

Who Gets Cited in Google AI Overviews? News Outlets, UGC Platforms, and Generative Search

Nick Hagar[✉]

nhagar@umn.edu

University of Minnesota
Minneapolis, Minnesota, USA

Nicholas Diakopoulos

Northwestern University
Evanston, Illinois, USA

Abstract

Google’s AI Overviews (AIOs) shift search from a ranked list of links toward a synthesized answer with a small set of citations. This change raises questions about which sources become visible as the basis for AI-generated search responses, especially for news organizations whose authority and audience access have long depended on search visibility. This article examines citations to selected news outlets and user-generated content (UGC) platforms in 351,568 AIOs collected by the National Internet Observatory. We compare citations to high-traffic U.S. news domains with citations to major UGC platforms, measuring prevalence, quarterly change, co-citation, and citation order. UGC platforms appeared in 24.15% of citation-bearing AIOs, compared with 5.94% for selected news outlets. The presence of citations to selected news outlets declined from 11.87% of AIOs in 2024 Q3 to 3.89% in 2026 Q1, while UGC citations rose sharply in the same time period. The two source types rarely appeared together (1.16% of AIOs). When they did co-occur, UGC platforms appeared before selected news outlets in 57.98% of cases. These findings suggest that UGC platforms function less as complements to selected news outlets than as substitutes within AIO citation lists. Our findings identify AI citations as a distinct mechanism of source visibility and show how generative search may marginalize news outlets not only through traffic loss, but through omission entirely.

Keywords

AI Overviews, generative search, search citations, news outlets, journalism and platforms

1 Introduction

Google Search has long mediated access to public knowledge. Search results are not neutral inventories of the web; they are ordered, filtered, and presented through ranking systems that shape which sources become visible and credible to users [12, 16, 26]. Yet the familiar search results page still organized that mediation around navigation. Users encountered links, snippets, publishers, and other cues, then decided which sources to open, compare, and evaluate. Google’s AI Overviews (AIOs) change that arrangement by generating a synthesized answer at the top of the results page and attaching a comparatively small set of cited links to that answer [14, 32]. This interface shifts search from retrieval toward response: Users are no longer asked only to choose among candidate sources, but are first given an answer that appears to have already integrated them.

This shift makes citation newly important and newly ambiguous. In human-authored work, a citation conventionally points to a source that the author has read, interpreted, and used to substantiate a specific claim. In an AIO, the relationship between answer

and source is produced by a proprietary retrieval, ranking, and generation system. Users cannot observe how candidate pages were selected, how conflicting sources were weighted, or whether a cited page supports the particular claim beside which it appears. Prior work on large language models and search has emphasized this provenance problem: Generative systems can produce fluent answers while obscuring the evidentiary chain that would allow users to verify them [19, 21, 35]. Recent studies of AI search and AI assistants have documented inaccurate, incomplete, or misleading citations, including cases in which systems attribute claims to sources that do not support them, fail to distinguish original reporting from republished material, or refer to AI-generated sources of suspect credibility [2, 11, 17]. The presence of a citation therefore should not be treated as equivalent to a transparent account of sourcing.

This ambiguity is especially consequential for news. News organizations have historically depended on search both as an audience pathway and as a public signal of relevance. AI-generated summaries intensify economic concerns that users may receive enough information from the results page and therefore not click through to the original publisher. Early evidence indicates that users are less likely to click links when an AI summary appears in Google results, while journalism industry surveys identify AI search and answer interfaces as a major source of anxiety about referral traffic and platform dependence [1, 6, 23]. However, the stakes are not only economic. AIO citations also create reputational relationships between AI systems and the sources they display. When an AI answer cites a news outlet, users may interpret the outlet as having supplied or endorsed the answer, even when the outlet had no control over the synthesis. Audience research on AI assistants suggests that users can assign some responsibility to cited news brands when AI-generated news responses are inaccurate [3]. The inverse problem has received less attention: What happens when news organizations are omitted from an authoritative AI-generated answer altogether?

Omission matters because AIO citations do not simply distribute clicks. They also signal which sources the system treats as relevant, credible, and useful for answering a question, extending the relationship between algorithmic ranking and source authority [5, 12]. If news outlets are excluded from those citations, their absence may weaken their public association with expertise in the domains they cover. At the same time, other source types may gain authority by appearing in their place. User-generated content (UGC) platforms are an important comparison because Google has explicitly designed search features to surface forum, Q&A, and social media content when queries may benefit from other people’s experiences, and to rank more content created from personal or expert points

of view [7]. UGC can be valuable precisely because it records situated perspectives that professional news may not capture. But UGC platforms and news organizations are not institutionally interchangeable. They differ in verification routines, professional norms, and editorial accountability [4, 8]. This distinction motivates our central question of whether UGC platforms may substitute for news outlets in Google AIO citations.

We examine this question using 351,568 AIO responses collected by the National Internet Observatory (NIO), a large-scale infrastructure project for studying online behavior in real-world browsing environments [22]. Our analysis asks how often AIOs generated in response to real user queries cite news outlets compared with UGC platforms, how the share of these citations changes over time, and how the two source types are ordered when they appear together. We find that UGC platforms are far more prominent than major news outlets in AIO citations. News citations appear relatively rarely, and their share decreases over the period we observe. By contrast, UGC platform citations become more prominent over time. The two source categories also rarely appear together in the same AIO. When they do, UGC sources tend to appear earlier in the citation list than news sources. Taken together, these patterns suggest that UGC platforms under real-world conditions are not merely supplementing news outlets in AIO citations; in many cases, they appear to be replacing them as visible sources for AI-generated answers.

The contribution of this study is threefold. First, it extends research on search and algorithmic authority by examining citations as a distinct form of platform-mediated source visibility. Second, the study contributes to research on journalism and platform dependence by identifying reputational omission as a complement to the more familiar concerns of traffic loss and misattribution. Third, the study offers empirical evidence about the source ecology of AIOs at scale and in response to real user queries.

2 Background

Search engines have long acted as algorithmic gatekeepers for news. They select, rank, summarize, and format information in ways that make some sources more visible than others [12, 13, 16, 37]. For news, visibility in search is also a form of public authority. Users often treat highly ranked search results as more relevant and credible, even when they know little about the underlying ranking process [26]. For news organizations, search visibility therefore affects more than referral traffic. It shapes which institutions appear available, authoritative, and relevant when users seek information about public affairs.

AI Overviews intensify this long-standing form of curation. Traditional search results pages ask users to navigate among a set of ranked sources. AIOs instead place a synthesized answer above or alongside those sources, with citations functioning as the visible trace of the system’s sourcing process [14, 32]. Google presents these AI features as an evolution of search, arguing that AI Overviews and AI Mode help users ask more complex questions and receive more useful answers with links to the web [33, 34]. However, the move from ranked links to synthesized answers changes the role of source selection. The sources cited in an AIO are not merely

options a user may consult; they become part of the evidentiary apparatus through which the answer is made credible.

What sources AIOs rely on to support that credibility remains an open empirical question. Google describes AI search features as grounded in web content and designed to show supporting links, but the company does not disclose the retrieval, ranking, or source-selection rules that determine which pages appear in a given AIO [14, 20]. Independent evidence is mixed. Some audits of AI search and AI assistants document citation errors, unsupported claims, and failures to distinguish original reporting from derivative or republished content [2, 11, 17, 19]. A recent large-scale study of Google AIOs found that, while cited domains were more credible than first-page results on average, a meaningful share of claims were unsupported by cited pages [39]. These findings suggest that AIO citations cannot be inferred from ordinary search ranking alone. They must be studied as a distinct source-selection layer.

There are reasons to expect news outlets to remain important within that layer. Professional journalism claims epistemic authority through routines of reporting, verification, and editorial accountability [4, 5, 8]. Those routines are especially relevant for current events, where information is contested, time-sensitive, and socially consequential. Google has also built news-specific ranking signals around topic authority, original reporting, and source reputation [15]. AIOs should therefore have a reason to keep news outlets visible, especially when queries touch on domains where professional reporting supplies timely evidence. Citations to recognizable news brands can help make an AI-generated response appear grounded and maintain the legitimacy of search as information infrastructure.

At the same time, the contemporary news environment gives Google reasons to rely heavily on user-generated content (UGC) platforms. The Reuters Institute’s 2026 Digital News Report identifies a milestone in news consumption: Across 48 markets, social media and video networks became the most widely used source of online news globally, ahead of news organizations’ own websites and apps [9]. YouTube, Instagram, TikTok, Facebook, Reddit, and other platforms now organize a large share of how users encounter public information. This shift does not mean that UGC is institutionally equivalent to journalism. Social media and video networks are associated with lower trust, misinformation concerns, and weaker accountability structures [9]. False or misleading information can spread rapidly in networked environments, and early work shows that AI systems are susceptible to manipulation via content on UGC platforms [18, 38, 40]. But UGC platforms also supply forms of situated, experiential, and community knowledge that search engines increasingly seek to surface. Google has explicitly redesigned parts of search to elevate social media posts and first-hand perspectives, and it has entered a data partnership with Reddit for this purpose [7, 27, 31].

These developments are further complicated by the economics and legality of AI-mediated news. News organizations have challenged AI companies over the use of copyrighted reporting in model training and AI-generated outputs, while publisher groups have criticized Google’s AI Mode and AI Overviews for using publisher content in ways that may reduce traffic and compensation [24, 25, 36]. At the same time, Google operates YouTube, one of the largest video platforms for news consumption [9]. The resulting incentives are uneven. News content can help AI systems answer

questions about current events, but licensing disputes and publisher resistance may make professional news a more contested input. UGC platforms, by contrast, offer large quantities of constantly updated content, often through platform relationships that can be operationalized through APIs and product integration.

Being cited in an AIO is also not straightforwardly beneficial for news outlets. Citation may create visibility, signal authority, and potentially route users to a publisher. But it can also expose a news brand to reputational risk if users interpret the outlet as responsible for an AI-generated synthesis it did not write or approve. Audience research suggests that users may assign some responsibility to cited news brands when AI assistants provide inaccurate news responses [3]. Non-citation presents the inverse risk: The outlet may be excluded from the visible evidentiary chain even when its reporting, expertise, or institutional authority would otherwise be relevant. This exclusion becomes more consequential as Google expands AI search [33, 34]. As these interfaces become more central to search, appearing or not appearing in AI citations may shape institutional visibility in ways that traditional click-through metrics only partially capture.

This makes citation a relational outcome rather than an all-or-nothing measure. A news outlet can be cited alone, cited alongside UGC, or omitted while UGC is cited instead. These patterns imply different models of how source categories operate in AI search. In a complementarity model, UGC and news sources appear together because they supply different kinds of evidence: Professional reporting may provide verification and institutional context, while UGC may provide lived experience, practical knowledge, or community interpretation. In a substitution model, UGC appears where news might otherwise be expected, either because the system treats UGC as sufficient evidence or because its retrieval and ranking processes favor UGC platforms over news outlets. Distinguishing these models requires examining not only whether news and UGC are cited, but also how often they co-occur, how their shares change over time, and where they appear in citation order.

We therefore ask:

RQ1: How often do Google AIOs cite news outlets compared with UGC platforms?

RQ2: How do news and UGC citation shares change over time?

RQ3: How often do news outlets and UGC platforms appear together in the same AIO?

RQ4: When news outlets and UGC platforms co-occur, how are they ordered in the citation list?

3 Data and Methods

This study uses AI Overview (AIO) responses collected by the National Internet Observatory (NIO) to examine how Google cites news outlets and user-generated content (UGC) platforms in AI-generated search responses [22]. We first describe the AIO citation corpus and the source-classification rules used to distinguish selected news outlets from UGC platforms. We then detail how we measure source composition, co-citation, and citation ordering.

3.1 AIO Citation Data

The raw data include AIO content records and AIO link records. Content records identify the Google search session, collection time,

search URL, extracted query, AIO text, and collection timeframe. Link records identify the cited URL, cited host, citation rank, link text, and description associated with the AIO. We deduplicate link records within each AIO by citation rank, URL, and normalized host, retaining the latest collected version.

351,568 AIOs had at least one cited host and form the primary analytic sample. The retained AIOs span 2024 Q3 through 2026 Q1. The median citation-bearing AIO cited three distinct domains. For analyses of co-citation, we further restrict the sample to AIOs with at least two distinct cited domains ($n = 328,785$), because a one-domain AIO has no opportunity to cite both news and UGC sources.

We normalized cited hosts by lowercasing domains, removing leading `www.`, stripping trailing punctuation, and extracting host-names either from the provided domain field or, when needed, from the cited URL. Domain matching uses a suffix rule that matches a root domain and its subdomains.

3.2 Source Classification

Each distinct AIO-host pair is coded for whether it matches a selected news domain, a UGC platform domain, both, or neither. The classification creates two auditable source measures that can be compared across AIOs.

The news measure comes from Press Gazette’s April 2026 ranking of the 50 biggest news websites in the United States by traffic, which Press Gazette compiles from Similarweb data [30]. We converted each listed site to a normalized root domain. Press Gazette’s April list includes `substack.com`; we removed that host from the news measure because it identifies a self-publishing platform rather than a specific news organization. The resulting measure therefore captures citations to 49 selected high-traffic news domains.

The UGC measure is also host based. We define a UGC-platform citation as a citation to a public social or content platform organized around user-created or user-contributed material. Our sample draws on Pew Research Center’s social-media/news and social-media-use fact sheets, which identify the major platforms through which U.S. adults use social media and receive news, and on the European Commission’s list of designated very large online platforms under the Digital Services Act [10, 28, 29]. The sample includes the principal public platform domains, common aliases, and short-link domains for YouTube, Reddit, TikTok, Facebook, Instagram, X/Twitter, Threads, Bluesky, Pinterest, Nextdoor, Snapchat, Rumble, and Truth Social.

3.3 Measures

The main unit of analysis is the AIO. For each citation-bearing AIO, we count the number of distinct cited domains, the number of news domains, the number of UGC-platform domains, and the number of remaining cited domains. We then assign each AIO to one of four source-composition states: *news only*, *UGC only*, *both*, or *neither*. These states provide the main descriptive measure for RQ1 and the quarterly trend measure for RQ2.

For co-citation analyses, we use only AIOs with at least two distinct cited domains. We define an AIO as co-citing news and UGC

when it contains at least one news domain and at least one UGC-platform domain. This measure supports RQ3 by distinguishing simple prevalence from relational source composition.

For ordering analyses, we use the citation rank supplied in the AIO link records. When an AIO contains both source types, we identify the first news citation and the first UGC-platform citation. We then calculate the percentage of mixed AIOs in which UGC appears before news, or where news appears before UGC. These measures address RQ4.

3.4 Analysis

The analysis proceeds in four steps, corresponding to the four research questions. First, we measure the overall prevalence of news and UGC-platform citations among citation-bearing AIOs. The main descriptive measure is the share of AIOs that cite news only, UGC only, both source types, or neither.

Second, we examine change over time by recalculating these source-composition shares within each collection quarter. This allows us to assess whether the citation environment shifts toward or away from news and UGC platforms over the observed period, using citation-bearing AIOs collected during each quarter as the denominator.

Third, we evaluate whether news and UGC platforms tend to appear together or apart. For this analysis, we restrict the denominator to AIOs with at least two distinct cited domains, because one-domain AIOs cannot co-cite source types. We then tabulate the frequency of source type co-occurrence, relative to all AIOs and to the appearance of each source type in isolation.

Fourth, we examine citation order within AIOs that cite both source types. For each mixed AIO, we compare the rank of the first news citation with the rank of the first UGC-platform citation. The main ordering measure is the share of mixed AIOs in which UGC appears before news, compared with the share in which news appears before UGC. This measure captures which source type receives earlier visible placement when both are present.

4 Results

The results show a clear asymmetry in how AIOs cite selected news outlets and UGC platforms. UGC platforms appear in a substantially larger share of citation-bearing AIOs than selected news outlets, and the two source types rarely appear together. When they do co-occur, UGC platforms more often appear earlier in the citation order.

UGC platforms appear more often than selected news outlets. RQ1 asked how often AIOs cite selected news outlets compared with UGC platforms. Among 351,568 citation-bearing AIOs, UGC platforms appear far more often than selected news outlets (24.15% and 5.94% of AIOs, respectively—Table 1). AIOs that cited UGC platforms without citing news outlets were also much more common than the inverse (22.99% and 4.78% of AIOs, respectively). The remaining 71.07% of AIOs cited sources from other categories, including health and medical resources (e.g., the Cleveland Clinic, Mayo Clinic), reference sites (e.g., Merriam-Webster, Britannica), government sites, and commercial sites (e.g., Amazon).

Table 1: Source composition of citation-bearing AI Overviews

Source composition	AIOs	% of citation-bearing AIOs
News only	16,794	4.78
UGC only	80,812	22.99
Both news and UGC	4,091	1.16
Neither selected news nor UGC	249,871	71.07
Any selected news	20,885	5.94
Any UGC platform	84,903	24.15

Note: Denominator is 351,568 citation-bearing AIOs. Rows in the lower panel are summary rows and are not mutually exclusive.

News citations decline while UGC citations grow. RQ2 asked how the share of news and UGC citations changes over time. Figure 1 shows the quarterly share of citation-bearing AIOs with any selected-news citation and any UGC-platform citation. Selected-news citations decline throughout the observed period, falling from 11.87% of citation-bearing AIOs in 2024 Q3 to 3.89% in 2026 Q1. UGC-platform citations initially decline between 2024 Q3 and 2024 Q4, but then rise sharply, increasing from 14.16% in 2024 Q4 to 37.33% in 2025 Q4 and 37.32% in 2026 Q1.

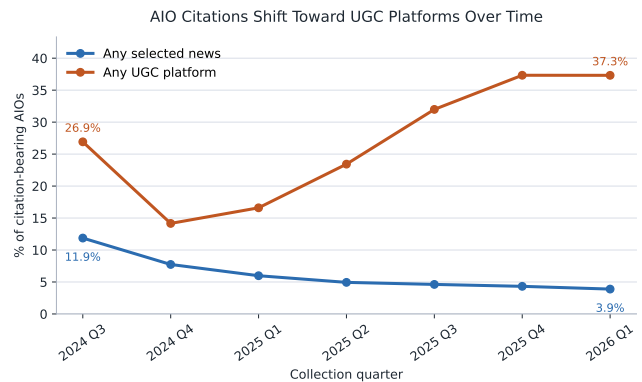


Figure 1: Quarterly prevalence of selected-news and UGC-platform citations among citation-bearing AIOs

Note: Points show the share of citation-bearing AIOs in each collection quarter. Quarterly denominators are 21,812, 63,342, 74,367, 75,647, 64,768, 48,288, and 3,344, respectively.

This trend suggests that the aggregate gap between news and UGC is growing: By the end of the period, AIOs were nearly ten times as likely to cite a UGC platform as a selected news outlet.

News and UGC rarely co-occur. RQ3 asked how often selected news outlets and UGC platforms appear together in the same AIO. They do so infrequently. Across all citation-bearing AIOs, only 1.16% cited both source types (Table 1). Restricting the comparison to AIOs with at least two distinct cited domains, where co-citation is possible, 4,091 AIOs (1.24%) cited both selected news and UGC.

UGC appears earlier when both source types are present. RQ4 asked how news and UGC are ordered when they co-occur. Among AIOs that cited both source types, UGC platforms appeared before selected news outlets in 57.98% of cases. Selected news outlets appeared before UGC platforms in 42.02% of cases.

This ordering result is narrower than the prevalence and co-citation results because it applies only to mixed AIOs. However, it reinforces the same overall pattern. Even in the relatively small set of AIOs where selected news outlets and UGC platforms both appear, UGC platforms more often receive earlier placement in the citation list. Together, these findings show that UGC platforms are not simply appearing as supplements to selected news citations. They are more prevalent overall, increasingly common over time, rarely co-occur with selected news outlets, and more likely to appear first when the two source types are both present.

5 Discussion

This study shows that AIO citations do not simply reproduce the source hierarchy that might be expected from professional news' epistemic authority over current events. Across more than 350,000 citation-bearing AIOs, UGC platforms appeared far more often than selected major news outlets, grew more prominent over time, rarely appeared alongside selected news sources, and tended to appear earlier when both source types were cited. These findings are most consistent with a substitution model of AIO citation. UGC platforms are not merely being added to news sources as complementary forms of evidence. In many AIOs, they appear to occupy citation space that selected news outlets do not.

This finding extends research on search and algorithmic authority by showing how generative search changes the object of source visibility. In traditional search, source authority is mediated through ranking: A news outlet may be more or less visible depending on where it appears in a list of links [12, 26]. In AIOs, visibility is mediated through a smaller and more selective citation layer attached to a synthesized answer. This changes the meaning of omission: A news outlet omitted from an AIO citation list is absent from the evidentiary frame of the answer itself. The results therefore suggest that AIO citations should be studied not only as traffic pathways, but as a distinct form of platform-mediated source recognition.

The findings also complicate the expectation that professional news will remain central in AI search because of its relationship to verification, original reporting, and public-interest information. Journalism's authority rests partly on institutional routines [4, 5, 8]. Google has also described topic authority and original reporting as important signals in news search [15]. Yet selected major news outlets appear in only a small share of AIO citations, and their presence declines over the observed period. This does not mean that AIOs never rely on journalism, nor does it establish that Google is intentionally replacing news with UGC. It does suggest, however, that the institutional authority of news does not automatically translate into visible citation authority in generative search.

For publishers, the results suggest that the risks of AI search extend beyond lost traffic. Existing concerns about AIOs often center on whether users will still visit publisher websites after receiving a summary on the search results page [6, 23]. That concern remains important, but our findings point to an additional problem of reputational omission. If AIOs answer news-relevant queries while citing UGC platforms instead of news outlets, publishers may lose not only traffic but also visible association with the topics they cover. This matters because search visibility has long helped define which institutions appear authoritative in public information

environments. Non-citation can therefore become a form of institutional marginalization, especially if users increasingly encounter information through AI-generated answers rather than through lists of navigable sources.

For search platforms, the findings point to a need for greater transparency around citation selection. AIOs present citations as evidence for a synthesized answer, but users and publishers have limited ability to assess why a source appeared, why another did not, or what role the cited source played in generating the response. This opacity is especially consequential when source types differ institutionally. One implication is that citation design should do more than provide links. Search platforms could make source type, source role, and provenance more legible, especially when an answer combines professional reporting (or analysis/opinion), platformed user content, official sources, and commercial pages. Another implication is that independent auditing of AI search should treat citations as structured outputs worthy of measurement in their own right. The source ecology of AI answers cannot be understood from answer text alone.

Limitations. Several limitations qualify these claims. First, the news measure captures a selected set of high-traffic U.S. news domains based on Press Gazette's ranking. This choice provides an auditable and substantively important sample of major news outlets, but it does not capture local news, niche publications, nonprofit outlets, wire services, international news organizations, or topic-specific expert sources that may also perform journalistic functions. The results therefore show the marginal position of selected major news outlets in AIO citations, not the full presence of all journalism.

Second, the UGC measure is platform-based at the domain level. A citation to a UGC platform may point to many different kinds of content. Domain-level classification is appropriate for measuring platform visibility, but it cannot determine the genre, producer, or evidentiary quality of each cited item. Future work should combine domain-level analysis with content-level coding to distinguish professional, creator, and ordinary-user sources within UGC platforms.

Third, the study examines a rapidly changing interface. Google continues to adjust AIOs, AI Mode, citation presentation, and ranking systems. The trends observed here may therefore reflect a particular stage in the development of AI search rather than a stable equilibrium. Continued measurement will be necessary to determine whether the rise of UGC citations and decline of selected-news citations persists as Google's AI search products mature.

Conclusion. AIOs are part of a broader reconfiguration of online sourcing in which platforms do not merely route users to information but increasingly synthesize, select, and authorize information on their behalf. The results of this study show that, within this reconfiguration, selected major news outlets occupy a relatively small and declining position in AIO citations, while UGC platforms occupy a much larger and growing one. This pattern matters because news and UGC are not epistemically interchangeable, even when they appear as equivalent links beneath an AI-generated answer. As search becomes more answer-oriented, citation becomes a site where platform design, source authority, and institutional visibility converge. Understanding who gets cited, who gets omitted, and how those choices change over time is therefore central to understanding the public role of generative search.

References

- [1] Saharsh Agarwal and Ananya Sen. 2026. Google AI Overviews and Publisher Traffic: Evidence from a Field Experiment. (2026). doi:10.2139/ssrn.6513059
- [2] Mowafak Allaham and Nicholas Diakopoulos. 2026. Synthetic Sources?: Auditing Generative Search Engine Citations for Evidence of AI-Generated Sources. *arXiv* (2026). arXiv:2605.23684 doi:10.48550/arxiv.2605.23684
- [3] BBC and Ipsos. 2025. *Audience Use and Perceptions of AI Assistants for News: Research Findings*. Technical Report. BBC and Ipsos. Accessed 2026-06-29. <https://www.bbc.co.uk/aboutthebbc/documents/audience-use-and-perceptions-of-ai-assistants-for-news.pdf>
- [4] Petter Bae Brandtzæg, Marika Lüders, Jochen Spangenberg, Linda Rath-Wiggins, and Asbjørn Følstad. 2016. Emerging Journalistic Verification Practices Concerning Social Media. *Journalism Practice* 10, 3 (2016), 323–342. doi:10.1080/17512786.2015.1020331
- [5] Matt Carlson. 2017. *Journalistic Authority: Legitimizing News in the Digital Era*. Columbia University Press, New York, NY. <https://cup.columbia.edu/book/journalistic-authority/9780231174459/>
- [6] Athena Chapekis and David Lieb. 2025. Google Users Are Less Likely to Click on Links When an AI Summary Appears in the Results. Pew Research Center. Accessed 2026-06-29. <https://www.pewresearch.org/short-reads/2025/07/22/google-users-are-less-likely-to-click-on-links-when-an-ai-summary-appears-in-the-results/>
- [7] Lauren Clark. 2023. Learn from Others' Experiences with More Perspectives on Search. Google Blog. Accessed 2026-06-29. <https://blog.google/products-and-platforms/products/search/google-search-perspectives/>
- [8] Mark Deuze. 2005. What Is Journalism? Professional Identity and Ideology of Journalists Reconsidered. *Journalism* 6, 4 (2005), 442–464. doi:10.1177/1464884905056815
- [9] Jim Egan. 2026. Overview and Key Findings of the 2026 Digital News Report. Reuters Institute for the Study of Journalism. Accessed 2026-06-29. <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2026/dnr-executive-summary>
- [10] European Commission. 2026. Supervision of the designated very large online platforms and search engines under DSA. <https://digital-strategy.ec.europa.eu/en/policies/list-designated-vlops-and-vloses>. Accessed 2026-06-30.
- [11] James Fletcher and Dorien Verckist. 2025. *News Integrity in AI Assistants: An International Public Service Media Study*. Technical Report. BBC and European Broadcasting Union. Accessed 2026-06-29. https://www.ebu.ch/files/live/sites/ebu/files/Publications/MIS/open/EBU-MIS-BBC_News_Integrity_in_AI_Assistants_Report_2025.pdf
- [12] Tarleton Gillespie. 2014. The Relevance of Algorithms. In *Media Technologies: Essays on Communication, Materiality, and Society*, Tarleton Gillespie, Pablo J. Boczkowski, and Kirsten A. Foot (Eds.). MIT Press, Cambridge, MA, 167–194. doi:10.7551/mitpress/9780262525374.003.0009
- [13] Eric Goldman. 2006. Search Engine Bias and the Demise of Search Engine Utopianism. *Yale Journal of Law and Technology* 8 (2006), 188–200. <https://yjl.org/search-engine-bias-and-demise-search-engine-utopianism>
- [14] Google Search Central. 2026. AI Features and Your Website. Accessed 2026-06-29. <https://developers.google.com/search/docs/appearance/ai-features>
- [15] Jen Granito. 2023. Understanding News Topic Authority. Google Search Central Blog. Accessed 2026-06-29. <https://developers.google.com/search/blog/2023/05/understanding-news-topic-authority>
- [16] Lucas D. Introna and Helen Nissenbaum. 2000. Defining the Web: The Politics of Search Engines. *Computer* 33, 1 (2000), 54–62. doi:10.1109/2.816269
- [17] Klaudia Jazwińska and Aishvarya Chandrasekar. 2025. We Compared Eight AI Search Engines. They're All Bad at Citing News. Columbia Journalism Review, Tow Center for Digital Journalism. Accessed 2026-06-29. https://www.cjr.org/tow_center/we-compared-eight-ai-search-engines-theyre-all-bad-at-citing-news.php
- [18] David M. J. Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, and Jonathan L. Zittrain. 2018. The Science of Fake News. *Science* 359, 6380 (2018), 1094–1096. doi:10.1126/science.aao2998
- [19] Nelson F. Liu, Tianyi Zhang, and Percy Liang. 2023. Evaluating Verifiability in Generative Search Engines. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, Houda Bouamor, Juan Pino, and Kalika Bali (Eds.). Association for Computational Linguistics, Singapore, 7001–7025. doi:10.18653/v1/2023.findings-emnlp.467
- [20] John Mueller. 2025. Top Ways to Ensure Your Content Performs Well in Google's AI Experiences on Search. Google Search Central Blog. Accessed 2026-06-29. <https://developers.google.com/search/blog/2025/05/succeeding-in-ai-search>
- [21] Pranav Narayanan Venkit, Philippe Laban, Yilun Zhou, Yixin Mao, and Chien-Sheng Wu. 2024. Search Engines in an AI Era: The False Promise of Factual and Verifiable Source-Cited Responses. arXiv:2410.22349 [cs.CL] doi:10.48550/arXiv.2410.22349
- [22] National Internet Observatory. 2026. National Internet Observatory. Accessed 2026-06-29. <https://nationalinternetobservatory.org/>
- [23] Nic Newman. 2026. *Journalism, Media, and Technology Trends and Predictions 2026*. Technical Report. Reuters Institute for the Study of Journalism. Accessed 2026-06-29. <https://reutersinstitute.politics.ox.ac.uk/journalism-media-and-technology-trends-and-predictions-2026>
- [24] News/Media Alliance. 2024. News/Media Alliance Calls on FTC, DOJ to Investigate Google's Misappropriation of Digital News Publishing and Stop Expansion of Generative AI Overviews Offering. Accessed 2026-06-29. <https://www.newsmediaalliance.org/release-news-media-alliance-calls-on-ftc-doj-to-investigate-googles-misappropriation-of-digital-news-publishing-stop-expansion-of-generative-ai-overviews-offering/>
- [25] News/Media Alliance. 2025. News/Media Alliance Statement on Google AI Mode. Accessed 2026-06-29. <https://www.newsmediaalliance.org/google-ai-mode-statement/>
- [26] Bing Pan, Helene Hembrooke, Thorsten Joachims, Lori Lorigo, Geri Gay, and Laura Granka. 2007. In Google We Trust: Users' Decisions on Rank, Position, and Relevance. *Journal of Computer-Mediated Communication* 12, 3 (2007), 801–823. doi:10.1111/j.1083-6101.2007.00351.x
- [27] Rajan Patel. 2024. An Expanded Partnership with Reddit. Google Blog. Accessed 2026-06-29. <https://blog.google/company-news/inside-google/company-announcements/expanded-reddit-partnership/>
- [28] Pew Research Center. 2025. Social Media and News Fact Sheet. <https://www.pewresearch.org/journalism/fact-sheet/social-media-and-news-fact-sheet/>. Accessed 2026-06-30.
- [29] Pew Research Center. 2025. Social Media Fact Sheet. <https://www.pewresearch.org/internet/fact-sheet/social-media/>. Accessed 2026-06-30.
- [30] Press Gazette. 2026. Biggest websites for news US: Top 50 updated each month. <https://pressgazette.co.uk/media-audience-and-business-data/media-metrics/most-popular-websites-news-us-monthly-3/>. April 2026 section; monthly ranking based on Similarweb data. Accessed 2026-06-30.
- [31] Reddit, Inc. 2024. Expanding Our Partnership with Google. Accessed 2026-06-29. <https://redditinc.com/news/reddit-and-google-expand-partnership>
- [32] Elizabeth Reid. 2024. Generative AI in Search: Let Google Do the Searching for You. Google Blog. Accessed 2026-06-29. <https://blog.google/products/search/generative-ai-google-search-may-2024/>
- [33] Elizabeth Reid. 2025. AI in Search: Going Beyond Information to Intelligence. Google Blog. Accessed 2026-06-29. <https://blog.google/products-and-platforms/products/search/google-search-ai-mode-update/>
- [34] Elizabeth Reid. 2026. A New Era for AI Search. Google Blog. Accessed 2026-06-29. <https://blog.google/products-and-platforms/products/search/search-io-2026/>
- [35] Chirag Shah and Emily M. Bender. 2022. Situating Search. In *Proceedings of the 2022 Conference on Human Information Interaction and Retrieval*. Association for Computing Machinery, New York, NY, 221–232. doi:10.1145/3498366.3505816
- [36] The New York Times Company. 2023. Complaint, The New York Times Company v. Microsoft Corporation et al., No. 1:23-cv-11195. U.S. District Court for the Southern District of New York. Accessed 2026-06-29. <https://www.courthousenews.com/wp-content/uploads/2023/12/new-york-times-microsoft-open-ai-complaint.pdf>
- [37] Daniel Trielli and Nicholas Diakopoulos. 2019. Search as News Curator: The Role of Google in Shaping Attention to News Information. *Proc. CHI* (2019).
- [38] Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The Spread of True and False News Online. *Science* 359, 6380 (2018), 1146–1151. doi:10.1126/science.aap9559
- [39] Haofei Xu, Umar Iqbal, and Jacob M. Montgomery. 2026. Measuring Google AI Overviews: Activation, Source Quality, Claim Fidelity, and Publisher Impact. arXiv:2605.14021 [cs.CY] <https://arxiv.org/abs/2605.14021>
- [40] Tingwei Zhang, Harold Trieman, and Vitaly Shmatikov. 2026. Deep-Research Agents Can Be Poisoned via User-Generated Content. arXiv:2605.24245 [cs.CR]. doi:10.48550/arXiv.2605.24245