Auditing Google's AI Overviews for Local Information Seeking

Anonymous Author(s)

Abstract

The quality of our local information ecosystem has degraded, leading communities to have poorer access to information. Now, credibly worded generative AI responses generated in local contexts stand to potentially obscure underlying data quality. In this study, we seek to understand the relationship between the localness of sources available on the web and their likelihood to be featured in locally relevant AI outputs. We conduct a large-scale audit of Google's AI Overviews feature, which provides searchers with a generative AI summary aimed to quickly answer queries in natural language text. We construct a corpus of local and civically relevant queries, which we simulate on Google across all 3,000+ counties in the United States. Our results show that Google frequently relies on non-local sources, including aggregate databases, to answer local queries. AI Overviews generated in response to opinion-based queries rely proportionally more on localized social media, as do queries about culture and crime. These findings suggest that without targeted investments in local data quality and journalism, AI-mediated information access may exacerbate existing inequalities in civic knowledge and community engagement.

Keywords

Local Information Ecosystems, Local Online Groups, Generative AI, Algorithmic Audit

ACM Reference Format:

Anonymous Author(s). 2018. Auditing Google's AI Overviews for Local Information Seeking. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX)*. ACM, New York, NY, USA, 5 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

The mental model for online search is changing. The introduction of generative AI into popular search engines, and particularly Google, is shifting the structural way that user queries are answered. With generative AI — a set of tools that allow for the realistic on-the-fly creation of new content — search engines can now generate human-like responses to any query. While historically, query responses were presented as an ordered set of links, new AI features enable search engines to increasingly surface one authoritative summary to answer user queries [22]. Experimental work on the topic found that this change in search engine behavior may lead users to alter the types of queries that they formulate, that users may perceive AI Overviews as more authoritative, may not apply the same critical thinking skills when evaluating information, and change their interaction behaviors [15, 25, 26].

We focus this study on an audit of local information ecosystems on Google Search. This area of focus is socially consequential, and likely to significantly vary in different communities. Due to lack of funding and increased siphoning of ad revenues to tech platforms, the last decades have gutted the U.S. local news industry, leaving many communities with no news at all [21]. These communities are also termed "news deserts" [1]. The closure of local news outlets is associated with many community-level harms, from increased corruption to stronger polarization [2, 12, 13, 19]. In the face of these changes to local news production, scholars have called to widen our definition of what is considered local news [**?**]. While scholarly work has documented the consequences of shuttering outlets on community-level outcomes, and tracked the production of local newspapers, this current effort is the first to undertake the mapping of online local information availability more broadly on Google Search across the U.S.

In this study, we seek to understand the health of the local information ecosystem on Google Search. In particular, since AI summaries tend to homogenize underlying information quality, our findings focus on the relationships between user queries and cited website attributes.

Specifically, we answer the following questions:

(1) What types of websites are cited by Google's AI Overviews in local contexts?
(2) How does the type of user query influence the AI Overview websites that Google cites?

To answer these questions, we conduct the largest audit of Google's new AI Overview feature to date. This work also represents one of the first search audits of local information on Google, and the first audit of localized generative AI performance. We construct a set of 115 queries which represent community critical information needs and reflect real user searches. We then use a distributed architecture to generate searches for these queries over a three-month period, across all 3,000+ counties in the United States. To evaluate the outcomes, we contribute a novel framework for judging the localness of websites by *content localness* and *source localness*.

Based on our analysis, we conclude that a majority of sources (51.5%) that Google relies upon to construct local AI Overviews are not created by local communities. We also find that more nuanced queries, and queries about safety, are particularly likely to cite localized social media websites. Finally, Google relies most on community-created content when users search about local "events and news," suggesting that there are still no true replacements for local news provision in our online ecosystem. This study points to the future of local information acquisition in an AI-mediated world.

Conference acronym 'XX, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/2018/06
https://doi.org/XXXXXXX.XXXXXXX

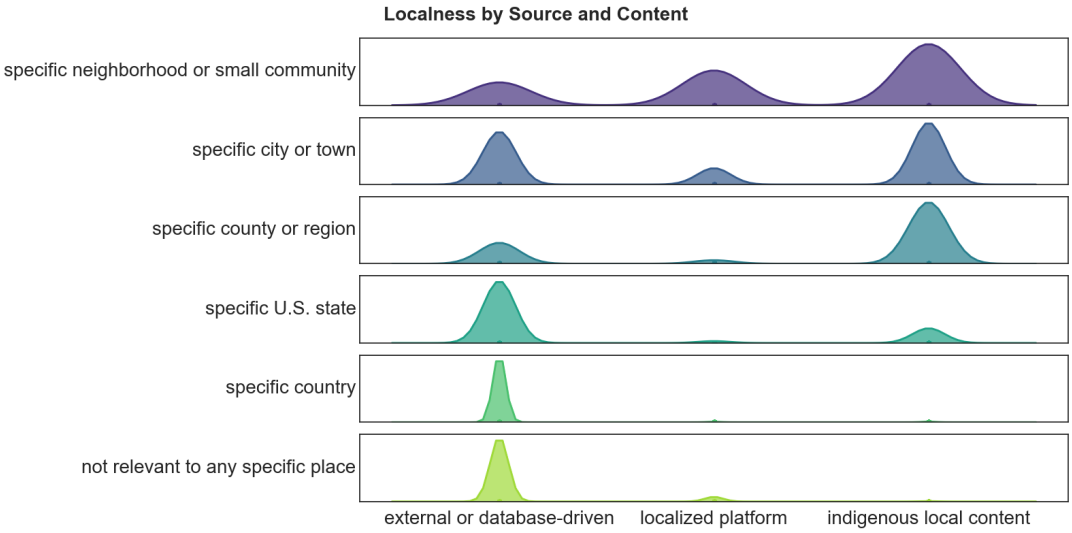

Figure 1: Ridge plots of the distributions of the source localness (x axis) vs the content localness (y axis) categories across all sources cited in AI Overviews.

2 Data and Methods

In this section, we describe our approach to auditing Google's AI Overviews, including how we gathered data and the details of the audit.

County-level Datasets

We leverage a few county-level datasets. First, we rely on past work by Fischer et al. [10] which makes available a list of 3,143 counties in the U.S. with their county seats and a latitude and longitude. We complement this geographic information with census data from the American Communities Survey (ACS). Finally, for the local news data, we rely on the Northwestern Local News Initiative, which remains the best available record of local paper newspapers [18].

Constructing local queries

Past auditing work has pointed to the necessity to gather multiple queries when conducting location-based search engine audits [20], and the need to gather realistic user queries [24]. We sought queries that balanced representing civically important needs, and those that represented real user requests.

Thus, we constructed a multi-step query generation pipeline. We begin with the eight key community information needs identified by Waldman [28]. For each community information need, we use the description to create a set of relevant keywords [11, 29]. We selected nine population-stratified cities/towns across the U.S. We then used Google Search's auto-complete function to recursively search each location with each identified keywords for each community information need. For example, we would search "Austin, TX weather emergency", retrieve all auto-complete results, and gather the auto-complete results for each of the returned suggestions. This auto-completion was conducted using the `suggests` package [23].

This search process returned 104,302 total queries, which we further filtered according to specific criteria. Since we know that AI Overviews are most commonly returned for queries posed as questions, we retain only queries that are phrased as questions (e.g. begin with *how*, *do*, or *are)*. Queries were de-duplicated, and only those including the searched-for location were maintained. Following this process, 1,284 queries remain.

Since many queries had overlapping topics, we further grouped these queries using hierarchical agglomerative clustering. This approach yielded 197 total clusters of queries. We manually removed clusters that included specific location markers that would not be appropriate for all counties. For each remaining cluster, we selected the query that was the most general, using an automatic concreteness metric [3]. In total, we finish this selection process with 115 queries that represent a broad set of community information needs.

Auditing AI Overviews

We then conduct a comprehensive audit of AI Overview results on Google Search. We leverage AWS lambda functions to programmatically search Google for each query for each county over a three-month period, from March 22nd to June 4th, 2025. The retrieval process was constructed such that search queries were conducted in all counties in a short time period, and any new features introduced to AI Overviews during the data collection time would not impact between-county differences. Each browser instance was launched using the county seat latitude and longitudinal parameters to emulate specific locations. The python packages `selenium`, `beautifulsoup`, and `websearcher` were used to directly download the HTML and store the AI Overview information. For each query, we retrieve whether an AI Overview was generated, the top three websites that the AI Overview cited, and the regular search engine results.

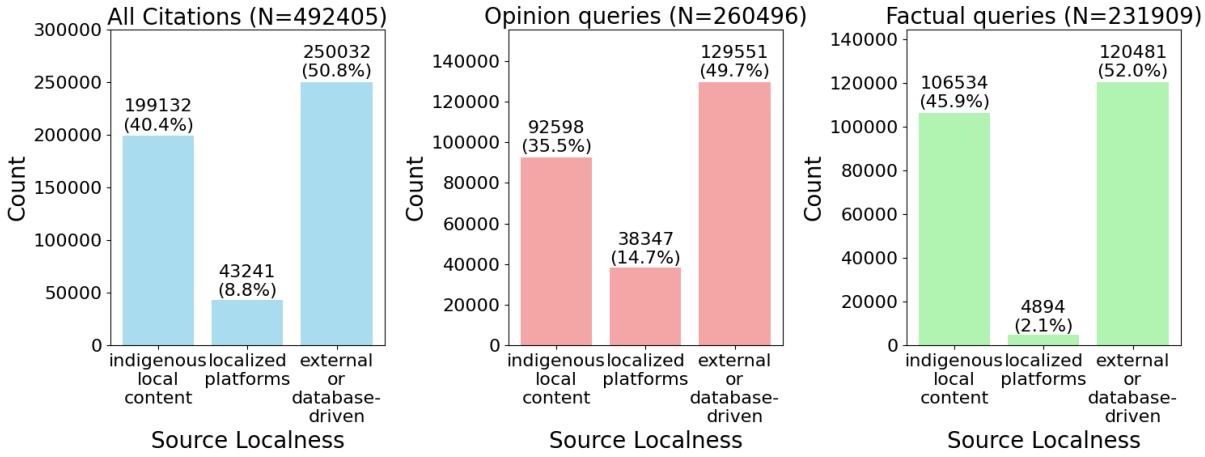

Figure 2: Bar graphs depict the source distributions for all (left), opinion-based (middle), and fact-based (right) queries. The x-axis is the source parameter, which represents whether content was created by local sources, filtered through a platform, or aggregated externally or through a dataset. Websites returned for opinion-based queries are significantly more likely to be localized social media platforms, and significantly less likely to be indigenous local content.

Localness Framework

AI Overviews include references to specific websites, which Google indicates support the information contained in the summaries. For each AI Overview, we additionally compute a set of labels associated with each source. These are aimed to capture two axes of localness. The first axis is *source localness*, which represents how the local information was created and by whom. We define three categories of *source localness*: "indigenous local content," "localized platforms and crowdsourcing," and "external or database-driven." The second category, *content localness*, defines the geographic scope of the content of a website. This category has six labels, ranging from "not relevant to any specific place" to relevant to a "specific neighborhood or small community."

Due to the large size of our dataset (with over 500,000 domains), we deploy LLMs to assign these two categories to each source. We used batch processing with gpt-4o-mini chat completions to label websites, and included in the prompt the website url, title, and text snippet for the specific query. For validation, one expert rater labeled a set of one hundred URLs according to their *source* and *content* localness. For both dimensions, exactly 90 out 100 websites were correctly tagged, indicating a 90% raw accuracy. Figure 1 shows how the two axes capture related, yet distinct concepts associated with localness. Higher *source* and *content* values demark more local sources.

3 Findings

In this section, we report on differences in the localness of sources that the AI overview cites, by different query types.

In total, 209,491 (66.7%) queries triggered an AI Overview. 7,856 (2.5%) queries triggered a banner saying "an AI Overview could not be generated for this search," and 7,345 (2.3%) queries did not load before the page timed out. The word count is bi-modally distributed with a long tail, a mean of 150 words and a median of 143 words.

AI Overviews by Query Type

Prior works have identified that query formulations significantly influence web search results. We thus conduct analyses to identify how two major query categories influence the sources cited in AI Overviews. The first variable we look at is the *query type*, which we manually label as *opinion* or *factual*, following prior work [7]. The second query-based variable we examine is *query category*. These query categories are grounded in the community information needs introduced by Waldman [28] and encompass *safety*, *culture*, *economy*, *civics*, *events*, *geography*, *health*, *education*, *politics*, and *transportation*.

Figure 2 shows all sources cited by the AI Overviews in our dataset, and their distributions by source localness. Half of sources cited by AI Overviews are external or database-driven($N = 250,032$ (50.8%)), followed by indigenous local content ($199,132 (40.4\%)$), and finally localized social platforms ($N = 43,241 (8.8\%)$) These distributions vary significantly for opinion-based vs factual questions. Specifically, AI Overviews that respond to opinion-based queries cite localized platforms much more (14.7% *vs* 2.1%) and indigenous local content much less (35.5% *vs* 45.9%), compared with factual questions. Notably, the fraction of sources returned that are external or database-driven remain relatively stable (51.0% *vs* 51.9%). A Mann-Whitney U test confirms that these differences in source localness between fact-based and opinion-based queries are statistically significant, $U = 31,351,337,916.5, p < 0.0001$.

AI Overviews by Query Category

Beyond the query type, we also evaluate how the source localness varies by query category. Figure 3 displays the ten query categories broken out by provenance source. AI Overviews in a few categories seem to specifically return different source distributions. The *economy* category contains the largest percentage of *external or database-driven* sources (69.5%), which is likely due to these queries requesting specific numbers or costs. The second category that

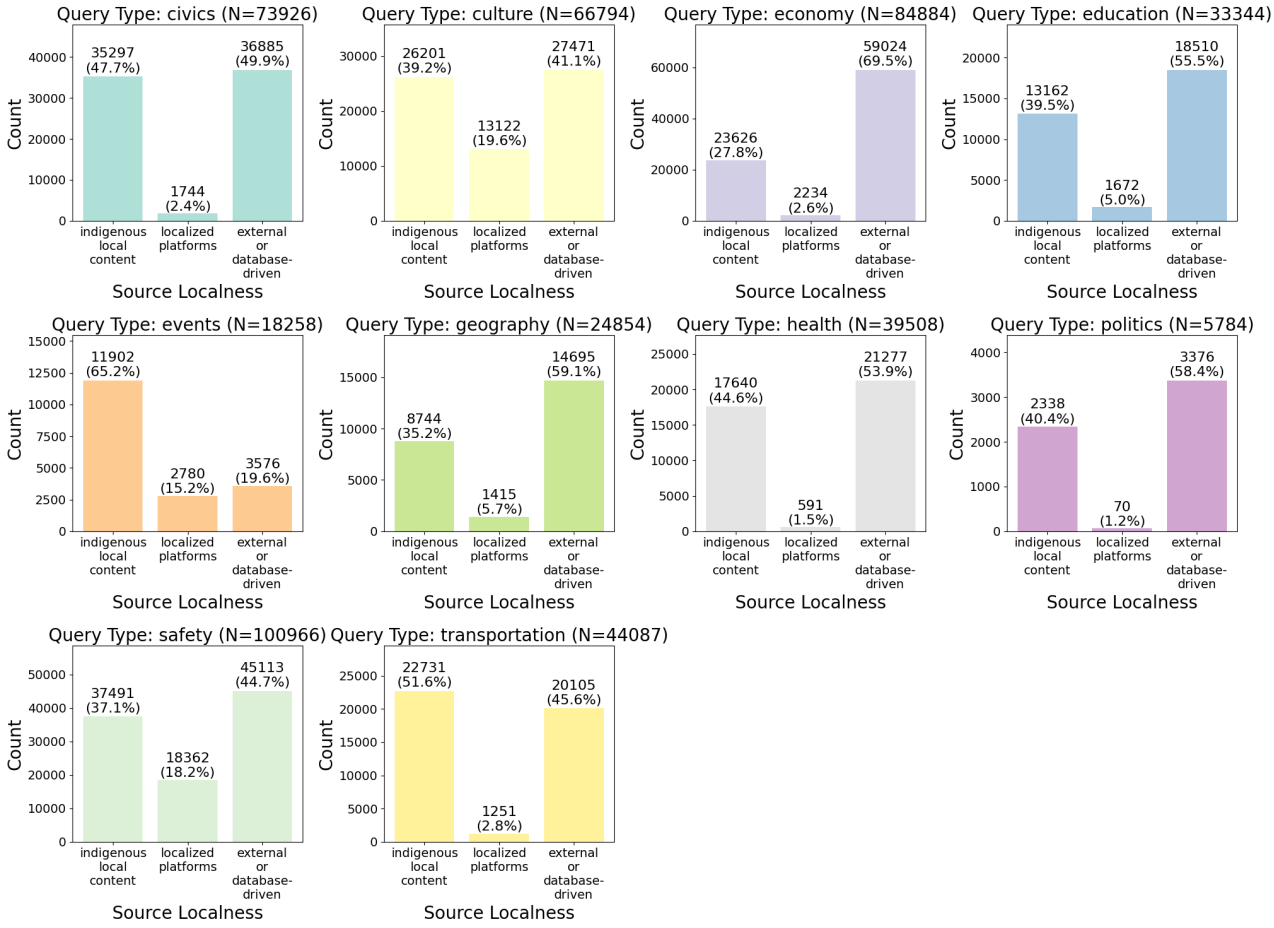

Figure 3: Bar graphs depict the source localness for each query category. The topic with the highest percentage of *external or database-driven* websites is *economy* (69.5%, top middle-right), the topic with the highest percentage of *indigenous local content* is *events* (65.2%, middle left), and the topics with the highest percentage of *localized social platforms* are *culture* (19.6%, top middle left) and *safety* (18.2%, bottom left).

stands out is the *events* category, which is the only category that is dominated by *indigenous local content* (65.2%). Finally, the *culture* and *safety* categories both comparatively cite higher percentages of localized social media platforms (19.6% and 18.2%, respectively).

4 Discussion

This study represents one of the first public audits of Google's AI Overview feature. The prevalence of AI Overviews was remarkably high, with an AI summary being generated for 66.7% of queries. While the queries we selected were specifically designed to be phrased as questions, and thus to be more likely to trigger AI Overviews, these results still exemplify the growing prevalence of AI Overviews on Google. This study thus joins the ranks of those that have measured AI prevalence across contexts including in crowdworking [27], reviewing [17], and across workplaces [4]. We find that generative AI summaries on Google are pervasive for question-based queries.

The current study also provides insights into the deployment of local generative AI systems. One major finding from this study is that half of citations that local AI Overviews rely on are produced by people outside of the location: they may be non-local sources or an aggregated database. Scholars observed people's increased reliance on data in local contexts, a trend that received attention during the Covid-19 pandemic [8, 30]. This finding suggests that local information exposure on Google is often unlikely to come from the communities themselves. Prior work found that Google News unfairly prioritizes national news over local news sources [10]: our analysis suggests this inequality extends beyond the news domain.

We also provide evidence of the pivotal role of localized social media in an AI-mediated information ecosystem. The percent of AI Overviews that cite a localized social media website is heavily dependent on the query type and category. In particular, our results suggest that local Facebook or Reddit discussions are more likely to be cited when people search opinion-based queries. Almost one fifth of total sources cited by AI Overviews about local culture or safety

and crime are localized social media platforms. This prevalence suggests that people looking for local information about crimes are highly likely to encounter discussions from localized social media. While these can be productive, and provide searchers an insight into local communities, prior work raised significant issues about the potential biases of such conversations [5, 6, 14].

Of all query types, we find only one where the most frequently cited websites are indigenous local content: the *events and news* category. This findings suggests that, while technological and scalable solutions exist for informing communities about many information needs, communities must continue to rely upon indigenous local content to receive up-to-date information about what is occurring in their communities. Kovach and Rosenstiel [16] argue that one of the key functions of journalism is *"the system societies generate to supply this information about what is and what's to come."* These findings suggest that while local news outlets have been financially displaced, scalable technologies offer limited viability to replace this function.

As AI systems increasingly mediate access to local information, our findings suggest that communities may receive much of their information from aggregate databases and non-local publishers rather than local sources. The homogenization of tone in Google AI Overviews is likely to obscure markers of answer quality, leading people to overrely and be less critical of lower-quality or external information [15, 25, 26]. This concern is particularly acute given that the presence of citations, rather than their quality, drive trust in AI generated answers [9]. This trend threatens to further weaken local information ecosystems and potentially limits communities' access to the indigenous local content necessary for civic engagement. Addressing these challenges will require coordinated investment in local information production and advocacy for tech companies to prioritize information quality.

References

[1] Penelope Muse Abernathy. 2020. *News Deserts and Ghost Newspapers: Will Local News Survive?* University of North Carolina Press.

[2] Charles Angelucci, Julia Cagé, and Michael Sinkinson. 2020. *Media Competition and News Diets.* Working Paper 26782. National Bureau of Economic Research. doi:10.3386/w26782

[3] Marianne Aubin Le Quéré and J. Nathan Matias. 2025. When Curiosity Gaps Backfire: Effects of Headline Concreteness on Information Selection Decisions. *Nature Scientific Reports* (2025).

[4] Jonathan Bright, Florence Enock, Saba Esnaashari, John Francis, Youmna Hashem, and Deborah Morgan. 2025. Generative AI is already widespread in the public sector: Evidence from a survey of UK public sector professionals. *Digital Government: Research and Practice* 6, 1 (2025), 1–13.

[5] Megan Brown, Zeve Sanderson, Sarah Graham, Minjoo Kim, Joshua Tucker, and Solomon Messing. 2024. Digital town square? Nextdoor's offline contexts and online discourse. *Journal of Quantitative Description: Digital Media* 4 (2024).

[6] Madiha Zahrah Choksi, Marianne Aubin Le Quéré, Travis Lloyd, Ruojia Tao, James Grimmelmann, and Mor Naaman. 2024. Under the (neighbor)hood: Hyperlocal Surveillance on Nextdoor. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24).* Association for Computing Machinery, New York, NY, USA, 1–22. doi:10.1145/3613904.3641967

[7] Alton YK Chua and Snehasish Banerjee. 2015. Measuring the effectiveness of answers in Yahoo! Answers. *Online Information Review* 39, 1 (2015), 104–118.

[8] Angel Desai, Pierre Nouvellet, Sangeeta Bhatia, Anne Cori, and Britta Lassmann. 2021. Data journalism and the COVID-19 pandemic: opportunities and challenges. *The Lancet Digital Health* 3, 10 (2021), e619–e621.

[9] Yifan Ding, Matthew Facciani, Ellen Joyce, Amrit Poudel, Sanmitra Bhattacharya, Balaji Veeramani, Sal Aguinaga, and Tim Weninger. 2025. Citations and trust in llm generated responses. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 39. 23787–23795.

[10] Sean Fischer, Kokil Jaidka, and Yphtach Lelkes. 2020. Auditing local news presence on Google News. *Nature Human Behaviour* 4, 12 (2020), 1236–1244.

[11] Lewis A Friedland. 2023. Taking it to the states: The origins of critical information needs. *The ANNALS of the American Academy of Political and Social Science* 707, 1 (2023), 21–28.

[12] Danny Hayes and Jennifer L. Lawless. 2018. The Decline of Local News and Its Effects: New Evidence from Longitudinal Data. *The Journal of Politics* 80, 1 (2018), 332–336. doi:10.1086/694105 Publisher: The University of Chicago Press.

[13] Danny Hayes and Jennifer L Lawless. 2021. *News hole: The demise of local journalism and political engagement.* Cambridge University Press.

[14] Waleed Iqbal, Vahid Ghafouri, Gareth Tyson, Guillermo Suarez-Tangil, and Ignacio Castro. 2023. Lady and the Tramp Nextdoor: Online Manifestations of Real-World Inequalities in the Nextdoor Social Network. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 17. 399–410. doi:10.1609/icwsm.v17i1.22155

[15] Carolin Kaiser, Jakob Kaiser, Rene Schallner, and Sabrina Schneider. 2025. A New Era of Online Search? A Large-Scale Study of User Behavior and Personal Preferences during Practical Search Tasks with Generative AI versus Traditional Search Engines. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems.* 1–7.

[16] Bill Kovach and Tom Rosenstiel. 2021. *The elements of journalism, revised and updated 4th edition: What newspeople should know and the public should expect.* Crown.

[17] Weixin Liang, Zachary Izzo, Yaohui Zhang, Haley Lepp, Hancheng Cao, Xuandong Zhao, Lingjiao Chen, Haotian Ye, Sheng Liu, Zhi Huang, et al. 2024. Monitoring ai-modified content at scale: A case study on the impact of chatgpt on ai conference peer reviews. *arXiv preprint arXiv:2403.07183* (2024).

[18] Local News Initiative. 2023. *The State of Local News 2023.* Technical Report. Medill School of Journalism, Northwestern University, Evanston, IL. https://localnewsinitiative.northwestern.edu/projects/state-of-local-news/2023/

[19] Ted Matherly and Brad N Greenwood. 2022. No news is bad news: The internet, corruption, and the decline of the fourth estate. *Corruption, and the Decline of the Fourth Estate (September 8, 2022)* (2022).

[20] Yelena Mejova, Tatiana Gracyk, and Ronald E Robertson. 2022. Googling for abortion: Search engine mediation of abortion accessibility in the United States. *arXiv preprint arXiv:2202.11760* (2022).

[21] Pew Research Center. 2021. *Trends and Facts on Newspapers | State of the News Media.* Technical Report. Pew Research Center. https://www.pewresearch.org/journalism/fact-sheet/newspapers/

[22] Elizabeth Reid. 2024. Generative AI in Search: Let Google do the searching for you. https://blog.google/products/search/generative-ai-google-search-may-2024/

[23] Ronald E Robertson, Shan Jiang, David Lazer, and Christo Wilson. 2019. Auditing Autocomplete: Recursive Algorithm Interrogation and Suggestion Networks. *Proc. of WebSci* (2019).

[24] Sebastian Schultheiß, Dirk Lewandowski, Sonja von Mach, and Nurce Yagci. 2023. Query sampler: generating query sets for analyzing search engines using keyword research tools. *PeerJ Computer Science* 9 (2023), e1421.

[25] Nikhil Sharma, Q Vera Liao, and Ziang Xiao. 2024. Generative echo chamber? effect of llm-powered search systems on diverse information seeking. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems.* 1–17.

[26] Sofia Eleni Spatharioti, David Rothschild, Daniel G Goldstein, and Jake M Hofman. 2025. Effects of LLM-based Search on Decision Making: Speed, Accuracy, and Overreliance. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems.* 1–15.

[27] Veniamin Veselovsky, Manoel Horta Ribeiro, Philip J Cozzolino, Andrew Gordon, David Rothschild, and Robert West. 2025. Prevalence and prevention of large language model use in crowd work. *Commun. ACM* 68, 3 (2025), 42–47.

[28] Steven Waldman. 2011. *The Information Needs of Communities:The Changing Media Landscape in a Broadband Age.* Carolina Academic Press.

[29] Matthew Weber, Peter Andringa, and Philip M Napoli. 2019. Local news on Facebook. *Sanford School of Public Policy, Duke University* (2019).

[30] Shangyuan Wu. 2021. Data "objectivity" in a time of coronavirus: uncovering the potential impact of state influence on the production of data-driven news. *Digital Journalism* 9, 9 (2021), 1303–1320.