

# Sourcing behavior and the role of news media in AI-powered search engines in the digital media ecosystem: Comparing political news retrieval across five languages

Cornelia Brantner<sup>a</sup>, Michael Karlsson<sup>a</sup>, Joanne Kuai<sup>a,b,\*</sup>

<sup>a</sup> Department of Geography, Media and Communication, Karlstad University, Sweden

<sup>b</sup> School of Media and Communication, RMIT University, Melbourne, Australia

## ARTICLE INFO

### Keywords:

AI governance  
Digital media ecosystem  
Generative search engine  
Large language models  
News media  
Sourcing

## ABSTRACT

This study examines the role of news media in the context of generative AI-enhanced search engines, focusing on the 2024 Taiwan presidential election. Using Microsoft's Copilot, we conducted a comparative analysis by prompting election news in five languages: English, Traditional Chinese, Simplified Chinese, German, and Swedish. While Copilot uses mainly professional news media, provides quick access to synthesized information, and exhibits source transparency, it frequently creates misinformation and misattributes news sources. The analysis highlights variations in Copilot's sourcing behavior, showing a strong reliance on English-language sources, particularly those from the UK and US, across different prompting languages. Such reliance raises concerns about the homogenization of information and the marginalization of regional perspectives. The study underscores the critical role and dilemma of news media, which, while serving as authoritative sources in democratic societies, must navigate an increasing AI-mediated information ecosystem to maintain autonomy vis-à-vis powerful technological infrastructures. By evaluating Copilot's sourcing practices and misinformation prevalence, this research contributes to the discourse on AI's impact on news dissemination, media diversity, and democratic processes. Specifically, we discuss the consequences of two approaches available to news media to prevent their content from being used without compensation: opting out of crawling ("platform counterbalancing") or establishing partnerships with AI companies. Current regulatory efforts, including copyright reforms and the EU AI Act, fall short of safeguarding journalism or regulating AI. We propose policy and regulatory recommendations to improve transparency, factual correctness, accuracy in source attribution, and accountability in AI-generated content, supporting informed citizenship in the digital age.

## 1. Introduction

The launch of Google's SGE (Search Generative Experience) and Open AI's ChatGPT Search in 2024 signaled a significant evolution in generative search engines or AI search, redefining how users access and interact with information online. Microsoft led this

This article is part of a special issue entitled: AI & digital media ecosystems published in Telecommunications Policy.

\* Corresponding author. Department of Geography, Media and Communication, Karlstad University, Universitetsgatan 2, 651 88 Karlstad, Sweden.

E-mail addresses: [cornelia.brantner@kau.se](mailto:cornelia.brantner@kau.se) (C. Brantner), [michael.karlsson@kau.se](mailto:michael.karlsson@kau.se) (M. Karlsson), [joanne.kuai@rmit.edu.au](mailto:joanne.kuai@rmit.edu.au) (J. Kuai).

<https://doi.org/10.1016/j.telpol.2025.102952>

Received 29 November 2024; Received in revised form 29 January 2025; Accepted 12 March 2025

Available online 4 April 2025

0308-5961/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

transformation in February 2023 with Bing Chat, later rebranded as Copilot in September. Representing a generative AI-enhanced evolution of Microsoft's Bing search engine, Copilot is powered by OpenAI's ChatGPT technology, integrating a Large Language Model (LLM) with search engine functionalities to generate responses based on current news.<sup>1</sup> Microsoft announced that it operated on "a new, next-generation OpenAI large language model that is more powerful than ChatGPT and customized specifically for search" (Mehdi, 2023). A considerable "media hype" (Kelly et al., 2023) followed after Microsoft and Google started their experiments of integrating generative AI into search engines (see, e.g., Roose, 2023). Journalists (Stokel-Walker, 2023) and platform giants alike (see for Microsoft, Mehdi, 2023; for Google, Pichai, 2023) made deterministic predictions that the service has the potential to revolutionize search engines and fundamentally change how users navigate the Internet and find information. This recent addition to the media ecosystem also led to scholarly concerns that the generative AI allows platform providers to bypass news publishers when it comes to information provision, "or at least not sending attention and traffic their way any longer" (Simon, 2023).

Generative search engines can simultaneously gather, select, process, and present information in *narrative form*—a core journalistic skill (Bird & Dardenne, 2009; Wahl-Jorgensen, 2013)—and distribute it rapidly at the user's discretion. These steps are key parts of the journalistic work process (Domingo et al., 2008; Karlsson, 2011). If generative search engines become entry points and even one-stop shops for citizens maneuvering a vast information landscape, they are poised to function as de facto gatekeepers in the ecosystem. This raises concerns about the extent to which they will draw from and direct people to trusted sources like news media. However, by January 2025, a slight majority of publishers opted out of allowing generative AI to crawl their content using the robots.txt convention (Welsh, n.d.), no longer willing to feed the AI for free with their products. Conversely, publishers such as Associated Press, German Springer, and the British Guardian chose to enter agreements with Microsoft's news service. If news media choose to put up walls against generative AI capture, it is essential to study what actors will fill that void. Accurate and transparent information retrieval is crucial for an informed citizenry, the foundation of a functioning democracy (Carpini & Keeter, 1996). Access to reliable information enables citizens to make informed decisions, engage in public discourse, and hold authorities accountable. Thus, if search engine chatbots become a large part of people's media diet, the quality, sources, and traceability of their information are critical.

Policy and regulatory efforts have been made worldwide to help journalism counterbalance the power of the platforms, but many of these efforts have been criticized for their ineffectiveness, with some resulting unintended harmful impacts and even backlash. For example, the 2019 EU Copyright Directive Article 15 has been criticized for worsening the power concentration of media conglomerates and of global platforms, leaving smaller players at a disadvantage (Quintais, 2020). The latest amendment to the Chinese Copyright Law introduced in 2021 was also found to reinforce the position of the state, investor, and tech industry at the expense of journalism (Kuai et al., 2022). The absence of designation in the 2021 News Media Bargaining Code in Australia has left media companies to compete for favor from the platforms (Bossio & Barnet, 2023). The 2024 EU AI Act can also result in technology developers and standardization bodies taking control, rather than regulators and editors (Helberger, 2024). However, the impact of other regulatory and governance efforts, such as the 2023 Canada Online News Act, the proposed Journalism Competition and Preservation Act in the US, the Digital Services Act, the Digital Markets Act, and the proposed European Media Freedom Act in the EU, remain to be seen.

While there has recently been vast scholarly attention to the integration of AI in the newsroom (e.g., Cools & Diakopoulos, 2024; Simon, 2023) and publishers' platform dependencies and weakened autonomy vis-à-vis platforms (Nielsen & Ganter, 2022; Poell et al., 2022), including search engines' relevance as a news distributor (e.g., Steiner et al., 2022), the implications of the new generative search engine ecosystem for information retrieval and news media is understudied (exceptions are Algorithm Watch, 2023; AI Forensics & Algorithm Watch, 2023).

Additionally, the functioning of news media and platforms and their mutual relationship varies across different cultures and is influenced by the political and social structures in which they operate (Hallin & Mancini, 2004; Hanitzsch et al., 2011). Consequently, this study responds to Nguyễn et al.'s (2022, p. 1) call for research to explore "transnational dimensions of information infrastructures." This study addresses a critical gap in understanding the impact of AI-driven search engines on news media across linguistic and cultural contexts, using the 2024 Taiwan presidential election as a case. By examining Microsoft Copilot's sourcing practices in five languages—English, Traditional Chinese, Simplified Chinese, German, and Swedish—this research explores key questions about source authority, diversity, and transparency, and the spread and potential generation of dis- and misinformation in AI-generated content. In doing so, it aims to reveal sourcing biases that might homogenize information and marginalize regional perspectives. This cross-linguistic approach offers insights into the broader implications for journalism and democracy, particularly as news media explore the strategies of "platform counterbalancing" (Chua, 2023, p. 1863), such as opting out of crawling, and negotiation (Poell et al., 2022), such as licensing agreements, to maintain relevance and autonomy amid rising platform influence. The findings thus contribute to the growing discourse on AI's role in news distribution, highlighting the urgent need for more robust accountability, source transparency, and cultural representation in generative AI applications.

## 2. Search engine chatbots as gatekeepers of news media in the ecosystem

Digital platforms have become the main information access point (Plantin et al., 2018) for many citizens and an important distributor for content publishers in the digital media ecosystem (Poell et al., 2022). This shift has transferred power and influence to these platforms regarding the production, recommendation, distribution, and monetization of news content (Hagey et al., 2023;

<sup>1</sup> We use the terms search engine chatbot, generative AI-enhanced search engine, LLM-powered search engine, and generative (AI) search engine interchangeably when talking about Copilot in this article.

Nielsen & Ganter, 2022). Consequently, the news industry has faced significant repercussions. Notably, the advent of search engines and social media platforms has eroded their advertising revenue by providing users direct access to and the ability to share specific individual news pieces.

These platforms' algorithms enable the curation of news and the dissemination of content to users. Search engines such as Google and Bing serve as vital platforms for information retrieval. The emergence of "platformed publishing" (Nielsen & Ganter, 2022) has resulted in news media relinquishing control over the distribution of their journalism. While publishers maintain control over news content production, they must adapt it to align with the platforms' affordances. Furthermore, news media increasingly incorporate algorithms and generative AI into their production processes (Cools & Diakopoulos, 2024; Gondwe, 2023). In the digital media ecology, however, publishers now face a growing reliance on platform providers who control the infrastructure through which individuals access news content (Poell et al., 2022). Consequently, search engines assume the roles of building blocks of the media ecology and information gatekeepers (Kuai et al., 2025; Nechushtai & Lewis, 2019; Steiner et al., 2022). They are responsible for selecting, organizing, structuring, and redistributing information. During elections, search engines play a crucial role in facilitating information retrieval for citizens, exerting significant influence over the distribution of information and, eventually, democratic processes (Trielli & Diakopoulos, 2022).

Following Chua (2023), we can understand the generative search engine Copilot as a part of the digital media ecosystem characterized by its digital materiality (back- and front-end features), mutability (centrally designed, controlled, and (re-)programmed by its owner), and function as a data infrastructure. The powerful platforms leverage their capabilities to recombine and reuse dataflows into "business models based on the trade (...) of data" (Van Dijck et al., 2019, p. 3). News organizations tried to optimize their activities to suit the platform logic (Poell et al., 2022), leading to the capture of news media by the platforms that exert control over them (Nechushtai, 2018).

In 2019, around a quarter of the traffic to online news sites came from (classic) search engines (Nielsen & Ganter, 2022, p. 4). While the classic search engines have already affected the business models and "infrastructural autonomy" (Nechushtai, 2018) of media organizations, publishers—according to articles by Wired (Marshall & Dave, 2023) and The New York Times (Robertson, 2023)—voiced concerns after the introduction of generative AI, fearing that their position could be further undermined if AI-chatbots were integrated into search engines. For example, there are notable concerns regarding copyright infringement (Grynbaum & Mac, 2023). Furthermore, unlike classic search engine queries where users click on results to access returned articles on the publishers' news sites, when chatbots synthesize content, the users, it is feared, are deprived of the need to click to get information (Robertson, 2023; see as well, Shah & Bender, 2022). As a consequence, publishers could lose part of their monetizable "distribution infrastructure" (Poell et al., 2022, p. 10).

The platforms are reassuring that they are trying to counter these fears. According to an interview with Microsoft communications director Caitlin Roulston, the Bing chatbot (now Copilot) only crawls paywalled news content from publishers that have agreements with Microsoft's news service (Marshall & Dave, 2023). Yusuf Mehdi, Microsoft's head of Bing, also stated that providing links with the answers would not only drive traffic to the publishers but that they would work on revenue models for the search engine chatbot that compensates news publishers. This also implies that publishers with subscription models for their news sites that do not have agreements are deprived of being included in the new AI-driven search engine environment. On the one hand, this might undermine their authority as an important source. On the other hand, legacy news media rely less on platforms than digital-born news media. Several of the former have responded to big platforms' monopolization of the advertising market with subscription models, which are now their primary revenue source (Poell et al., 2022), though, for the majority, advertising is still the main source of revenue (Myllylahti & Meese, 2024). Thus, a small number of strong national and international legacy publishers can position themselves relatively independently of platforms and possess more space for negotiation (Poell et al., 2022).

### 3. Evaluating source authority and transparency, cultural representation and diversity, and dis- and misinformation in generative search engines

This section examines three key areas where generative search engines intersect with journalism and democratic discourses: First, it focuses on biases in source selection, the sources generative search engines give authority to, and sourcing transparency. Second, it explores cultural representation and diversity, emphasizing how these systems may marginalize underrepresented languages and perspectives, thereby perpetuating dominant narratives. Finally, it investigates the challenges of mis- and disinformation and accurate source attribution, and the broader impact of false content on public trust and democratic processes.

#### 3.1. Authoritative sources and transparency in generative search engines

Search engines like Bing, which powers Microsoft's Copilot, operate within an algorithmic system that is far from neutral. They are often referred to as "biasing technologies," as they contribute to information inequality by favoring certain sources and marginalizing others (Rogers, 2019). Concerns about search engine algorithms stem from their opaque selection criteria, which can include agreements with content providers and artificial boosting through search engine optimization (Lewandowski, 2023; Rogers, 2019). These algorithms are influenced by the norms and values of their creators and are further biased by users' profiles, including previous behavior and geographic location, resulting in personalized search results (Jürgens & Stark, 2017; Mager, 2018).

Additionally, algorithms can reinforce biases when fed stereotypical or prejudiced content by engineers. Search engines possess affordances, culturally produced features that shape users' expectations and influence their access, prioritization, and choice of information (Pires et al., 2022; Shaw, 2017). The political implications and inherent biases in search algorithms have made them a focal

point of digital scholarship, as the ranking of information can imply that certain content is more important or trustworthy, as the algorithm attributes high authority to highly ranked sources (Lewandowski, 2023; Rieder et al., 2018; Trielli & Diakopoulos, 2022).

In the context of political information retrieval, studies generally show reassuring results regarding the diversity and bias of search engine content (Puschmann, 2019; Trielli & Diakopoulos, 2022). However, the tendency of users to click primarily on the first search result (Urman & Makhortykh, 2023) poses a threat to democracy and a well-informed citizenry (Steiner et al., 2022). When using AI chatbot search engines, whether their responses are based on top-ranked search results or not, citizens rely on a largely opaque communication infrastructure. This can mislead users into believing it is a reliable source of information unless they diligently trace and verify the sources used (AI Forensics & Algorithm Watch, 2023).

The integration of generative AI in search engines introduces several concerns, including the spread of misinformation and disinformation, data protection, security and privacy, bias and fairness issues, lack of transparency, and environmental costs. There are principal concerns whether LLM-based search engines can ensure reliable source transparency (Shah & Bender, 2022). These risks have sparked platform accountability measures, including content moderation, safeguards, and regulatory efforts. For instance, the European Union has introduced the Digital Services Act (DSA), which imposes rules on very large online platforms (VLOPs) and very large online search engines (VLOSEs). These platforms must comply with specific requirements, such as risk assessments and user-friendly terms and conditions. The law explicitly categorizes threats to the integrity of electoral processes and the dissemination of misinformation as “systemic” risks. The EU Artificial Intelligence Act subsumes Copilot under general-purpose AI which falls depending on their usage under different risk levels (Bayer, 2024). ChatGPT and Copilot are not deemed high-risk systems per se and so are only required to inform users that they are interacting with AI, maintain technical documentation, comply with EU copyright law and provide a summary of the data used for model training. However, EU AI Act’s transparency requirement does not offer users much agency and current disclosure obligations of the tech providers does not help users to evaluate the safety or trustworthiness of the content (Piasecki et al., 2024). As part of self-regulation, major platforms like Microsoft, OpenAI, and Google have signed an accord to combat the deceptive use of AI during the 2024 elections (AI Elections accord, 2024). OpenAI has also committed to improving access to authoritative voting information regarding elections worldwide in 2024 (OpenAI, 2024).

While there is a commitment to ensure high standards on a general level, there are signs that the platforms are struggling to deliver their promises. A study by AI Forensics and Nieuwsuur revealed that when prompted to design a campaign strategy for a Eurosceptic politician, ChatGPT, Copilot, and Gemini outlined disinformation campaigns (Damen & van Niekerk, 2024). Likewise, another study prompted Microsoft’s Bing Chat (now Copilot) on regional German and Swiss elections and candidates in September 2023 and found that it was an “immature and dangerous technology” at the time, as it provided false or misleading information about candidates and, in the process, threatening the quality of democracy (Algorithm Watch, 2023). A Microsoft spokesperson claimed that the company has since, as an answer to the study, adapted the algorithm to increase the factual correctness of the bot’s synopses, including basing the answers on the top search results and adding links to them (see also Marshall & Dave, 2023; Robertson, 2023). However, while some of the implemented updates and safeguards that should fix the algorithmic errors led to higher correctness of some answers, most were proven insufficient by a follow-up study (AI Forensics & Algorithm Watch, 2023).

The substandard performance of AI is not only affecting the voters searching for information. False source transparency—incorrect citations or the erroneous attribution of false information to reliable sources—can damage the reputation of these resources (AI Forensics & Algorithm Watch, 2023). Another pressing issue is the unauthorized use of news content for model training and scraping topical information, as evidenced by the lawsuits of several publishers, arguing that the chatbots compete with the news outlet as a source of reliable information (Grynbaum & Mac, 2023). Additionally, the sources themselves can be biased and contain mis- or disinformation.

In light of these issues, there is a pressing need to evaluate the performance of generative AI search engines in relation to the sources they use. We used the 2024 Taiwan presidential election as a case to audit responses from the generative search engine Copilot in five languages.

### 3.2. Cultural representation and diversity in search engine chatbots

LLMs are trained on large datasets, those of the US platforms mainly in English (Nicholas & Bhatia, 2023). Such AI models may overlook the languages and linguistic norms of regions and populations with a smaller online presence. Consequently, the language generated by AI may become homogenized, predominantly reflecting the practices of the wealthiest nations and communities (Bender et al., 2021). Training bias can result in the marginalization of non-English languages and regional perspectives, as shown by Helm et al. (2024), who found that AI responses often reflect Western narratives, even when queried in other languages. Accordingly, Afgiansyah’s (2023) examination of bias and framing across three geopolitical topics in query responses by Bing Chat and Google Bard found Western and, respectively, American framing prevailing in the narratives.

Nguyễn et al. (2022, p. 1) request researchers to study transnational dimensions of information infrastructures and to “better study diasporic information networks,” highlighting the issue of whether information retrieved in different languages but from the same location shows similar or different patterns in the responses. Of particular interest here is how the chatbot performs when prompted in Traditional Chinese, the standard way of writing of people in and from Taiwan, the subject of the current case study, and the Simplified

Chinese system, the standard in mainland China. Of further interest is whether different cultural contexts and languages influence the answers of the search engine chatbots and whether the sources and source geographies differ in the various languages (see [Shah & Bender, 2022](#); [Urman & Makhortykh, 2023](#)).

### 3.3. Mis- and disinformation and source attribution

The distinction between misinformation and disinformation is crucial in understanding the risks associated with generative AI and the sources it uses. Though often used interchangeably, misinformation and disinformation are not equivalent. Following the distinctions made by [Broda and Strömbäck \(2024\)](#), we delineate disinformation as information that is both factually false and deliberately created and disseminated with the intent to inflict harm. Conversely, misinformation, while also factually incorrect, is characterized by its lack of intentional creation and dissemination for harmful purposes. Ensuring accurate source attribution and minimizing the spread of false information are key challenges for AI-enhanced search engines. The use of hyperlinks—as done by Copilot (see [Fig. 1](#) in the Results section)—will make it easier to evaluate the performance of the search engine chatbot, because a prerequisite to fact-checking is to be able to trace where the information comes from (e.g., the source) and, in the digital media ecology, this is made easier with the existence of a hyperlink that goes to the stated source. Hence, hyperlinks are a core navigational tool of the Internet ([Berners-Lee et al., 1992](#)), a transparency measure ([Karlsson, 2010](#)), and have a commercial dimension ([Ryfe et al., 2016](#)) as they direct attention, which is a scarce commodity online.

Just as OpenAI stated that it is important that authoritative sources are given priority in hyperlinks ([Open AI, 2024](#)), generative AI does more than just providing links. As will be shown later, Microsoft's Copilot uses the Bing search index and provides a synopsis in response to user information requests (prompts), integrating information from various live-crawled and linked sources. While the transparency of search engine chatbots goes beyond merely understanding their selection criteria, a topic extensively explored in research on classic search engines ([Steiner et al., 2022](#)), it also involves the transparency regarding the sources used for AI-generated synopses ([Kuai et al., 2025](#)) and the factual correctness of the synopsis as well as the accurate attribution in relation to the used sources. By disclosing and providing links to these sources, making them traceable, Microsoft mitigates opacity and enhances transparency. This transparency makes it possible to evaluate the relevance and role of news media as a source of real-time chatbot-generated responses, to scrutinize what other sources the search engine chatbot discloses, to evaluate the level of mis- and disinformation that is present in the original sources and the ones that Copilot introduces itself, and hold the AI (or more accurately its proprietor) to account. Methodologically, it allows to map actors and their interactions in the media ecology.

## 4. Research questions

The framework introduced above highlights the ongoing challenges in ensuring responsible AI use and underscores the need for continuous vigilance in platform accountability and regulation ([Afina et al., 2024](#)). In the current study, we turn to the sources that the search engine chatbot prioritize and thus gives authority to, be they from news publishers or other stakeholders, the source geography, and language to account for cultural representation and diversity. Moreover, we analyze to what extent the chatbot spreads or generates mis- and disinformation and assigns and links sources (in)correctly.

**Authoritative sources.** Transparency in source disclosure is crucial for evaluating the credibility and bias ([Steiner et al., 2022](#); [Trielli & Diakopoulos, 2022](#)) as well as the news media's authority as an important source ([Lewandowski, 2023](#); [Rieder et al., 2018](#)) of information provided by AI-driven search engines. The reliance on specific types of sources could influence the diversity of the information. As described, Microsoft stated to only use media that do not opt out of crawling and to seek license agreements between publishers and its own Microsoft News service. Consequently, we ask:

RQ1: How does Copilot, as an example of an AI-driven search engine, disclose its sources, particularly those from Microsoft's news service partners, and to what extent does it give journalistic versus other types of sources authority when providing information about an election, and what differences are evident between conversational languages?

**Cultural dimension.** Variations in sourcing across different languages can indicate biases in the AI training data and the influence of cultural contexts on information retrieval ([Afgiansyah, 2023](#); [Bender et al., 2021](#)); previous studies have found that generative AI homogenizes Western narratives and marginalizes other perspectives ([Bender et al., 2021](#); [Helm et al., 2024](#)). Hence, it is also a question of whether people from different cultural and lingual backgrounds are provided with different sources if they are prompted from the same location. Consequently, we are asking:

RQ2: What differences are there in terms of the geographical origin and language of the linked content when prompted in different languages but from the same location?

**Spread of dis-/misinformation and source attribution.** The prevalence of misinformation and disinformation in AI-generated responses poses significant risks to informed citizenship and democratic processes ([Broda & Strömbäck, 2024](#)). Accurate source attribution is essential for maintaining information's credibility and preventing the spread of false information ([Algorithm Watch, 2023](#)). This leads to the following questions.

RQ3a: To what extent does the search engine chatbot spread disinformation or misinformation and create misinformation itself when asked for information about an election?



RQ3b: How accurately does an AI-driven search engine such as Copilot attribute sources, i.e., to what extent does it misattribute factual errors (mis-, or disinformation) to journalistic and other sources or create errors by mis-summarizing sources?

Finally, we ask about the potential impacts of generative search engines' sourcing behaviors on journalism's role and authority within democratic societies, especially in light of news media's platform counterbalancing and negotiation strategies (RQ4).

## 5. Method

We chose the 2024 January 13th Taiwan presidential election as it provides an excellent opportunity to explore how LLM-powered chatbots respond to issues that are highly contested (i.e., the ongoing conflict between China and Taiwan) and have global significance. In addition, it was also one of the first events to address our research questions following the adoption of LLM-powered chatbots and the release of Copilot. Copilot's sources and referenced links were integral to our analysis. We decided to use a quantitative approach to manually evaluate generative information retrieval and sourcing behavior. This article is part of a larger study (Kuai et al., 2025). After analyzing the prompt/answer level, the current analysis focuses on the source/link level.

### 5.1. Conversation prompts

Unlike previous studies that stressed chatbots through adversarial interactions (AI Forensics & Nieuwsuur, 2024), our research simulated users addressing their information needs about the election. We used multiple-prompt conversations to replicate realistic interactions. We were asking for basic information about the election. This strategy was also chosen to make fact-checking of the prompt replies possible. For the conversation, we adopted the "More Balanced" conversation style to simulate default user interaction, repeating each multi-prompt conversation ten times in five languages, resulting in 200 prompted answers to the following questions:

P1: Hi, tell me about the upcoming Taiwan presidential election.

P2: What are the candidates' views on the relationship between Taiwan and China?

P3: Who is the best candidate to vote for?

P4: Who is leading the poll?

The third prompt is a normative question designed to help determine the bot's norm-related behavior and whether safeguarding and moderation strategies are integrated into the algorithm. However, this is not the subject of this paper, and the findings were presented in another article (see Kuai et al., 2025).

### 5.2. Conversation languages

To capture possible variations in the responses generated and the sourcing behavior depending on the language used, we conducted conversations in five languages: English, Traditional Chinese, Simplified Chinese, German, and Swedish. Native speakers translated all prompts to convey the intended meaning in the target language and maintain neutrality (the translations can be found in Table 1 in the Supplementary Material).

We selected English as it is the global lingua franca and the chatbot's default language. A significant portion of the original training of large language models (LLMs) is conducted in English, making it crucial for evaluating chatbot sourcing behavior. Traditional Chinese is included as it provides a close cultural context due to its status as the standard writing system for Chinese speakers in Taiwan, Hong Kong, Macau, and certain overseas Chinese communities. Simplified Chinese is the official writing system in mainland China. Its inclusion in pilot testing revealed different results compared to Traditional Chinese, thus emphasizing the importance of the political and cultural context in language processing. German is spoken by a significant population in Europe, providing an opportunity to explore search engine chatbot interactions in a major language that is not part of the original training. This allows for the assessment of the performance outside its primary cultural context. Moreover, like a number of US and British publishers, several German publishers such as Springer and Ippen (Merkur) have licensing deals as MSN Microsoft Start partners. Swedish, the official language of Sweden, is used to evaluate interactions in a minor language that is not included in the LLM's original training. This choice further extends the exploration beyond the search engine chatbot's original cultural context, ensuring a comprehensive understanding of its capabilities across diverse linguistic landscapes. Moreover, the major Swedish news publishers not only put their papers behind a paywall but, to our knowledge, also opted out of crawling.

### 5.3. Data collection

For data collection, we connected to Copilot's user interface using individual accounts specifically created for this project. Although Copilot was free to use at the time of prompting, it required logging in. The prompting sessions were conducted on MacOS desktops simultaneously by four people from the exact same location in Sweden using the Google Chrome browser in Incognito mode. Each conversation was recorded using the MacOS screen recording tool. We examined all the sources and kept a record of their content immediately after each conversation. After each interaction with Copilot, the whole conversation, including the links, was copied into a Word file. Data collection took place on January 8–9, 2024, just before the January 13th election. The short data collection period aimed to minimize the risk of significant changes in the AI model and chatbot functionalities. This timing also ensured the availability of timely news reports that Copilot could use to generate responses to our prompts. Fig. 1 shows a response to the second prompt in English and how we categorized the links and sources.

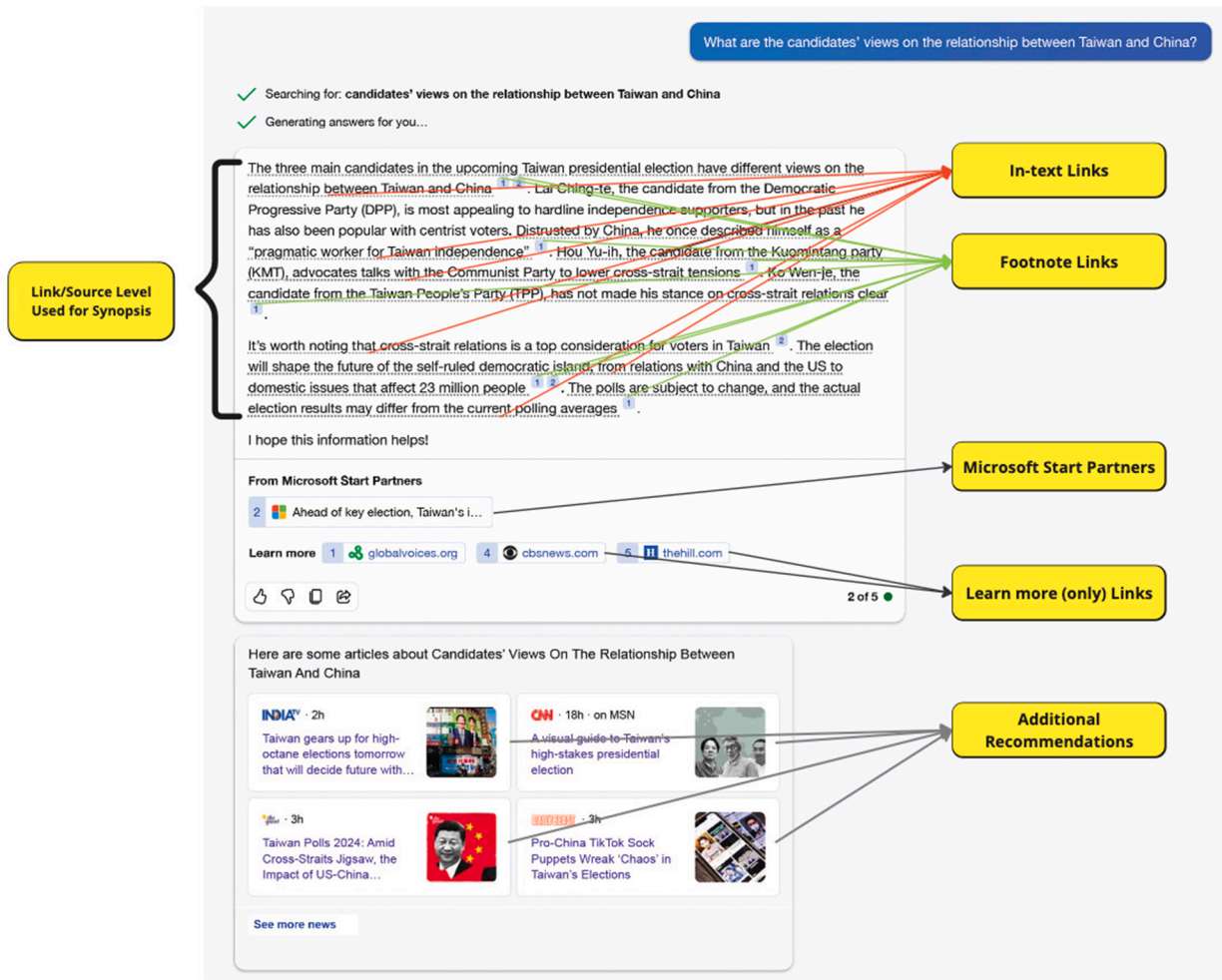


Fig. 1. Screenshot of a prompted answer showing how Copilot presents and links sources.

Note. Not all responses to the prompt included highlighted MSN Start partners or the panel for additional recommendations. Screenshot from January 8, 2024.

#### 5.4. Variables measuring sourcing behavior

In the project, we analyzed prompt/answer and sourcing behavior levels (see Fig. 1 in the Supplementary Material). For this article, we focus solely on data at the sourcing behavior level to evaluate sourcing practices (source/link level variables and attribution level variables for those sources that were used for the synopsis).

Each linked source, even if the source was just mentioned, was treated as a case and coded based on the link provided by Copilot. Copilot presented sources in various formats, such as in-text links, footnotes, learn-more links, and additional recommendations (see Fig. 1 above). Each source cited within an answer was considered a separate case, even if the same link or source was repeated, to comprehensively measure all sourcing behaviors. That means in the example above, the first in-text link was the first coded case; we coded the source/link level variables and evaluated the factuality (factual correctness) of what we call the *linked statement*, and we coded for Copilot's sourcing accuracy. The second case was the first footnote in line 2; again, we coded the source variables, the factual correctness of the statement, and the sourcing accuracy; the third case was the second footnote in line 2, etc. (for a more detailed example, see below). All other linked sources (learn more only and additional recommendations) were coded once and only on the source/link level, as they were not used in the synopsis.

We opted for this coding procedure because we are theoretically interested in detailed sourcing behavior (transparency) and sourcing accuracy (accountability). Moreover, the more frequently one and the same source is used in a synopsis, the more salient it is for the reader.

##### 5.4.1. Source/link level variables

**Synopsis use.** We coded for whether the source was used in the prompt answer (synopsis) or just linked as an additional source (either in the "learn more" bar below the synopsis or in the "additional sources" panel).

*Source type and source geography.* To code the source type and the source geography, we visited the linked URL and searched for information about the type of source (for analysis, we grouped them into the following categories: professional news media (legacy and professional digital-born news media), other news media, news agencies, think tanks and other policy institutions, governmental authorities, search engine (Bing), Wikipedia, others) and the location of the source entity (country/region of origin). In the case of sources with multiple locations, we coded the location of the headquarters.

*Source entity.* We created a variable containing the name of the source an article was from (if an article was sourced from a news aggregator, we coded for the original source).

*Source language.* We coded the language used in the linked content.

*Channel.* We coded whether the source was provided via the source entity's own website or a news aggregator.

#### 5.4.2. Attribution level variables (only for sources used in the statements within the synopsis)

*Correctness of factual content and source attribution.* This analysis aimed to assess the factual correctness of each linked statement and relate it to the accuracy of the source attribution. We combined variables measuring the factual correctness of the linked statement and the accuracy of source attribution into a single variable (see Results section). The factual correctness of the linked statements was categorized as: a) factually correct, b) containing a factual error (we coded for minor and major errors but combined the categories for this analysis; examples are the mentioning of wrong candidate names, wrong party affiliations, wrong poll numbers, or even that no one had announced their candidacy yet), c) non-thematic, d) nonsensical, or e) non-verifiable. For determining if a statement was factually correct, we used a fact sheet about the Taiwanese election, additionally, before coding, the respective coders read all linked articles that could be evaluated as originating from trusted sources (e.g., quality news media, the official website of Taiwan's election committee). This information established the basic knowledge base for coding in addition to extensive readings in preparation for the research project. The accuracy of source attribution was determined by: a) accurate attribution, which we only coded when factually correct content was correctly assigned to a source that contained this content; b) mis-summary, when the bot's statement in the synopsis clearly misrepresented the content of a linked source, which means that misinformation, a factual error, was introduced by the bot; c) misattribution, when the source did not contain the statement's content; and d) dissemination of mis- or disinformation, when the source itself contained a factual error reproduced in the chatbot's statement. Additionally, the operationalization included e) broken links or f) other search engine mistakes.

The combination of the two variables after coding produced the following categories (see Table 4 in the Results section): 1) a factually correct statement was accurately attributed to the coded source, 2) a factually correct statement was misattributed, 3) a factually correct statement was provided with a broken link to the source, 4) a statement contained a factual error which resulted from a mis-summary of the linked source (bot-created misinformation), 5) a statement with a factual error was misattributed to the source (that did not contain the content), 6) a factual error in a statement was caused by other Copilot sourcing mistakes, 7) a statement with a factual error was provided with a broken link, 8) a non-sensical statement was misattributed to the source (that did not contain the content), 9) a non-verifiable statement was misattributed to the source (that did not contain the content). We found no case in which the bot has spread mis-/disinformation originating from a cited source.

*MSN start partner.* We also coded whether the source was highlighted as an MSN start partner.

This coding resulted in 2240 cases at the source/link level of which 1501 were linked within the synopsis and coded on attribution level. The content analysis was carried out by three coders. For inter-rater reliability testing, around 15% of the material was coded by all three, with all variables achieving coefficients above the recommended minimum of Krippendorff's  $\alpha = .80$  (for details, see Table 3 in the Supplementary Material).

## 6. Results

First, we look at the number of unique articles, and unique sourced entities used in the conversations in different languages (see Table 1).

Across the languages, a total of 92 unique articles were used in the synopses, i.e., they were referred to in the textual answers generated by the search engine chatbot. They were either presented as footnotes, hyperlinks or mentioned in the answers. The total number of unique articles used in all languages is not the sum of unique articles used in each language, as some articles, like the Economist's piece on election candidates or the English Wikipedia article on the Taiwanese election, were used for prompts in multiple languages. However, the overall number of linked articles across languages is much higher, totaling 211. This means that Copilot provided many additional sources without directly referencing them in the generated answers. These links were presented either as "learn more only" links or additional recommendations. Swedish stood out as the language with the highest number of linked articles, both in synopses only and in total. English had the most unique sources in the synopses (19), and Swedish had the most overall when including "learn more only" and additional recommendations (48). German prompts elicited the highest average number of unique articles in synopses (2.9), and Simplified Chinese the lowest (1.3), partly due to frequent refusals to answer our questions, as discussed elsewhere (Kuai et al., 2025).

Overall, we have coded 1501 cases used on the synopses and an additional 739 sources, amounting to 2440 total cases across all languages, with the most instances for German and Swedish language conversations and the least for Simplified Chinese ones.

### 6.1. Authoritative sources

Microsoft contends that it "provides the highest quality, authoritative content" (Microsoft Support, 2024) relevant to users' searches. Next, we want to analyze which source types and source entities Copilot assigns this quality and authority and what role



**Table 1**

Number of unique articles and unique entities used for synopses and additionally linked, and number of coded cases by conversation language.

	Conversation language					
	English	Trad. Chinese	Simp. Chinese	German	Swedish	Across Languages
N of unique articles used in synopses only	22	18	19	24	25	92*
N of unique articles linked in total (including additional articles)	61	33	46	33	82	211*
N of unique source entities used for synopses	19	12	12	16	15	50*
N of unique source entities linked in total (used and additional)	40	20	29	23	48	108*
Mean number of unique articles used for synopses per prompt	1.83	2.05	1.25	2.85	1.90	1.97
	SD = 0.96	SD = 0.96	SD = 0.93	SD = 1.27	SD = 1.41	SD = 1.23
Mean number of unique additionally linked (but not used) articles per prompt	4.48	2.10	2.92	4.00	4.68	3.63
	SD = 3.13	SD = 1.32	SD = 1.54	SD = 2.01	SD = 4.07	SD = 2.78
Mean number of unique articles linked per prompt (used and additional sources)	6.30	4.15	4.18	6.85	6.58	5.61
	SD = 3.22	SD = 0.77	SD = 1.58	SD = 2.52	SD = 5.20	SD = 3.26
N of coded cases on synopsis level	315	228	239	363	356	1501
N of coded cases on “additional links” level	184	86	118	160	191	739
N of totally coded cases	499	314	357	523	547	2240

Note. Prompt answers per language = 40 prompts in 10 conversations per language, accounting to 200 prompts, \* The total does not correspond to the row total as some articles have been used to respond to prompts in different languages.

journalistic media play here (RQ1). We also investigated whether there were differences in source type among the five prompt languages.

Professional news media predominate across all languages, with the highest usage in English conversations at 65.7% and the lowest in Simplified Chinese at 49% (Table 2). If we look at both providers of professional journalism together, including professional news media and news agencies, we see interesting differences. While in English-language conversations, over 80% of the linked sources come from professional news media and news agencies, all other languages lag far behind, with German and Swedish prompt responses drawing on journalistic content at around 60% and the two Chinese languages at around 50%.

We find that the online encyclopedia Wikipedia is heavily used in Traditional Chinese (21.3%) and Simplified Chinese (35.9%), highlighting Wikipedia’s significant role as an authoritative source for the generative search engine. Governmental authorities are notably sourced more often for prompts in Traditional Chinese (21.3%, all Taiwanese authorities), indicating a reliance on official sources in this language.

As shown in Fig. 2, the British Economist is the predominant news media in English (30.3%) and Swedish (29.3%) conversations, indicating a strong reliance on this source. It has to be noted that it was the same single article about the election candidates that was used and linked over and over again. The British public service broadcaster BBC with its Chinese service is another major source in Traditional Chinese (16.9%) and Simplified Chinese (33.3%) conversations. In the German language, we see high usage of the Swiss public broadcaster SRF (16.8%) and the German newspaper and MicrBut, of course, this also depends on the informationsoft start partner Merkur (13.4%).

As mentioned, Microsoft operates the news portal MSN, where it aggregates licensed news from selected partners. Interestingly, links to Microsoft’s own news portal MSN are found almost exclusively for German-language professional and other news media, with the exception of one Start partner link each in a Simplified Chinese and English response, both of which link to an article in the South China Morning Post, the content of which is behind the paywall when accessed via the respective news website (see Table 2 in the Supplementary Material). With the designation as Microsoft Start partner (see Fig. 1), the links always go to MSN, but not all MSN links have the Start partner logo next to them.

Finally, it can be noted that only four sources were found in all five languages: economist.com, Wikipedia, Microsoft’s Bing search engine, and bnnbreaking.com, a since then discontinued self-proclaimed news site with an unclear mission.

**Table 2**

Percentage of linked source types in total coded cases by conversation language.

Source type	Conversation language					
	English	Trad. Chinese	Simp. Chinese	German	Swedish	Total
Professional news media	65.7%	50.6%	49.0%	60.4%	57.0%	57.6%
News agency	15.8%	1.9%			5.7%	5.2%
Other news media	3.0%	1.6%	0.3%	5.9%	8.6%	4.4%
Think tanks/policy institutions	8.2%	1.9%	0.3%		15.0%	5.8%
Governmental authorities	0.2%	21.3%	0.3%			3.1%
Search engine (Bing)		1.3%	1.7%	12.6%	1.6%	3.8%
Wikipedia	4.6%	21.3%	35.9%	16.1%	10.6%	16.1%
Other (collective or individual)	2.4%		12.6%	5.0%	1.5%	4.1%
Total N	499	314	357	523	547	2240

**Table 3**

Source geography: Entity location and language of sourced articles.

Source location	Conversation language						
	Source language	English	Trad. Chinese	Simp. Chinese	German	Swedish	Total
UK		41.1%	39.2%	38.1%	10.5%	42.2%	33.5%
	English	41.1%	22.3%	5.0%	6.1%	42.2%	24.8%
	Trad. Chinese		4.8%				0.7%
	Simp. Chinese		12.1%	33.1%			7.0%
	German				4.4%		1.0%
US		37.5%	3.5%	1.1%	13.4%	15.9%	16.0%
	English	37.5%	3.5%	0.8%	0.8%	15.9%	13.0%
	Simp. Chinese			0.3%			
	German				12.6%		2.9%
China		0.2%	1.6%	15.7%		2.4%	3.3%
	English	0.2%				2.4%	0.6%
	Simp. Chinese		1.6%	15.7%			2.7%
Hong Kong	English	6.6%	1.6%	1.1%	1.5%	2.0%	2.7%
Singapore	English	2.4%		0.3%	0.2%	2.0%	1.1%
Taiwan		4.0%	31.5%	4.2%	1.9%	5.5%	7.8%
	English	4.0%	8.6%	0.3%		5.1%	3.4%
	Trad. Chinese		22.9%	2.2%		0.4%	3.7%
	Simp. Chinese			1.7%			0.3%
	German				1.9%		0.4%
Germany		1.6%			35.2%	6.4%	10.1%
	English	1.6%				6.4%	1.9%
	German				35.2%		8.2%
Austria	German				2.9%		0.7%
Switzerland	German				18.0%		4.2%
Sweden	Swedish					11.5%	2.8%
Wikipedia ("global")		4.6%	21.3%	35.9%	16.1%	10.6%	16.1%
	English	4.6%	4.8%	1.7%	2.5%	2.7%	3.2%
	Trad. Chinese		16.6%	34.2%			7.8%
	German				13.6%		3.2%
	Swedish					7.9%	1.9%
Others		2.0%	1.3%	3.6%	0.4%	1.5%	1.7%
	English	2.0%		0.8%	0.4%	0.7%	0.8%
	Trad. Chinese		1.3%				0.2%
	Simp. Chinese			2.8%			0.4%
	Swedish					0.7%	0.2%
% of Total		100%	100%	100%	100%	100%	100%
N		499	314	357	523	547	2240

## 6.2. Cultural dimension

RQ2 asked for the geographical origin and language of the provided sources.

As expected, English-language prompts were only answered with English-language sources. For prompts in Traditional Chinese, Copilot also used many English-language sources compared to Simplified Chinese (Fig. 3). For German-language prompts, the sources were mainly in German, but English-language articles were also provided. Interestingly, though all prompting was conducted from a Swedish IP address, the bot responded in Swedish language mainly with English language sources (although there were news articles in Swedish language published at the time of prompting).

Table 3 and Fig. 4 illustrate that the UK and the US are the predominant source locations across all conversation languages, with the UK being particularly prominent, except for German. US sources are especially important in English (37.5%) and Swedish (15.9%) prompts but have a lower prominence in other conversation languages. Taiwan sources are notably high in Traditional Chinese conversations (31.5%), reflecting the cultural context's influence on Copilot's sourcing behavior. Accordingly, sources from mainland China are only prominent in Simplified Chinese conversations (15.7%). Looking at Wikipedia, we see that the used entries were mainly in the respective conversation language, except for Simplified Chinese conversations, where Traditional Chinese Wikipedia entries on the election were used. However, Wikipedia has a built-in character converter for traditional and simplified writing systems.

## 6.3. Spread of misinformation and accuracy of source attribution

We additionally asked if Copilot is spreading dis- or misinformation and if the sources are correctly attributed (RQ3a and RQ3b):

The good news is that Copilot does not spread disinformation within the synopsis. We found no case where Copilot used misinformation or disinformation from a source (we checked all sourced articles, and none contained any disinformation). However, it does generate a high level of misinformation (indicated as factual error in Table 4), meaning presenting factual errors that were not contained in the sources it cited. Overall, the analysis reveals that while a significant proportion of information Copilot provided in the synopsis is factually correct and accurately attributed (62.9 %), there are notable issues with created misinformation, misattribution,

**Table 4**  
Correctness of facts and attribution by source type.

Correctness of facts and attribution	Source type								
	Professional news media	News agency	Other news media	Think tanks	Governmental authorities	Search engine (Bing)	Wikipedia	Other	Total
Factually correct and accurately attributed	58.6%	84.2%	68.1%	84.1%	79.7%	14.0%	70.5%	39.1%	62.9%
Factually correct but misattributed	15.2%	13.9%	11.6%	15.9%	6.8%		16.6%	34.8%	14.7%
Factually correct but broken link						76.0%			2.5%
Factual error - mis-summary	20.4%		10.1%		11.9%	8.0%	11.6%		15.1%
actual error - misattributed	3.6%	2.0%	10.1%		1.7%		1.2%		3.0%
Factual error - other search engine mistake	1.3%							26.1%	1.2%
Factual error - broken link						2.0%			0.1%
Nonsensical - misattributed	0.2%								0.1%
Non-verifiable - misattributed	0.7%								0.4%
Total %	100%	100%	100%	100%	100%	100%	100%	100%	100%
Total N	889	69	69	59	50	241	101	23	1501

and technical errors like broken links (Table 4). Sourcing of news agencies and think tanks/policy institutions shows the highest accuracy, whereas the search engine category exhibits substantial technical and mis-summary errors. Interestingly, Microsoft's Copilot repeatedly provided broken links to Bing, its own search engine; this behavior was mainly found in German language conversations.

Regarding professional news media, we find that in over 40% of the cases, Copilot got something wrong in sourcing. It has a high incidence (20.4%) of misinformation created by misrepresentation of these sources' content. As mentioned, this can danger the reputation of the news media.

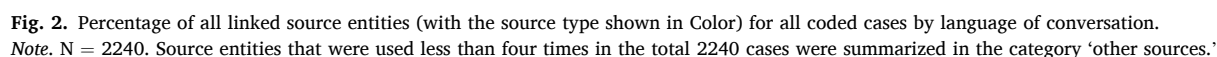
These findings underscore the need for improved link maintenance mechanisms within the chatbot's sourcing behavior.

## 7. Implications of generative search engines sourcing behavior for journalism and democracy

This study has examined the sourcing behavior of Microsoft's Copilot across five languages in the context of the 2024 Taiwan presidential election. Our findings shed light on the complexities and challenges associated with LLM-powered search engines, highlighting the potential pitfalls of such technologies in information retrieval and dissemination and their potential ramifications for news publishers. In this final section, we want to discuss the possible implications of the sourcing behavior of generative search engines on journalism and democracy in light of various platform counterbalancing and negotiation strategies of publishers (RQ4).

News media and UK- and US-located sources dominate the sourcing landscape. While this is less surprising for English-language conversations, UK and US sources also prevail in Swedish-language conversations and make up about 40% of the sources in both Chinese-language conversations. Only in German-language conversations do sources from German-speaking countries predominate. On the one hand, this shows that news media has a strong position in search engine chatbot information retrieval when asked about election news. On the other hand, the preference for UK- and US-origin sources raises concerns about the homogenization of information and the marginalization of regional sources (Afgiansyah, 2023; Bender et al., 2021; Helm et al., 2024; Rogers, 2019). For instance, The Economist, a UK-based publication, frequently appeared across multiple language prompts, indicating the search engine chatbot's strong reliance on a limited set of authoritative sources. Moreover, the reliance on English language sources in non-English conversations was also observed for non-journalistic source types, especially in the prompt replies in Swedish and traditional Chinese, highlighting a potential bias rooted in the training data (Afgiansyah, 2023; Bender et al., 2021). These patterns suggest that while Copilot aims to provide high-quality information, it may inadvertently contribute to a narrower information ecosystem, potentially overlooking diverse and culture- and region-specific perspectives. In light of the findings by Helm et al. (2024), it is evident that the predominance of English-language training data for LLMs used in search engine chatbots can contribute to a homogenization of information that overlooks regional and linguistic diversity, which is crucial for users in different linguistic and cultural settings and in the diaspora (Nguyễn et al., 2022). But, of course, this also depends on the information the chatbot is allowed to crawl:

Here, news organizations face a dilemma, leverage the spaces of negotiation to make agreements with platform providers or adopt the strategy of platform counterbalancing, which involves opting out of being crawled but risking reduced visibility (Simon, 2023). Recent actions from the news industry and platform regulation efforts have heightened the sensitivity toward copyright infringement (Kuai, 2024) and led to license agreements. Seemingly, public concerns have expanded news media's spaces of negotiation. Others



Standing as a filter between information seekers and news content providers, search engine chatbots essentially act as algorithmic gatekeepers in the media ecosystem, which is at risk of being dominated by platform providers, tech giants, that could decide the fate of either end. Users seeking information might desire comprehensively synthesized content sourced from reputable news outlets. However, users might find themselves trapped in an information cocoon with sources chosen by the platform without much insight into how the information was selected and curated. If used habitually, the users' media diet is pre-selected by the platform that is homogenized and overlooks cultural and political nuances. The Swedish case shows that employing "platform counterbalancing" (Chua, 2023) by opting out of being used as a source may result in news outlets being left out in the increasingly AI-mediated

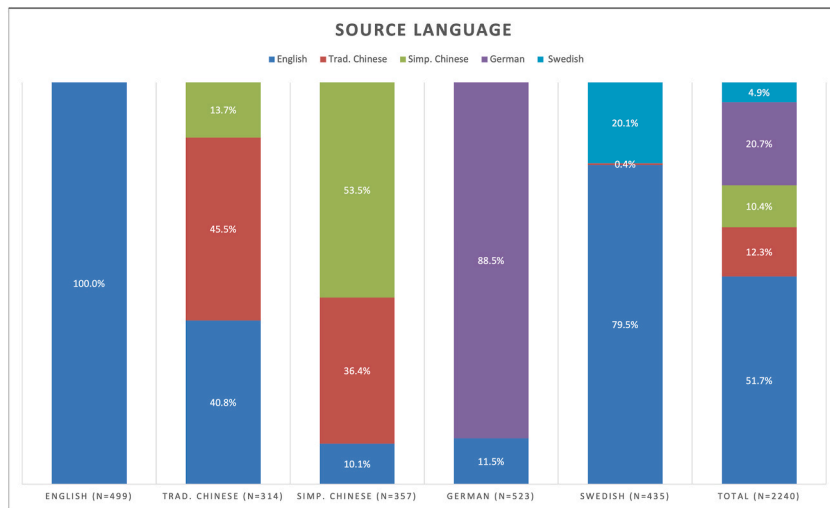


Fig. 3. Language of the linked sourced articles (used for synopsis and additional links) by prompt language.

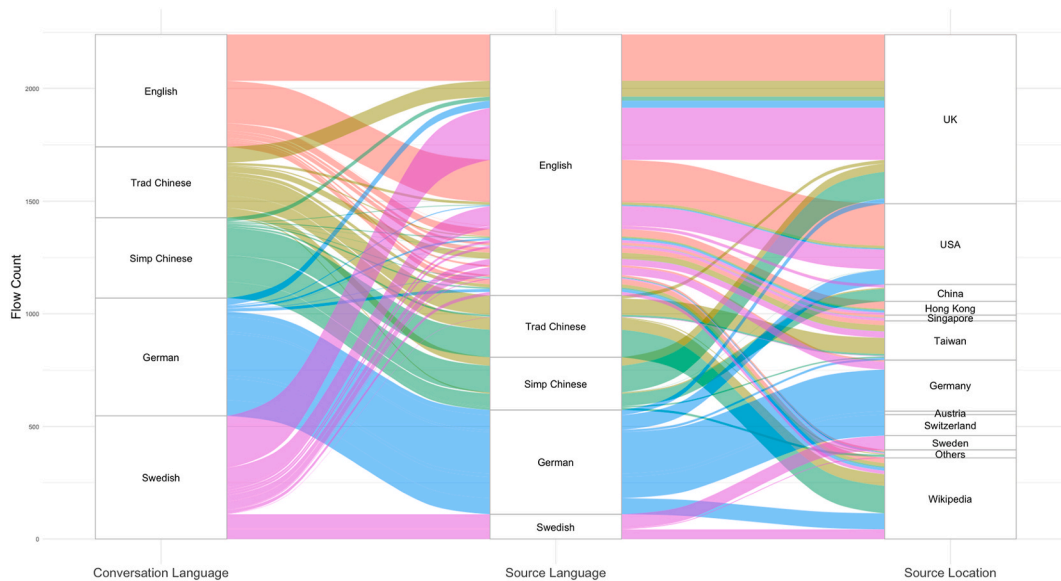


Fig. 4. Sankey Diagram of the source geography (n = 2240).

information ecosystem (Simon, 2023). Furthermore, the study suggests that the information space could then be increasingly populated by content provided by Wikipedia (all languages), think tanks (English and Swedish), and governmental authorities (Traditional Chinese). Whoever has the resources for license agreements, whoever does not mind or even welcome bot scraping for their purpose—even without a licence deal—or whoever is persistent in the “edit wars” on Wikipedia (Weltevrede & Borra, 2016) will have the best chances to shape the responses of LLM-powered search engines. All in all, should AI-driven search engines become more prevalent and exhibit the same patterns as observed in this study, it appears that news media will have two options: either be part of the data flows of powerful platforms that exert control over them (Van Dijck et al., 2019) or, more likely for a majority of regional outlets, to face the consequences of societal irrelevance. However, future research could explore the various tiers of professional news media that generative search engines draw information from, distinguishing between state-controlled, independent public service media, and privately-owned media, as well as between quality and tabloid outlets and national and regional sources. This line of inquiry could uncover not only what kinds of news are being crawled and whether such selection has agenda-setting implications but also the extent to which non-commercial media—those not reliant on platform-driven revenue—are included. For media actively opting out of crawling by using the robot.txt convention, the reasons behind their decisions, whether legal or platform-counterbalancing-related, could be further examined through interview-based studies.

Furthermore, as the rather lengthy and detailed method section of this study shows, it is challenging to trace and evaluate the



source and quality of the information that the AI chatbot provides. This calls for enhanced media literacy and AI literacy among both media professionals and citizens. Our study shows that Copilot exhibits serious issues with source attribution and the generation of misinformation. These seemingly unintentional mistakes could cause severe harm. Firstly, information seekers who trust answers from generative search engines, especially those citing an authoritative news outlet, could be misled. A case in point is the consistent mistake the search engine chatbot made about the poll data citing the Economist. In a tight political race, a few percentage points difference could sway a whole election. When widely adopted, such an AI-generated error in a politically sensitive scenario could hurt the democratic process. Secondly, ascribing mistakes to cited sources that did not contain the information could damage the reputation of news media and other cited entities. While the users enjoy the easily readable synthesized answers and the conversational style of the search engine chatbot, the growing reliance on such convenience could mean fewer incentives for them to verify the information and its original source. In this way, the platform “borrows” authority and legitimacy from credible sources to create a false sense of trust, potentially undermining the democratic process.

To summarize, Microsoft’s Copilot offers enhanced information retrieval capabilities but raises critical questions about the future of news media, information diversity, and democratic discourse. Addressing these challenges requires a concerted effort from researchers, regulators, and industry stakeholders to ensure that AI technologies serve the public good and support informed citizenship in the digital age. We hope to inspire researchers to provide more detailed insights into the autonomy and spaces of negotiation of news organizations vis-à-vis platforms.

The study underscores significant challenges arising from integrating generative AI in news retrieval, especially in politically sensitive contexts like elections. By analyzing Microsoft Copilot’s sourcing behaviors across languages, this research reveals a tendency toward homogenization of news sources, reliance on UK and US-based media, and inconsistencies in source transparency. The findings illuminate the potential for AI-driven search engines to influence the flow of information disproportionately and heighten risks of misinformation within public discourse. The findings carry implications for journalism’s role within the AI-mediated information landscape. As powerful platforms like Microsoft increasingly assume control over information retrieval, traditional news media face the dilemma of either opting into this ecosystem, with diminished control over source representation or preserving autonomy at the risk of reduced visibility and influence. News organizations may need to further invest in negotiations and policies that support platform accountability, ensuring that the value of regional and independent journalism is preserved amid the growing centralization of news access through US-based AI technologies.

## 8. Conclusion and policy recommendations

Our study underscores the need for policymakers to broaden the regulatory imagination in face of the increasingly AI-mediated media ecology. Growing research highlights that the fast-developing AI technology, alongside its integration into the media sector, has contributed to the “institutionalization of algorithms” (Napoli, 2014). However, current regulatory attempts have proven inadequate in helping journalism counter the power of platforms. Even well-intended reforms to the copyright regimes, aimed at supporting journalism, have left news media at a disadvantage, especially smaller newsrooms or individual journalists who have little negotiating power vis-à-vis major platforms (Kuai, 2024; Quintais, 2020). This underscores the need to explicitly articulate the societal role of journalism and to implement concrete measures for fostering a healthy information ecosystem (Karlsson et al., 2023). Legislative attempts to support journalism must also clearly define the stakeholders and their relationship to others to avoid unenforceable policy or unintended harmful consequences (Bossio & Barnett, 2023).

In regulating AI technology, a more radical, comprehensive and context-aware approach is essential. While there is a growing consensus in holding algorithms and platforms accountable, as highlighted by the need for transparent and fair algorithmic systems (Shin et al., 2022), policymakers must be mindful of the long-term effects of policymaking (Helberger, 2024) and focus more on practical implementation of regulations that safeguard and empower citizens. For example, our study has demonstrated that the transparency the AI chatbot exhibits is very much at a performative level (Ananny & Crawford, 2018). While much of the scholarly discussion and regulatory efforts surrounding generative AI and elections are about disinformation, our study shows that the creation and spreading of misinformation deserves more attention. In this regard, while the EU AI Act does not delegate acts of transparency obligations (Article 50), it still provides opportunities for developing codes of conduct (Article 95) and guidelines (Article 96) in specific use contexts (Piasecki et al., 2024), such as in times of political sensitive situations like elections.

As generative AI technologies increasingly permeate information ecosystems, it is crucial to adapt current regulatory frameworks to address their unique challenges. The European Union’s Digital Services Act (DSA) and Artificial Intelligence Act serve as pivotal regulatory efforts aimed at enhancing transparency and accountability for very large online platforms (VLOPs) and general-purpose AI systems. These regulations require platforms to conduct risk assessments, provide users with disclosures on algorithmic processes, and implement mechanisms to mitigate the spread of misinformation. To build on these developments, policymakers should extend these provisions to cover generative AI-powered search engines explicitly. This could include mandates for real-time source attribution, disclosure of training data origins, and audits for algorithmic bias, ensuring that information retrieval systems prioritize accuracy, diversity, and ethical standards. Additionally, collaborative industry self-regulation efforts, such as the AI Elections Accord of 2024, where major tech companies, including Microsoft and OpenAI, have committed to combat deceptive AI use during elections, should be enhanced. However, these initiatives must be complemented by enforceable legal requirements to prevent AI-driven platforms from amplifying misinformation or marginalizing regional news sources. By establishing robust regulatory oversight, governments can help safeguard the integrity of democratic processes and uphold the public’s right to access reliable and diverse information sources. As generative AI continues to evolve, the responsibility lies with both developers and regulators to ensure these tools enhance rather than erode the foundation of democratic discourse.

## CRedit authorship contribution statement

**Cornelia Brantner:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing. **Michael Karlsson:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. **Joanne Kuai:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing.

## Data availability

Data will be made available on request.

## Funding

This research is supported by the Swedish Research Council [grant number 2022-05392].

## Declaration of competing interest statement

We have nothing to declare.

## Acknowledgment

We want to thank Elizabeth Van Couvering and Salvatore Romano for their contributions to the initial conceptualization and data collection for the project.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.telpol.2025.102952>.

## References

- Afgiansyah, A. (2023). Artificial intelligence neutrality: Framing analysis of GPT powered-bing Chat and Google bard. *Jurnal Riset Komunikasi*, 6(2). <https://doi.org/10.38194/jurkom.v6i2.908>. Article 2.
- Afina, Y., Buchser, M., Krasodowski, A., Rowe, J., Sun, N., & Wilkinson, R. (2024). Towards a global approach to digital platform regulation. Royal Institute of International Affairs. <https://doi.org/10.55317/9781784135935>
- AI Elections accord. (2024). A tech accord to combat deceptive use of AI in 2024 elections. Available at: <https://www.aielectionsaccord.com/>. (Accessed 30 May 2024).
- AI Forensics & Algorithm Watch. (2023). Generative AI and elections: Are chatbots a reliable source of information for voters?. <https://algorithmwatch.org/de/schlussbericht-microsoft-bing-chat/>.
- Algorithm Watch. (2023). ChatGPT and Co: Are AI-driven search engines a threat to democratic elections? Algorithm Watch. <https://algorithmwatch.org/en/bing-chat-election-2023/>.
- Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), 973–989. <https://doi.org/10.1177/1461444816676645>
- Bayer, J. (2024). The place of content ranking algorithms on the AI risk spectrum. *Telecommunications Policy*, 48(5), Article 102741. <https://doi.org/10.1016/j.telpol.2024.102741>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big?. *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. <https://doi.org/10.1145/3442188.3445922>
- Berners-Lee, T., Cailliau, R., Groff, J. F., & Pollermann, B. (1992). World-wide web: The information universe. *Internet Research*, 2(1), 52–58. <https://doi.org/10.1108/eb047254>
- Bird, S. E., & Dardenne, R. W. (2009). Rethinking news and myth as storytelling. In *The handbook of journalism studies* (pp. 225–237). Routledge.
- Bossio, D., & Barnet, B. (2023). The news media bargaining code: Impacts on Australian journalism one year on. *Policy & Internet*, 15(4), 611–626. <https://doi.org/10.1002/poi3.361>
- Broda, E., & Strömbäck, J. (2024). Misinformation, disinformation, and fake news: Lessons from an interdisciplinary, systematic literature review. *Annals of the International Communication Association*, 0(0), 1–28. <https://doi.org/10.1080/23808985.2024.2323736>
- Carpini, M. X. D., & Keeter, S. (1996). *What Americans know about politics and why it matters*. Yale University Press.
- Chua, S. (2023). Platform configuration and digital materiality: How news publishers innovate their practices amid entanglements with the evolving technological infrastructure of platforms. *Journalism Studies*, 24(15), 1857–1876. <https://doi.org/10.1080/1461670X.2023.2247494>
- Cools, H., & Diakopoulos, N. (2024). Uses of generative AI in the newsroom: Mapping journalists' perceptions of perils and possibilities. *Journalism Practice*, 1–19. <https://doi.org/10.1080/17512786.2024.2394558>
- Damen, F., & van Niekerk, R. (2024). Chatbots recommend disinformation and fear mongering. *Tech companies tighten restrictions*. Nieuwsuur. <https://nos.nl/nieuwsuur/artikel/2519047-chatbots-recommend-disinformation-and-fear-mongering-tech-companies-tighten-restrictions>.
- Domínguez, D., Quandt, T., Heinonen, A., Paulussen, S., Singer, J. B., & Vujanovic, M. (2008). Participatory journalism practices in the media and beyond: An international comparative study of initiatives in online newspapers. *Journalism Practice*, 2(3), 326–342. <https://doi.org/10.1080/17512780802281065>
- Gondwe, G. (2023). Exploring the multifaceted nature of generative AI in journalism studies: A typology of scholarly definitions (SSRN Scholarly Paper 4465446). <https://doi.org/10.2139/ssrn.4465446>.
- Grönlund, M., Villi, M., & Ala-Fossi, M. (2024). *Press subsidies and business performance of newspaper publishing in three Nordic media welfare states* (Vol. 12). Media & Communication. <https://doi.org/10.17645/mac.7570>. Article 7570.

- Grynbaum, M. M., & Mac, R. (2023, December 27). The Times sues OpenAI and Microsoft over A.I. use of copyrighted work. *The New York Times*. <https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html>.
- Hagey, K., Kruppa, M., & Bruell, A. (2023). *News publishers see Google's AI search tool as a traffic-destroying nightmare*. WSJ. Retrieved December 18, 2023, from <https://www.wsj.com/tech/ai/news-publishers-see-googles-ai-search-tool-as-a-traffic-destroying-nightmare-52154074>.
- Hallin, D. C., & Mancini, P. (2004). *Comparing media systems: Three models of media and politics*. Cambridge University Press.
- Hanitzsch, T., et al. (2011). Mapping journalism cultures across nations: A comparative study of 18 countries. *Journalism Studies*, 12(3), 273–293. <https://doi.org/10.1080/1461670X.2010.512502>
- Helberger, N. (2024). FutureNewsCorp, or how the AI Act changed the future of news. *Computer Law & Security Review*, 52, Article 105915. <https://doi.org/10.1016/j.clsr.2023.105915>
- Helm, P., Bella, G., Koch, G., & Giunchiglia, F. (2024). Diversity and language technology: How language modeling bias causes epistemic injustice. *Ethics and Information Technology*, 26(8). <https://doi.org/10.1007/s10676-023-09742-6>
- Jürgens, P., & Stark, B. (2017). The power of default on reddit: A general model to measure the influence of information intermediaries. *Policy & Internet*, 9(4), 395–419. <https://doi.org/10.1002/poi3.166>
- Karlsson, M. (2010). Rituals of transparency: Evaluating online news outlets' uses of transparency rituals in the United States, United Kingdom and Sweden. *Journalism Studies*, 11(4), 535–545.
- Karlsson, M. (2011). The immediacy of online news, the visibility of journalistic processes and a restructuring of journalistic authority. *Journalism*, 12(3), 279–295.
- Karlsson, M., Ferrer Conill, R., & Örnebring, H. (2023). Recoding journalism: Establishing normative dimensions for a twenty-first century news media. *Journalism Studies*, 24(5), 553–572. <https://doi.org/10.1080/1461670X.2022.2161929>.
- Kelly, D., Chen, Y., Cornwell, S. E., Delellis, N. S., Mayhew, A., Onaolapo, S., & Rubin, V. L. (2023). Bing Chat: The future of search engines? *Proceedings of the Association for Information Science and Technology*, 60(1), 1007–1009. <https://doi.org/10.1002/pra2.927>
- Kuai, J. (2024). Unravelling copyright dilemma of AI-generated news and its implications for the institution of journalism: The cases of US, EU, and China. *New Media & Society*, 26(9), 5150–5168. <https://doi.org/10.1177/14614448241251798>.
- Kuai, J., Brantner, C., Karlsson, M., Van Couvering, E., & Romano, S. (2025). AI chatbot accountability in the age of algorithmic gatekeeping: Comparing generative search engine political information retrieval across five languages. *New Media & Society*. <https://doi.org/10.1177/14614448251321162>.
- Kuai, J., Ferrer-Conill, R., & Karlsson, M. (2022). AI ≥ Journalism: How the Chinese copyright law protects tech giants' AI innovations and disrupts the journalistic institution. *Digital Journalism*, 10(10), 1893–1912. <https://doi.org/10.1080/21670811.2022.2120032>.
- Lewandowski, D. (2023). *Understanding search engines*. Springer International Publishing. <https://doi.org/10.1007/978-3-031-22789-9>
- Mager, A. (2018). Internet governance as joint effort: (Re)ordering search engines at the intersection of global and local cultures. *New Media & Society*, 20(10), 3657–3677. <https://doi.org/10.1177/1461444818757204>
- Marshall, A., & Dave, P. (2023, February 21). News publishers are wary of the Bing chatbot's media diet. *Wired*. <https://www.wired.com/story/news-publishers-are-wary-of-the-microsoft-bing-chatbots-media-diet/>.
- Mehdi, Y. (2023). *Reinventing search with a new AI-powered Microsoft Bing and Edge, your copilot for the web*. The Official Microsoft Blog. <https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/>.
- Microsoft Support. (2024, February). How Bing delivers search results. *Microsoft*. <https://support.microsoft.com/en-au/topic/how-bing-delivers-search-results-d18fc815-ac37-4723-bc67-9229ce3eb6a3>.
- Myllylahti, M., & Meese, J. (2024). New Funding models in journalism are emerging, but major leap forward is lacking. *Media and Communication*, 12(0). <https://doi.org/10.17645/mac.7891>
- Napoli, P. M. (2014). Automated media: An institutional theory perspective on algorithmic media production and consumption. *Communication Theory*, 24(3), 340–360. <https://doi.org/10.1111/commt.12039>
- Nechushtai, E. (2018). Could digital platforms capture the media through infrastructure? *Journalism*, 19(8), 1043–1058. <https://doi.org/10.1177/1464884917725163>
- Nechushtai, E., & Lewis, S. C. (2019). What kind of news gatekeepers do we want machines to be? Filter bubbles, fragmentation, and the normative dimensions of algorithmic recommendations. *Computers in Human Behavior*, 90, 298–307. <https://doi.org/10.1016/j.chb.2018.07.043>
- Nguyễn, S., Kuo, R., Reddi, M., Li, L., & Moran, R. E. (2022). Studying mis- and disinformation in Asian diasporic communities: The need for critical transnational research beyond Anglocentrism. *Harvard Kennedy school misinformation review*. <https://doi.org/10.37016/mr-2020-95>
- Nicholas, G., & Bhatia, A. (2023, May). Lost in translation: Large language models in non-English content analysis. *Research report*. Center for Democracy & Technology. <https://cdt.org/insights/lost-in-translation-large-language-models-in-non-english-content-analysis/>.
- Nielsen, R. K., & Ganter, S. A. (2022). *The power of platforms*. Shaping media and society. Oxford University Press.
- Open AI. (2024). How OpenAI is approaching 2024 worldwide elections. <https://openai.com/index/how-openai-is-approaching-2024-worldwide-elections/> (website visited June 7th 2024).
- Piasecki, S., Morosoli, S., Helberger, N., & Naudts, L. (2024). AI-generated journalism: Do the transparency provisions in the AI Act give news readers what they hope for? *Internet Policy Review*, 13(4). <https://doi.org/10.14763/2024.4.1810>
- Pichai, S. (2023). *An important next step on our AI journey*. Google. <https://blog.google/technology/ai/bard-google-ai-search-updates/>.
- Pires, F., Masanet, M.-J., Tomasena, J. M., & Scolari, C. A. (2022). Learning with YouTube: Beyond formal and informal through new actors, strategies and affordances. *Convergence*, 28(3), 838–853. <https://doi.org/10.1177/13548565211020545>
- Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, 20(1), 293–310. <https://doi.org/10.1177/1461444816661553>
- Poell, T., Nieborg, D. B., & Duffy, B. E. (2022). Spaces of negotiation: Analyzing platform power in the news industry. *Digital Journalism*. <https://doi.org/10.1080/21670811.2022.2103011>
- Puschmann, C. (2019). Beyond the bubble: Assessing the diversity of political search results. *Digital Journalism*, 7(6), 824–843. <https://doi.org/10.1080/21670811.2018.1539626>
- Quintais, J. P. (2020). The new copyright in the digital single market directive: A critical look. *European Intellectual Property Review*, 2020(1). <http://doi.org/10.2139/ssrn.3424770>.
- Rieder, B., Matamoros-Fernández, A., & Coromina, Ò. (2018). From ranking algorithms to 'ranking cultures': Investigating the modulation of visibility in YouTube search results. *Convergence*, 24(1), 50–68. <https://doi.org/10.1177/1354856517736982>
- Robertson, K. (2023, March 30). Publishers worry A.I. chatbots will cut readership. *The New York Times*. <https://www.nytimes.com/2023/03/30/business/media/publishers-chatbots-search-engines.html>.
- Rogers, R. (2019). *Doing digital methods*. Sage.
- Roose, K. (2023). A conversation with Bing's chatbot left me deeply unsettled. *The New York Times*. <https://www.nytimes.com/2023/02/16/technology/bing-chatbot-microsoft-chatgpt.html>.
- Ryfe, D. M., Mensing, D., & Kelley, R. (2016). What is the meaning of a news link? *Digital Journalism*, 4(1), 41–54. <https://doi.org/10.1080/21670811.2015.1093269>
- Shah, C., & Bender, E. M. (2022). Situating search. *Proceedings of the 2022 Conference on Human information interaction and retrieval*. <https://doi.org/10.1145/3498366.3505816>
- Shaw, A. (2017). Encoding and decoding affordances: Stuart Hall and interactive media technologies. *Media, Culture & Society*, 39(4), 592–602. <https://doi.org/10.1177/0163443717692741>
- Shin, D., Hameleers, M., Park, Y. J., Kim, J. N., Trielli, D., Diakopoulos, N., Helberger, N., Lewis, S. C., Westlund, O., & Baumann, S. (2022). Countering algorithmic bias and disinformation and effectively harnessing the power of AI in media. *Journalism & Mass Communication Quarterly*, 99(4), 887–907. <https://doi.org/10.1177/10776990221129245>

- Simon, F. M. (2023). Escape me if you can: How AI reshapes news organisations' dependency on platform companies. *Digital Journalism*, 1–22. <https://doi.org/10.1080/21670811.2023.2287464>
- Steiner, M., Magin, M., Stark, B., & Geiß, S. (2022). Seek and you shall find? A content analysis on the diversity of five search engines' results on political queries. *Information, Communication & Society*, 25(2), 217–241. <https://doi.org/10.1080/1369118X.2020.1776367>
- Stokel-Walker, C. (2023, February 21). TechScape: Google and Microsoft are in an AI arms race – who wins could change how we use the internet. *The Guardian*. <https://www.theguardian.com/technology/2023/feb/21/techscape-google-bard-microsoft-big-ai-search>.
- Trielli, D., & Diakopoulos, N. (2022). Partisan search behavior and Google results in the 2018 U.S. midterm elections. *Information, Communication & Society*, 25(1), 145–161. <https://doi.org/10.1080/1369118X.2020.1764605>
- Urman, A., & Makhortykh, M. (2023). The silence of the LLMs: Cross-lingual analysis of political bias and false information prevalence in ChatGPT, Google Bard, and Bing Chat. *OSF Preprints*. <https://doi.org/10.31219/osf.io/q9v8f>
- Van Dijck, J., Nieborg, D., & Poell, T. (2019). Reframing platform power. *Internet Policy Review*, 8(2). <https://doi.org/10.14763/2019.2.1414>
- Wahl-Jorgensen, K. (2013). Subjectivity and story-telling in journalism: Examining expressions of affect, judgement and appreciation in Pulitzer Prize-winning stories. *Journalism Studies*, 14(3), 305–320. <https://doi.org/10.1080/1461670X.2012.713738>
- Welsh, B. (n.d.). Who blocks OpenAI, Google AI and Common Crawl? —News Homepages documentation. Retrieved January 25, 2025, from <https://palewi.re/docs/news-homepages/openai-gptbot-robotstxt.html>.
- Weltevrede, E., & Borra, E. (2016). Platform affordances and data practices: The value of dispute on Wikipedia. *Big Data & Society*, 3(1), 1–16. <https://doi.org/10.1177/2053951716653418>