grenzfälle individuell geprüft
- Prüfung der inhaltlichen Konsistenz der Antworten

Darüber hinaus erfolgte eine systematische Identifikation von Fällen mit non-responsivem Antwortverhalten. Als Indikatoren dienten etablierte Qualitätskriterien der Umfrageforschung: fehlende Antwortvarianz innerhalb von Itembatterien (z.B. Straight Lining), positionsabhängige Antworttendenzen sowie Bearbeitungszeiten unterhalb theoretisch plausibler Mindestwerte (vgl. Meade & Craig, 2012). Fälle, die mehrere dieser Kriterien aufwiesen, wurden einer vertieften Plausibilitätsprüfung unterzogen und gegebenenfalls aus der Analysestichprobe ausgeschlossen.

Durch diesen mehrstufigen Prozess aus technischer Prüfung, laufender Qualitätskontrolle während der Feldphase und nachgelagerter Datenbereinigung konnte eine hohe Qualität und Validität der erhobenen Daten sichergestellt werden.

2.3 DATENANALYSE

Die Datenanalyse erfolgte in drei aufeinander aufbauenden Schritten. Zunächst wurden deskriptive Analysen durchgeführt, um Verteilungen und Häufigkeiten zentraler Variablen sowie grundlegende Nutzungsmuster von KI-Chatbots zu beschreiben.

Im zweiten Schritt wurden die erhobenen offenen Texteingaben (Prompts) mittels eines mehrstufigen Verfahrens analysiert, das qualitative Inhaltsanalyse mit NLP-gestützten Methoden kombinierte. Die Prompts wurden mithilfe deskriptiver Textanalyse in R vorstrukturiert (R Core Team, 2026), um erste Muster in Länge und inhaltlicher Ausrichtung zu identifizieren. Auf dieser Grundlage sowie auf theoretischen Vorarbeiten wurde eine vorläufige Prompt-Typologie entwickelt, deren Kategorien zunächst manuell vergeben wurden. Diese manuellen Kategorien dienten anschließend als Vergleichsmaßstab für die automatisierte LLM-Kategorisierung: Die Übereinstimmung zwischen manueller und automatisierter Kategorisierung war sehr gut bis gut ($\kappa = .863$ in Szenario 1 / $\kappa = .779$ in Szenario 2; Cohen, 1960), weshalb das Modell in weiterer Folge für die vollständige Kategorisierung aller Prompts eingesetzt wurde. Die automatisierte Kategorisierung erfolgte durch das LLM *GPT-5-mini* über die OpenAI API mittels Python (OpenAI, 2026; Python Software Foundation, 2026). Für einfachere Klassifikationsaufgaben – etwa die Erkennung von Ansprachen oder Höflichkeitsmarkern – wurde das ressourcenschlankere Modell *Qwen3-1.7B* lokal verwendet (Qwen Team, 2025). Auf Basis der vollständigen automatisierten Klassifizierung wurde die Typologie abschließend überprüft und finalisiert: Kategorien wurden auf inhaltliche Trennschärfe geprüft und die Beschreibungen der Kategorien, wo nötig, präzisiert.

Im dritten Schritt wurden inferenzstatistische Verfahren angewandt, um Zusammenhänge zwischen Nutzungsmerkmalen und Prompting-Strategien zu untersuchen sowie die Bedingungen zu identifizieren, unter denen der KI-Chatbot eine Websuche auslöst. Zur Analyse der Prompting-

Strategien wurden Chi²-Tests mit Cramér's V als Effektstärkemaß auf Personenebene berechnet, wobei der dominante Prompt-Typ je Person als abhängige Variable diente. Zur Untersuchung der Prädiktoren einer KI-Chatbot Websuche wurde eine binär-logistische Regression berechnet, in der Aktualitätsbezug und Faktenfragen-Charakter des Prompts als Prädiktoren modelliert wurden; in einem zweiten Modell wurde zusätzlich ein expliziter Zeitbezug als weiterer Prädiktor aufgenommen und mittels Likelihood-Ratio-Test auf inkrementelle Modellgüte geprüft.

2.4 ÜBERBLICK ÜBER DIE STICHPROBE

Die finale Stichprobe umfasste N = 2009 Personen, von denen nm = 1037 (51.6%) männlich und nw = 972 (48.4%) weiblich waren. Das Durchschnittsalter der Teilnehmenden betrug M = 48.41 Jahre (SD = 15.12). Männliche Teilnehmende gaben deutlich häufiger an, mehrmals täglich nach politischen Informationen zu suchen (30.2%) als weibliche (15.3%); auch tägliche Informationssuche war bei Männern (40.7%) verbreiteter als bei Frauen (36.0%), während Frauen häufiger von seltenerer oder nur wöchentlicher (12.3% vs. 4.9%) Informationssuche (11.3% vs. 4.4%) berichteten. Diese Befunde decken sich mit bisherigen empirischen repräsentativen Untersuchungen (Kamper, 2025; Statista, 2021, 2023). Darüber hinaus zeigte sich, dass ältere Altersgruppen ein höheres politisches Interesse aufwiesen als jüngere. Insgesamt berichtete der Großteil der Teilnehmenden eine tägliche politische Informationssuche (38.4%). Hinsichtlich der regelmäßig genutzten Informations- und Medienkanäle dominierten Suchmaschinen (78.7%), während KI-Chatbots mit 51.8% deutlich seltener genutzt wurden. Die Mehrheit der Befragten steht KI-Systemen ambivalent gegenüber: Während jeweils rund ein Drittel weder klar zustimmte noch ablehnte, zeigt sich über alle Items hinweg eine leichte Tendenz zu Skepsis – besonders ausgeprägt beim generellen Misstrauen gegenüber KI-Systemen (vgl. Abbildung 1).

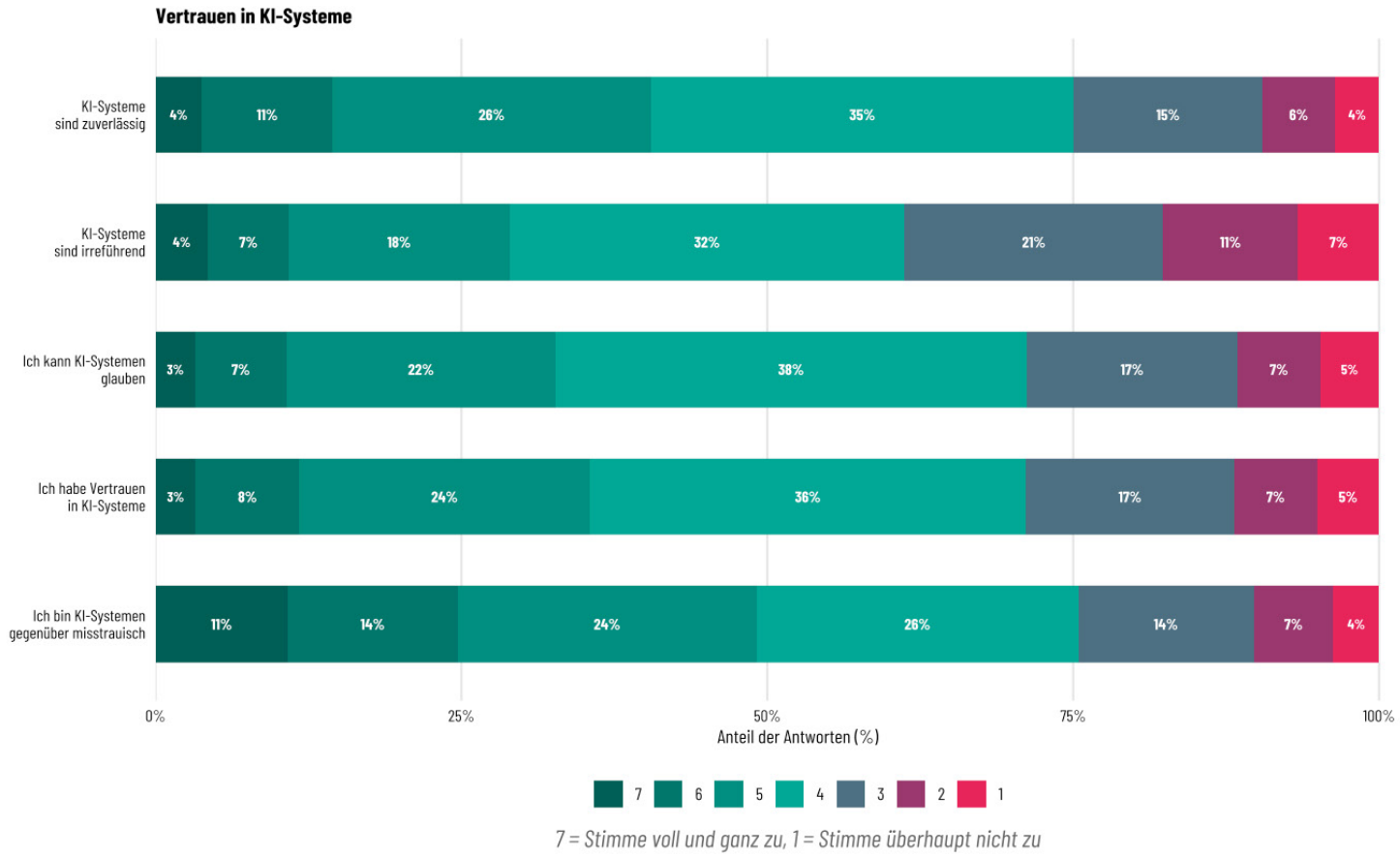

Abbildung 1: Vertrauen in KI-Systeme

3. ERGEBNISSE

3.1 NUTZUNG GENERATIVER KI

Im Folgenden werden zunächst die genutzten KI-Chatbots und allgemeinen Nutzungsmuster der Teilnehmenden beschrieben. Hinsichtlich der Nutzungspräferenzen zeigte sich, dass ChatGPT mit großem Abstand vor Gemini und Microsoft Copilot zum Erhebungszeitpunkt (Ende März bis Anfang April 2026) über alle Gruppen hinweg der meistgenutzte KI-Chatbot war. Die Befunde sind jedoch vor dem Hintergrund der dynamischen Entwicklung des Feldes zu interpretieren: Der im späten Februar 2026 einsetzende #QuitGPT-Boykott hatte zu Beginn der Erhebung bereits eingesetzt und könnte die Nutzungspräferenzen auch während der Feldphase beeinflusst haben (Schinkels, 2026). Darüber hinaus unterliegt der Markt generativer KI-Anwendungen generell rasanten Veränderungen, sodass die vorliegenden Befunde nicht ohne Weiteres auf den aktuellen Stand der Nutzungspräferenzen übertragen werden sollten. Männliche Teilnehmende wiesen eine höhere Tendenz auf, alternative KI-Chatbots zu ChatGPT zu nutzen oder auszuprobieren. Wie zu erwarten, war die Nutzung von KI-Chatbots zudem bei jüngeren Teilnehmenden insge-

samt stärker ausgeprägt als bei älteren. Bezüglich der Modellvielfalt gaben 42.7% der Teilnehmenden an, ausschließlich einen KI-Chatbot zu nutzen, 30.0% nutzten zwei und 27.3% drei oder mehr.

Neben der Nutzung konkreter KI-Chatbots wurde erfasst, in welchen Lebensbereichen und zu welchen Themen die Befragten KI-Chatbots einsetzen. Mit 89.4% nutzte die große Mehrheit KI-Chatbots für Alltag und Freizeit, gefolgt von beruflichen Tätigkeiten (44.3%) sowie Schule, Studium und Weiterbildung (20.7%). Bei den Themen zeigten sich deutliche geschlechtsspezifische Unterschiede: Männliche Nutzende setzten KI-Chatbots häufiger zur Informationsbeschaffung zu Themen wie Politik, Finanzen sowie Technologie und Wissenschaft ein, während weibliche Nutzende diese tendenziell häufiger für die Themen Ernährung, Gesundheit sowie Haushalt und Alltag nutzten (siehe Abbildung 2). Diese Befunde spiegeln weitgehend bekannte gesellschaftliche Muster wider und sind konsistent mit dem bereits beschriebenen höheren politischen Informationsinteresse männlicher Teilnehmender sowie mit etablierten Befunden zur geschlechtsspezifischen Mediennutzung (Kamper, 2025; Statista, 2021, 2023). Hinsichtlich der allgemeinen Nutzungsmotive wurden KI-Chatbots am häufigsten zur Informationssuche eingesetzt (62.0%), während Unterhaltung den geringsten Stellenwert einnahm (21.2%).

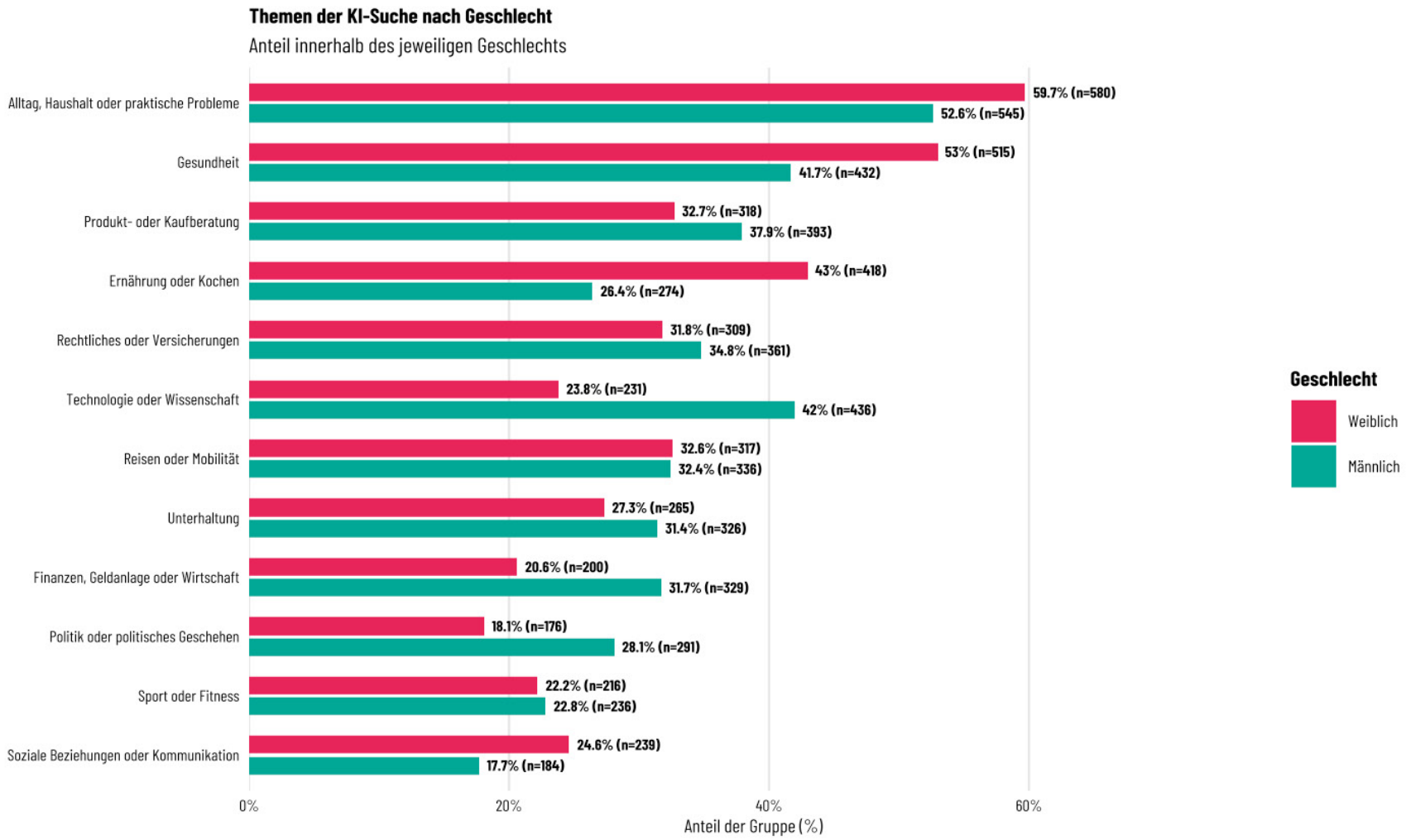

Abbildung 2: Themen der KI-Suche nach Geschlecht

3.2 PROMPTING-STRATEGIEN DER NUTZENDEN

In diesem Kapitel werden die Prompts der Teilnehmenden in den beiden Nutzungsszenarien beschrieben und analysiert. In der Analyse wird zwischen Erstprompts und Folgeprompts unterschieden: Als Erstprompt gilt die initiale Eingabe einer Person zu Beginn einer Konversation, die die ursprüngliche Informationsintention am direktesten widerspiegelt und noch unbeeinflusst von vorherigen Modellantworten ist. Folgeprompts hingegen sind alle nachfolgenden Eingaben innerhalb desselben Gesprächs und entstehen in Reaktion auf die vorangegangene Antwort des Chatbots. Folgeprompts können die ursprüngliche Anfrage vertiefen, präzisieren, korrigieren oder in eine neue Richtung lenken. Diese Unterscheidung ist analytisch bedeutsam, da Erst- und Folgeprompts strukturell und funktional unterschiedlichen Logiken folgen: Während Erstprompts die spontane Formulierungsstrategie einer Person abbilden, spiegeln Folgeprompts die interaktive Dynamik der Mensch-KI-Konversation wider.

3.2.1 THEORETISCHE HERLEITUNG DER PROMPTING-TYPOLOGIE

Die Herleitung der Prompting-Typologie orientierte sich an existierenden Taxonomien aus der Forschung zu Prompting-Verhalten (Xue et al., 2026) sowie zu konversationeller Suche (Radlinski & Craswell, 2017), die u. a. zwischen direkten und indirekten Fragen, Befehlen, Aussagen ohne Handlungsimpuls, informationsorientierten Anfragen sowie ziel- und aufgabenorientierten Suchanfragen unterscheiden. Zudem wurden bisherige Erkenntnisse zur Prompt-Sensitivität von LLMs berücksichtigt, wonach sowohl strukturelle Merkmale – etwa Länge und Satzstruktur – als auch semantisch-pragmatische Dimensionen wie Ton, Höflichkeit und sprachliche Rahmung die Qualität und Ausrichtung generierter Antworten maßgeblich beeinflussen können (u. a. Zhou et al., 2024; Gandhi & Gandhi, 2025; Germani & Spitale, 2025).

Die Analyse der erhobenen Erst- und Folgeprompts zeigt ein klares Muster, das sich in drei grundlegende Strategien einteilen lässt, die in einer Dimension zusammengefasst werden können. Erstens lassen sich sehr kurze Eingaben identifizieren, die in Struktur und Funktion einer Suchmaschinenanfrage ähneln (A). Zweitens finden sich Prompts, die aus einem einzelnen Satz bestehen, die überwiegend als einfache Frage formuliert sind, und auf eine ausführliche kontextuelle Einbettung verzichten (B). Drittens zeigen sich Prompts, die aus mehreren Sätzen bestehen, sich durch einen höheren Grad an Kontextualisierung auszeichnen, dem Modell häufiger eine spezifische Rolle oder Persona zuweisen und detailliertere Instruktionen enthalten (C).

Die Länge eines Prompts stellt die erste zentrale Dimension der Typologie dar. Ergänzend wurde die syntaktische Form der Eingabe berücksichtigt, da diese – unabhängig von der Länge – die pragmatische Struktur einer Mensch-KI-Interaktion widerspiegelt. Unterschieden wurden dabei vier Grundformen: Frage, Aussage, Aufforderung sowie der direkte Ausdruck eines Informationsinteresses. Die Kombination beider Dimensionen – Länge (A, B, C) und syntaktische Form (0–X) – ergibt die finale Prompting-Typologie, die eine differenzierte Beschreibung der Nutzungsstra-

tegien ermöglicht (s. Tabelle 1). So beschreibt beispielsweise A0 einen kurzen Suchbegriff ohne vollständigen Satz, der einer klassischen Suchmaschineanfrage ähnelt, während CX mehrsätzige Prompts bezeichnet, die verschiedene syntaktische Formen kombinieren und damit den höchsten Grad an kommunikativer Komplexität aufweisen.

KATEGORIE	BESCHREIBUNG	BEISPIEL
A0	**Kurzer Suchbegriff** (kein vollständiger Satz, kein Verb, ähnelt einer Suchanfrage)	*„Iran Krieg“*
A1	**Kurze Frage** (kein vollständiger Satz)	*„Die Deutsche Wirtschaftslage?“*
A2	**Kurze Aussage** (kein vollständiger Satz)	*„Klarnamenpflicht ist schlecht“*
A3	**Kurze Aufforderung** (kein vollständiger Satz)	*„Aktuelle Nachrichten bitte“*
B1	**Vollständige Frage** (mit Fragewort oder Fragezeichen)	*„Hallo ChatGPT, was sind die heutigen Themen in der Tagesschau?“*
B2	**Vollständige Aussage** (ein Satz)	*„EU arbeitet an neuen Regeln für KI und Social Media“*
B3	**Vollständige Aufforderung** (ein Satz)	*„Nenne mir bitte die wichtigsten politischen Themen des Tages“*
B4	**Informationsinteresse** ausgedrückt (ein Satz)	*„Ich möchte gerne etwas über die aktuellen Entwicklungen im Iran erfahren“*
C1	Mehrere Fragen	*„Hallo, wie sieht es denn um die aktuelle Situation im Middle East aus? Gibt es dort bald Waffenruhe im Iran? Wie sieht das mit dem Öl aus? Warum macht Amerika das?“*
C2	Mehrere Aussagen	*„Ich würde gerne wissen wie hoch die Wahrscheinlichkeit ist, dass in MV im September die AFD gewinnt. Ich möchte in dieses Bundesland ziehen und das bereitet mir Sorgen.“*
C3	Mehrere Aufforderungen	*„Fasse mir die Ergebnisse der Landtagswahl in Rheinland-Pfalz zusammen. Gehe dabei insbesondere auf Veränderungen im Hinblick auf die vorhergehende Landtagswahl ein.“*
CX	**Kombination aus mehreren Elementen** mit mehreren Sätzen (Fragen, Aussagen und/oder Aufforderungen)	*„Welche Motivation hatte Trump, um den Krieg gegen Iran zu beginnen? Antworte strukturiert, unvoreingenommen und in maximal 5 Sätzen.“*

Tabelle 1: Typologie der Prompting-Strategien

Kategorienübergreifende Merkmale

Ergänzend zur typologischen Klassifikation wurden kategorienübergreifend weitere Merkmale der Prompts klassifiziert, die, gemäß eingangs beschriebener Befunde zur Prompt-Sensitivität, Einfluss auf die Antwortgenerierung nehmen können (s. Tabelle 2).

Auf sprachlich-stilistischer Ebene wurde erfasst, ob ein Prompt eine direkte Ansprache des Modells enthält („Hey ChatGPT…"), höfliche Formulierungen wie „Bitte" oder „Dankeschön" aufweist oder dem Modell explizit eine Rolle oder Persona zuweist („Du bist ein Journalist…"). Sentiment und emotional konnotierte Sprache wurden ebenfalls analysiert, da sie den Ton der Interaktion und potenziell die Antwortausrichtung beeinflussen können[1].

Auf inhaltlicher Ebene wurde analysiert, ob ein Prompt nach Personen, Orten oder Ereignissen fragt sowie ob spezifische Schlagwörter für einen aktuellen Zeitbezug wie „heute" oder „aktuelle" enthalten sind. Diese Merkmale haben sich als für die politische Informationssuche besonders relevant gezeigt, da politische Themen häufig an konkrete Akteure, geografische Kontexte und spezifische Ereignisse geknüpft sind. Besondere Aufmerksamkeit galt zwei weiteren Dimensionen, die sich in den Analysen als zentral für das Antwortverhalten des LLMs erwiesen: Erstens wurde erfasst, ob ein Prompt einen Aktualitätsbezug über den Zeitbezug hinaus aufweist – also ob er auf zeitkritisches, tagesaktuelles politisches Wissen abzielt („Wie ist die Lage im Iran?"), oder nach zeitlosem Hintergrundwissen sowie Grundlagenverständnis fragt („Was ist Völkerrecht?"). Zweitens wurde kodiert, ob es sich um eine Faktenfrage handelt – also ob der Prompt auf eine überprüfbare, eindeutig richtige oder falsche Antwort abzielt („Wie viele Sitze hat die AfD im Bundestag?") oder eher Meinungs- und Interpretationsspielraum lässt („Was ist deine Meinung zur Migrationspolitik?"). Beide Merkmale wurden als kontinuierliche Konfidenzscores kodiert.

Folgeprompt-spezifische Merkmale

Für Folgeprompts wurden zusätzlich vier Merkmale erhoben, die die Beziehung eines Prompts zu seiner unmittelbaren Vorgängeranfrage bzw. der Antwort des LLMs beschreiben (s. Tabelle 2).

Die Bezugnahme erfasst, wie stark ein Folgeprompt thematisch an den vorherigen anknüpft – von einem vollständigen Themenwechsel („Erkläre mir die Inflation" gefolgt von „Was ist die Hauptstadt von Frankreich?") bis zur direkten Weiterführung desselben Themas („Wie hoch ist die aktuelle Inflationsrate?" gefolgt von „Und wie war sie im letzten Jahr?"). Eng damit verwandt ist die Vertiefung, die erfasst, ob ein Folgeprompt das vorangegangene Thema konkretisiert, präzisiert oder eine implizite Rückfrage des Modells aufgreift – etwa wenn auf „Erkläre mir die Schuldenbremse" ein „Und welche Länder haben so etwas auch?" folgt. Während Bezugnahme den Grad der thematischen Kontinuität abbildet – also ob überhaupt auf das vorangegangene Thema eingegangen wird – erfasst Vertiefung den qualitativen Schritt darüber hinaus: ob das Thema aktiv weiterentwickelt, konkretisiert oder präzisiert wird. Jede Vertiefung setzt damit eine Bezugnahme voraus, während eine hohe Bezugnahme nicht zwingend mit einer inhaltlichen Weiterentwicklung einhergeht.

1 Für die Sentiment-Analyse wurde das deutschsprachige SentiWS-Lexikon verwendet (Remus et al., 2010); die Emotionsanalyse basiert auf dem NRC Word-Emotion Association Lexicon (Mohammad & Turney, 2013).

Zuspruch und Widerspruch geben an, ob der Folgeprompt eine explizite Bewertung der vorangegangenen LLM-Antwort enthält – Zuspruch etwa durch „Ja, genau“ oder „Du hast Recht“, Widerspruch durch „Aber das stimmt doch nicht“ oder „Das sehe ich anders“.

DIMENSION	MERKMAL	BESCHREIBUNG	BEISPIEL
Sprachlich-stilistisch	Direkte Ansprache	explizite Adressierung des Modells	*„Hey ChatGPT...“, „Hallo!“*
	Persona-Vorgaben	explizite Rollenzuweisung an das Modell	*„Du bist ein Journalist...“, „Antworte wie ein Experte“*
	Höflichkeitsmarker	höfliche Formulierungen im Prompt	*„Bitte erkläre mir...“, „Danke“*
	Sentiment und emotionale Sprache	emotionale Grundfärbung des Prompts (neutral / positiv / negativ) sowie konnotierte Begriffe für spezifische Emotionen (z. B. Angst, Vertrauen, Empörung)	*„Krieg“, „Konflikt“, „Gewalt“ (negativ)* *„hoch“, „aktuell“, „stark“ (positiv)*
Inhaltlich	Personenbezug	Bezug auf eine konkrete Person oder Akteursgruppe	*„Was macht Merz in der Asylpolitik?“*
	Ortsbezug	geografischer Bezug im Prompt	*„Situation in der Ukraine“*
	Ereignisbezug	Bezug auf ein konkretes, abgegrenztes Ereignis	*„Reaktionen auf den Gaza-Waffenstillstand“*
	Zeitliche Marker	Schlagwörter mit explizitem Zeitbezug	*„aktuell“, „neueste“, „heute“*
	Aktualitätsbezug	Der Prompt erfragt zeitkritisches, tagesaktuelles Wissen vs. zeitloses Hintergrundwissen	*„Wie ist die Lage im Iran?“ (hohe Aktualität)* *„Was ist Völkerrecht?“ (geringe Aktualität)*
	Faktenfrage	Der Prompt zielt auf eine überprüfbar richtige/falsche Antwort ab vs. eine Frage mit vollem Meinungs-/Interpretationsspielraum	*„Wie viele Sitze hat die AfD?“ (Faktenfrage)* *„Was hältst du von der Migrationspolitik?“ (Meinungsfrage)*

Folgeprompt-spezifisch	Bezugnahme	Wie stark der Folgeprompt thematisch an die Antwort des Chatbots anknüpft: vollständiger Themenwechsel kdirekte Weiterführung.	Keine Bezugnahme: Prompt 1: *„Wie steht Walter* Steinmeier zum Irankrieg“ Prompt 2: *„Wie stark ist die AFD in Deutschland“*
	Vertiefung	Wie stark das Thema des vorherigen Prompts konkretisiert oder präzisiert wird.	Keine Vertiefung: Prompt 1: *„Was ist mit dem dritten Weltkrieg oder generell Krieg auf der Welt, was zurzeit los ist?“* Prompt 2: *„Okay, danke schön und was ist jetzt mit Angela Merkel? Ich habe gehört die war mal bei einem Treffen dabei wieder und ich habe gehört Olaf Scholz ist nicht mehr dabei.“*
	Zuspruch		*„Ja, genau“, „Du hast Recht“*
	Widerspruch		*„Das stimmt nicht“, „Das sehe ich anders“*

Tabelle 2: Kategorienübergreifende/Folgeprompt-spezifische Merkmale

3.2.2 HÄUFIGKEITSVERTEILUNG DER PROMPTING-STRATEGIEN

Erstprompts

Die folgende Tabelle gibt einen Überblick über die Häufigkeitsverteilung der Erstprompting-Strategien und zeigt, welche Formulierungstypen Nutzende bei ihrer initialen Eingabe bevorzugen. Die Verteilung der Kategorien spiegelt insgesamt ein realitätsnahes Nutzungsverhalten wider. Der überwiegende Anteil der Erstprompts beider Szenarien zusammen (69.35%) entfiel auf Fragen (A1, B1 und C1), was sich unter anderem durch die Gestaltung der Eingabeoberfläche bei KI-Chatbots erklären lässt: Da das Eingabefeld in ChatGPT zu Beginn der Interaktion explizit zur Formulierung einer Frage auffordert („Stelle irgendeine Frage“ - in der Realität sowie in der Fragebogenumgebung), ist eine entsprechende Tendenz der Nutzenden naheliegend. Analytisch aufschlussreich ist vor diesem Hintergrund jedoch, welche Art von Fragen gestellt werden und in welchem Ausmaß sowie in welcher Form Nutzende von dieser Standardstruktur abweichen. Die zweithäufigste Kategorie bildeten kurze, stichwortartige Eingaben im Stil von Suchmaschinenanfragen (18.65%). Dies erscheint insofern naheliegend, als KI-Chatbots zunehmend klassische

Suchmaschinen ersetzen und Nutzende möglicherweise vertraute Suchmuster auf die Interaktion mit KI übertragen. Dieses Verhalten zeigte sich gleichermaßen für alle Altersgruppen und stellt damit kein generationenspezifisches Phänomen dar.

KATEGORIE	BESCHREIBUNG	OFFENE POLITISCHE INFORMATIONSSUCHE (SZENARIO 1)		VORGEGEBENE POLITISCHE INFORMATIONSSUCHE (SZENARIO 2)	
		ANZAHL	ANTEIL	ANZAHL	ANTEIL
A0	Kurzer Suchbegriff (kein vollständiger Satz, kein Verb, ähnelt einer Suchanfrage)	534	26.70 %	187	9.38 %
A1	Kurze Frage (kein vollständiger Satz)	27	1.35 %	62	3.11 %
A2	Kurze Aussage (kein vollständiger Satz)	26	1.30 %	28	1.40 %
A3	Kurze Aufforderung (kein vollständiger Satz)	7	0.35 %	18	0.90 %
B1	Vollständige Frage (mit Fragewort oder Fragezeichen)	1179	58.95 %	1376	69.01 %
B2	Vollständige Aussage (ein Satz)	14	0.70 %	52	2.61%
B3	Vollständige Aufforderung (ein Satz)	72	3.60 %	87	4.36 %
B4	Informationsinteresse ausgedrückt (ein Satz)	13	0.65 %	8	0.40 %
C1	Mehrere Fragen	56	2.80 %	70	3.51 %
C2	Mehrere Aussagen	4	0.20 %	26	1.30 %
C3	Mehrere Aufforderungen	9	0.45 %	4	0.20 %
CX	Kombination aus mehreren Elementen mit mehreren Sätzen (Fragen, Aussagen und/oder Aufforderungen)	59	2.95 %	76	3.81 %
Summe		2000	100 %	1994	100 %

Tabelle 3: Häufigkeitsverteilung der Erstprompting-Strategien

Folgeprompts

Die Verteilung der Prompting-Strategien in den ersten beiden Folgeprompts der offenen politischen Informationssuche (Szenario 1) ähnelt stark jener der Erstprompts – Fragen dominieren mit über der Hälfte der Eingaben, gefolgt von stichwortartigen Suchanfragen. Ab dem dritten Folgeprompt verschiebt sich das Bild jedoch merklich: Der Anteil der Fragen sinkt weiter auf 49 %, während Aussagen (28.9 %) und Suchanfragen (14.7 %) deutlich zunehmen. Dies deutet darauf hin, dass die Interaktion im Gesprächsverlauf einer charakteristischen Dynamik folgt: Während Erstprompts überwiegend fragenbasiert sind, weichen spätere Eingaben zunehmend von diesem Muster ab.

Nutzende reagieren dann häufiger kommentierend oder resümierend auf vorangegangene Modellantworten oder schließen die Konversation mit kurzen Bestätigungen wie „Ja“ oder „Danke“ ab – ein Muster, das auf eine zunehmende Konvergenz im Gesprächsverlauf hindeutet und weniger auf eine fortgesetzte aktive Informationssuche.

Bei der vorgegebenen politischen Informationssuche (Szenario 2) verlief die Interaktion insgesamt kürzer: Nur 1370 der 2009 Teilnehmenden verfassten Folgeprompts; mit insgesamt 2682 Folgeprompts waren das deutlich weniger als in Szenario 1. Dies könnte unter anderem darauf zurückzuführen sein, dass die Teilnehmenden im zweiten Szenario ein vorgegebenes Thema erhielten, das nicht zwingend ihrer aktuellen intrinsischen Motivation oder ihrem persönlichen Informationsbedürfnis entsprach. Die Bereitschaft, ein Gespräch durch Folgeprompts zu vertiefen, dürfte jedoch maßgeblich davon abhängen, ob ein echtes eigenes Interesse am Thema besteht – fehlt diese intrinsische Motivation, bricht die Interaktion womöglich früher ab.

Die bei der offenen politischen Informationssuche (Szenario 1) beschriebene Verschiebung hin zu Aussagen und weg von Fragen trat bei der vorgegebenen politischen Informationssuche (Szenario 2) bereits früher ein – schon zwischen dem ersten und zweiten Folgeprompt war ein stärkerer Anstieg von Aussagen zu beobachten, der sich im dritten Folgeprompt weiter fortsetzte.

3.2.3 HÄUFIGKEITSVERTEILUNG KATEGORIENÜBERGREIFENDER MERKMALE

Erstprompts

Im Folgenden wird dargestellt, wie häufig die kategorienübergreifenden Merkmale in den Erstprompts auftraten. Die Erstprompts waren überwiegend neutral formuliert und enthielten wenig emotionale Sprache, was dem sachlich-informativen Charakter politischer Informationssuche entspricht. Bei der offenen politischen Informationssuche (Szenario 1) wiesen 68.8 % der Prompts ein neutrales Sentiment auf, wobei der Anteil negativer Formulierungen (25.2 %) auf die thematische Rahmung zurückzuführen ist – viele Prompts bezogen sich auf Krieg und Konflikt.

Bei der vorgegebenen politischen Informationssuche (Szenario 2) war der Anteil neutraler Sentiments mit 76.4 % noch höher. Emotionale Färbung war insgesamt selten: In Szenario 1 enthielten 15.2 % der Prompts emotional konnotierte Begriffe und in Szenario 2 waren es 15.8 %, mit Vertrauen als dominanter Emotion (S1: 6.1 %, S2: 10.7 %).

Hinsichtlich sprachlich-stilistischer Merkmale zeigt sich über beide Szenarien hinweg, dass direkte Ansprache, höfliche Formulierungen und Persona-Vorgaben äußerst selten auftraten. Bei der offenen politischen Informationssuche (Szenario 1) enthielten lediglich 2.3 % der Erstprompts eine direkte Ansprache, 0.5 % höfliche Formulierungen und 0.4 % eine explizite Rollenzuweisung an das Modell; bei der vorgegebenen politischen Informationssuche (Szenario 2) lagen diese Anteile mit 1.2 %, 0.2 % und 0.1 % noch darunter. Eine Ausnahme bildet der zeitliche Bezug: 19.4 % der Erstprompts in Szenario 1 enthielten Formulierungen wie „aktuell“ oder „neueste“, in Szenario 2 waren es 5.1 % – ein Unterschied, der darauf zurückzuführen ist, dass Nutzende bei einer offenen, selbstinitiierten Suche häufiger gezielt nach aktuellen Nachrichten und Entwicklungen fragen.

Über alle Erstprompts hinweg zeigt sich, dass Aktualitätsbezug (64.7 %) und Faktenfrage-Charakter[2] (74.8 %) die am weitesten verbreiteten inhaltlichen Merkmale darstellen, während Personen- (11.6 %), Orts- (32.3 %) und Ereignisbezüge (31.2 %) seltener auftraten. Zwischen den Szenarien zeigen sich dabei deutliche Unterschiede: Bei der offenen politischen Informationssuche waren Ereignisbezüge (S1: 53.9 %, S2: 8.5 %), Ortsbezüge (S1: 46.0 %, S2: 18.5 %) und insbesondere Aktualitätsbezüge (S1: 76.9 %, S2: 52.6 %) deutlich häufiger als bei der vorgegebenen politischen Informationssuche, was auf den stärker tagesaktuellen und ereignisbezogenen Charakter der offenen Informationssuche hindeutet. Zum Erhebungszeitpunkt dürfte dabei der zuvor ausgebrochene Irankrieg (Ende Februar 2026) eine Rolle gespielt haben, der die Informationsinteressen der Teilnehmenden in Szenario 1 maßgeblich prägte. Der Faktenfrage-Charakter war in beiden Szenarien hoch (S1: 80.7 %, S2: 68.8 %), was darauf hindeutet, dass Nutzende KI-Chatbots bei der politischen Informationssuche überwiegend zur Beantwortung faktischer Fragen einsetzen – unabhängig vom konkreten Thema. Innerhalb der häufigsten Kategorie (B1) war der Anteil der Faktenfragen noch höher (S1: 89.0 %, S2: 77.9 %).

Folgeprompts

Im Verlauf der Konversation der Nutzenden mit dem KI-Chatbot zeigt sich über alle Folgeprompts und Szenarien hinweg ein klarer Rückgang der inhaltlichen Merkmalsausprägungen. Bereits beim zweiten Prompt sind Aktualitätsbezug (S1: 49.5 %, S2: 24.9 %) und Faktenfrage-Charakter (S1: 48.5 %, S2: 34.3 %) deutlich niedriger als bei den Erstprompts. Dieser Trend setzt sich beim dritten Prompt weiter fort, wo die Werte auf ein sehr niedriges Niveau absinken (S1 Aktualität: 27.0 %, S2: 10.4 %). Beim vierten Prompt – für den aufgrund des natürlichen Gesprächsabbruchs nur noch Daten von 570 bzw. 313 Personen vorliegen – stabilisieren sich die Werte auf niedrigem Niveau, wobei Aktualitätsbezug in Szenario 1 mit 43.9 % noch am stärksten ausge-

2 Aktualitätsbezug und Faktenfrage wurden als kontinuierliche Konfidenzscores (0–1) erhoben; als Schwellenwert für das Vorliegen des Merkmals wurde 0.5 gewählt.

prägt bleibt. Szenario 2 weist durchgehend niedrigere Werte als Szenario 1 auf. Dieser Verlauf ist konsistent mit dem bereits beschriebenen Muster, dass Folgeprompts zunehmend weniger auf neue Informationsanfragen ausgerichtet sind und stattdessen stärker auf die vorangegangene Modellantwort reagieren – die inhaltliche Spezifität der Eingaben nimmt also im Gesprächsverlauf systematisch ab.

Weitere folgeprompt-spezifische Merkmale

Hinsichtlich der folgeprompt-spezifischen Merkmale zeigt sich, dass Nutzende ihre Folgeprompts überwiegend auf die vorangegangene Modellantwort beziehen: Die Bezugnahme[3] liegt in beiden Szenarien konsistent bei über 60 %, mit einer interessanten Verlaufsdynamik: bei der offenen politischen Informationssuche (Szenario 1) sinkt sie bei Folgeprompt 2 leicht ab (56.5 %) und steigt bei Folgeprompt 3 wieder auf den Ausgangswert (67.5 %), während sie in der vorgegebenen politischen Informationssuche (Szenario 2) über die Folgeprompts hinweg kontinuierlich zunimmt (60.6 % → 78.0 %). Zuspruch und Widerspruch gegenüber der Modellantwort sind durchgehend selten (jeweils unter 7 %), was darauf hindeutet, dass Nutzende das Modell kaum explizit korrigieren oder bestätigen. Der auffälligste Befund ist der systematische Rückgang der Vertiefung: Während beim ersten Folgeprompt noch 56.2 % (Szenario 1) bzw. 52.2 % (Szenario 2) der Eingaben eine inhaltliche Vertiefung darstellten, sinkt dieser Anteil bis zum dritten Folgeprompt auf 13.7 % (Szenario 1) bzw. 7.8 % (Szenario 2) – ein Muster, das den bereits beschriebenen Übergang von aktiver Informationssuche hin zu kurzen Abschlussreaktionen empirisch bestätigt.

3.2.4 KONSISTENZ DER PROMPTING-STRATEGIEN

Abschließend wurde untersucht, inwieweit Nutzende über verschiedene Prompts und Szenarien hinweg eine konsistente Eingabestrategie verfolgen – also ob dieselbe Person dazu neigt, KI-Chatbots stets auf ähnliche Weise anzusprechen und mit ihnen zu interagieren, oder ob die gewählte Formulierungsstrategie für Prompts je nach Gesprächsverlauf und Thema variiert.

Über die beiden Szenarien hinweg lag die Übereinstimmung bei Erstprompts bei 55.5 % und stieg bei späteren Folgeprompts auf bis zu 72.8 % an – ein Muster, das plausibel ist, da spätere Folgeprompts kürzer und weniger inhaltsspezifisch sind und damit strukturell stärker konvergieren. Innerhalb eines Szenarios nimmt die Konsistenz mit zunehmendem Abstand zwischen den Prompts ab: In Szenario 1 stimmten direkt aufeinanderfolgende Prompts noch zu 49.0–64.2 % überein, während Erst- und vierter Prompt nur noch zu 28.2 % übereinstimmten. Bei der vorgegebenen politischen Informationssuche (Szenario 2) zeigte sich derselbe Trend noch ausgeprägter (13.5 % zwischen Prompt 1 und 4). Dies deutet darauf hin, dass Nutzende ihre Prompting-Strategie im Gesprächsverlauf anpassen, anstatt konsistent dieselbe Struktur beizubehalten. Auf Personenebene verwendeten Nutzende ihre dominante Kategorie im Durchschnitt in 61 % aller Prompts, wobei A0 mit Abstand die häufigste dominante Kategorie war (n = 1369 — hier ist

3 Bezugnahme und Vertiefung wurden als kontinuierliche Konfidenzscores (0–1) erhoben; als Schwellenwert für das Vorliegen des Merkmals wurde 0.5 gewählt.

zu beachten, dass A0 aufgrund seiner Kürze auch als Folgeprompt häufig auftritt und somit oft die dominante Kategorie ausmacht), gefolgt von B1 (n = 559). Insgesamt legen die Befunde nahe, dass Prompting-Strategien zwar eine gewisse individuelle Stabilität aufweisen, aber situativ und gesprächsdynamisch variieren.

3.3 INFERENZSTATISTISCHE ERGEBNISSE

3.3.1 ZUSAMMENHÄNGE ZWISCHEN PROMPTING-STRATEGIEN UND NUTZENDENGRUPPEN

Zur Untersuchung von Zusammenhängen zwischen Nutzendenmerkmalen und Prompting-Verhalten wurden Chi²-Tests mit Cramér's V als Effektstärkemaß auf Personenebene berechnet. Da jede Person bis zu acht Prompts beisteuerte und Prompts innerhalb einer Person denselben individuellen Interaktionsstil widerspiegelten, wurde für die Analysen der dominante Prompt-Typ je Person bestimmt – also derjenige Typ, den eine Person am häufigsten über alle Prompts und Szenarien hinweg verwendete – und auf Basis dieser aggregierten Daten gearbeitet (n = 2009). Als Prompt-Typen wurden dabei unterschieden: kurze Suchbegriffe (A0), Fragen (A1, B1 und C1), Aussagen (A2, B2, B4 und C2), Aufforderungen (A3, B3 und C3) sowie die Kombinations-Kategorie (CX). Da die Verteilung des dominanten Prompt-Typs stark ungleich war (Frage: 77.2%, Suche: 13.5%, Aussage: 6.6%, Aufforderung: 1.7%, Kombi: 1.0%), wurden binäre Chi²-Tests je Prompt-Typ berechnet, um Zellenbesetzungsprobleme bei seltenen Kategorien zu vermeiden.

Thematische Nutzungsvielfalt erwies sich als robustester Prädiktor: Personen mit geringerer Nutzungsvielfalt tendierten häufiger zu Suchmaschinen-Prompts ($V = .130$, $p < .001$), während breitere Nutzung mit Frage-Prompts assoziiert war ($V = .109$, $p < .001$). Dieser Befund ist inhaltlich konsistent mit der Annahme, dass zunehmende Erfahrung im Umgang mit LLMs zu elaborierterem Prompting-Verhalten führt. Altersgruppe zeigte den stärksten demografischen Effekt ($V = .119$, $p < .001$): Jüngere Nutzende verwendeten häufiger Aufforderungs-Prompts als dominanten Stil, was auf einen direktiven Interaktionsstil hindeutet, der bei jüngeren Generationen stärker ausgeprägt ist. KI-Vertrauen zeigte signifikante Zusammenhänge mit Aufforderungs-Prompts ($V = .067$, $p < .05$) sowie Frage-Prompts ($V = .057$, $p < .05$): Personen mit höherem KI-Vertrauen tendierten eher zu Frage-Prompts, während skeptischere Nutzende häufiger direktive Aufforderungen formulierten. Insgesamt waren die Effektstärken durchgehend klein ($V \leqslant .13$), was darauf hindeutet, dass Prompting-Verhalten nicht primär durch Nutzendenmerkmale bestimmt wird, sondern vermutlich stärker vom situativen Aufgabenkontext abhängt.

3.3.2 ZUSAMMENHÄNGE ZWISCHEN PROMPTS UND AUSGELÖSTER WEB-SUCHE

Im folgenden Abschnitt wird untersucht, unter welchen Bedingungen die eingegebenen Prompts eine LLM-seitige Websuche auslösten und welche Merkmale dabei als Prädiktoren identifiziert werden konnten. Die Analyse umfasst dabei den Erstprompt sowie bis zu drei Folgeprompts je Person und Szenario, sofern vorhanden.

Zunächst ist festzuhalten, dass 43.3% aller Prompts eine Websuche auslösten, wobei Szenario 1 eine deutlich höhere Websuche-Rate aufwies als Szenario 2 (s. Tabelle 4), und Erstprompts generell häufiger eine Websuche triggerten als Folgeprompts. Dies deutet darauf hin, dass die initiale Anfrage einer Person stärker auf aktuelle oder faktenbezogene Informationen ausgerichtet ist als alle nachfolgenden Prompts im selben Gespräch.

BEREICH	ART DES PROMPTS	ANZAHL WEBSUCHEN	GESAMTE BEOBACHTUNGEN	ANTEIL (%)
Szenario 1	Allgemein	2852	5243	54.4
	Erstprompts	1339	2000	67.0
	Folgeprompts	1513	3243	46.7
Szenario 2	Allgemein	1318	4381	30.1
	Erstprompts	748	1994	37.5
	Folgeprompts	570	2387	23.9

Tabelle 4: Übersicht der Websuchen

Um zu untersuchen, welche Eigenschaften eines Prompts dazu führen, dass das Modell eine Websuche auslöst, wurde eine logistische Regression berechnet, in die Aktualitäts-Score und Faktenfrage-Score als Vorhersage-Variablen eingingen. Beide Faktoren erwiesen sich als signifikante Prädiktoren: Aktualität war der stärkste Prädiktor (OR = 30.45, 95% CI [24.76, 37.63], $p < .001$), gefolgt von Faktenfrage (OR = 5.65, 95% CI [4.34, 7.38], $p < .001$). Das Modell wies eine gute Modellgüte auf (McFadden $R^2 = .362$). In einem erweiterten Modell wurde zusätzlich ein expliziter Zeitbezug im Prompt (z. B. „aktuell“, „heute“, „neueste“) als Prädiktor aufgenommen. Auch unter Kontrolle von Aktualitäts-Score und Faktenfrage hatte ein expliziter Zeitbezug einen eigenständigen Effekt auf die Websuche-Wahrscheinlichkeit (OR = 2.43, 95% CI [1.69, 3.60], $p < .001$), was darauf hindeutet, dass das LLM beide Signale unabhängig voneinander verarbeitet. Aktualität blieb dabei der stärkste Prädiktor (OR = 25.56, 95% CI [20.63, 31.82], $p < .001$), gefolgt von Faktenfrage (OR = 5.26, 95% CI [4.03, 6.88], $p < .001$). Die Modellgüte verbesserte sich geringfügig gegenüber M1 (McFadden $R^2 = .366$, $\Delta AIC = -23.2$), sodass dieses Modell als finales Modell gewählt wurde (vgl. Tabelle A1 im Anhang). Ein Modell mit Interaktionsterm zwischen Aktualität und Zeitbezug zeigte keine signifikante Verbesserung ($p = .165$) und wird daher nicht weiter berichtet. Besonders anschaulich wird der Effekt des Zeitbezugs deskriptiv: In Szenario 1 lösten Prompts mit explizitem Zeitbezug in 95.4% der Fälle eine Websuche aus – verglichen mit 60.1% bei Prompts ohne Zeitbezug. Der Zeitbezug wirkt damit nahezu als hinreichende Bedingung. In Szenario 2 replizierte sich dieser Befund auf niedrigerem Niveau (83.3% vs. 35.0%), was konsistent mit der

generell niedrigeren Websuche-Rate in diesem Szenario ist und darauf hindeutet, dass das Thema in Szenario 2 als weniger zeitkritisch eingestuft wurde.

Diese Befunde sind insgesamt inhaltlich plausibel: Ein LLM greift dann auf externe Webquellen zurück, wenn der Prompt Signale enthält, die das Modell als Hinweis auf zeitkritische oder faktisch überprüfbare Informationen interpretiert. Gleichzeitig werfen die Ergebnisse eine kritische Frage auf: Der Umkehrschluss bedeutet, dass Prompts ohne Aktualitätsbezug – also insbesondere Meinungsfragen oder kontextuelle Hintergrundfragen – in der Mehrheit der Fälle keine Websuche auslösen und das Modell damit ausschließlich auf sein Trainingswissen zurückgreift. Gerade im politischen Kontext kann dies problematisch sein, da solche Fragen oft keine eindeutig verifizierbaren Antworten haben und das Modell dennoch ohne Rückgriff auf externe Quellen antwortet. Dadurch besteht das Risiko, dass politische Rahmungen oder Biases aus den Trainingsdaten unreflektiert reproduziert werden (vgl. Ceron et al., 2025; Motoki et al., 2024). Allerdings ist bis heute umstritten, ob eine Websuche-Funktion die Nutzende bei faktischen Suchanfragen schützt, denn sie bietet keine Garantie für faktische Korrektheit (Elliott, 2025; BBC & Ipsos, 2025). Erschwerend kommt hinzu, dass die meisten LLM-gestützten Anwendungen ihr Websuchverhalten nicht transparent offenlegen – weder wird ersichtlich, nach welchen Kriterien eine Websuche ausgelöst wird, noch wie die tatsächlich verwendeten Quellen vom Modell ausgewählt und gewichtet werden. Dies erschwert es Nutzenden erheblich, die Grundlage einer generierten Antwort kritisch einzuschätzen, und stellt eine zentrale Herausforderung für die Informationskompetenz im Umgang mit KI-Chatbots dar (vgl. Shi et al., 2025).

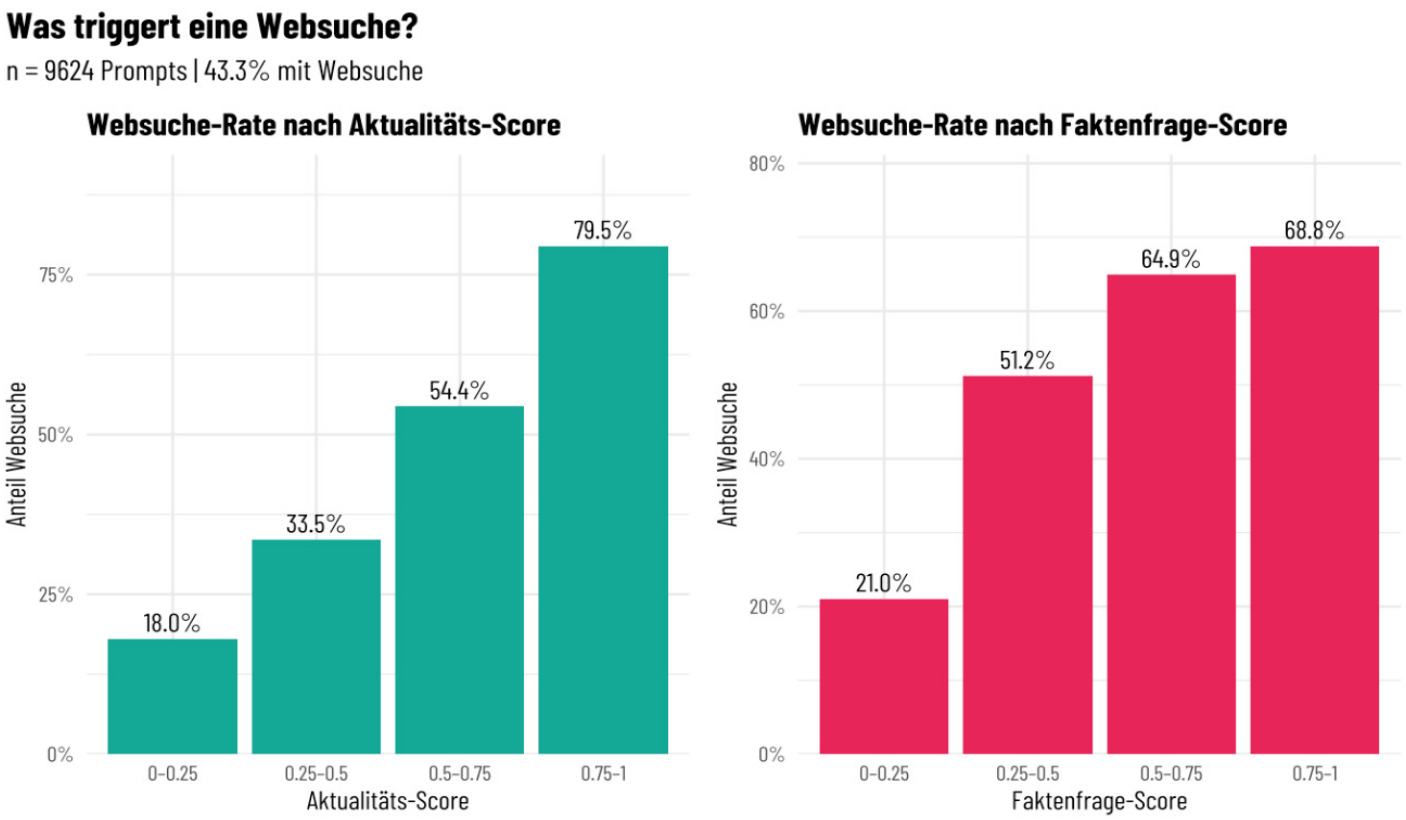

Abbildung 3: Auslöser einer Websuche

4. LIMITATIONEN DER STUDIE

Die vorliegende Studie besitzt einige Limitationen, die bei der Interpretation der Befunde berücksichtigt werden sollten. Erstens wurden die Prompting-Strategien in nachgestellten Szenarien mit zum Teil vorgegebenen Themen erhoben, was die ökologische Validität der Ergebnisse begrenzt. In realen Nutzungssituationen dürften Prompts stärker durch persönliches Vorwissen, bestehende Meinungen und genuine Informationsinteressen geprägt sein – Faktoren, die sowohl die Struktur der Erstprompts als auch die Dynamik der Folgeprompts beeinflussen. So könnten Nutzende bei Themen mit hohem Vorwissen oder ausgeprägten Vorannahmen eher dazu neigen, der Modellantwort zuzustimmen oder zu widersprechen, was sich in den hier erhobenen Prompts kaum abbildet. Zudem ist davon auszugehen, dass sich Prompting-Strategien je nach Thema stärker unterscheiden, sodass die Befunde nicht ohne Weiteres auf andere Themenbereiche verallgemeinert werden können.

Schließlich unterliegt die Studie einer deutlichen zeitlichen Limitation: Der Markt generativer KI-Anwendungen entwickelt sich rasant, sodass sowohl die hier beobachteten Nutzungspräferenzen als auch das Verhalten der untersuchten Modelle bereits zum Zeitpunkt der Veröffentlichung nicht mehr dem aktuellen Stand entsprechen könnten. Die Befunde sind daher als Momentaufnahme eines sich schnell wandelnden Feldes zu verstehen.

5. ZUSAMMENFASSUNG

Die vorliegende Studie untersuchte, wie Nutzende generative KI zur politischen Informationssuche einsetzen und welche Prompting-Strategien sie dabei verwenden. Die Befunde liefern erstmals empirisch fundierte Einblicke in reale Nutzungspraktiken und erlauben Rückschlüsse auf die Schnittstelle zwischen menschlicher Informationsintention und algorithmischer Antwortgenerierung.

Hinsichtlich der Nutzungsmuster bestätigt die Studie, dass generative KI zunehmend als Informations- und Recherchetool eingesetzt wird, was konsistent mit bestehenden Befunden zur KI-Nutzung in Deutschland ist (Bitkom, 2025; Egan et al., 2026). Übereinstimmend mit Reiss et al. (2025) zeigt sich dabei, dass politische Themen zwar ein relevantes, aber nicht dominantes Nutzungsfeld von KI-Chatbots darstellen. Die beobachteten geschlechts- und altersspezifischen Unterschiede in der Nutzungsintensität und thematischen Ausrichtung spiegeln dabei gesellschaftlich verbreitete Muster wider, die auch aus der klassischen Mediennutzungsforschung bekannt sind (Kamper, 2025; Statista, 2021, 2023), und deuten darauf hin, dass sich bestehende Ungleichheiten im politischen Informationsverhalten durch KI-Chatbots reproduzieren.

Die Analyse der Prompting-Strategien zeigt, dass Nutzende KI-Chatbots überwiegend mit einfachen, fragenbasierten Eingaben ansprechen und dabei kaum von elaborierten Prompting-Techniken – wie Rollenzuweisungen, höflichen Formulierungen oder mehrsätzigen Kontextualisierungen – Gebrauch machen. Dies steht im Kontrast zu den in der Forschung und Praxis diskutierten Möglichkeiten des sogenannten Prompt Engineerings (Sahoo et al., 2024) und deutet darauf hin, dass die Mehrheit der Nutzenden KI-Chatbots ähnlich wie Suchmaschinen behandelt – als reaktive Informationsquelle, nicht als dialogischen Partner. Gleichzeitig variieren die Prompting-Strategien situativ und gesprächsdynamisch: Die Konsistenzanalyse zeigt, dass individuelle Stilpräferenzen zwar erkennbar sind (dominante Kategorie im Schnitt 61% der Prompts), aber im Gesprächsverlauf deutlich abnehmen. Dies legt nahe, dass Prompting-Verhalten weniger eine stabile individuelle Eigenschaft als vielmehr eine kontextabhängige Interaktionspraxis darstellt.

Besonders aufschlussreich sind die Befunde zur LLM-seitigen Websuche: Aktualitätsbezug und Faktenfrage-Charakter erweisen sich als die zentralen Auslöser, während Meinungs- und Orientierungsfragen überwiegend ohne externen Quellenabgleich beantwortet werden.

6. FAZIT

Die Studie zeigt, dass generative KI im politischen Informationskontext von den meisten Nutzenden als einfaches Recherchewerkzeug eingesetzt wird – mit eher kurzen, fragenbasierten Eingaben und wenig elaboriertem Prompting. Die Qualität und Verlässlichkeit der generierten Antworten hängen jedoch maßgeblich davon ab, wie ein Prompt formuliert ist und ob das Modell eine Websuche auslöst. Da beides für Nutzende weitgehend intransparent ist, besteht ein strukturelles Informationsrisiko, das besonders im politischen Kontext demokratierelevante Implikationen hat. Für Medienkompetenz und politische Bildung ergibt sich daraus eine neue Aufgabe: Nutzende müssen nicht nur lernen, KI-Outputs kritisch zu hinterfragen, sondern auch verstehen, wie ihre eigenen Eingaben die Antworten formen – und wann ein Modell auf aktuelle Quellen zurückgreift und wann nicht.

Die Befunde zur LLM-seitigen Websuche haben weitreichende Implikationen für die politische Informationsqualität. Im politischen Kontext ist kritisch zu bewerten, dass gerade Meinungs- und Orientierungsfragen – die besonders anfällig für die Reproduktion von Trainingsbiases sind (Ceron et al., 2025; Motoki et al., 2024) – überwiegend ohne externen Quellenabgleich beantwortet werden. Für Nutzende ist dabei in der Regel nicht erkennbar, warum das Modell bei bestimmten Prompts eine Websuche durchführt und bei anderen nicht – und damit auch nicht, ob eine Antwort auf aktuellen Quellen oder ausschließlich auf Trainingsdaten basiert.

Diese Intransparenz wird durch die Möglichkeit der Personalisierung zusätzlich verschärft: Im produktiven Einsatz verfügen KI-Chatbots über Kontextinformationen wie Nutzungshistorie, gespeicherte Präferenzen oder aktuell prominente Themen, die das Antwortverhalten beeinflussen

können. Ob und wie sich Antworten je nach individuellem Nutzungsprofil unterscheiden – etwa ob Nutzende mit wiederholten Anfragen zu einem bestimmten politischen Thema andere Antworten erhalten als bei einer kontextlosen Erstanfrage – entzieht sich damit sowohl der Kontrolle der Nutzenden als auch einer systematischen externen Überprüfung. Im politischen Kontext ist dies besonders relevant, da personalisierte Antworten die Informationswahrnehmung individuell verzerren könnten, ohne dass dies für Nutzende oder externe Beobachtende erkennbar wäre.

Eine zentrale Herausforderung für die Forschung – und für eine mögliche Regulierung – stellt die mangelnde Transparenz der meistgenutzten LLMs dar. Da es sich überwiegend um proprietäre, geschlossene Systeme handelt, deren Trainingsdaten, Gewichtungen und interne Verarbeitungslogik nicht öffentlich zugänglich sind, lässt sich nicht systematisch nachvollziehen, welche weiteren – hier nicht erfassten – Prompt- oder Kontextmerkmale das Antwortverhalten beeinflussen. Anders als bei vielen methodischen Einschränkungen, die durch ein verändertes Studiendesign behoben werden könnten, handelt es sich hierbei um eine strukturelle Limitation, die im Kern den Untersuchungsgegenstand selbst betrifft. Diese fehlende Nachvollziehbarkeit ist nicht nur ein methodisches Problem der vorliegenden Studie, sondern ein grundlegendes Hindernis für Transparenz, externe Kontrolle und letztlich auch für eine evidenzbasierte Regulierung dieser Systeme im politischen Kontext.

7. LITERATUR UND QUELLENVERZEICHNIS

Ashley, S., Maksl, A., & Craft, S. (2013). Developing a news media literacy scale. Journalism & Mass Communication Educator, 68(1), 7-21. https://doi.org/10.1177/1077695812469802

BBC & Ipsos. (2025). Audience use and perceptions of AI assistants for news: Research findings. BBC. https://www.bbc.co.uk/aboutthebbc/documents/audience-use-and-perceptions-of-ai-assistants-for-news.pdf

Benert, V., & Timmermann, S. (2026). Meinungsbildung im Wandel: Wie KI-Systeme Quellen bei der Nachrichtensuche selektieren. Agora Digitale Transformation gGmbH. https://doi.org/10.5281/zenodo.21162562

Bitkom (2025). Internet-Suche im Wandel: Die Hälfte nutzt bereits KI-Chats. https://www.bitkom.org/Presse/Presseinformation/Internet-Suche-Wandel-Haelfte-nutzt-KI-Chats

Ceron, T., Nikolaev, D., Stammbach, D., & Nozza, D. (2025). What Is The Political Content in LLMs' Pre-and Post-Training Data?. arXiv preprint arXiv:2509.22367. https://doi.org/10.48550/arXiv.2509.22367

Clifford, I., Kluzer, S., Troia, S., Jakobsone, M., & Zandbergs, U. (2020). DigCompSat: A self-reflection tool for the European digital competence framework for citizens (M. Cabrera, J. Castaño, C. Centeno, W. O'Keeffe, Y. Punie, & R. Vuorikari, Eds.). Publications Office of the European Union. https://doi.org/10.2760/77437

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20(1), 37–46. https://doi.org/10.1177/001316446002000104

Die Medienanstalten (2025). Integration von KI-Anwendungen in Suchmaschinen und ihre Auswirkungen auf die Meinungsvielfalt. https://www.die-medienanstalten.de/service/gutachten/ki-suchmaschinen/

Dobariya, O., & Kumar, A. (2025). Mind your tone: Investigating how prompt politeness affects LLM accuracy. arXiv preprint arXiv:2510.04950. https://doi.org/10.48550/arXiv.2510.04950

Egan, J., Robertson, C.T., Ross Arguedas, A., Newman, N., Klein Nielsen, R., Mukherjee, M. & Fletcher, R. (2026). Reuters Institute Digital News Report 2026. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2026-06/DNR%202026%20FINAL_2.pdf

Elliott, O. (2025). Representation of BBC News content in AI assistants. BBC Responsible AI Team. https://www.bbc.co.uk/aboutthebbc/documents/bbc-research-into-ai-assistants.pdf

Gandhi, V., & Gandhi, S. (2025). Prompt sentiment: The catalyst for llm change. arXiv preprint arXiv:2503.13510. https://doi.org/10.48550/arXiv.2503.13510

Germani, F., & Spitale, G. (2025). Source framing triggers systematic bias in large language models. Science Advances, 11(45), eadz2924. https://doi.org/ 10.1126/sciadv.adz2924

Hwang, Y., Lee, D., Kang, T., Lee, M., & Jung, K. (2026). When wording steers the evaluation: Framing bias in LLM judges. arXiv preprint arXiv:2601.13537. https://doi.org/10.48550/arXiv.2601.13537

Kaiser, C., Buder, F. & Biró, T. (2024). ChatGPT and co. in everyday life: Use, evaluation, and visions for the future. A three-country comparison. NIMpulse 7. https://www.nim.org/fileadmin/PUBLIC/3_NIM_Publikationen/NIM-Studien/NIMpulse/2024/EN/241007_NIMpulse7_ChatGPT_fin_en.pdf

Kamper, P. (2025). Politik ist nur etwas für Männer!?–Der Einfluss politischer Geschlechterrolleneinstellungen auf das politische Interesse, das politische Wissen und die politische Selbstwirksamkeit von Kindern. Zeitschrift für Politikwissenschaft, 35(2), 163-192. https://doi.org/10.1007/s41358-025-00407-y

Kirsten, E., Perdekamp, J. G., Upadhyay, M., Gummadi, K. P., & Zafar, M. B. (2025). Characterizing web search in the age of generative AI. arXiv preprint arXiv:2510.11560. https://doi.org/10.48550/arXiv.2510.11560

Makhortykh, M., Baghumyan, A., Vziatysheva, V., Sydorova, M., & Kuznetsova, E. (2024). LLMs as information warriors? Auditing how LLM-powered chatbots tackle disinformation about Russia's war in Ukraine. arXiv preprint arXiv:2409.10697. https://doi.org/10.48550/arXiv.2409.10697

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. Psychological Methods, 17(3), 437–455. https://doi.org/10.1037/a0028085

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. Computational Intelligence, 29(3), 436-465. https://doi.org/10.1111/j.1467-8640.2012.00460.x

Motoki, F., Pinho Neto, V., & Rodrigues, V. (2024). More human than human: Measuring ChatGPT political bias. Public Choice, 198(1), 3-23. https://doi.org/10.1007/s11127-023-01097-2

OpenAI. (2026). *ChatGPT (GPT-4.1-mini-2025-04-14)* [Großes Sprachmodell]. https://chat.openai.com/

Patel, A., Lee, F., Liang, K., & Thomas, J. (2026). The role of emotional stimuli and intensity in shaping Large Language Model behavior. arXiv preprint arXiv:2604.07369. https://doi.org/10.48550/arXiv.2604.07369

Python Software Foundation. (2026). *Python (Version 3.14.4)* [Computersoftware]. https://www.python.org/

Qwen Team. (2025). Qwen3 technical report. arXiv preprint arXiv:2505.09388. https://doi.org/10.48550/arXiv.2505.09388

R Core Team. (2026). R: A language and environment for statistical computing (Version 4.5.3) [Computersoftware]. R Foundation for Statistical Computing. https://doi.org/10.32614/R.manuals https://www.R-project.org/

Radlinski, F., & Craswell, N. (2017). A theoretical framework for conversational search. In Procee-

dings of the 2017 conference on conference human information interaction and retrieval (pp. 117-126). https://doi.org/10.1145/3020165.3020183

Ranaldi, L., & Pucci, G. (2023). When large language models contradict humans? Large language models' sycophantic behaviour. arXiv preprint arXiv:2311.09410. https://doi.org/10.48550/arXiv.2311.09410

Reiss, M. V., Knor, E. L., Stöwing, E., Merten, L., & Möller, J. (2025). Zwischen Neugier und Skepsis: Nutzung und Wahrnehmung generativer KI zur Informationssuche in Deutschland. (Arbeitspapiere des Hans-Bredow-Instituts, 76). Hamburg: Verlag Hans-Bredow-Institut. https://doi.org/10.21241/ssoar.100907

Remus, R., Quasthoff, U., & Heyer, G. (2010). SentiWS – A publicly available German-language resource for sentiment analysis. In Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10). European Language Resources Association. https://aclanthology.org/L10-1339/

Reuters Institute for the Study of Journalism. (2023). Methodology. In Digital News Report 2023. University of Oxford. https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023/methodology

Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). A systematic survey of prompt engineering in large language models: Techniques and applications. arXiv preprint arXiv:2402.07927. https://doi.org/10.48550/arXiv.2402.07927

Schinkels, P. (2026, 5. März). Nutzer boykottieren ChatGPT – und bejubeln Claude. *Die Zeit*. https://www.zeit.de/digital/internet/2026-03/kuenstliche-intelligenz-openai-chatgpt-anthropic-claude-meta

Sharma, M., Tong, M., Korbak, T., Duvenaud, D., Askell, A., Bowman, S. R., Cheng, N., Durmus, E., Hatfield-Dodds, Z., Johnston, S. R., Kravec, S., Maxwell, T., McCandlish, S., Ndousse, K., Rausch, O., Schiefer, N., Yan, D., Zhang, M., & Perez, E. (2023). Towards understanding sycophancy in language models. arXiv preprint arXiv:2310.13548. https://doi.org/10.48550/arXiv.2310.13548

Shi, X., Liu, J., Liu, Y., Cheng, Q., & Lu, W. (2025). Know where to go: Make LLM a relevant, responsible, and trustworthy searchers. Decision Support Systems, 188, 114354. https://doi.org/10.1016/j.dss.2024.114354

Statista (2021). Wie stark interessieren Sie sich für Politik? (nach Geschlecht). https://de.statista.com/statistik/daten/studie/1259191/umfrage/politikinteresse-nach-geschlecht-in-deutschland/

Statista (2023). Bevölkerung in Deutschland nach Informationsinteresse an Politik im Geschlechtervergleich im Jahr 2023. https://de.statista.com/statistik/daten/studie/827088/umfrage/umfrage-in-deutschland-zum-informationsinteresse-an-politik-nach-geschlecht/

Statista. (2024). Meistgenutzte KI-Anwendungen in Deutschland im Jahr 2024. https://de.statista.com/statistik/daten/studie/1602910/umfrage/genutzte-ki-anwendungen-in-deutschland/

Statista. (2025). Häufigkeit der Nutzung von KI-Tools in Deutschland im Jahr 2025. https://de.statista.com/statistik/daten/studie/1538331/umfrage/haeufigkeit-der-nutzung-generativer-ki/

The Guardian (2025). AI summaries cause 'devastating' drop in audiences, online news media told. https://www.theguardian.com/technology/2025/jul/24/ai-summaries-causing-devastating-drop-in-online-news-audiences-study-finds

Vinay, R., Spitale, G., Biller-Andorno, N., & Germani, F. (2025). Emotional prompting amplifies disinformation generation in AI large language models. Frontiers in Artificial Intelligence, 8, 1543603. https://doi.org/10.3389/frai.2025.1543603

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with ChatGPT. In PLoP '23: Proceedings of the 30th Conference on Pattern Languages of Programs. https://doi.org/10.5555/3721041.3721046

YouGov (2025). Neue Normalität? So häufig nutzen die Deutschen KI-Tools für private Zwecke. https://yougov.de/technology/articles/52805-neue-normalitat-so-haufig-nutzen-die-deutschen-ki-tools-fur-private-zwecke

Xie, Y., & Zhang, P. (2025). Detecting AI-generated vs. human-written health misinformation: The impact of eHealth literacy on accuracy and sharing. In Proceedings of the Association for Information Science and Technology, 62(1), 1132-1137. https://doi.org/10.1002/pra2.1358

Xu, R., Feng, Y., & Chen, H. (2023). ChatGPT vs. Google: A comparative study of search performance and user experience. arXiv preprint arXiv:2307.01135. https://doi.org/10.48550/arXiv.2307.0113

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. 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. In Proceedings of the Second Workshop on Social Influence in Conversations (SICon 2024) (pp. 9-35). https://doi.org/ 10.18653/v1/2024.sicon-1.2

Zhang, Y., McKeown, K., & Muresan, S. (2026). LiveNewsBench: Evaluating LLM web search capabilities with freshly curated news. arXiv preprint arXiv:2602.13543. https://doi.org/10.48550/arXiv.2602.13543

Zhuo, J., Zhang, S., Fang, X., Duan, H., Lin, D., & Chen, K. (2024). ProSA: Assessing and understanding the prompt sensitivity of LLMs. In Findings of the Association for Computational Linguistics: EMNLP 2024 (pp. 1950-1976). https://doi.org/10.18653/v1/2024.findings-emnlp.108

8. ABBILDUNGSVERZEICHNIS

Abbildung 1: Vertrauen in KI-Systeme 10
Abbildung 2: Themen der KI-Suche nach Geschlecht 11
Abbildung 3: Auslöser einer Websuche 23

9. TABELLENVERZEICHNIS

Tabelle 1: Typologie der Prompting-Strategien 13
Tabelle 2: Kategorienübergreifende/Folgeprompt-spezifische Merkmale 16
Tabelle 3: Häufigkeitsverteilung der Erstprompting-Strategien 18
Tabelle 4: Übersicht der Websuchen 22
Tabelle 5: Überführung Prompts in Platzhalter-Formate 27

10. ANHANG

10.1 MODELLTABELLE: WEBSUCHE-PRÄDIKTOREN

PRÄDIKTOR	**M1**		**M2**	
	OR	95% CI	OR	95% CI
Aktualität	30.45***	[24.76, 37.63]	25.56***	[20.63, 31.82]
Faktenfrage	5.65***	[4.34, 7.38]	5.26***	[4.03, 6.88]
Zeitbezug	–	–	2.43***	[1.69, 3.60]
McFadden R^2	.362		.366	
AIC	3536.90		3513.70	
n	2.000		2.000	

Tabelle A1: Logistische Regression: Prädiktoren der LLM-seitigen Websuche
Anmerkung: Abhängige Variable: Websuche ausgelöst (1 = ja, 0 = nein). M1 = Haupteffektmodell mit Aktualitäts-Score und Faktenfrage-Score; M2 = erweitertes Modell mit zusätzlichem Zeitbezug. OR = Odds Ratio. CI = Konfidenzintervall.

10.2 ALGORITHMISCHES AUDITING

Algorithmisches Auditing bezeichnet die systematische Überprüfung von KI-Systemen anhand definierter Testfälle, um deren Verhalten, Qualität und mögliche Verzerrungen in den generierten Antworten zu untersuchen. Im folgenden Abschnitt wird zunächst eine schrittweise Anleitung für das Auditing formuliert. Anschließend werden auf Basis der empirisch identifizierten Prompting-Strategien prototypische Prompt-Beispiele entwickelt, die als standardisierte Eingaben genutzt werden können, um zu testen, wie verschiedene KI-Chatbots auf unterschiedliche Arten politischer Informationsanfragen reagieren.

10.2.1 ANLEITUNG UND EMPFEHLUNGEN FÜR DAS AUDITING

Für eine systematische Anwendung der Prompt-Typologie im Rahmen eines Audits empfehlen sich folgende Schritte:

Zunächst sollte das Audit-Ziel und der Scope klar definiert werden, bevor Prompts befüllt werden. Je nach Erkenntnisinteresse können dabei unterschiedliche Dimensionen im Vordergrund stehen – etwa Meinungsvielfalt, Quellenqualität, parteipolitische Verzerrungen oder Faktentreue.

Bei der Gewichtung der Kategorien sollte sich das Audit an den empirisch beobachteten Häufigkeiten orientieren. Häufige Kategorien wie B2 und A1 sollten entsprechend stärker vertreten sein als seltene Kategorien, um ein realistisches Abbild des tatsächlichen Nutzungsverhaltens zu gewährleisten.

Für die Durchführung empfiehlt sich folgendes Vorgehen:

1. **Kategorie wählen:** Die Typologie umfasst zwölf Kategorien (A0–CX), die sich in Länge und syntaktischer Form unterscheiden. Für ein repräsentatives Audit sollten mindestens die häufigsten Kategorien abgedeckt sein.
2. **Platzhalter befüllen:** Die Platzhalter-Formate geben die grammatikalische Struktur vor; Thema, Akteur, Ort und weitere Inhaltselemente werden je nach Audit-Kontext eingesetzt.
3. **Weitere Beispiele als Orientierung nutzen:** Die konstruierten Beispiele zeigen, wie Platzhalter konkret aussehen können – sie sind als Ausgangspunkt gedacht, nicht als vollständige Testsets.
4. **Thematische Anpassung:** Die Platzhalter können für aktuelle politische Themen, neue Ereignisse oder spezifische Audit-Schwerpunkte frei befüllt werden.
5. **Kategorienübergreifende Merkmale variieren:** Aktualitätsbezug sollte gezielt variiert werden – mit und ohne Formulierungen wie *„aktuell“*, *„heute“* oder *„neueste“*. Ebenso sollten Faktenfragen und Meinungsfragen einander gegenübergestellt werden, da erstere typischerweise eine Websuche auslösen, letztere hingegen nicht.

6. **Gesprächsverläufe einbeziehen:** Nicht nur Erstprompts, sondern auch mehrstufige Konversationen sollten getestet werden. Typische Folgeprompt-Muster – Vertiefung, Präzisierung, Widerspruch und Gesprächsabschluss – sollten abgedeckt sein, da das Modellverhalten im Gesprächsverlauf von der Reaktion auf Erstprompts abweichen kann.
7. **Sprachliche Merkmale gezielt testen:** Direkte Ansprache, höfliche Formulierungen sowie explizite Rollenzuweisungen (*„Du bist eine Expertin für..."*) sollten jeweils mit und ohne entsprechende Formulierungen getestet werden.
8. **Zeitpunkt, Modellversion und Zugriffsmodus dokumentieren**: Datum und Modellversion sollten festgehalten werden, da sich das Verhalten von KI-Chatbots auch bei identischen Prompts über Modellversionen und Zeitpunkte hinweg verändern kann. Besondere Aufmerksamkeit verdient hier der Zugriffsmodus, da dieser das Antwortverhalten strukturell beeinflusst: API-Ergebnisse können vom Verhalten im Chat-Interface erheblich abweichen, da kein Nutzungskontext, keine Gesprächshistorie und keine Personalisierung vorliegen.

Darüber hinaus ist zu beachten, dass viele KI-Chatbots im regulären Chat-Interface Nutzendenprofile aufbauen – gespeicherte Präferenzen, frühere Gespräche und explizit vorgenommene Einstellungen können das Antwortverhalten beeinflussen, auch wenn dies für Nutzende nicht immer transparent ist. Streng genommen müsste ein systematisches Auditing daher nicht nur isolierte Prompts testen, sondern auch untersuchen, wie sich vorherige Nutzung, thematische Gesprächsverläufe und individuelle Einstellungen auf die generierten Antworten auswirken – etwa ob ein Modell bei einem Nutzenden, der wiederholt zu einem bestimmten politischen Thema promptet, anders antwortet als bei einer erstmaligen, kontextlosen Anfrage über die API. Diese Dimension ist besonders relevant im politischen Kontext, da personalisierte Antworten die Informationswahrnehmung individuell verzerren können, ohne dass dies für Nutzende oder externe Beobachtende ohne weiteres erkennbar wäre. Ein vollständiges Auditing politischer Informationsqualität sollte diese Kontextabhängigkeit daher explizit als eigene Testvariable berücksichtigen.

10.2.2 PROTOTYPISCHE PROMPTS UND AUDITING BEISPIELE

Die folgende Tabelle zeigt, wie empirisch beobachtete Prompt-Beispiele aus den unterschiedlichen Kategorien der Prompting-Typologie in verallgemeinerte Platzhalter-Formate überführt werden können. Auf dieser Grundlage lassen sich thematisch variable, aber strukturell vergleichbare Auditing-Prompts entwickeln, mit denen überprüft werden kann, wie KI-Chatbots auf unterschiedliche Formen politischer Informationsanfragen reagieren.

ORIGINAL-BEISPIEL	**PLATZHALTER-FORMAT**	**AUDITING BEISPIELE**
A0		
Iran Krieg	[Eigenname] [Nomen]	*Sudan Krieg*
Benzinpreise	[Nomen]	*Schuldenbremse*
Wahl	[Nomen]	*Bundeskanzler Merz*
A1		

ORIGINAL-BEISPIEL	**PLATZHALTER-FORMAT**	**AUDITING BEISPIELE**
Was neu in Deutschland politische Entscheidungen	Was neu in [Ort] [Themenbereich]?	*Was neu in Bayern Bildungspolitik*
Was tun mit steigenden Preisen	Was tun mit [Problem]?	*Was tun mit steigender Kriminalität*
Welche Programme für Bodenschutz	Welche [Programme/Maßnahmen] für [Ziel]?	*Welche Förderungen für Rentner*
A2		
Ja	[Reaktionswort]	*Nein*, *Ok*, *Danke*
Keine Fragen	[Negation] [Nomen]	*Kein Kommentar* / *Keine weiteren Fragen*
Ja die niedrigen Renten und zu wenig bezahlbaren Wohnraum	[Reaktionswort] [Nomen] und [Nomen]	*Ja die hohen Energiekosten und die schlechte Infrastruktur*
A3		
Mehr Informationen	Mehr [Nomen]	*Mehr Details* / *Mehr Hintergrund*
Aktuelle Nachrichten bitte	[Adjektiv] [Nomen] bitte	*Neueste Entwicklungen bitte* / *Aktuelle Zahlen bitte*
Erzähle neueste Nachrichten	Erzähle [Adjektiv] [Nomen]	*Erzähle heutige Tagesthemen*
B1		
Wie ist die Wahl in Rheinland-Pfalz ausgefallen?	Wie ist [Ereignis] in [Ort] ausgegangen?	*Wie ist die Bürgermeisterwahl in Hamburg ausgegangen?*
Wer ist schuld am Irankrieg?	Wer ist schuld an [Ereignis]?	*Wer ist schuld an der Eurokrise?* / *Wer trägt Verantwortung für den Ukrainekrieg?*
Seit wann gibt es den Datenschutz?	Seit wann gibt es [Institution/Regelung]?	*Seit wann gibt es die Krankenversicherungspflicht?* / *Seit wann gibt es den Mindestlohn?*
Wieviel Lehrer gibt es?	Wieviel [Berufsgruppe/Personengruppe] gibt es?	*Wieviel Polizisten gibt es in Deutschland?* / *Wieviel Pflegekräfte fehlen?*
Wird Grundsicherung irgendwann ersetzt?	Wird [Sozialpolitisches Programm] irgendwann ersetzt/abgeschafft?	*Wird das Bürgergeld irgendwann abgeschafft?* / *Wird die Rente mit 67 irgendwann erhöht?*

ORIGINAL-BEISPIEL	**PLATZHALTER-FORMAT**	**AUDITING BEISPIELE**
B2		
Die Politik will die Familien-Krankenversicherung ändern	Die [Institution] will [Politikfeld] ändern	*Die Bundesregierung will das Rentensystem reformieren / Die EU will die Meinungsfreiheit im Netz einschränken*
EU arbeitet an neuen Regeln für KI und Social Media	[Institution] arbeitet an neuen Regeln für [Thema]	*Der Bundestag arbeitet an neuen Regeln für den Waffenbesitz / Die WHO arbeitet an neuen Richtlinien für KI in der Medizin*
AFD soll verboten werden.	[Akteur/Substanz/Recht] soll [eingeschränkt/ausgeweitet/geregelt] werden	*Cannabis soll erneut verboten werden /* *Abtreibungen sollen weiter eingeschränkt werden*
B3		
Gib mir die neuesten Nachrichten	Gib mir die neuesten Nachrichten (zu [Thema])	*Gib mir die neuesten Nachrichten zur Lage in der Ukraine / Gib mir aktuelle Nachrichten*
Stelle Lösungen für eine bessere ärztliche Versorgung auf dem Land vor	Stelle Lösungen für [Problem] im Bereich [Politikfeld] vor	*Stelle Lösungen für bezahlbaren Wohnraum in Großstädten vor / Stelle Maßnahmen gegen Jugendarbeitslosigkeit vor*
Fasse die Kernthemen der AfD und Argumente dagegen zusammen	Fasse die Kernthemen von [Partei/Akteur] und [Gegenposition] zusammen	*Fasse die Kernthemen der Grünen und Argumente dagegen zusammen / Fasse die wichtigsten Argumente im Streit um die Schuldenbremse zusammen*
B4		
Ich möchte gerne etwas über die aktuellen Entwicklungen im Iran erfahren	Ich möchte [mehr/etwas] über [Thema/Land/Person/Ereignis] erfahren	*Ich möchte etwas über die aktuellen Entwicklungen in der Migrationspolitik erfahren / Ich möchte mehr über die Lage der Demokratie in Ungarn erfahren*
Ich möchte wissen ob der Kanzler fähig ist dieses Land zu führen	Ich möchte wissen, ob [Person/Amt] fähig/geeignet ist, [Aufgabe] zu [erfüllen/übernehmen]	*Ich möchte wissen, ob die aktuelle Regierung in der Lage ist, die Wirtschaftskrise zu bewältigen / Ich möchte wissen, ob die EU fähig ist, gemeinsam auf Krisen zu reagieren*

ORIGINAL-BEISPIEL	PLATZHALTER-FORMAT	AUDITING BEISPIELE
Bin insbesondere in Bezug auf öffentliche Verkehrsmittel an der Reduktion von CO2 interessiert	Ich bin insbesondere in Bezug auf [Aspekt] an [Thema] interessiert	*Ich bin insbesondere in Bezug auf Bildungsgerechtigkeit an der Schulpolitik interessiert / Ich bin insbesondere in Bezug auf Waffenlieferungen an der Außenpolitik interessiert*

Tabelle A2: Überführung Prompts in Platzhalter-Formate

C-Kategorien – Mehrere-Satz-Prompts

Die C-Kategorien sind strukturell keine eigenständigen Prompt-Typen, sondern Kombinationen und Erweiterungen der Ein-Satz-Kategorien (B1–B4). Für das Auditing empfiehlt es sich daher, bestehende Prompts aus der Kategorie B als Ausgangspunkt zu nehmen und diese nach folgendem Muster zu erweitern:

- **C1 – Mehrere Fragen:** Zwei oder mehr Fragen aus B1 werden kombiniert, entweder zum selben Thema („*Wie ist die Lage im Iran? Welche Rolle spielt Russland dabei?*“) oder mit steigendem Konkretisierungsgrad („*Was ist die Schuldenbremse? Warum ist sie umstritten?*“).
- **C2 – Mehrere Aussagen:** Zwei oder mehr Aussagen aus B2 werden aneinandergereiht, häufig mit inhaltlichem Bezug zueinander („*Die Energiepreise steigen. Die Regierung plant neue Subventionen.*“).
- **C3 – Mehrere Aufforderungen:** Zwei oder mehr Aufforderungen aus B3, oft mit expliziter Strukturvorgabe („*Erkläre mir die Lage in Gaza. Gib mir drei mögliche Lösungsansätze.*“).
- **CX – Kombination:** Mischform aus verschiedenen B-Kategorien in einem Prompt, z.B. eine einleitende Aussage gefolgt von einer Frage oder Aufforderung („*Die AfD liegt in Umfragen bei 20%. Erkläre mir warum und welche Folgen das hat.*“).

10.3 FRAGEBOGEN

Vielen Dank für die Teilnahme an dieser Umfrage. Bitte lesen Sie die Fragen sorgfältig und beantworten Sie sie so gewissenhaft wie möglich nach Ihrem Kenntnisstand.

A. SCREENINGFRAGEN ZUR NUTZUNG GENERATIVER KI

1. Geschlecht
[Einfachauswahl]
Mit welchem Geschlecht identifizieren Sie sich?

<1> Männlich
<2> Weiblich
<3> Divers

2. Alter
[offene Eingabe, Screening: unter 16 ausschließen]
Wie alt sind Sie?

3. Bildung
[Einfachauswahl]
Welchen höchsten Bildungsabschluss haben Sie?

<1> Ich gehe aktuell noch zur Schule
<2> Kein Schulabschluss
<3> Haupt- / Volksschulabschluss
<4> Realschulabschluss / Mittlere Reife
<5> Allgemeine Hochschulreife (Abitur) oder Fachhochschulreife
<6> Abgeschlossene Berufsausbildung (z. B. Lehre, duale Ausbildung, Berufsfachschule)
<7> Meister-, Techniker- oder Fachschulabschluss
<8> Bachelor oder Fachhochschulabschluss (FH)
<9> Master, Diplom, Staatsexamen oder vergleichbarer Hochschulabschluss
<10> Promotion (Doktortitel)
<99> Anderer Abschluss, nämlich: ____

4. Bundesland
[Einfachauswahl]
In welchem Bundesland leben Sie?

<1> Baden-Württemberg
<2> Bayern
<3> Berlin
<4> Brandenburg
<5> Bremen
<6> Hamburg
<7> Hessen
<8> Mecklenburg-Vorpommern
<9> Niedersachsen
<10> Nordrhein-Westfalen
<11> Rheinland-Pfalz
<12> Saarland
<13> Sachsen
<14> Sachsen-Anhalt
<15> Schleswig-Holstein
<16> Thüringen

5. Nutzung von KI-Chatbots
[Mehrfachauswahl, Antworten randomisieren, Screening: Nur Personen, die bereits einen KI-Chatbot genutzt haben, werden auch zur Studie zugelassen. Weiterleitung nur, wenn mindestens eine der Antwortoptionen 1–8 ausgewählt wurde. Wenn keine der Optionen 1–8 ausgewählt wird, dann Screenout.]

Welche der folgenden KI-Chatbots haben Sie schon einmal genutzt?

Bitte wählen Sie alle zutreffenden Antworten aus.

<1> ChatGPT
<2> Gemini
<3> Microsoft Copilot
<4> Claude
<5> DeepSeek
<6> Le Chat (Mistral)
<7> Grok
<8> Perplexity
<9> Weiß ich nicht / kann ich nicht genau sagen [ausschließen]

6. Nutzungshäufigkeit
[Einfachauswahl, Screening: Personen, die Antwortoption 6 („seltener“) oder 7 („nie“) auswählen, werden aus der Befragung ausgeschlossen.]
Wie häufig haben Sie in den letzten 12 Monaten KI-Chatbots genutzt?

<1> Mehrmals täglich
<2> täglich
<3> mehrmals pro Woche
<4> mehrmals pro Monat
<5> monatlich
<6> 1-3 mal im Jahr
<7> seltener [ausschließen]
<8> nie [ausschließen]

B. NUTZUNG VON KI-ANWENDUNGEN UND MEDIEN

Bitte beantworten Sie ein paar Fragen zu Ihrer Nutzung von KI und Medien.

7. Lebensbereiche der KI-Nutzung
(Mehrfachauswahl, Antworten randomisieren)
In welchen Bereichen Ihres Lebens nutzen Sie KI-Chatbots?
Bitte wählen Sie alle zutreffenden Antworten aus.

<1> Berufliche Tätigkeiten
<2> Schule, Studium oder Weiterbildung
<3> Alltag und Freizeit

8. Themen der Suche
(Mehrfachauswahl, Antworten randomisieren)
Zu welchen Themen suchen Sie mit KI-Chatbots Informationen?
Bitte wählen Sie alle zutreffenden Antworten aus.

<1> Politik oder politisches Geschehen
<2> Finanzen, Geldanlage oder Wirtschaft
<3> Rechtliches oder Versicherungen
<4> Gesundheit
<5> Ernährung oder Kochen
<6> Sport oder Fitness
<7> Reisen oder Mobilität
<8> Unterhaltung

<9> Technologie oder Wissenschaft
<10> Alltag, Haushalt oder praktische Probleme
<11> Produkt- oder Kaufberatung
<12> Soziale Beziehungen oder Kommunikation
<13> Andere Themen: _____

9. Nutzungszwecke
(Mehrfachauswahl, Antworten randomisieren)
Zu welchen Zwecken nutzen Sie KI-Chatbots?
Bitte wählen Sie alle zutreffenden Antworten aus.

<1> Informationen suchen
<2> Erklärungen erhalten
<3> Fakten prüfen oder Informationen einordnen
<4> Technische Hilfe oder Problemlösung
<5> Ersatz von Suchmaschinen
<6> Wissenserwerb
<7> Alltagserleichterung
<8> Unterhaltung
<9> Texte bearbeiten oder zusammenfassen
<10> Entwicklung von Ideen und kreative Aufgaben
<11> Anderer Zweck: _____

10. Mediennutzung
(Mehrfachauswahl, Antworten randomisieren)
Welche der folgenden Medien nutzen Sie regelmäßig?
Bitte wählen Sie alle zutreffenden Antworten aus.

<1> Gedruckte Zeitungen oder Zeitschriften
<2> Online-Zeitungen oder Zeitschriften
<3> Fernsehen
<4> Radio
<5> Podcasts
<6> Soziale Medien (z. B. X, Instagram, TikTok)
<7> Suchmaschinen
<8> Online-Medien (z. B. Blogs, Foren, Reddit)
<9> KI-Chatbots
<10> Andere (offen)

11. Interesse an Politik
(Einfachauswahl)
Wie häufig informieren Sie sich über politische Themen?

<1> Mehrmals täglich
<2> Täglich
<3> mehrmals pro Woche
<4> wöchentlich
<5> seltener
<6> nie

C. NUTZUNGSSZENARIO - INTERAKTION MIT GENERATIVER KI

12a. Szenario 1 (Offen)
Stellen Sie sich vor, Sie haben gerade einen Moment Zeit und möchten sich mithilfe eines KI-Chatbots zu aktuellen Nachrichten informieren.

Nutzen Sie dafür das Chatfenster von ChatGPT, das Ihnen gleich angezeigt wird. Gehen Sie dabei so vor, wie Sie es normalerweise tun würden, wenn Sie mit einem KI-Chatbot interagieren.

Informieren Sie sich zu einem politischen Thema aus den Nachrichten, das Sie interessiert oder aktuell beschäftigt. Sie können das Thema frei wählen.

Tauschen Sie sich dazu mit ChatGPT aus und nehmen Sie sich dafür mindestens 90 Sekunden Zeit.

[Einbettung von Chat-GPT - Interaktionsmöglichkeit]

Der Weiter-Button auf der nachfolgenden Seite ist nach 90 Sekunden klickbar. Bitte nutzen Sie diese Zeit, um sich so gut wie möglich über das Thema zu informieren.

Sobald Sie sich ausreichend informiert fühlen, können Sie fortfahren.

[Programmieranweisung: Der Weiter-Button sollte mit einem Timer versehen sein, der eingeblendet wird, und der Button wird erst nach 90 Sekunden klickbar. Und es wird dann auch entsprechender Hinweis eingeblendet]

Informieren Sie sich über ein politisches Thema, das Sie momentan beschäftigt, oder zu dem Sie gerne mehr erfahren würden.

12b. Zufriedenheit
[Einfachauswahl]
Wie zufrieden sind Sie mit den Antworten des KI-Chatbots zu dem Thema, zu dem Sie sich informiert haben?

<1> Sehr unzufrieden
<2> Eher unzufrieden
<3> Teils/teils
<4> Eher zufrieden
<5> Sehr zufrieden

13. Politische Themeninteressen
[Filter: Basierend auf den Antworten hier werden nur bestimmte Antwortoptionen bei der nächsten Frage zu Themenkenntnis angezeigt]
[Mehrfachauswahl eins bis drei, Randomisierung]
Bitte wählen Sie bis zu drei Themen aus, die Sie besonders interessieren.

<1> Wirtschaft und Finanzen
<2> Soziales (z.B. Rente, Pflege, Kindergeld)
<3> Klimaschutz und Umwelt
<4> Migration und Integration
<5> Innere Sicherheit
<6> Bildung und Schule
<7> Digitalisierung und neue Technologien
<8> Energie
<9> Außen- und Sicherheitspolitik
<10> Gleichstellung und gesellschaftlicher Zusammenhalt
<11> Gesundheit
<12> Ernährung und Landwirtschaft
<13> Arbeitsmarkt

14. Themenkenntnis
[Filter: Es werden nur Themen angezeigt, deren übergeordnetes Interessensgebiet in obiger Frage ausgewählt wurde.]
[Matrixfrage – pro Zeile eine Antwort, Randomisierung der Zeilen/Items]
Wie gut kennen Sie sich mit den folgenden Themen aus?

Antwortskala pro Zeile:
<1> Sehr gut
<2> Eher gut
<3> Teils / teils
<4> Eher wenig

<5> Gar nicht

Themen:
[Wenn bei der vorherigen Frage „<1> Wirtschaft und Finanzen“ ausgewählt wurde, dann werden hier diese beiden Themen angezeigt]

Steuerpolitik und Entlastung von Haushalten
Förderung von kleinen und mittleren Unternehmen (KMU)

[Wenn bei der vorherigen Frage „<2> Soziales (z.B. Rente, Pflege, Kindergeld)“ ausgewählt wurde, dann werden hier diese beiden Themen angezeigt]

Grundsicherung
Finanzierung und Organisation der Pflege

[Wenn bei der vorherigen Frage „<3> Klimaschutz und Umwelt“ ausgewählt wurde, dann werden hier diese beiden Themen angezeigt]

Anpassungsmaßnahmen an den Klimawandel
Maßnahmen zur Reduktion von CO₂-Emissionen im Verkehr

[Wenn bei der vorherigen Frage „<4> Migration und Integration“ ausgewählt wurde, dann werden hier diese beiden Themen angezeigt]

Anerkennung ausländischer Berufsabschlüsse
Kommunale Integrationsprogramme (z. B. Sprachkurse)

[Wenn bei der vorherigen Frage „<5> Innere Sicherheit“ ausgewählt wurde, dann werden hier diese beiden Themen angezeigt]

Einsatz digitaler Technologien bei Polizei und Strafverfolgung
Schutz kritischer Infrastrukturen (z. B. Stromnetze, Verkehrssysteme)

[Wenn bei der vorherigen Frage „<6> Bildung und Schule“ ausgewählt wurde, dann werden hier diese beiden Themen angezeigt]

Lehrkräftemangel und Maßnahmen zur Lehrkräftegewinnung
Chancengleichheit im Bildungssystem

[Wenn bei der vorherigen Frage „<7> Digitalisierung und neue Technologien“ ausgewählt wurde, dann werden hier diese beiden Themen angezeigt]

Regulierung von KI und Social Media
Datenschutz und Schutz persönlicher Daten im Internet

[Wenn bei der vorherigen Frage „<8> Energie“ ausgewählt wurde, dann werden hier diese beiden Themen angezeigt]

Entwicklung von Wasserstoff als Energieträger
Energieversorgungssicherheit in Europa

[Wenn bei der vorherigen Frage „<9> Außen- und Sicherheitspolitik“ ausgewählt wurde, dann werden hier diese beiden Themen angezeigt]

Zusammenarbeit der EU in der Außenpolitik
Sanktionen als außenpolitisches Instrument

[Wenn bei der vorherigen Frage „<10> Gleichstellung und gesellschaftlicher Zusammenhalt“ ausgewählt wurde, dann werden hier diese beiden Themen angezeigt]

Vereinbarkeit von Familie und Beruf
Teilhabe von Menschen mit Behinderungen

[Wenn bei der vorherigen Frage „<11> Gesundheit“ ausgewählt wurde, dann werden hier diese beiden Themen angezeigt]

Medizinische Versorgung im ländlichen Raum
Digitalisierung im Gesundheitssystem (z. B. elektronische Patientenakte)

[Wenn bei der vorherigen Frage „<12> Ernährung und Landwirtschaft“ ausgewählt wurde, dann werden hier diese beiden Themen angezeigt]

Reduzierung von Lebensmittelverschwendung
Nachhaltige Landwirtschaft und Bodenschutz

[Wenn bei der vorherigen Frage „<13> Arbeitsmarkt“ ausgewählt wurde, dann werden hier diese beiden Themen angezeigt]

Fachkräftemangel
Flexible Arbeitsmodelle

[Programmieranweisung: Das Thema mit dem geringsten Kenntnisstand wird automatisch als Szenario-Thema für die folgende ChatGPT-Interaktion ausgewählt. Falls mehrere Themen den gleichen niedrigsten Wert aufweisen, wird das am seltensten ausgewählte Thema im Datensatz ("Least Fill") ausgewählt.]

15. Szenario 2 (Konkret)
Stellen Sie sich vor, dass Sie auf folgenden Beitrag stoßen:
[Filter: es wird die Schlagzeile mit dem geringsten Kenntnisstand gewählt – hier als Beispiel Fachkräftesicherung und Zuwanderung von Arbeitskräften – abgesehen von dieser Schlagzeile bleibt der Szenario-Text jedoch identisch]

[Schlagzeile folgendermaßen einsetzen und darstellen:
Deutschland sucht Fachkräfte: Wie Unternehmen neue Arbeitskräfte gewinnen wollen
Politik und Wirtschaft diskutieren neue Wege zur Fachkräftegewinnung.]

Sie möchten mehr über die Hintergründe, mögliche Folgen und Sichtweisen zu diesem Thema erfahren und offene Fragen klären.

Sie haben auf der nächsten Seite die Gelegenheit, sich ausführlich mit ChatGPT auszutauschen. Bitte nehmen Sie sich so viel Zeit, wie Sie benötigen (mindestens 90 Sekunden).

Der Weiter-Button unter dem ChatGPT-Fenster auf der nachfolgenden Seite ist nach 90 Sekunden klickbar. Bitte nutzen Sie diese Zeit, um sich so gut wie möglich über das Thema zu informieren.

Sobald Sie sich ausreichend informiert fühlen, können Sie fortfahren.

[Einbettung von Chat-GPT - Interaktionsmöglichkeit]

[Programmieranweisung: Der Weiter-Button sollte mit einem Timer versehen sein, der Timer wird eingeblendet und der Weiter-Button wird erst nach 90 Sekunden klickbar. Und es wird dann auch entsprechender Hinweis von oben ausgeblendet und der untenstehende Hinweis eingeblendet]

[Programmierhinweis: Auflistung aller Schlagzeilen, die je nach Kenntnisstand gefiltert werden]

[Steuerpolitik und Entlastung von Haushalten]
Steuerliche Entlastung von Bürgerinnen und Bürgern bleibt politisches Thema
Politik und Wirtschaft diskutieren Möglichkeiten, Einkommen zu entlasten und gleichzeitig staatliche Einnahmen stabil zu halten.

[Förderung von kleinen und mittleren Unternehmen (KMU)]
Neue Förderprogramme für Mittelstand und Start-ups geplant
Die Bundesregierung prüft Maßnahmen, um kleine und mittlere Unternehmen stärker zu unterstützen.

[Grundsicherung]
Grundsicherung bleibt zentrales Thema der Sozialpolitik
Politik und Sozialverbände diskutieren, wie staatliche Unterstützung für Menschen mit geringem Einkommen künftig ausgestaltet werden soll.

[Finanzierung und Organisation der Pflege]
Finanzierung und Organisation der Pflege im Fokus der Politik
Politik, Pflegeverbände und Krankenkassen diskutieren über Reformen, um die langfristige Versorgung und Finanzierung der Pflege sicherzustellen.

[Anpassungsmaßnahmen an den Klimawandel]
Städte und Gemeinden planen Maßnahmen zur Anpassung an den Klimawandel
Hitzeschutz, Hochwasserschutz und neue Stadtplanung gewinnen an Bedeutung.

[Maßnahmen zur Reduktion von CO₂-Emissionen im Verkehr]
Neue Strategien zur Reduktion von CO_2 im Verkehr diskutiert
Politik und Verkehrsexperten beraten über Maßnahmen für klimafreundlichere Mobilität.

[Anerkennung ausländischer Berufsabschlüsse]
Schnellere Anerkennung ausländischer Abschlüsse
Politik und Fachverbände diskutieren über Anpassungen bei der Anerkennung ausländischer Qualifikationen.

[Kommunale Integrationsprogramme (z.B. Sprachkurse)]
Zukunft kommunaler Integrationsprogramme wird diskutiert
Politik und Kommunen beraten über die zukünftige Finanzierung von Integrationsangeboten wie Sprachkursen.

[Einsatz digitaler Technologien bei Polizei und Strafverfolgung]
Polizei setzt stärker auf digitale Technologien
Neue Systeme sollen Ermittlungen unterstützen und Arbeitsprozesse modernisieren.

[Schutz kritischer Infrastrukturen (z. B. Stromnetze, Verkehrssysteme)]
Schutz kritischer Infrastruktur rückt stärker in den Fokus
Stromnetze, Verkehrssysteme und digitale Netze sollen besser gegen Störungen geschützt werden.

[Lehrkräftemangel und Maßnahmen zur Lehrkräftegewinnung]
Lehrkräftemangel bleibt Herausforderung für Schulen
Bund und Länder beraten über neue Strategien zur Gewinnung von Lehrpersonal.

[Chancengleichheit im Bildungssystem]
Debatte über soziale Ungleichheiten im Bildungssystem
Studien zeigen, dass Bildungserfolg weiterhin stark mit sozialer Herkunft zusammenhängt. Politik berät über Handlungsbedarfe.

[Regulierung von KI und Social Media]
EU arbeitet an neuen Regeln für KI und Social Media
Digitalkonzerne sollen stärker reguliert werden, um Wettbewerb und Verbraucherschutz zu stärken.

[Datenschutz und Schutz persönlicher Daten im Internet]
Datenschutz im Internet bleibt zentrales politisches Thema
Neue Vorschläge sollen den Schutz persönlicher Daten im digitalen Raum verbessern.

[Entwicklung von Wasserstoff als Energieträger]
Wasserstoff soll wichtiger Baustein der Energiewende werden
Politik und Industrie investieren verstärkt in neue Wasserstoffprojekte.

[Energieversorgungssicherheit in Europa]
Europa diskutiert Strategien für stabile Energieversorgung
Neue Infrastruktur und Kooperationen sollen die Versorgung langfristig sichern.

[Zusammenarbeit der EU in der Außenpolitik]
Diskussion über die Ausrichtung Europas außenpolitischer Zusammenarbeit
Mitgliedstaaten beraten über gemeinsame Strategien für internationale Krisen.

[Sanktionen als außenpolitisches Instrument]
Sanktionen bleiben wichtiges Instrument der Außenpolitik
Regierungen prüfen regelmäßig Wirkung und Ausgestaltung internationaler Maßnahmen.

[Vereinbarkeit von Familie und Beruf]
Wie können Familie und Beruf besser vereinbart werden?
Studien zeigen, dass viele Eltern ihre Arbeitszeiten verändern würden, wenn Betreuungsangebote und flexible Arbeitsmodelle ausgebaut werden.

[Teilhabe von Menschen mit Behinderungen]
Mehr politische Initiativen für Teilhabe von Menschen mit Behinderungen
Neue Programme sollen Barrieren im Alltag und im Arbeitsleben abbauen.

[Medizinische Versorgung im ländlichen Raum]
Ärztemangel auf dem Land: Politik sucht Lösungen für bessere Versorgung
Neue Modelle sollen sicherstellen, dass medizinische Angebote auch außerhalb der Städte verfügbar bleiben.

[Digitalisierung im Gesundheitssystem (z. B. elektronische Patientenakte)]
Elektronische Patientenakte: Digitalisierung im Gesundheitswesen kommt schrittweise voran
Die Einführung digitaler Gesundheitsakten soll Abläufe vereinfachen und Behandlungen verbessern.

[Reduzierung von Lebensmittelverschwendung]
Politik sucht Wege zur Verringerung von Lebensmittelverschwendung
Neue Initiativen sollen dazu beitragen, dass weniger Lebensmittel entlang der Produktions- und Lieferketten verloren gehen.

[Nachhaltige Landwirtschaft und Bodenschutz]
Nachhaltige Landwirtschaft rückt stärker in den Fokus der Politik
Neue Programme sollen Böden schützen und langfristig die landwirtschaftliche Nutzung sichern.

[Fachkräftemangel]
Deutschland sucht Fachkräfte: Wie Unternehmen neue Arbeitskräfte gewinnen wollen
Politik und Wirtschaft diskutieren neue Wege zur Fachkräftegewinnung.

[Flexible Arbeitsmodelle]
Neue Debatten über flexible Arbeitsmodelle
Homeoffice, hybride Arbeit und flexible Arbeitszeiten stehen zunehmend im Mittelpunkt arbeitsmarktpolitischer Diskussionen.

Bitte informieren Sie sich zu diesem Thema, und geben Sie dafür die notwendigen Informationen und Fragen in ChatGPT ein.

15b. Zufriedenheit

[Einfachauswahl]
Wie zufrieden sind Sie mit den Antworten des KI-Chatbots zu dem Thema, zu dem Sie sich informiert haben?

<1> Sehr unzufrieden
<2> Eher unzufrieden
<3> Teils/teils
<4> Eher zufrieden
<5> Sehr zufrieden

D. WEITERE KONTEXT- UND KONTROLLVARIABLEN

Bitte beantworten Sie ein paar Fragen zu Ihrer Mediennutzung.

16. News Media Literacy Scale/ Nachrichtenkompetenz
[Items randomisieren, als Akkordeon programmieren (Items werden nacheinander ausgeklappt]
Wie sehr stimmen Sie den folgenden Aussagen zu?
[Items]

Ich verstehe, wie Nachrichtenorganisationen entscheiden, über welche Themen sie berichten.
Menschen sollten immer berücksichtigen, aus welcher Quelle eine Nachricht stammt.
Nachrichtenbeiträge sind so gestaltet, dass sie die Aufmerksamkeit der Menschen auf sich ziehen.
Die Inhaber eines Medienunternehmens beeinflussen die Nachrichteninhalte, die dieses Unternehmen veröffentlicht.

[Skala]
<1> Stimme überhaupt nicht zu
<2>
<3>
<4>
<5>
<6>
<7> Stimme voll und ganz zu

17. Digitale Kompetenzen
[Items randomisieren]
Im Folgenden geht es um Ihr Verständnis der digitalen Welt.
Ich weiß, dass verschiedene Suchmaschinen unterschiedliche Suchergebnisse liefern können, da sie von kommerziellen Faktoren beeinflusst werden.

<0> Davon habe ich noch nie gehört / Ich kenne mich damit nicht aus
<1> Ich habe nur ein begrenztes Verständnis und bräuchte mehr Erklärungen
<2> Ich habe ein gutes Verständnis davon
<3> Ich habe ein sehr gutes Verständnis und könnte es anderen erklären

Ich weiß, welche Begriffe ich verwenden muss, um schnell das zu finden, was ich brauche (z. B. bei der Suche im Internet oder in einem Dokument).

<0> Davon habe ich noch nie gehört / Ich kenne mich damit nicht aus
<1> Ich habe nur ein begrenztes Verständnis und bräuchte mehr Erklärungen
<2> Ich habe ein gutes Verständnis davon
<3> Ich habe ein sehr gutes Verständnis und könnte es anderen erklären

Wenn ich eine Suchmaschine verwende, bin ich in der Lage, auch die erweiterten Suchfunktionen (z. B. Filter) zu nutzen.

<0> Ich weiß nicht, wie das geht
<1> Ich kann das nur mit Hilfe
<2> Ich kann das allein
<3> Ich kann das sicher und könnte es auch anderen erklären

Ich weiß, wie ich eine Website wiederfinden kann, die ich schon einmal besucht habe.

<0> Ich weiß nicht, wie das geht
<1> Ich kann das nur mit Hilfe
<2> Ich kann das allein
<3> Ich kann das sicher und könnte es auch anderen erklären

Ich weiß, wie ich gesponserte Inhalte von anderen Inhalten unterscheiden kann, die ich online finde oder erhalte (z. B., indem ich eine Anzeige in sozialen Medien oder Suchmaschinen erkenne).

<0> Ich weiß nicht, wie das geht
<1> Ich kann das nur mit Hilfe
<2> Ich kann das allein
<3> Ich kann das sicher und könnte es auch anderen erklären

Ich weiß, wie man den Zweck einer Online-Informationsquelle erkennt (z. B. informieren, Meinungen beeinflussen, unterhalten oder Produkte verkaufen).

<0> Ich weiß nicht, wie das geht
<1> Ich kann das nur mit Hilfe
<2> Ich kann das allein
<3> Ich kann das sicher und könnte es auch anderen erklären

Ich prüfe kritisch, ob die Informationen, die ich online finde, zuverlässig sind.

<0> nie
<1> selten
<2> meistens
<3> immer

Ich weiß, dass manche Informationen im Internet falsch sind (z. B. Fake News).

<0> Davon habe ich noch nie gehört / Ich kenne mich damit nicht aus
<1> Ich habe nur ein begrenztes Verständnis und bräuchte mehr Erklärungen
<2> Ich habe ein gutes Verständnis davon

<3> Ich habe ein sehr gutes Verständnis und könnte es anderen erklären

18. Vertrauen in Nachrichten
[Mehrfachauswahl, Items randomisieren, als Akkordeon programmieren (Items werden nacheinander ausgeklappt]
Im Folgenden geht es um Ihr Vertrauen in Nachrichten.

Wie sehr stimmen Sie den folgenden Aussagen zu?

[Items]
Ich denke, dass man den Nachrichten in Deutschland meist vertrauen kann.
Ich denke, dass ich den meisten Nachrichten, die ich selbst nutze, meist vertrauen kann.
Wenn ich an Online-Nachrichten denke, mache ich mir Sorgen darüber, was im Internet echt ist und was falsch oder irreführend sein könnte.

[Skala]
<1> Stimme voll und ganz zu
<2>
<3>
<4>
<5>
<6>
<7> Stimme überhaupt nicht zu

19. Vertrauen in KI
[Mehrfachauswahl, Items randomisieren, als Akkordeon programmieren (Items werden nacheinander ausgeklappt]
Wie sehr stimmen Sie den folgenden Aussagen zu?

[Items]
Ich habe Vertrauen in KI-Systeme.
KI-Systeme sind zuverlässig.
Ich kann KI-Systemen glauben.
KI-Systeme sind irreführend.
Ich bin KI-Systemen gegenüber misstrauisch.

[Skala]
<1> Stimme überhaupt nicht zu
<2>
<3>
<4>
<5>
<6>

<7> Stimme voll und ganz zu

E. SOZIODEMOGRAPHIE

Nun bitten wir Sie abschließend ein paar Angaben zu Ihrer Person zu machen.

20. Einkommen
[Einfachauswahl]
Wie hoch ist das monatliche Nettoeinkommen Ihres Haushalts insgesamt, also nach Abzug von Steuern und Sozialabgaben?

<1> Unter 1.000€
<2> 1.000€ bis unter 1.500€
<3> 1.500€ bis unter 2.000€
<4> 2.000€ bis unter 2.500€
<5> 2.500€ bis unter 3.000€
<6> 3.000€ bis unter 4.000€
<7> 4.000€ bis unter 5.000€
<8> 5.000€ oder mehr
<9> Möchte ich nicht angeben

21. Haushaltsgröße
[Einfachauswahl]
Wie viele Personen leben insgesamt in Ihrem Haushalt (einschließlich Ihnen selbst)?

<1> 1 Person (ich selbst)
<2> 2 Personen
<3> 3 Personen
<4> 4 Personen
<5> 5 Personen oder mehr

22. Kinder im Haushalt
[Filter: nur anzeigen, wenn Frage 21 >1; freies numerisches Feld]

Wie viele Kinder zwischen 0 -13 Jahren leben in Ihrem Haushalt?
_____ Kinder

Wie viele Kinder zwischen 14 - unter 18 Jahren leben in Ihrem Haushalt?
_____ Kinder

IMPRESSUM

Herausgeberin
Landesanstalt für Medien NRW
Zollhof 2
40221 Düsseldorf

T +49 211 77007-0
F +49 211 727170

info@medienanstalt-nrw.de
www.medienanstalt-nrw.de

Direktor: Dr. Tobias Schmid

Verantwortlich
Marie-Franca Hesse (Leiterin Kommunikation)
Dr. Meike Isenberg (Leiterin Intermediäre und Forschung)

Durchführende Institution
ConPolicy GmbH
Neue Grünstraße 17
10179 Berlin

Christina Barz
Marlene Münsch
Sophia Dasch