

A comparison of online search engine autocompletion in Google and Baidu

Geng Liu, Pietro Pinoli, Stefano Ceri, Francesco Pierrri

Dipartimento di Elettronica, Informazione e Bioingegneria
Politecnico di Milano
Contact: name.surname@polimi.it

Abstract

Warning: This paper contains content that may be offensive or upsetting.

Online search engine auto-completions make it faster for users to search and access information. However, they also have the potential to reinforce and promote stereotypes and negative opinions about a variety of social groups. We study the characteristics of search auto-completions in two different linguistic and cultural contexts: Baidu and Google.

We find differences between the two search engines in the way they suppress or modify original queries, and we highlight a concerning presence of negative suggestions across all social groups. Our study highlights the need for more refined, culturally sensitive moderation strategies in current language technologies.

Introduction

Online search engines provide individuals with easy access to various information and content. Billions of queries and searches are made daily through Google (Feng and Shah 2022) or Baidu¹. Considering these volumes, the interactions between search engines and online users can significantly influence how people perceive and view the world (Azzopardi 2021; Lin et al. 2023).

Auto-suggestions or auto-completions in search engines are ubiquitous, streamlining user queries and accelerating the search process (Jiang 2014; Sullivan 2018; Gog, Pibiri, and Venturini 2020; Lin et al. 2023). However, these functionalities also carry the risk of reinforcing stereotypes (Allport 1979; Olteanu, Diaz, and Kazai 2020; Rogers 2023), and they could act as a mirror to society, drawing from user query or browsing history, which reflect people's attitudes and biases (Wang et al. 2018; Hazen et al. 2020; Lehmann and Buschek 2021; Cai, De Rijke et al. 2016; Robertson, Lazer, and Wilson 2018).

Research has explored auto-completions in English-language search queries that involve social groups, but there is limited information about such phenomenon in the context of China, as Western-based popular search engines, such as Google, are not accessible there (Leidinger and Rogers 2023; Rogers 2023; Miller

and Record 2017; Lin et al. 2023; Baker and Potts 2013; Jiang 2014).

We explore auto-completions of online searches in Baidu, a prominent Chinese-based search engine, and compare it to the most used Western-based search engine, Google. We base our analyses on two previous studies. The first study by (Choenni, Shutova, and van Rooij 2021) introduces a novel dataset of stereotypical attributes across various social groups and proposes an unsupervised method to elicit stereotypes and biases encoded by pre-trained language models. This research reveals the evolution of stereotypes with training data modifications during model fine-tuning. The second study by (Leidinger and Rogers 2023) investigates stereotype moderation in search engine autocompletion, results by Western search engines (Google, DuckDuckGo, and Yahoo). The authors highlight different moderation strategies across search engines and the impact on stereotype perpetuation, emphasizing the need for nuanced moderation and transparent policies in NLP.

Regarding the moderation of online search results, it is not completely clear how Google operates. They claim that:

"[their] systems aim to prevent policy-violating predictions from appearing. But if any such predictions do get past our systems, and we're made aware (such as through public reporting options), our enforcement teams work to review and remove them as appropriate. In these cases, we remove both the specific prediction in question and often use pattern-matching and other methods to catch closely related variations." (Sullivan 2020)

Baidu, instead, does not report any auto-completion content policy in public. In light of this, our study adopts the definition by (Leidinger and Rogers 2023) to categorize the suppression and prevalence of inconsistent results as (potential) moderation practices implemented by search engines.

We compare the moderation practices of Baidu with those of Google to identify similarities and differences in stereotype representation. This exploration is relevant, given Baidu's unique linguistic, cultural, and regulatory context, which may lead to distinct patterns of stereotype representation compared to its Western counterparts (Nadeem, Bethke, and Reddy 2021). Specifically, we formulate the following research questions:

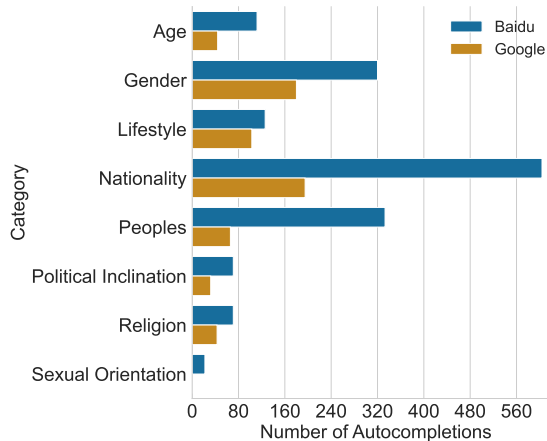


Figure 1: Number of auto-completions obtained in each category, for Baidu and Google, in the former case counted after removing the duplicates introduced by the two templates.

RQ1 What is the amount of search moderation/suppression on Baidu versus Google?

RQ2 What is the sentiment of search engine auto-completion on Baidu versus Google?

Our contributions are the following. We collected over 2K auto-completions for 146 unique social groups describing eight categories of individuals (age, gender, lifestyle, political inclination, racial group, nationality, religion, and sexual orientation) on Baidu and Google. We looked at the proportion of unresponded queries and inconsistent suggestions generated by the engine as a signal of moderation mechanisms. We also employed GPT-4 to calculate the sentiment of auto-completions and compared results between the two search engines. Our study aims to provide a better understanding of the dynamics of stereotypes on Chinese digital platforms, thereby contributing to the broader discourse on AI ethics and bias in language technologies.

Methods

Data Collection

We collected search auto-completion data from Google, the most widely adopted Western search engine, and Baidu, the main Chinese search engine. According to its guidelines², Google generates its auto-complete predictions based on several features (language, location, trending interests, and past searches). According to its Wiki (Baidu Baike), Baidu generates auto-completions based on hundreds of millions of user search terms every day³.

We followed the methodology employed in (Choenni, Shutova, and van Rooij 2021; Leidinger and Rogers 2023) for querying both engines' auto-

²<https://support.google.com/websearch/answer/7368877?hl=en>

³<https://baike.baidu.com/item/çŽ¿åž¶äŸŸŒæŒL/7139864>

completion services⁴ using the `requests` Python library. Specifically, we simulated an anonymous user querying auto-completions in November 2023.

We used the same search template as in (Choenni, Shutova, and van Rooij 2021; Leidinger and Rogers 2023):

- "Why are [social group] so [attribute]?"

Where [social group] represents the name of a specific social group, and [attribute] acts as a placeholder for characteristics suggested by the search engine's autocomplete function. We notice that each query for each group can return zero or more results.

For Google, we used directly the English template. In the case of Baidu, we translated queries using Baidu's translation API⁵ to ensure they carried the same meaning in Chinese. In particular, we specified two templates in Chinese (Elliott 1965), one *formal* and one *conversational*:

- **Formal Translation:** "为什么[Social Group] 这么[Attribute]?"
- **Conversational Translation:** "[Social Group] 为什么这么 [Attribute]?"

To build our queries, we translated English social group terms from (Leidinger and Rogers 2023) into Chinese, utilizing Baidu's translation API. The translations were then verified by three bilingual Master's students and a linguistic expert for contentious terms. Our analysis focused on identifying cultural differences and ensuring a fair comparison of stereotypes in search engine auto-completions between China and Western countries. We refined our term selection by excluding those without direct Chinese equivalents (e.g., 'frat boys', 'sorority girls', 'gingers') and focusing on complete names (e.g., 'teenagers', 'black people', 'white people') and specific familial terms with more search relevance in Baidu, e.g., '爷爷 (paternal grandfather)' and '奶奶 (paternal grandmother)' over '外公 (maternal grandfather)' and '外婆 (maternal grandmother)'. See Table 1 for some examples.

We find some differences in the results returned by Baidu using the two templates, which return respectively 1179 and 1176 results each, with only 579 common results; we leave a more in-depth investigation of differences for future work. In this paper, we refer to Baidu data simply as the union of the results returned by both templates, totalling 1776 unique auto-completions. We also conducted tests to check whether Baidu auto-completion differs depending on the location of the user. We queried the search engine using IP addresses located in China, Italy and Singapore, obtaining the same results in all cases. This suggests that Baidu provides a uniform delivery of auto-completions and that our findings are irrespective of geographical position. Figure 1 shows the total number of results obtained for each category and each search engine. We provide the full data and code to replicate our analyses in the repository associated with this paper⁶.

⁴<https://support.google.com/websearch/answer/7368877?hl=en>

⁵<https://fanyi-api.baidu.com/>

⁶https://github.com/leolepsyche/stereotype_in_baidu_

Category	No. groups	Social Groups	Social Groups Chinese
Age	8	Boomers, Children, Millennials, ...	婴儿潮, 儿童, 千禧一代, ...
Gender	23	Women, Men, Boys, Girls, ...	妇女, 男人, 男孩, 女孩, ...
Lifestyle	12	Geeks, Hippies, Celebrities, ...	极客, 嬉皮士, 名人, ...
Nationality	47	Americans, British, Chinese, ...	美国人, 英国人, 中国人, ...
Peoples	30	Africans, Asians, Europeans, ...	非洲人, 亚洲人, 欧洲人, ...
Political Inclination	8	Capitalists, Communists, Liberals, ...	资本家, 共产主义者, 自由主义者, ...
Religion	11	Buddhists, Christians, Muslims, ...	佛教徒, 基督徒, 穆斯林, ...
Sexual Orientation	7	Bisexuals, Gay people, Lesbians, ...	双性恋者, 同性恋者, 女同性恋, ...
Total	146		

Table 1: Number of unique social groups per category, with some examples in English and Chinese.

Sentiment Analysis

To quantify sentiment, we scored the sentiment of each full auto-completion using `gpt-4-1106-preview` model from OpenAI⁷. We chose this model as it shows superior performance to existing methods of automated sentiment analysis, and it achieves relatively high accuracy across many languages. Meanwhile, it is easy to use with simple prompts (Rathje et al. 2023; Prinzing and Fredrickson 2023).

We used OpenAI’s template for the prompt, where a system and a user role are specified⁸, and we followed (Rathje et al. 2023) to specify these roles:

- `system`: "Act as careful and accurate zero-shot text classifier."
- `user`: "Is the sentiment of this text positive, neutral, or negative? Please respond only with a number: (1) for positive, (2) for neutral, and (3) for negative."

For sentiment analysis, in both Google and Baidu, we considered the composition of the query plus the auto-completion, as in Chinese, the sentiment extracted for the auto-completion alone may significantly differ from the sentiment of the auto-completion in the context of the enclosing query.

Results

RQ1: Suppressed results in search engine auto-completion

We first look at the number of queries that do not return any suggestions from Baidu and Google and compute the proportion for each category.

As shown in Figure 2, we can observe that Google exhibits a much higher proportion of suppressed results compared to Baidu. In particular, no auto-completion is suggested by Google for all 7 Sexual Orientation groups, while Baidu does not return results for two groups only ("pansexual people" and "queer people"). Besides, Baidu returns suggestions for every group in the Nationality category and almost all people groups, while Google provides auto-completions respectively for less than half of Nationality (24 out of 47) and $\sim 1/3$ of People (9 out of 30).

google

⁷<https://platform.openai.com/docs/models/gpt-4-and-gpt-4-turbo>

⁸<https://platform.openai.com/docs/quickstart?context=python>

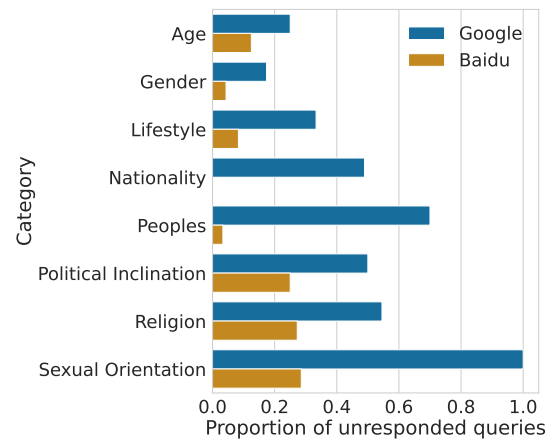


Figure 2: Proportion of unresponded queries for different categories, for Baidu and Google.

A manual check of unresponded queries for Baidu (e.g., boomers, hipsters, blondes) suggests that these groups are most likely not very familiar to Chinese people. For what concerns Google, it appears that unresponsiveness (e.g., homeless people, Black Americans, Asian kids, etc.) is probably due to the intention of avoiding promoting cultural, social and political stereotypes⁹.

Our observations reveal significant disparities between Baidu’s and Google’s moderation practices, particularly in reducing or suppressing auto-completion results. The moderation practices of Google are more conservative, with outputs significantly limited in each social category, especially in sexual orientation. Meanwhile, Baidu returns more suggestions across different categories than Google. Yet, auto-completions are not always consistent with the original query, as detailed in the next section.

RQ1: Inconsistent results in search engine auto-completion

We investigate the extent to which Baidu and Google respond to queries inconsistently, i.e., if they satisfy one of the two following conditions:

- (a) The responded query is completely different from

⁹https://github.com/leolepsyche/stereotype_in_baidu_google

the original template, e.g., "Why are schoolgirls so ... ?" → "Why are there more female students in American universities?"

- (b) The responded query follows the template, but it does not contain the original social group, e.g., "Why are schoolgirls so ...?" → "Why are teenagers so depressed?"

Overall, we find 53% (389 out of 734) inconsistent results for Google and 41% (685 out of 1659) for Baidu.

We compute the proportion of inconsistent results for each category and show them in Figure 3, highlighting the fraction of results satisfying condition (b) which is a subset of (a) by definition. We can observe that both search engines exhibit similarly high rates of inconsistent results across all categories, with larger values for Political Inclination, Religion and Sexual Orientation categories; for the last category, we notice that there are 0 results for Google, while for Baidu they are all inconsistent.

For Google, most of the time, social groups in the auto-completion do not match the one in the original query (cf. the hatched bar in Figure 3), while for Baidu, this occurs only for some categories. A manual check shows that the query template is either converted to different questions or changed to synonyms of the social group in the original query. Some examples for Sexual Orientation groups in Baidu are:

- "Why are bisexual people so ...?" → "Why can't we talk about bisexuality?"
- "Why are asexual people so ...?" → "Why do some people have asexuality?"

Compared with the previous analysis, these findings indicate that Baidu applies less content suppression to its searches, but it significantly alters many of the results (Google also exhibits this behaviour). We notice that we cannot discern whether inconsistency derives only from moderation policies or if it is rather a result of users' search behaviour, which is used by Baidu to generate auto-completions.

RQ2: Sentiment of search engine auto-completion

We now look at the sentiment of suggestions returned by both search engines to understand whether they convey negative views and might contribute to promoting stereotypes and bias. We compute the sentiment only for consistent results, i.e., suggestions that respect the original query template (and that contain the same social group).

As shown in Figure 4, we can observe that both Baidu and Google suggest negative auto-completions over 50% of the time across all categories (except Nationality, where only Baidu exceeds this number), with the highest rates for Age and Gender categories. Overall, Baidu provides a higher number of negative results for most categories, especially Lifestyle, Nationality, Political Inclination and Religion. We notice that there are no results for Sexual Orientation because all auto-completions returned by Baidu are inconsistent (while there are none for Google). We provide some exam-

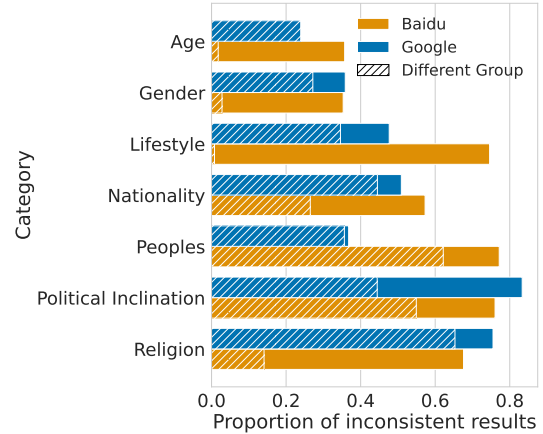


Figure 3: Proportion of inconsistent results (out of all responded queries) for different categories for each search engine. Hatched bars represent the proportion of results that satisfy condition (b). Note that we do not show the Sexual Orientation category as it contains no results for Google; for Baidu, it provides 100% inconsistent results, 60% of which do not contain the same social group.

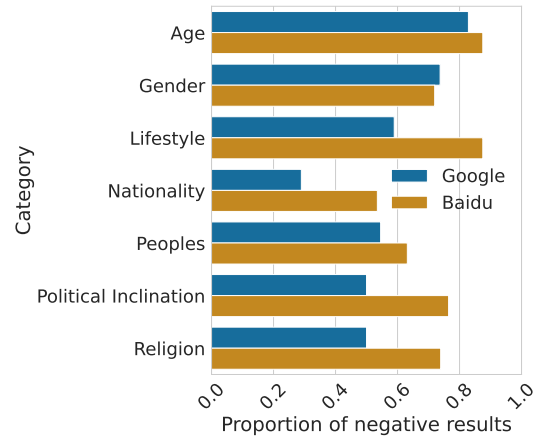


Figure 4: Proportion of negative results for consistent results of responded queries of different categories in Baidu and Google.

ples of negative auto-completions in Tables 3 and 2 for Baidu and Google, respectively.

Our analysis reveals a pervasive and worrisome negative bias in the auto-completions of both search engines, which is more pronounced for Baidu across categories compared to Google.

Investigating biases in Chinese-based LLMs

In this section, we briefly draw a picture of our ongoing research on examining biases in Chinese-based language models, which will leverage the data presented in this paper.

Conversational search is anticipated to be the next generation of search paradigms, as LLMs can enhance

Group	Examples
Asian parents	strict, controlling, harsh, strict about grades, judgemental
millennials	soft, lonely, poor, difficult to work with, hard to work with
females	emotional, materialistic, defensive, aggressive, immature

Table 2: Examples of negative sentiment auto-completions in the Google Dataset for the top 3 social groups in terms of number of auto-completions.

Group	Examples
poor people	difficult, tired, busy, many, difficult loans
old men	verbose, bad, hard to serve, annoying, stubborn
old people	stubborn, annoying, selfish, stupid, silly

Table 3: Examples of negative sentiment auto-completions in the Baidu Dataset, translated in English for the top 3 social groups in terms of number of auto-completions.

the indexing process of information retrieval systems (Mao et al. 2023; Sarkar 2024). Bing had already integrated GPT models in customized searches in February 2023 (Mehdi 2023). However, LLMs can expose online users to social biases or stereotypes in their outputs stemming from the training data and model architecture (Navigli, Conia, and Ross 2023; Dong et al. 2024). Previous research on bias evaluation and mitigation has been primarily focused on the English language, often within a US-centric context, and has predominantly addressed gender bias (Ducel, Néveol, and Fort 2023). However, biases are not universal; they differ across languages and cultures and are not confined solely to gender.

We aim to perform a comparative analysis of Western-based language models such as the OpenAI series¹⁰, and their Chinese-based counterparts, notably the ERNIE series¹¹. Through this comparative analysis, we aim to contribute to understanding cultural and linguistic biases in LLMs, fostering more fairness in large language models. In addition to analysing the sentiment of the output of those models, similar to what is presented in this paper, we will explore links between stereotypes and underlying emotions to understand the dimensions of prejudice in LLMs further. For instance, we could use the NRC Emotion Lexicon (Mohammad and Turney 2013) to quantify emotional responses to different social groups. We also plan to investigate the variance in toxicity levels between persona-aligned and non-persona LLM outputs within most commercial Chinese language models (Deshpande et al. 2023).

¹⁰<https://openai.com/chatgpt>

¹¹<http://research.baidu.com/Blog/index-view?id=185>

Discussion

We performed a comparative analysis of search engine auto-completions in two major search engines, Baidu and Google, which pertain to two different cultural and linguistic settings.

Our findings indicate that these differences are likely reflected in suggestions returned by the two engines to their users in terms of unresponded queries and inconsistent results. They also indicate a significant presence of negative auto-completions in both settings, suggesting that current moderation practices may not be sufficiently countering negative biases. Both engines exhibit the potential to perpetuate and promote derogatory and negative stereotypes across a wide range of social groups and categories.

Our research has certain limitations. We adapted social groups from a U.S.-centric study, which may not fully represent global perspectives. Search engine auto-completions are subject to change over time and vary according to each search engine’s policies. Lastly, for assigning a sentiment label to Chinese text, we employed a GPT-based model, which might not be fully accurate.

Despite these limitations, we believe our research is significant in better understanding content moderation policies in under-studied settings such as the Chinese language. To our knowledge, this is the first study to investigate stereotypes moderation policies between two major Western and Chinese search engines. Our findings highlight the danger of stereotype perpetuation by commercial language technology providers and call for more transparent moderation processes. In the future, we aim to extend this approach to measure the amount of bias in Chinese-based LLMs, along the lines of (Choenni, Shutova, and van Rooij 2021), and compare it to Western-centric foundation models.

Ethical Considerations

Our study aims to better understand the dynamics of stereotypes in two major search engines (Baidu and Google) and their impact on stereotype representation without violating social norms or privacy. It is important to note that some examples shown in the paper may be considered offensive; however, they do not reflect the author’s values and beliefs. Moreover, this study primarily focuses on publicly accessible data from search engines, employing a completely automated and anonymized data collection process. The specific stereotypical attributes and/or search queries cannot be traced back to individual users. We also strongly encourage future studies to consider the ethical aspects of biases in large language models from the outset of study design through to the final research outputs.

References

- Allport, G. W. 1979. The Nature of Prejudice. Unabridg. *Reading: Perseus Books*.
- Azzopardi, L. 2021. Cognitive Biases in Search: A Review and Reflection of Cognitive Biases in Information Retrieval. In *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval*,

- CHIIR '21, 27–37. New York, NY, USA: Association for Computing Machinery. ISBN 9781450380553.
- Baker, P.; and Potts, A. 2013. ‘Why do white people have thin lips?’ Google and the perpetuation of stereotypes via auto-complete search forms. *Critical discourse studies*, 10(2): 187–204.
- Cai, F.; De Rijke, M.; et al. 2016. A survey of query auto completion in information retrieval. *Foundations and Trends® in Information Retrieval*, 10(4): 273–363.
- Choenni, R.; Shutova, E.; and van Rooij, R. 2021. Stepmothers are mean and academics are pretentious: What do pretrained language models learn about you? In Moens, M.-F.; Huang, X.; Specia, L.; and Yih, S. W.-t., eds., *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 1477–1491. Online and Punta Cana, Dominican Republic: Association for Computational Linguistics.
- Deshpande, A.; Murahari, V.; Rajpurohit, T.; Kalyan, A.; and Narasimhan, K. 2023. Toxicity in chatgpt: Analyzing persona-assigned language models. In Bouamor, H.; Pino, J.; and Bali, K., eds., *Findings of the Association for Computational Linguistics: EMNLP 2023*, 1236–1270. Singapore: Association for Computational Linguistics.
- Dong, X.; Wang, Y.; Yu, P. S.; and Caverlee, J. 2024. Disclosure and Mitigation of Gender Bias in LLMs. *arXiv preprint arXiv:2402.11190*.
- Ducel, F.; Néveol, A.; and Fort, K. 2023. Bias Identification in Language Models is Biased. In *Workshop on Algorithmic Injustice*.
- Elliott, D. E. 1965. INTERROGATION IN ENGLISH AND MANDARIN CHINESE.
- Feng, Y.; and Shah, C. 2022. Has ceo gender bias really been fixed? adversarial attacking and improving gender fairness in image search. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, 11882–11890.
- Gog, S.; Pibiri, G. E.; and Venturini, R. 2020. Efficient and effective query auto-completion. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2271–2280.
- Hazen, T. J.; Olteanu, A.; Kazai, G.; Diaz, F.; and Gołębiewski, M. 2020. On the Social and Technical Challenges of Web Search Autosuggestion Moderation. *First Monday*, 27.
- Jiang, M. 2014. The business and politics of search engines: A comparative study of Baidu and Google’s search results of Internet events in China. *New media & society*, 16(2): 212–233.
- Lehmann, F.; and Buschek, D. 2021. Examining Autocompletion as a Basic Concept for Interaction with Generative AI. *i-com*, 19(3): 251–264.
- Leidinger, A.; and Rogers, R. 2023. Which Stereotypes Are Moderated and Under-Moderated in Search Engine Autocompletion? In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, 1049–1061.
- Lin, C.; Gao, Y.; Ta, N.; Li, K.; and Fu, H. 2023. Trapped in the search box: An examination of algorithmic bias in search engine autocomplete predictions. *Telematics and Informatics*, 85: 102068.
- Mao, K.; Dou, Z.; Mo, F.; Hou, J.; Chen, H.; and Qian, H. 2023. Large Language Models Know Your Contextual Search Intent: A Prompting Framework for Conversational Search. In Bouamor, H.; Pino, J.; and Bali, K., eds., *Findings of the Association for Computational Linguistics: EMNLP 2023*, 1211–1225. Singapore: Association for Computational Linguistics.
- Mehdi, Y. 2023. Reinventing search with a new AI-powered Microsoft Bing and Edge, your copilot for the web. *Official Microsoft Blog*, 7.
- Miller, B.; and Record, I. 2017. Responsible epistemic technologies: A social-epistemological analysis of autocompleted web search. *New Media & Society*, 19(12): 1945–1963.
- Mohammad, S. M.; and Turney, P. D. 2013. Nrc emotion lexicon. *National Research Council, Canada*, 2: 234.
- Nadeem, M.; Bethke, A.; and Reddy, S. 2021. StereoSet: Measuring stereotypical bias in pretrained language models. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 5356–5371.
- Navigli, R.; Conia, S.; and Ross, B. 2023. Biases in Large Language Models: Origins, Inventory, and Discussion. *J. Data and Information Quality*, 15(2).
- Olteanu, A.; Diaz, F.; and Kazai, G. 2020. When are search completion suggestions problematic? *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2): 1–25.
- Prinzing, M.; and Fredrickson, B. 2023. Reaching the Gold Standard: Automated Text Analysis with Generative Pre-trained Transformers Matches Human-Level Performance.
- Rathje, S.; Mirea, D.-M.; Sucholutsky, I.; Marjeh, R.; Robertson, C.; and Van Bavel, J. J. 2023. GPT is an effective tool for multilingual psychological text analysis. *PsyArXiv*.
- Robertson, R. E.; Lazer, D.; and Wilson, C. 2018. Auditing the Personalization and Composition of Politically-Related Search Engine Results Pages. In *Proceedings of the 2018 World Wide Web Conference, WWW '18*, 955–965. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee. ISBN 9781450356398.
- Rogers, R. 2023. Algorithmic probing: Prompting offensive Google results and their moderation. *Big Data & Society*, 10(1): 20539517231176228.
- Sarkar, D. 2024. Navigating the Knowledge Sea: Planet-scale answer retrieval using LLMs. *arXiv preprint arXiv:2402.05318*.
- Sullivan, D. 2018. How Google autocomplete works in Search. Retrieved November, 22: 2018.
- Sullivan, D. 2020. How Google autocomplete predictions are generated. Retrieved October, 8: 2020.

Wang, P.; Mi, X.; Liao, X.; Wang, X.; Yuan, K.; Qian, F.; and Beyah, R. A. 2018. Game of Missuggestions: Semantic Analysis of Search-Autocomplete Manipulations. In *NDSS*.