How Users Prompt Generative Search Engines for News and Political Information: An Empirical Taxonomy

Dorian Tsolak[a], Leona Murray[a,b]

[a] *Media Authority of North Rhine-Westphalia, Düsseldorf, Germany*
[b] *University of Münster*

1. Introduction

Generative search engines (GSEs) substantially alter how people look for information online (Haider & Sundin, 2019; Lewandowski, 2025). Current research often studies this phenomenon through systematic prompting and evaluation of responses. We posit that the prompts researchers craft are heavily influenced by research design decisions that lead to prompts not necessarily reflecting realistic scenarios of how people interact with GSEs. We contribute to solving this problem by providing an empirically-grounded taxonomy on how people prompt for news and political information in practice.

GSEs are often discussed in their function as news recommender systems (Schatto-Eckrodt et al., 2025). There are multiple concerns about the use of current GSEs as news recommenders such as lack of veracity of information (Kuznetsova et al., 2024; Kuai et al., 2025), political biases (Kuai et al., 2025; Yang et al., 2025) or lack of variety of sources (Gleason et al., 2023; Yang, 2025). Research examining these subjects is often conducted through systematic prompting and response evaluation while varying only certain characteristics of the prompt. From a research design perspective, this is understandable: One wants to measure the variation in the response, while only modifying the prompt in a minimal way. While data sets such as WildChat (Zhao et al., 2026) exist to provide some guidance, they often do not contain enough prompts on certain categories of information (e.g., news and political information) to infer realistic prompting behavior. Consequently, prompts are usually created by researchers ad hoc without being grounded in empirical evidence on actual patterns of user interaction with GSEs. Since prompts are known to heavily influence GSEs-responses - even though models are trying to improve prompts under the hood through guessing user intent - we consider additional research on users' prompting behavior essential for conducting future studies that audit the output of GSEs.

We conducted a survey experiment with slightly more than 2000 German GSE-users. The core of our survey consists of two information tasks (vignettes) where users are asked to interact with a GSE (gpt-4.1-2025-04-14) that is integrated into the survey. In the first scenario, users are asked to acquire information on a political or news topic they choose themselves. In the second scenario, users are given one of 26 news headlines and are asked to acquire information about the presented topic (we survey users' interest and knowledge on these topics beforehand and assign users accordingly). As controls, we survey sociodemographics, users' attitude towards AI systems as well as media literacy.

Data analysis was conducted using a multi-step approach. First, the data was coded qualitatively, and the coding scheme was iteratively refined. Afterwards, the prompts of the respondents were classified by instructing GPT-5-mini according to the developed coding scheme. The coding scheme differentiates between three broad categories of prompts: A) very short prompts which are not a whole sentence, B) a single sentence and C) multi-sentence prompts. For each category, sub-categories are defined that capture if the prompt consists of only of keywords, a question, a statement or a request.

SEASON 2026, September 15–17, 2026, Hamburg, Germany
EMAIL: dorian.tsolak@medienanstalt-nrw.de
ORCID: 0000-0001-6695-6169 (Dorian Tsolak); 0009-0006-4602-4595 (Leona Murray).
DOI: https://doi.org/10.48441/4427.3699

Additionally, each prompt has been checked for cross-category features and labelled accordingly. These include the sentiment, if the model has been addressed directly, if polite phrases have been used, if the model has been instructed to follow a persona, if a current affair has been addressed, if a relation to a place has been made and if a factual question has been asked.

Descriptively we find that 24 per cent of users state that they use AI-Chatbots when looking for political information, though this value is higher for men (28 per cent, than for women 18 per cent). Other popular uses include asking AI-Chatbots about household-related everyday problems or practical problems (56 per cent), about health (48 per cent), product advice (36 per cent) or diet and cooking advice (36 per cent). Furthermore, 67 per cent of initial prompts trigger a web search in the first scenario (free choice of topic) and 38 per cent in the second scenario (given topic). We find that most users rely on asking a single sentence question (Scenario 1: 59 per cent; scenario 2: 69 per cent). However, up to 27 per cent (Scenario 1) of users rely on very short prompts of 1 to 3 words, showing patterns of interactions known from traditional search engines. The resulting taxonomy of typical prompting patterns represents a central contribution of this study and examines in greater depth how prompts are generally structured and which syntactic structures are favored by users. Beyond the taxonomy itself, it includes examples that other researchers can adapt for their own studies to conduct prompt and response evaluation, using more realistic prompts than is currently possible.

2. References

Gleason, J., Hu, D., Robertson, R. E. & Wilson, C. (2023). Google the Gatekeeper: How Search Components Affect Clicks and Attention. *Proceedings Of The International AAAI Conference On Web And Social Media*, *17*, 245–256. https://doi.org/10.1609/icwsm.v17i1.22142

Haider, J., & Sundin, O. (2019). *Invisible Search and Online Search Engines: The Ubiquity of Search in Everyday Life* (1st ed.). Routledge. https://doi.org/10.4324/9780429448546

Kuai, J., Brantner, C., Karlsson, M., Van Couvering, E. & Romano, S. (2025). THE DARK SIDE OF LLM-POWERED CHATBOTS: MISINFORMATION, BIASES, CONTENT MODERATION CHALLENGES IN POLITICAL INFORMATION RETRIEVAL. *AoIR Selected Papers Of Internet Research*. https://doi.org/10.5210/spir.v2024i0.13977

Kuznetsova, E., Makhortykh, M., Vziatysheva, V., Stolze, M., Baghumyan, A. & Urman, A. (2024). In generative AI we trust: can chatbots effectively verify political information? *Journal Of Computational Social Science*, *8*(1). https://doi.org/10.1007/s42001-024-00338-8

Lewandowski, D. (2025). KI-Chatbots als Suchsysteme, Suchmaschinen als KI-Bots? Veränderungen, Begriffe und die Rolle des Datenbestands. In: M. Eibl (Hrsg.): *Datenströme und Kulturoasen - Die Informationswissenschaft als Bindeglied zwischen den Informationswelten*. Proceedings des 18. Internationalen Symposiums für Informationswissenschaft (ISI 2025), Chemnitz, Deutschland, 18.—20. März 2025. Glückstadt: Verlag Werner Hülsbusch, S. 150-167. https://doi.org/10.5281/zenodo.14925636

Schatto-Eckrodt, T., Liebig, L., Reiss, M., Geislinger, R., Schaetz, N., Merten, L., Schröder, J. T., Königslöw, K. K., Laugwitz, L., Knor, E. L., Semmann, M., Behre, J. & Lischka, J. A. (2025). ChatGPT as a news recommender system: Measuring source types and diversity across different interfaces. *SocArXiv*. https://doi.org/10.31235/osf.io/wjzp3_v3

Yang, J., Han, X. & Baldwin, T. (2025). Benchmarking Gender and Political Bias in Large Language Models. *ArXiv.org*. https://doi.org/10.48550/arxiv.2509.06164

Yang, K. (2025). News Source Citing Patterns in AI Search Systems. *ArXiv.org*. https://doi.org/10.48550/arxiv.2507.05301

Zhao, W., Ren, X., Hessel, J., Cardie, C., Choi, Y. & Deng, Y. (2024). WildChat: 1M ChatGPT Interaction Logs in the Wild. *ArXiv.org*. https://doi.org/10.48550/arxiv.2405.01470