Postprint: A Study on Fake News Governance in Search Engines

Authors: Cui Xinyuan, Hu Yiquan, Fu Yu

Date: 2023-10-08T00:00:00+00:00

Abstract

Fake news represents a prominent topic in the field of journalism and communication studies. In existing research on fake news governance, scholars have predominantly focused on the agency of social media platforms, while overlooking the significant role of search engines. In fact, the ranking mechanisms of search engines exert considerable impetus on the dissemination of fake news. Since 2016, foreign search engines have intensified their governance efforts against fake news, encompassing technological optimization, improvements to search result quality, the establishment of collaborative fact-checking networks, and the implementation of media literacy programs for key populations. In contrast, in China, the governance of fake news and rumor debunking relies primarily on a few mainstream media platforms and government agencies, with relatively limited investment from search engines. From a long-term perspective, beyond drawing upon foreign governance measures, building trust constitutes the primary objective for search engine platforms in governing fake news and achieving sustainable development.

Full Text

Preamble

Exploring Fake News Governance in Search Engines

Cui Xinyuan, Hu Yiquan, Fu Yu
(Hangzhou Dianzi University, Hangzhou, Zhejiang 310018)

Abstract: Fake news represents a hot-button issue in journalism and communication studies. Existing research on fake news governance has predominantly focused on the role of social platforms as primary actors, while largely overlooking the critical importance of search engines. In reality, search engine ranking mechanisms significantly amplify the spread of fake news. Since 2016,

foreign search engines have intensified their governance efforts, including optimizing technology to improve search result quality, establishing collaborative fact-checking networks, and implementing media literacy programs for key populations. In contrast, China' s fake news governance and debunking efforts rely primarily on a handful of mainstream media platforms and government agencies, with relatively minimal investment from search engines. In the long term, beyond drawing lessons from international governance measures, building trust should be the primary objective for search engine platforms in governing fake news and achieving sustainable development.

Keywords: Journalism and Communication; Platform; Search Engine; Fake News; Governance
Chinese Library Classification: G210
Document Code: A
Article ID: 1671-0134(2021)08-022-04
DOI: 10.19483/j.cnki.11-4653/n.2021.08.004

Citation Format: Cui Xinyuan, Hu Yiquan, Fu Yu. Exploring Fake News Governance in Search Engines[J]. China Media Technology, 2021(08): 22-25.

The term "fake news" became a globally visible concept after the 2016 U.S. presidential election. During the election, a large volume of political disinformation spread unchecked across online platforms, eroding American public trust in media and government. In 2017, "fake news" became one of Google' s most popular search terms, with polls showing that 45% of citizens attributed primary responsibility for fake news governance to the U.S. government and public officials, while 42% believed social media and search engines should bear the main responsibility [1]. Consequently, fake news governance emerged as a heated topic in journalism and communication studies. However, while scholars have extensively focused on disinformation spread and governance on social media [2-3], they have neglected the crucial role of search engines, necessitating a deeper examination and understanding of current search engine practices in fake news governance.

1. The Relationship Between Search Engines and Fake News Causes

Search engines are not merely platforms; they frame and shape how people view the world. Precisely for this reason, search engines play a vital role in fake news governance. The generation of fake news through search engines can be attributed to three main aspects.

First, the free ranking mechanism. For websites not using paid advertising, search engines employ their own ranking rules and algorithmic mechanisms that assign weights to various metrics. Generally, a website' s click volume, dwell time, and frequency of being linked by third parties serve as important criteria for priority recommendation. Regarding click volume, when a link is clicked, the search engine algorithm assigns relevance between that link and the search query.

When enough users click the link, its relevance index rises, boosting its search ranking. This positive feedback loop generally helps optimize search results. However, the problem arises when a link containing false news is repeatedly clicked and presented to audiences, ultimately increasing the likelihood that fake news will be mistaken for truth and spread widely. Additionally, from the perspective of news characteristics, sensational, conflicting, and controversial information more easily attracts audiences, increasing the risk that fake news spreads faster than real news.

Second, the search engine advertising ranking mechanism. Generally, both domestic and foreign search engines have advertising ranking mechanisms to increase revenue. Through real-time bidding, search engines sell user behavior data to advertisers, allowing advertising fees to help websites rank at the top of search results. Different search engines have different bidding rules, requiring internet companies to set regulations and conduct manual gatekeeping. However, corporate commercial logic ultimately requires choosing between company interests and public interests. Over-prioritizing the former easily introduces low-quality and false text or video links. The 2016 Wei Zexi incident brought domestic search engines' paid search ranking rules into the spotlight of public opinion, also reflecting how false information spreads under this mechanism.

Third, user-generated content and spam indexing. Spam indexing violates search engine rules. Creating and formatting content on websites is relatively easy for people with basic reading and writing skills. However, creating high-quality content on online platforms—such as regularly updating information, providing accurate sources, ensuring readability, and collecting materials—often requires significant time, effort, and money. Spam indexers can bypass these cost barriers. Using special software, they produce similar but different content from a core piece while keeping the webpage links unchanged, rapidly generating thousands of fake webpages that evade search engine detection. These fake news operations have mature gray industrial chains, making search paths fraught with traps. For example, many fake websites confuse audiences by using URLs or website identifiers similar to mainstream media, such as Abcnews.com.co, whose logo closely resembles the real ABC News website. Similarly, Baidu' s recommendations contain much fake news content masquerading as news while 夹杂 advertising.

In the big data era, algorithm-driven news distribution has become increasingly common. While this satisfies public demand for personalization, it also creates "filter bubbles"that isolate users from diverse perspectives, creating "information silos." If users accidentally click on fake news once, the probability of receiving similar news 推送 increases during future browsing.

2.1 Technology-Led Algorithmic Governance

In most cases, netizens are not direct producers of fake news but rather play the role of disseminators. Since technology can boost fake news, it can also

block it. When using technology for governance, the main approaches involve improving existing algorithmic mechanisms and utilizing new technologies like AI to automatically identify fake news.

Platforms have revised their recommendation algorithms. For instance, Google and Bing have downranked low-credibility information. Google uses its unique "PageRank" algorithm to redefine search result recommendation rules. The new search result ranking assigns weight to three key aspects: "relevance" between information and queries, "authority" of content sources, and "freshness" of information, while downgrading misleading and false online articles. Meanwhile, human teams have recruited over 10,000 volunteers globally to objectively evaluate search results generated by Google' s algorithm according to guidelines exceeding 160 pages, measuring content source authority [4]. In 2020, the COVID-19 pandemic highlighted the dangers of fake news. In response, Google announced a $6.5 million investment to fund global fact-checkers focused on detecting novel coronavirus misinformation. Google' s YouTube also introduced special policies to handle all content contradicting WHO recommendations. Additionally, to combat health misinformation (COVID-19), Microsoft (Bing) created COVID-19 information centers in 53 global markets with an experienced editorial team.

AI technology has also played an important role in fake news identification. Since 2016, Google has used AI algorithms to clean up machine-generated fakes. Researchers at the University of Michigan developed an automated fake news detection system based on linguistic algorithms by analyzing news grammar structure, punctuation, and vocabulary choices, achieving 76% accuracy. This algorithm helps platforms like search engines effectively filter fake news, and Google has collaborated on this technology. Google also set up an extension program in its Google Chrome browser to automatically grade news as safe, pending verification, or unsafe.

Although current efforts to combat fake news using algorithms and other intelligent technologies remain in the exploratory stage and cannot fundamentally eradicate fake news—struggling to accurately identify false content such as out-of-context quotes, unverifiable propaganda, and claims—the speed and sensitivity of technology can at least slow the spread of fake news.

2.2 Multi-Party Collaboration to Establish Fact-Checking Mechanisms

Fact-checking represents an effective method for fake news governance, encompassing both internal news verification tools and collaboration with third parties.

Internally, Google added news authenticity checking tools to its search pages. When users conduct news-related searches, authenticity check options from PolitiFact or Snopes appear below the results. Snopes and PolitiFact are specialized online databases for fake news identification. Through this cross-platform content exchange and verification mechanism, they can verify news facts and determine whether trending news is authentic.

Beyond general users, Google also provides targeted information verification services for journalists. The "Reporting" module launched in Google News Lab in 2015 offers tools including My Google Maps (allowing users to create personalized interactive maps), YouTube, Google Earth Pro, Google Media Tools Alerts (enabling journalists to track breaking events and reporting progress on maps in real time), and Citation Permissions (similar to academic citations, protecting copyrights of some reports). Google Maps helps media professionals effectively verify information by analyzing searched information and its geographic location. The "Reverse Image Search" function helps verify whether images are original, allowing journalists to find sources of images and videos. This module packages these tools to provide journalists with as much news background information as possible, helping them better judge retrieved information.

In third-party collaboration, Google partners with 115 international professional organizations including Snopes and PolitiFact, launching dedicated fact-checking labels on user search pages to present news verification basis. Google News Lab, together with organizations like the Nieman Journalism Lab, Pew Research Center, Facebook, Twitter, and foundations, established the non-profit network alliance "First Draft News," gathering relevant parties including media, third-party fact-checking agencies, and independent media on its platform.

In 2016, First Draft News established an "Online Practice Community" with nearly 100 news organizations and journalism schools globally, focusing on exploring and discussing false information in elections. In 2017, First Draft News collaborated with news site ProPublica to launch a "multi-party verification" project for the French election, monitoring false and misleading statements circulating online during the 10 weeks before the presidential election through multi-institutional cooperation. In 2020, Google provided $6.5 million to fact-checkers and non-profit organizations worldwide to combat COVID-19-related fake news, including First Draft News, Brazil' s collaborative verification project Comprova, Europe' s Full Fact and Maldita.es verification organizations, Germany' s Corrective organization, and Spanish-language verification organization Latamchequa [5].

In summary, search engine companies represented by Google mostly leverage their influence and network relationships to organize multiple media institutions to establish verification alliances for fact-checking projects. This approach integrates the strength of fact-checking institutions nationwide, better promoting fake news governance.

2.3 Implementing Media Literacy Training Programs to Enhance Public Information Identification Capacity

Media literacy education in the West began in the 1950s-1960s, evolving from initially emphasizing student empowerment to now empowering the general public, with increasingly broad coverage, and 被视为 an effective tool for building an inclusive society [6]. Promoters of media literacy include governments, enter-

prises, NGOs, and schools. The prevalence of fake news requires participation from multiple social forces to jointly enhance society's overall information identification capacity. As the world' s largest search engine, Google provides different programs annually for different population groups.

First is professional training for journalists. In March 2018, Google launched a project called "Google News Initiative," allowing news organizations to partner with Google and earn revenue through the "Subscribe with Google" function to support quality journalism reporting. Google also developed an open-source tool called Outline for news organizations, enabling journalists to access more secure network links.

Google India selected 200 journalists from various cities to participate in a five-day "train the trainers" boot camp aimed at improving their fact-checking and training abilities. Trained journalists become trainers who can train more journalists through half-day, one-day, or two-day workshops. Since India uses multiple languages, Google added Hindi, Tamil, Telugu, Bengali, Marathi, and Kannada to the training beyond English. In April 2018, Google opened a boot camp in the Middle East and North Africa region, planning to train 4,000 journalists within a year, while journalists could register online for Google News Lab to receive training in research, writing, editing, and fact-checking. Additionally, Google partnered with the World Bank and Code For Africa to train 6,000 journalists in 12 African cities.

In 2019, Google collaborated with Taiwan' s FactCheck Center—a third-party verification organization certified by the International Fact-Checking Network (FCN)—to hold a three-day "Tackling Misinformation Deep Fact-Checking Training Camp" at Google' s Taipei office. This training camp specially invited senior international media professionals to teach active news media workers and journalism academics diverse verification methods and techniques, aiming to continue implementing and promoting deep fact-checking in Taiwan' s media environment and jointly discuss how to enhance digital literacy, media literacy, and civic participation.

Second, since 2018, Google and other companies have frequently launched media literacy programs for children and adolescents. In 2019, Google launched the two-year "Interland Be Internet Awesome" program for children, encompassing six media literacy activities: understanding AI and robots, avoiding phishing attacks, verifying information credibility, evaluating information sources, identifying online false information, and discovering fake URLs. The program collaborates with national media literacy education associations to develop curriculum systems, which are available online for teachers and families in eight languages including English and Spanish [7].

For teenagers (middle and high school students), Google collaborated with Stanford University' s Graduate School of Education on the "MediaWise" project, aiming to enhance citizens' ability to identify "fake news" in digital environments. Launched in 2018, the project targets reaching 1 million teenagers. Through

online and offline teaching, it introduces teenagers to techniques for identifying fake news, such as using reverse Google image search to verify image sources and using fact-checking websites. The project also works with teenagers to create educational materials on social media to help more adolescents improve their media literacy [8]. These new media literacy programs primarily adopt participatory learning methods, placing greater emphasis on cultivating digital literacy skills such as retrieving, utilizing, evaluating, and discerning information.

3. Purifying Cyberspace: Comparative Reflections on Chinese Search Engines

Baidu search engine ranks in the same tier as Google. Although Google' s advertising promotion mechanism and Baidu' s bidding ranking mechanism are similar models, Baidu's bidding ranking places greater emphasis on the payment capacity of corporate clients—meaning that regardless of whether enterprise information is authentic, as long as they pay enough, their product and commodity advertisements can be ranked at the top of search results. The consequence of such ranking has triggered a series of bidding ranking scandals, from CCTV' s "News 30 Minutes" exposure of Baidu' s bidding ranking 内幕 in 2008, to "Economy and Law" program' s renewed exposure in 2011, to the 2016 Wei Zexi incident, and Xinhua News Agency' s renewed exposure of Baidu' s bidding ranking problems in 2018. In 2019, an article titled "Search Engine Baidu is Dead" from a self-media lab went viral, accusing Baidu of directing more than half of its search results to its self-media platform Baijiahao, which is not only filled with low-quality marketing content but also promotes fake news with readership reaching hundreds of thousands [9].

Although Baidu, like foreign search engines, uses AI technology and manual patrol methods to jointly 挖掘 obscene and pornographic content, gambling-related harmful information, and has rejected hundreds of millions of 违规 medical advertisements, it also launched a public participation-based "Duchacha Self-Discipline Committee" in 2020 to combat vulgar and false information through 吸纳 social forces. However, the number of 违规 information reports from netizens remains limited [10]. Compared with Google, Baidu' s handling of fake news only represents a small portion of false information or advertisements, focusing more on governing vulgar content and advertising content, with obviously insufficient investment in fake news governance. Among internet webpages, quality webpages account for only 7.4%, while poor-quality websites account for 21% [11].

In China, specialized fake news debunking entities include Xinhua News Agency, The Paper, and other professional news institutions, as well as relevant government departments for authoritative investigation and release. A small portion of news relies on netizens' questioning and spontaneous investigation, representing a top-down governance approach. Social media platforms, self-media accounts, and commercial media are increasingly 介入 in fake news production and creation. These links, through search engines, generate information flow

driven by click volume [12]. Relying solely on media organizations and government departments is far from sufficient; search engines should 理应 become an important force in fake news governance.

The key difficulty in combating fake news lies in the conflict between platform business models and fake news governance. The advertising business model essentially extends user time on platforms and encourages user engagement by stimulating attractive content, catering to user preferences. In other words, platforms send targeted advertisements or false information to different users based on algorithmic logic, triggering clicks to generate revenue, and the proliferation of fake news stems precisely from platforms' "click economy." Low-quality fake news content can bring false prosperity to platforms in a short time, making it difficult to replace. Foreign search engines' measures to "hinder fake news economic incentives" by prohibiting fake news websites from using advertising services and canceling false accounts also have problems. On one hand, they sacrifice platform economics; on the other hand, virtual websites can return to platforms under new guises, and the powerful industrial chains behind them place fake news in a governance 困境 of "growing back after being cut." In the long term, gaining audience trust is the path to development for online platform companies; the key to continuously attracting audience attention lies in mutual trust rather than eye-catching information [13]. Chinese search engines need to optimize algorithmic ranking technology, intensify fact-checking, collaborate with media to establish fact-checking mechanisms, and build a healthy online environment for the public. Quality journalism can bring high-quality services to search engine users. Search engines joining fake news governance to reduce fake news spread is essential for rebuilding social trust.

References

[1] Xinhua News Agency. Post-truth era: Europe and America take measures, the battle between true and fake news never ends [EB/OL]. Xinhuanet, 2017-02-16. http://www.xinhuanet.com/zgjx/2017-02/16/c_{136060574}.htm.

[2] Shao Guosong. How Social Media Affects U.S. Presidential Campaigns [J]. People' s Tribune • Academic Frontier, 2020(15): 83-93.

[3] Zhang Chao. Algorithmic Governance of Fake News on Social Platforms: Logic, Limitations, and Collaborative Governance Models [J]. Journalism and Mass Communication, 2019(11): 19-28+99.

[4] Alex Hern. Google acts against fake news on search engine [EB/OL]. The Guardian, 2017-04-25. https://www.theguardian.com/technology/2017/apr/25/google-launches-major-offensive-against-fake-news.

[5] Alexios Mantzarlis. COVID-19: $6.5 million to help fight coronavirus misinformation [EB/OL]. Google Blog, 2020-04-02. https://blog.google/outreach-initiatives/google-news-initiative/covid-19-65-million-help-fight-coronavirus-misinformation/.

[6] Zhang Kai. From Grassroots Movement to Policy Promotion: Global Media Literacy Education Moving Toward Rational Development [J]. Modern Distance Education, 2012(04).

[7] Beth Staats. Interland: Google' s media literacy program teaches kids about digital safety [EB/OL]. Minitex, 2019-06-27. https://www.minitex.umn.edu/news/2020-07/interland-googles-media-literacy-program-teaches-kids-about-digital-safety.

[8] Kerry Flynn. How Google-backed MediaWise is teaching teens media literacy [EB/OL]. Digiday, 2019-01-07. https://digiday.com/media/google-backed-mediawise-teaching-teens-media-literacy/.

[9] Fang Kecheng. Search Engine Baidu is Dead [EB/OL]. The Paper, 2019-01-24. https://www.thepaper.cn/newsDetail_{{{forward}}_{{2897939}}}.

[10] Southern Metropolis Daily. Baidu Establishes Duchacha Committee, Cooperates with All Sectors to Combat Online Harmful Information [EB/OL]. Sohu, 2020-11-30. https://www.sohu.com/a/435315389_{161795}.

[11] Baidu. Baidu Search Engine Webpage Quality White Paper [EB/OL]. Baidu Resources, 2019-09-18. https://ziyuan.baidu.com/college/articleinfo?id=1337&page=3.

[12] Annual Fake News Research Group, Bai Hongyi, Zhang Tian, Chen Bin. 2019 Fake News Research Report [J]. Shanghai Journalism Review, 2020(1): 22-33.

[13] Zhang Cong. Why Fake News Periodically Resurrects—Notes from the World Internet Conference Interview No. 2 [EB/OL]. People' s Daily, 2016-12-07. http://opinion.people.com.cn/n1/2016/1207/c1003-28931860.html.

Author Biographies:

Cui Xinyuan (1999-), female, from Chongqing, majoring in Communication Studies at Hangzhou Dianzi University, research focus: new media communication.

Hu Yiquan (2000-), female, from Hangzhou, Zhejiang, majoring in Communication Studies at Hangzhou Dianzi University, research focus: new media communication.

Fu Yu (1991-), female, from Xinyang, Henan, lecturer, Ph.D., at Hangzhou Dianzi University, research focus: media technology studies, media literacy.

(Responsible Editor: Li Jing)

Note: Figure translations are in progress. See original paper for figures.