WINNING THE WEB

How Beijing exploits search results to shape views of Xinjiang and COVID-19

Jessica Brandt Bret Schafer Elen Aghekyan Valerie Wirtschafter Adya Danaditya

B | Foreign Policy at BROOKINGS



GOVERNANCE | MAY 2022

WINNING THE WEB

How Beijing exploits search results to shape views of Xinjiang and COVID-19

Jessica Brandt, Bret Schafer, Elen Aghekyan, Valerie Wirtschafter, and Adya Danaditya

EXECUTIVE SUMMARY

As the war in Ukraine unfolds, Russian propaganda about the conflict has gotten a boost from a friendly source: government officials and state media out of Beijing. In multiple languages and regions around the world, China's "wolf warrior" diplomats and state media routinely amplify Kremlin conspiracy theories rationalizing President Vladimir Putin's invasion, and undermining the credibility and appeal of the United States, NATO, and independent media even as China declines to endorse the Kremlin's adventurism wholesale. This spring, for example, China's messengers promoted the baseless Russian claim that the United States has been supporting a biological weapons program in Ukraine - at times, more aggressively than Russia itself.

Because Russian state media have been deamplified or banned by multiple Western social media platforms, Beijing's messaging could play an outsized role in channeling Kremlin talking points to audiences around the world. These narratives do not just spread on social media. Beijing's state-funded publishers have considerable success in a domain that has received comparatively little attention: search results.

For months, our team has been tracking how China has exploited search engine results on Xinjiang and COVID-19, two subjects that are geopolitically salient to Beijing — Xinjiang, because the Chinese government seeks to push back on condemnation of its rights record; COVID-19, because it seeks to deflect criticism for its early mishandling of the pandemic. In both cases, Beijing is quite focused on positioning itself as a responsible global leader and softening perceptions to the contrary.

To evaluate these concerns, we compiled daily data over a 120-day period on 12 terms related to Xinjiang and COVID-19 from five different sources: (1) Google Search; (2) Google News; (3) Bing Search; (4) Bing News; and (5) YouTube. We found that:

- Chinese state media are remarkably effective at influencing the content returned for the term "Xinjiang" across several search types. "Xinjiang," which is among the most neutral terms in our data set, regularly returned state-backed content across news searches, with at least one Chinese state-backed news outlet appearing in the top 10 results in 88% of searches (106 out of 120 days searched). On YouTube, state media appeared among the top 10 results in searches for "Xinjiang" in 98% of searches (118 out of 120 days searched).
- Consistent with past research, search results for conspiratorial terms across all search types yielded a high volume of statedriven content. Take, for example, the term "Fort Detrick" - a military base in Maryland that housed the United States' biological weapons program from 1943 to 1969 and has become a central figure in China's efforts to spread disinformation about the origins of the coronavirus outbreak. On YouTube, searches for "Fort Detrick" regularly returned state-backed content, with 619 observations of videos from Chinese state media outlets appearing in the top 10 search results during our study (or around five per day). Similarly, "Unit 731," a biological and chemical weapons research unit located in Japanoccupied China during World War II and a subplot in China's efforts to connect the origins of the coronavirus outbreak to Fort Detrick, appeared on the first page of search results for news searches every single day of data collection.
- Both news search (Google News and Bing News) and YouTube search are much more likely to disseminate Chinese state media than web search. Chinese state media accounted for roughly 22% of the observed pages and 25% of observed channels in search results for queried topics related to Xinjiang and the origins of the coronavirus on news and YouTube searches, respectively. By comparison, Chinese state media accounted for only 6% of results for the same queried topics on Google and Bing web search.

- Terms tied to the pandemic were less likely to return state-backed content than terms tied to Xinjiang, likely due to the considerable attention platforms have paid to moderating COVID-19 content.
- Because of China's extensive content hosting and influencer arrangements, our research likely underestimates the prevalence and prominence of Chinese state media in search results, given the challenges of identifying this content republished in seemingly independent sources. In our dataset, at least 19 different sources that are not officially affiliated with the Chinese government but regularly republished Chinese state media content verbatim (e.g., the Helsinki Times) appeared in top search results. Including observations from only those 19 sources would increase the total number of Chinese state media occurrences in search results by nearly 10%.

To address these findings, we propose that companies:

- Address hosting, reposting, and syndication, recognizing that agreements between international news outlets and Chinese state media are a significant avenue for the proliferation of Beijing's narratives, including misleading and conspiratorial content. Potential remedies include clear labels and links to the original source.
- Expand the practice of labeling the websites of state media, agencies, and officials in search results, which provides important context to users.
- Provide notice to users when result quality is suspect, as Google has done for breaking news events, including for searches for contested terms or topics that are a battleground for search result dominance.
- Provide information on how ranking decisions regarding state content are assessed and made — including whether factors that lead to deranking (as in the case of Russian state media and Google) may have implications for content produced by other states.

- Contribute to public education about how engines find, rank, and surface content. User trust in engines like Google remains high, but there is little evidence that users have a strong understanding of the factors that determine results.
- Collaborate and share information with other search engines, as they have in the past, to improve the performance of their technologies. The goal should not necessarily be the formation of a new institution or mechanism, but rather information exchange on how to address common vulnerabilities.

Other actors can play a role too. In particular:

 Content creators who cover issues related to or of importance to Beijing – including research organizations, government officials, and activists – should develop an understanding of how audiences are searching for their stories and maintain awareness of efforts to subvert them. Considering audience behavior when tagging, using keywords, and developing communication strategies can contribute to countering misleading stories. Authoritative outlets should reconsider syndication agreements with state media that do not have appropriate controls to ensure editorial independence. At a minimum, they should enhance disclosure and labels to better inform audiences.

By taking these steps, companies, content creators, and authoritative outlets can ensure that Beijing is not able to dominate search results for terms related to its geopolitical interests, and that users have the information they need to contextualize the propaganda they encounter.

PART 1. INTRODUCTION

In research and discourse about information manipulation by autocrats, social media platforms have received the bulk of attention, with good reason. Russian and Chinese state media and officials use social media to disseminate their preferred, often distorted narratives around strategic issues through their own accounts and those of sympathetic influencers and amplifiers. But social media platforms are just one component of the information ecosystem that autocrats use to spread propaganda. Search engines are also a vector for influence.

The Kremlin has repeatedly capitalized on search results to disseminate multiple, at times conflicting conspiracy theories to deflect blame for its misdeeds and seed the idea that there is no such thing as objective truth. After Russianbacked separatists in eastern Ukraine downed flight MH17 in 2014, Russian state media spread multiple false claims - including that the airliner was shot down by Ukrainians who mistakenly thought they were targeting Vladimir Putin's plane.¹ For weeks, these claims showed up in articles from Russian state-controlled outlets on the front page of Google News results for searches of "MH17" through the site's "Top Stories" function.² A similar phenomenon took place after the poisoning of Russian defector Sergei Skripal and his daughter Yulia in Salisbury in the United Kingdom in 2018, when Kremlinbacked narratives denying evidence of Russian wrongdoing and promoting alternative theories of culpability regularly appeared in search results for the term "Skripal."3 Most recently, in the early days after Russia's invasion of Ukraine, Russian propaganda performed surprisingly well in searches for key terms related to the conflict.⁴ Beijing, which has adopted other elements of Moscow's information manipulation playbook, seems to have benefited from the same dynamic. For at least two months last year, Google News results for the term "Fort Detrick" – a U.S. Army research lab in Maryland that has been the target of a disinformation push casting it as the source of COVID-19 – were dominated by CGTN and Global Times, two Chinese state-run outlets that are central to Beijing's information operations.⁵

Auditing the prominence of propaganda content in search results is particularly important given consistently high levels of public trust in search engines.⁶ Surveys and sociological studies suggest that users tend to view search engines as neutral conduits of information, believing that searchers are in control of what they find.7 The commonly repeated trope of "do your own research" - which helped fuel misinformation during the coronavirus pandemic - is far more likely to lead people to search engines than the public library. Yet as the previous examples demonstrate, the search results users encounter can also be distorted to suit a preferred narrative and may not provide the most authoritative content⁸

This report focuses on the performance of Chinese state media in search results related to two topics that are critical to Beijing's information agenda: Xinjiang and the COVID-19 pandemic. The objectives of this research are twofold: first, to enhance our understanding of how search engines can expose users to false or misleading information, and second, to contribute to discourse about the role and responsibilities of search engines in our information infrastructure.

PART 2. CHINESE STATE NARRATIVES AND STRATEGIC TOPICS

The Chinese government commands a robust communication machine, including traditional global media outlets available in dozens of languages, China's so-called "wolf warrior" diplomats, pro-government trolls, and paid social media influencers. This multilayered system allows China to shape and control narratives both overtly and covertly, reinforcing and legitimizing China's strategic messaging. Once largely defensive in nature and designed to insulate China from supposed information threats from abroad, China's efforts to assert narrative dominance have increasingly been waged through external propaganda aimed at foreign audiences.⁹

To that end, Beijing has invested heavily in its global media presence over the past decade, developing websites, news agencies, and broadcast networks with the stated goal of "telling China's story well."10 In less public statements. Chinese officials have been more explicit, referring to state media as the "front line of external propaganda" whose mission is "to win an acknowledged place [for China] in the global public opinion competition."11 CGTN, the international arm of state-run China Central Television (CCTV), operates a network of international bureaus and produces news in English, Spanish, Arabic, and four other languages.¹² State news agency Xinhua has content hosting agreements with outlets around the world, including in the United States.¹³ News websites like China Daily and Global Times provide additional avenues for the promotion of state narratives online, and China Radio International broadcasts content in more than 44 languages through radio stations across the world,¹⁴ some of which obscure their connection to the Chinese government.¹⁵

The mission of "telling China's story well" is primarily one of propaganda: state outlets endeavor to soften the image of China's authoritarian regime, suppress critical coverage of domestic policies, and promote Beijingfriendly narratives about both China and the world. A number of strategic topics recur within these narratives, including Taiwan, Hong Kong, Xinjiang, and, as of 2020, the COVID-19 pandemic. Beijing's information apparatus generates a large volume of content around these topics online. Chinese state-run outlets, journalists, and state officials amplify this content on social media platforms — an information loop that circulates user attention back to state media.

Beijing uses specific framing and terminology to distinguish its preferred narratives. For example, international media and human rights organizations have described China's mass internment and forced indoctrination of Uyghurs in Xinjiang as a genocide,¹⁶ while Chinese state media¹⁷ and government officials¹⁸ have labelled genocide allegations as "the lie of the century" and flooded genocide hashtags with content that instead highlights the genocide of Native Americans.



From left to right: Examples of tweets from China's consul general in Durban, South Africa, and China's embassy in the Netherlands highlighting the genocide of Native Americans and referring to genocide allegations as "the lie of the century" (February 9, 2022, and December 2, 2021).



Hua Chunying 华春莹

@SpokespersonCHN

China government official

The **#US** keeps calling for transparency & investigation. Why not open up Fort Detrick & other bio-labs for international review? Why not invite **#WHO** & int'l experts to the US to look into **#COVID19** source & response?

11:27 AM \cdot May 8, 2020 \cdot Twitter Web App

603 Retweets 117 Quote Tweets 3,651 Likes

Tweet from China's Foreign Ministry Spokesperson Hua Chunying suggesting a connection between Fort Detrick and the outbreak of the coronavirus (May 8, 2020).

This phenomenon is related to the role of search engines in the state media information loop. Search engines rely on users to conceptualize a query and retrieve information, and the search terms that an individual uses are influenced by a broad linguistic universe, defined by ideology, prior exposure to the term, and other factors.¹⁹ We suspect that exposure to Beijing's narratives on social media can influence how a user searches for information about those topics. Prior research has shown that Chinese state media can rank high for Google and Bing searches for terms that regularly appear in Chinese state media — like "Fort Detrick."²⁰

To better understand how primed keywords can lead search engine users to state media, this report focuses on terms surrounding two topics of importance to the Chinese government: Xinjiang and COVID-19. These topics were

selected in part because of their prominence in Chinese messaging over the past year. Though COVID-19 was obviously a top topic among all global media outlets and governments in 2021, it received disproportionate attention from Chinese diplomats and state media. According to data collected by the Alliance for Security Democracy (ASD)'s Hamilton 2.0 dashboard, which tracks outputs of Russian, Iranian, and Chinese stateaffiliated media and government officials on social media, Chinese official accounts were nearly three times more likely to mention "covid" on Twitter than comparable Russian accounts in 2021.²¹ And #covid19 and #Xinjiang were the two most frequently used hashtags on Twitter by all official Chinese accounts in 2021, with Xinjiang registering as both the top keyword and the top hashtag used by Chinese diplomats and government officials (Figure 1).

FIGURE 1

Most frequent hashtags used by Chinese government officials and state media accounts in 2021



SOURCE: Hamilton 2.0 dashboard

Beyond the volume of content produced by Chinese state media on these topics, both Xinjiang and COVID-19 are also associated with conspiracy narratives linked to specific keywords, unlike topics such as Hong Kong and Taiwan. The unique vocabulary used by Chinese officials to discuss Xinjiang and COVID-19 diverges from language used by most international media outlets, subject matter experts, and other credible sources of information to discuss the same topics.

PART 3. DATA ANALYSIS AND FINDINGS

METHODOLOGY AND DATA COLLECTION

To explore the prominence of Chinese state narratives across search results, we compiled daily data on 12 terms from five different sources: (1) Google Search; (2) Google News; (3) Bing Search; (4) Bing News; and (5) YouTube. We began collecting data on the 12 terms on November 1, 2021. We finished the data collection for all terms on February 28, 2022.²²

SEARCH TERMS

The 12 terms for which we collected data focus on two topics: Xinjiang and the COVID-19 pandemic. These terms were identified for their strategic importance through extensive monitoring of Beijing's narratives about Xinjiang and COVID-19. For more information about why these topics and terms were chosen, refer to Part 2.

Xinjiang-related terms

Xinjiang: The name of a region in China that is home to a number of ethnic minority groups, including the Uyghur population.

Uyghur: An ethnic and religious minority group living in Xinjiang.

Kashgar: A major city in the Xinjiang region.

Urumqi: A major city in the Xinjiang region.

Adrian Zenz: A U.S.-based anthropologist who studies Xinjiang and is frequently maligned by Beijing-backed outlets and individuals.

Xinjiang Debunked: A primed term commonly promoted by Chinese state-linked accounts on social media to discredit international narratives around Chinese treatment of the Uyghur population in Xinjiang. **Xinjiang Terrorism**: A primed term commonly promoted by Chinese state-linked accounts on social media to justify Chinese action in the Xinjiang region.

Coronavirus-related terms

Fort Detrick: A U.S. military facility at the center of Beijing's COVID-19 origin conspiracy theory.

Unit 731: A biological and chemical weapons research unit located in Japanese-occupied China during World War II that Chinese messengers have connected to Fort Detrick to cast further doubt about the lab's role in the outbreak of COVID-19.

EVALI Virologist: A primed term based on Chinese state media attempts to tie outbreaks of EVALI (e-cigarette or vaping use-associated lung injury) to biological research in the United States.

EVALI Coronavirus: A primed term based on Chinese state media attempts to tie outbreaks of EVALI (e-cigarette or vaping use-associated lung injury) to biological experimentation in the United States.

Huanan Seafood Market: The location of the earliest believed cases of COVID-19, in Wuhan, China.



Liu Xin刘欣 @LiuXininBeijing 旨 China state-affiliated media

People talk about cultural genocide in #Xinjiang, I wonder what they r talking about? From the Urumqi airport to our hotel, every road sign has at least two languages if not three: Putonghua(standard Chinese), Uyghur & English! Lie No. 1 debunked. Beautiful clean air BTW.





Video: Document reveals rare footage of terrorist attacks on civilians and why #Xinjiang is China's key region to fight against terrorism #Antiterrorism



Examples of Chinese state media tweets "debunking lies" about Xinjiang and covering terrorist attacks in Xinjiang. From CGTN anchor Liu Xin and CGTN's official Twitter account (March 31, 2021 and December 8, 2019).



Examples of Chinese diplomatic and state media tweets referencing Unit 731 and Fort Detrick, respectively, and suggesting a connection between COVID-19 and an outbreak of EVALI in 2019. From China's Ministry of Foreign Affairs Spokesperson Zhao Lijian and state media outlet People's Daily (June 23, 2021, and August 5, 2021).

Where possible, we collected data, including title, source, and URL, from the first page of search results for each term in our analysis, which we refer to throughout the report as "top search results."23 We also noted the rank of the article on the page and the date the information was collected. On average, we collected 583 observations daily across the five sources and 12 terms.²⁴ The data set for the extended period totals 70.042 observations. Following the completion of the data collection process, we classified each search engine result based on whether or not it directly originated from a Chinese government-backed outlet or organization. To complete this classification process, we relied on the outlets and accounts tracked by the Alliance for Securing Democracy's Hamilton 2.0 Dashboard.²⁵ Across all search terms and search processes, we classified 11,191 observations (16%) as state-backed media. A breakdown of observations by source and classification is included in Table 1.

For our analysis, we examine three metrics: (1) the raw number of search results per term stemming from Chinese state media; (2) the number of days Chinese state media featured among the top search results (out of a total of 120 days); and (3) the page order ranking of state and non-state media in top search results. Where applicable, we combine Google News and Bing News together as "news search" or simply "news" and Google Search and Bing Search together as "web search" or simply "search."

TABLE 1

Search engine	Total observations	Chinese state media observations	Not Chinese state media observations	Percentage Chinese state media
Bing News	12,477	2,260	10,217	18%
Google News	14,310	3,525	10,785	25%
Bing Search	15,840	762	15,078	5%
Google Search	13,255	1,076	12,179	8%
YouTube	14,160	3,568	10,592	25%
TOTAL	70,042	11,191	58,851	16%

16% of top search data collected over 120 days returned Chinese state media content

HOW DO SEARCH ENGINES WORK?

To understand the performance of state media in search rankings, it is useful to understand how search works. Search engines function through three basic steps: (1) crawling the web to find pages, (2) indexing these pages based on unique features like keywords and images, and (3) sorting through the index in response to a query to deliver the most relevant results, which the engine ranks based on factors that include the "freshness" or "authoritativeness" of content. Engines may personalize results to a certain degree based on language, location, and past content consumption, and generally function according to their unique systems of algorithms. On social media platforms like YouTube, search functions in similar ways, though within the boundaries of the content on the platform.²⁶

DATA ANALYSIS

Broadly, we find that both the nature of the search process and the substance of the search term play a critical role in the prevalence of statebacked content across search results, both in terms of raw numbers and days where state content featured among top results. As Table 1 shows, both news and YouTube search results are much more likely to disseminate Chinese state media than web search results. However, consistent with past research, search results for more conspiratorial terms across all types of search yielded a high volume of state-driven content.

Terms tied to COVID-19 were less likely to return state-backed content than terms tied to Xinjiang, likely due to an abundance of frequently published authoritative content about the pandemic, as well as concerted efforts by the companies to protect the quality of healthrelated information on their platforms. Across search terms, the frequency with which Chinese state-backed media appeared in the top search results remained fairly constant over time. This is particularly notable, given that the 2022 Winter Olympics in Beijing occurred during the period in question. Across all terms, we do not register an obvious decline (tied potentially to a crackdown by platforms) or increase (tied potentially to a concerted effort by Beijing to "flood the zone" with favorable narratives) in state-backed content performance in search results, despite the attention paid to China during this period.

In subsequent sections, we detail additional results tied to the thematic focuses of our terms – the treatment of Uyghur populations in Xinjiang and the coronavirus pandemic. We also explain that these results are likely an undercount of the prevalence of state-backed content across search processes and terms. This is due to content hosting and influencer agreements with international news outlets and content creators across news and YouTube searches, respectively, which are widely prevalent but not counted in our classified results.

Search platforms boost state-backed media for both loaded and neutral terms tied to Xinjiang

In examining the performance of Chinese state media for terms tied to the treatment of the Uyghur population in Xinjiang, we find that statebacked content tends to regularly feature in top search results across the board. In total, 9,005 out of 41,405 search results (21%) originated from state-backed media, with four of the seven queried terms in our dataset generating over 1,000 state media results each during this period.²⁷ The only term that returned fewer than 500 state media results across search types was "Uyghur." Figure 2 illustrates these results by term and search platform.

Chinese state media frequently appeared in top search results for both neutral and loaded terms related to Xinjiang



■ bing news ■ bing search ■ google news ■ google search ■ youtube

We also found that more loaded terms like "Xinjiang debunked" and "Xinjiang terrorism" returned state-backed content across all three types of searches. As noted in Figure 3, across news, web, and YouTube searches, Chinese state media appeared in the top search results for these loaded terms in all but four of the 120-day period during which we monitored search performance. This is not particularly surprising given past research on the quality of search performances tied to conspiratorial terms in the United States.²⁸

What is surprising is how well state media performed in the top searches for more neutral terms like "Adrian Zenz," the name of an anthropologist who studies the region; "Kashgar," a city in Xinjiang; or even just "Xinjiang," the name of the region where the Uyghur population is concentrated (Figure 4). This suggests that regardless of whether users enter biased search queries based on previous exposure to China's preferred narratives about Xinjiang, they are still likely to encounter Chinese state media content about the region.

FIGURE 3

Across all search types, Chinese state media appear in the top results nearly every day for loaded terms related to Xinjiang²⁹

Out of 120 days of data collection



Search Term: Xinjiang Terrorism



Search Term: Xinjiang Debunked



FIGURE 4

Chinese state media frequently appeared in the top results on news and YouTube searches for neutral terms related to Xinjiang

Out of 120 days of data collection

First Page Not on First Page

Search Type: News Search



Search Type: Web Search



Search Type: YouTube Search



Across all search types, the term "Adrian Zenz" surfaced state-backed content almost every day. Chinese sources appeared in the top search results for at least 118 out of 120 days across news, web, and YouTube search (Figure 4, column 1). The term "Kashgar" returned fewer state media articles in web search results than the term "Adrian Zenz," but state media did still appear in at least 120 and 118 days of news and YouTube results, respectively (column 2). The term "Urumgi" surfaced less state content on YouTube than most other terms, but state media did appear in top search results every day across news-based searches (column 3). "Uyghur" is the only term in our list that did not regularly return Chinese state media content, with only 10 total state-backed news results in 120 days.

These findings are consistent with prior research about how a user's framing of search terms can influence results. Francesca Tripodi, a sociologist and media scholar at the University of North Carolina at Chapel Hill, has explored how ideology shapes the keywords that individuals use to research divisive topics, finding stark contrasts in how search engines surface different results for ideologically biased or value-laden keywords.³⁰ For example, Tripodi shows that guerying Google Search with "illegal alien" and "undocumented worker" return very different results. ³¹ In the case of the former, the top results in her query included content from conservative think tanks and media, as well as links to the websites of immigration agencies; in the case of the latter, top results led to news from more liberal news outlets and legal resources.³² Ideological beliefs are only one set of factors that can affect the keywords that people use to search for information - and, subsequently, the information they reach. In the case of Xinjiang and COVID-19, contested geopolitics play a role, shaping the way these issues are conceptualized and described online.

The surprising performance of state media for the term "Xinjiang"

Among the most surprising results in our analysis is how effective state media was at influencing the content returned for the term "Xinjiang" across several search types. "Xinjiang," which is among the most neutral terms on our list, regularly returned state-backed content across news (106 days) and YouTube (118 days) searches (Figure 4, column 5). A closer look at the daily result in Figure 5 demonstrates that this was not a simple case of a single search result among a sea of authoritative content, but rather part of a larger patchwork of state-backed media surfaced over time. Across Bing News and YouTube in particular, state-backed media regularly occupied a position in the top three search results, and frequently appeared more than once in search results on any given day. The composition of these results did not shift over time-across the 120 days for which we collected data, state media remained a regular fixture in search results for the term "Xinjiang," particularly on Bing News and YouTube.

The neutral term "Xinjiang" regularly surfaced Chinese state media in top results across news and YouTube searches

Out of 120 days of data collection







A typical search result for the term "Xinjiang" on Bing News (November 9, 2021)

Search results recreated using daily data. State-backed media highlighted in orange.



An atypical day for the term "Xinjiang" on Google News (December 14, 2021)

Search results recreated using daily data. State-backed media highlighted in orange.



For YouTube, our data for December 28, 2021, showed the second, third, and sixth results all came from known Chinese state media accounts (highlighted in red). Bing News on November 9, 2021, returned state-backed content in the second, third, fifth, and sixth positions (similarly highlighted in red). And finally, on Google News, the fourth, sixth, and seventh results for December 14, 2021, funneled users to Chinese state content. These visualizations highlight a fairly standard day for Bing News and YouTube (there are 26 other days in our study that feature Chinese state content equally or more prominently for Bing News and 14 for YouTube) and a somewhat atypical one for Google News (there are only five other days where state content is equally or more visible). The examples demonstrate how easy it is for users

to inadvertently stumble across state-backed content in top search results, even for a neutral term like "Xinjiang."

The freshness vs. novelty trade-off for state-backed content tied to COVID-19

For terms tied to the pandemic, we find that the performance of Chinese state media across search results is highly variable, both by term and source. In total, 2,186 search results out of 28,455 (7.6%) originated from state-backed media, with two of the five terms generating over 700 state media search results each in 120 days.³³ In Figure 9, we provide some examples of state-backed media content found in our search results.

FIGURE 9

Examples of Chinese state media headlines in our search results



As is evident in Figure 10, Chinese state media are far less likely to surface in top search results for terms tied to coronavirus than for terms tied to Xinjiang (Figure 2). On average, one day of search yielded approximately 11 Chinese state-backed media results for each Xinjiangrelated term, as compared to an average of approximately four results for terms related to the pandemic. In other words, over 120 days, results across all search types would include on average 11 results from state-run sources per term, or approximately two results per search type, per day - around four times the number of state-backed results for COVID-related terms. This remained the case even after removing more loaded terms tied to the ongoing human rights catastrophe in Xinjiang, such as "Xinjiang debunked" and "Xinjiang terrorism." On average, each neutral Xinjiang term generated 6.5 total Chinese state media results per day across all search types, or nearly double the state-backed results for coronavirus terms (3.6).

While we can only speculate as to why this may be the case, one post-hoc explanation is that due to the ongoing pandemic, coronavirus-related topics have garnered more global attention from a diverse range of trustworthy media over the past two years - and thus generated more content, making it more difficult for statebacked outlets to shape this narrative through search results. Another possibility is that due to widespread efforts to combat coronavirus misinformation, technology companies have altered search and recommender algorithms and content moderation policies to prioritize authoritative content, but have not made the same adjustments for content related to Xinjiang.34

FIGURE 10

Chinese state media appear in top search results less frequently for coronavirus terms than for Xinjiang terms



For coronavirus-related terms, however, Figure 10 also highlights that the prevalence of state media in top search results varies both by term and search type. For the term "Unit 731," Chinese state media appeared in top Google News results 596 times (or around five results per day). The remaining four search types accounted for the other 204 state media hits. On YouTube, the term "Fort Detrick" regularly returned state-backed content, with 619 hits, also around five per day. During the pandemic, Chinese state media have consistently pushed the conspiracy theory that the coronavirus originated at Fort Detrick. Across all other search types, state media appeared in top search results for the term "Fort Detrick" a total of 99 times (or just 14% of all state media hits for this term).

Despite this variance by both term and search type, Beijing-backed content does appear to have maintained a regular presence across some top search results. For news searches (Google News and Bing News), Chinese state media appeared for 55 days (out of a total of 120 days) in the top search results for the term "Huanan seafood market" — believed to be the starting point of the pandemic — and for 57 days for the term "Fort Detrick." For the term "Unit 731," state media appeared in news results every day of our search, especially on Google News (Figure 11, row 1).

By contrast, web search (Google Search and Bing Search) returned less state-backed content than news search for terms like "Unit 731," "Fort Detrick," and "Huanan seafood market," but included a large number of Chinese state media for other terms that we examined, including "EVALI coronavirus" and "EVALI virologist" which draw on conspiratorial links between a respiratory illness tied to vaping and symptoms that mimic those of COVID-19 (Figure 11, row 2). In days 81 and 111, respectively, state-backed media appeared in top web search results for these terms. Across web searches, we recorded no days where the term "Unit 731" returned statebacked content, despite its high prevalence across news searches.

FIGURE 11

News and web search surface state media in top results for different types of coronavirus terms, whereas YouTube regularly surfaces state-backed content across nearly all terms related to the pandemic

Out of 120 days of data collection

105 Days



118 Days

115 Days

117 Days

These differences between web and news results seem to be driven by the freshness of the content and the novelty of the topic. Terms like "Unit 731" and "Fort Detrick" are connected to well-established topics with longer histories than, for example "EVALI." For these terms, web searches tend to return more authoritative content. However, due to the fact that news searches seem to prioritize the freshness of the content, they are more prone to returning state media attempts to construct new narratives around terms with long-standing roots. By contrast, terms like "EVALI virologist" and "EVALI coronavirus" may be akin to more traditional data voids, where the absence of authoritative content allows search results to be co-opted by actors attempting to shape a term's meaning in this case, by linking coronavirus to the rise in vaping.35

Due to the freshness vs. novelty trade-off, statebacked content performed differently across news and search, with some terms regularly surfacing Chinese state sources in one but not the other search type, and vice versa. However, on YouTube, which is largely tied to user-generated video content, pandemic-related terms regularly returned Chinese state media. For four of the five terms in our dataset, state media appeared in top search results for at least 105 out of 120 days (Figure 11, row 3). The only term that did not regularly surface state-backed media on YouTube was "EVALI virologist," perhaps due to a dearth of content related to the term more broadly. Despite an ASD report on Chinese state media's capture of the Fort Detrick coronavirus origin conspiracy, published one month before our data collection began, Beijing-backed content on the topic performed well on YouTube throughout the period we examined (Figure 12).³⁶ On average, more than half of all top search results for "Fort Detrick" on YouTube included Chinese state media channels, and Chinese state media were the first result for 102 of 120 days. This is in stark contrast to the performance of the term across web and news searches (and in particular, other Google products like Google News and Google Search), where Chinese state media rarely featured in top results.

FIGURE 12

Despite a clear debunking of the Chinese state-backed narrative linking COVID-19 to Fort Detrick, state media regularly surface in top search results on YouTube



📕 Chinese State Media 📃 Not Chinese State Media

One video on YouTube, titled "How terrifying is the history of U.S. Fort Detrick lab?" describes a "quiet town in Maryland" called Fort Detrick, "where the U.S. government conducts experiments on dangerous pathogens including close siblings of the novel coronavirus."³⁷ This video, which suggests a possible link between Fort Detrick and the pandemic, appeared in the top position 78 times and has been viewed nearly 27,000 times with approximately 1,200 likes and 700 comments, including many calling for an investigation into the research facility (Figure 13).

FIGURE 13

Still from the YouTube video, "How terrifying is the history of U.S. Fort Detrick lab?"



A network of syndicated content and influencers boosting Chinese state media

Although the results detailed in the previous section paint a complex picture of search dynamics across a variety of engines, topics, and terms, they likely fail to capture the full extent of Beijing's efforts to influence search results. Prior analyses focused on known media outlets directly linked to the Chinese government; however, known state-affiliated domains and news outlets are not the only avenues through which Beijing-backed content appears in search results. In some cases, state-backed content can be masked by opague hosting or influencer agreements, allowing Chinese state media to be laundered through seemingly independent sources. This not only makes these narratives more prominent across search results but also more difficult to identify.

How content hosting agreements facilitate the dissemination of state-backed narratives

Content hosting agreements, or web syndication, entail one entity providing content for publication and promotion on another's website. This common marketing tactic enables creators to increase their exposure and distributors to attract more users. Chinese state media have adopted this practice. For example, the staterun press agency Xinhua has signed content hosting agreements with international news outlets, including major news aggregators like MSN³⁸ as well as state news agencies like ANSA in Italy³⁹ and NAN in Nigeria.⁴⁰ As part of these agreements, international news outlets often repost content word-for-word from Xinhua. In effect, these agreements allow state narratives to spread in search results through broad-reaching news aggregators and less easily identifiable domains.41

Syndication of Chinese state media across news outlets (Top 10)

Rank	URL of Outlet	Search Term	Source Search Process	Unique State Media Article Reposted	No. of Days in Top Search Results	Highest Search Position
	Finland					
1			bing search	1	108	2
	helsinkitimes.fi	xinjiang debunked	google search	2	115	1
	UAE					
	bignewsnetwork.com	kashgar ⁻	bing news	2	14	1
			bing search	1	1	4
		urumqi	bing news	24	33	1
2			google news	1	1	8
		xinjiang	bing news	7	7	1
		xinjiang debunked	bing news	1	14	2
		xinjiang terrorism	bing news	2	15	1
	Global/USA					
		urumqi	bing news	4	13	1
3 m		xinjiang	bing news	10	15	1
	msn.com	xinjiang debunked	bing news	1	3	3
			bing search	1	2	1
		xinjiang terrorism	bing news	1	7	2
	Pakistan					
4	dailytimes.com.pk	kachgar	bing news	1	16	1
4	uanytimes.com.pk	kashgar ·	google news	1	2	2
	Chad					
5	alwihdainfo.com	xinjiang terrorism	bing news	1	13	1
	India					
6	indianexpress.com	kashgar ·	bing news	1	10	6
0			google news	1	2	3
	Pakistan					
7	bolnews.com	urumqi	google news	4	6	4
	Ukraine					
	ukranews.com	kashgar	google news	1	1	6
8		urumqi ·	bing news	1	2	3
			google news	1	2	4
		xinjiang	bing news	1	1	4
	Bangladesh					
9	tbsnews.net	xinjiang terrorism	google news	1	5	1
	Zambia					
10	iol.co.za	urumqi	bing news	1	2	1
			google news	1	2	6

Using an exact title match between a corpus of Chinese state media articles and article titles in our dataset, we find that reposted content frequently features in our collected search results across Google Search, Google News, Bing Search, and Bing News for keywords related to Xinjiang. The sources reposting Chinese state media content are diverse – 19 different news outlets from 16 different countries for a total of 65 unique headlines. Including these sources in our analysis would increase the total number of Chinese state media occurrences in search results by nearly 10%. In Table 2, we detail the top ten sources of syndicated content from Chinese state media. As is evident, some outlets are more prolific than others. For example, Helsinki Times, an English-language Finnish news outlet that has a separate section dedicated to reposting China Daily, appeared in the top search results for the term "Xinjiang debunked" on 119 of 120 days of web search (Figure 14). One article from the Helsinki Times that appears in our top web search results nearly every day, titled "Witnessing the real Xinjiang, foreign diplomats debunk lies," seeks to discredit conventional wisdom about what is happening in Xinjiang and to bolster Beijing's narratives.⁴²

FIGURE 14

Fri Sat HELSINKI TIMES Ö HELSINKI -2°C 4 4 4 4 FINLAND light snow 1°C 1°C -3°C 1°C -3°C -5°C -5°C CULTURE CHINA NEWS FINLAND LIFESTYLE EAT AND DRINK THEMES BUSINESS COLUMNS WORLD 中文 OVID - 19 ESS TRAVEL + 🖧 = 🔐 COVID-Witnessing the real Xinjiang, foreign This page is provided in cooperation with People's Daily Online. diplomats debunk lies 09 APRIL 2021 people.cn CHINA NEWS ZONE HELSINKI TIMES does not exercise PREVIOUS ARTICLE is not responsible for the topics and How Le Monde pretended a French reporter didn't exist Lies are stopped by truths content of this section TOOLS -- From March 30 to April 2, a delegation comprising Vladimir Noroy, Daily secretary-general of the Shanghai Cooperation Organization, and PRIN more than 30 diplomats from some 21 countries, visited Xinijang, E-mail TYPOGRAPHY -- They talked with locals from various sectors of society, visited local MEDIUM schools and enterprises, and debunked rumors about Xinjiang, a DEFAULT region whose image is frequently distorted by some Western READING MO politicians, organizations and media. SHARE THIS -- From 2014 to 2019, the GDP of Xinijang increased from 919,59 billion yuan to 1,36 trillion MORE NEWS ABOUT CHINA 😏 f yuan, with an average annual growth rate of 7.2 percent. Under the current standards, 3.06 A 500-million-dollar business! America's million rural residents in the region have been lifted out of poverty, 3,666 poverty-stricken state-sponsored anti-China propaganda villages have withdrawn from the poverty list, and all 35 poverty-stricken counties are now Disinformation on China in full bloom in off the poverty list. US URUMQI, April 3 (Xinhua) - Savoring a piece of naan sprinkled with spices and chatting with a Why China is still attractive to foreign naan master at the International Grand Bazaar in Urumqi, Vladimir Norov was amazed at how investors this bread-like staple food has boosted the development of northwest China's Xinjiang Uvgur Commentary: Beijing Paralympics set to Autonomous Region

A Helsinki Times article that appeared on web search results on 119 of 120 days monitored

Another outlet that appears in our dataset for syndicated reposting of Chinese state media is the Big News Network. In 2020, the Big News Network was a focus of a broader investigation by the EU DisinfoLab into a disinformation apparatus tied to India that targeted the European Union and the United Nations.⁴³ This outlet shared at least 28 different state media articles that appeared in our top search results for "Xinjiang terrorism," "Urumqi," "Kashgar," "Xinjiang debunked," and "Xinjiang" on a total of 60 of 120 days of observation.

The syndication network for Chinese state media also often cuts across several outlets that may repost the same stories: in one case, a narrative about one Uyghur family's generational prosperity in Xinjiang appeared in our top search results for the term "Urumqi" from four different news outlets (Figure 15).

FIGURE 15

Four different outlets shared the same Chinese state content



State-affiliated influencers increase the reach of Beijing-backed content by at least 27% on YouTube

Content hosting agreements are not the only place where state-backed media may flourish. Specific to YouTube. Chinese state narratives are also spread by influencers who create content about topics of interest to the Chinese government, including Xinjiang and the COVID-19 pandemic. The Chinese government has leveraged the content created by these influencers as part of a global propaganda push by the state to counter critical reporting from authoritative news outlets.44 Beijing's relationships with influencers have an amplification effect similar to syndication but may be even more difficult to identify because these relationships, as well as possible agreements formalizing them, are generally opague.

Often, these influencers share similar characteristics: they have a Western upbringing but have lived in China, and they frequently portray themselves as independent. Despite the influencers' professed independence, their content regularly aligns with state narratives and is frequently shared by Chinese state officials on social media. They may also appear on Chinese state-backed news outlets as correspondents.

In 2021, the Australian Strategic Policy Institute (ASPI) published research into a network of Beijing-friendly influencers on YouTube, with a specific focus on channels producing content about Xinjiang.⁴⁵ The study identified influencers based on how frequently Chinese state media or officials reposted their content, as well as the influencers' appearances on or contributions to state media channels. Figure 16 shows an example of one channel identified in the ASPI report, including a screenshot of a video and the "About" page, which describes a "Dad and Son, from England but kickin' it in China." This channel has more than 40 million views.

FIGURE 16

An example of Beijing-linked influencer content (left) and account details (right) on YouTube



ASPI found that influencer content generally relied on two approaches to challenging established coverage of Xinjiang, including: (1) promoting a positive and vibrant image of life in the region by highlighting local customs and culture; and (2) calling out "Western injustices" and biases to counter allegations of forced labor and detention centers. These attempts at reframing narratives of Xinjiang are similar to those found across search results in our dataset, yet due to the nature of veiled connections to Beijing, they were undetected in our primary analysis. Drawing on the list of influencers compiled in the ASPI report, we found that 16 accounts with Beijing connections frequently appear in YouTube search results across both Xinjiang and COVID-19 terms (Figure 17).

FIGURE 17

Accounting for Beijing-linked influencers in top YouTube search results increases the number of total state media hits by 27%



If we account for these influencers in our findings, they increase the total number of state media hits across YouTube's top search results by 27%, from 3,568 to 4,522 videos with known state affiliations. It is likely that the

network documented by ASPI still represents an undercount of Chinese-backed influencers across YouTube, given the difficulty of identifying these relationships.

RESOURCES AND INTENT

These findings highlight Beijing's unique edge in seeding the internet with their preferred narratives. A key driver of this information power lies in resources - state media can produce content about strategic topics beyond the constraints of newsworthiness.⁴⁶ The resource advantage has three consequences. First, it enables a consistent stream of state-backed narratives to flow into the broader information environment through state websites, where it can be amplified by state actors on social media. Second, the continuous publication of content optimizes performance, as a number of search engine algorithms prioritize freshness. Independent media are beholden to news cycles that are defined by current events, public attention, and resource availability. Even when credible independent coverage does not lose relevance for a result, it may become stale and thus outperformed by a steady stream of state media content. Finally, the resources of state media outlets allow them to operate without paywalls and without dependence on subscriptions. This may attract more readers than content that requires a paid subscription.

A matter that is less clear, though equally relevant to understanding Chinese state media rank in search results, is that of intent. Whether Chinese state outlets proactively manipulate search engine vulnerabilities to dominate search results for strategic terms, or whether the remarkable performance of their content is a byproduct of volume and freshness, remains unknown. However, while this report has not established that Chinese state media make deliberate efforts to capture strategic search terms, the consequences for audiences are largely unchanged. Propaganda and state narratives, by design, are imbued with political intent. Even if prominence among search results is an unintentional consequence of state domains' resource advantage and strategic use of language, the negative impact on the integrity of search environments remains.

PART 4. MOVING TOWARD SOLUTIONS

CURRENT AND FUTURE RESEARCH

Recent scholarship has explored a range of ways in which search engines contribute to the pollution of the information environment, whether due to manipulation of results by malign actors or general algorithmic performance. The most relevant to this report is research about data voids, which are search engine vulnerabilities shaped by missing data, algorithmic logic, and media manipulation. Defined by Michael Golebiewski and danah boyd as search environments in which results are "limited, nonexistent, or deeply problematic," data voids present opportunities for manipulation through search engine optimization.47 Golebiewski and boyd define different types of data voids. Breaking news data voids, for example, occur when users search for news-related keywords before journalistic coverage catches up.48 These data voids are shaped by an information supply chain problem that is temporary. Another type of data void manipulation involves fragmented concepts - a way to segment users who are searching for the same concept (but through different frames) into distinct information environments.⁴⁹ For example, users searching for "social justice" are led to a far more liberal well of information than those searching for "wokeness" or "cancel culture," the terms more commonly used by conservative outlets. The resulting information environment is related to what Eli Pariser calls a filter bubble – a state of information isolation that arises from content personalization.⁵⁰

There has also been significant progress in investigating intentional manipulation of search engine results, largely focusing on content

generated by nonstate actors - extremist groups, conspiracy theorists, hate groups, and junk news outlets among those with political or social motivations, as well as commercial actors utilizing the rapidly growing search engine optimization (SEO) industry for financial gain. Samantha Bradshaw, for example, conducted a three-year study of how junk news websites optimize content discoverability on Google Search through basic SEO and marketing strategies.⁵¹ Golebiewski and boyd's seminal research on data voids illuminated campaigns by white nationalists, fringe political groups, and conspiracy theorists to use search engines as tools to amplify problematic content perhaps exemplified most tragically by the white supremacist Dylan Roof, who, after murdering nine African Americans in a church in South Carolina, stated that he had been radicalized after Googling "black on white crime."52 Other researchers have explored how far-right groups specifically target DuckDuckGo to spread politically motivated content, and how the quality of results for contested topics differs between search engines like Google and Bing.53

A number of organizations have built on this research to expand our understanding of search engines' role in information consumption, the risks of limited or low-quality data, and pathways to stronger information integrity. In September 2020, in response to problematic information flows surrounding the COVID-19 pandemic, First Draft and the University of Sheffield introduced a framework for monitoring the demand and supply of information related to public health, based on signals like search trends and fact checking reports.⁵⁴ The project introduced the idea of data deficits, situations defined by high interest in a topic and low availability of credible

information about the topic, as a way to capture mismatches in the information supply chain. Also in 2020, researchers at Harvard's Berkman Klein Center conducted a project to identify the quantifiable dimensions of data voids, and to create a framework for understanding the harms posed by poor search performance.⁵⁵

Companies running search engines have also signaled strong awareness of the power and responsibilities of their tools in information infrastructure. In February 2021, Google unrolled a Search feature that provides users with context about the results they see. The feature, in beta at the time of writing, is essentially an "about this result" pop-up that provides information about when a site was first indexed, whether the result is an ad, whether connection to the site is secure, and, when available, a description of the site from Wikipedia.⁵⁶ In June 2021, Google unrolled a warning notice for search results that are rapidly changing and may not contain reliable sources — a move that addresses risks associated with breaking news events, when public interest in a topic may outpace journalistic coverage.⁵⁷ Google has also expanded the regional availability of its Questions Hub, a tool that identifies content gaps (including by asking users what questions they were unable to answer through Search) and helps coordinate efforts to fill them.

EVOLVING PLATFORM POLICIES RELATED TO STATE MEDIA

In the early months of 2022, Western technology companies, including search engines, enacted measures to limit the reach of Russian state media. Google removed Russian state-funded publishers from Google News results. Google-owned YouTube blocked Kremlin media globally, after initially banning RT and Sputnik in Ukraine and the EU. Microsoft announced it would remove RT news apps from the Windows app store and tweak Bing so that it only returns RT and Sputnik links "when a user clearly intends to navigate to those pages."58 These measures reflect the extraordinary nature of the crisis in Ukraine and the intense pressure technology companies faced from governments and the public to take action. Some of these changes will apply to Chinese state media, which are actively distorting narratives around another crisis – the one simmering in Xinjiang. Many will not.

Notable platform policies regarding state media

- Account labeling: Twitter, Facebook, and YouTube label state media accounts, including Russian and Chinese outlets. YouTube began the practice of providing publisher context for statefunded media in 2018, and Twitter and Facebook followed in 2020.⁵⁹
- Source context: In 2021, Google rolled out a beta feature that provides users information about each search result, including a brief description of the source, the web page's language, and whether the result is an ad.⁶⁰

Notable policy updates

- Demonetization: Google, YouTube, Twitter, Meta, and Microsoft demonetized Russian state media.⁶¹
- Targeted deranking or deamplification: Google and Microsoft enacted policies to de-rank Russian state media content in search results.62 (Google has previously claimed to derank Russian state media, albeit not in an official company statement.⁶³) YouTube blocked access to channels linked to Russian state media, initially in Europe and then globally.64 Twitter implemented a new policy that limits the recommendation of state-run accounts from a country that "blocks or limits access to online services within their state, undercutting the public's voice and ability to freely access information, but continues to use online services for their own communications."65 Some have interpreted that this new policy may also include China.
- Extension of state media labels: Twitter has attached a content warning label to any tweet sharing links from state-affiliated media from a number of countries, including Russia and China, regardless of the account sharing the content.⁶⁶

RECOMMENDATIONS

While these efforts undoubtedly contribute to information integrity, much remains to be done to address the vulnerabilities of search engines to surfacing state-backed or manipulated content – and to protect the role of search engines as an indispensable part of today's information infrastructure. The problems detailed in this report are nots for search engines to address alone. Rather, responsibilities should be shared by content creators and search engines:

- Address hosting, reposting, and syndication: As this research has suggested, content hosting and syndication agreements between international news outlets and Chinese state media are a potentially significant avenue for the proliferation of Beijing's narratives, including misleading and conspiratorial content. Alarmingly, outlets that host content from agencies like Xinhua do not always clearly label the original source. Authoritative outlets should reconsider their agreements with state media that do not have appropriate controls to ensure editorial independence. At a minimum, outlets should enhance disclosure and labels to better inform audiences about sources of information. Search engines, in turn, can consider ways to add labels to syndicated content, or to clearly link to original sources.
 - Enhance labeling and context: Some search engines have already begun providing information about the domains that populate results — for example, Google is testing an "about this result" feature in Search. To enable users to better navigate their search for information, search engines should formalize the practice of providing context to users and consider including clear labels for state domains.
- Provide clarity and transparency around ranking decisions: Search engines have a responsibility to provide more information about ranking decisions surrounding state media. In 2017, Alphabet chairman Eric Schmidt announced that RT and Sputnik were to be deranked by Google in an effort to curb the spread of false information by the Russian

state outlets.⁶⁷ Details about the decision and its implementation have been sparse, and the issue of Russian state outlets' performance in search results resurfaced again in March 2022, following Russia's invasion of Ukraine, when the EU requested that Google cease returning RT and Sputnik results for users within the EU.⁶⁸ As companies consider such measures about rank and inclusion of state domains, it is imperative that they provide information on how these decisions are made — including in how authoritativeness and quality of state media are assessed, and whether ranking decisions regarding one state's domains have implications for others'.

- Provide notices about result quality: Sometimes, low quantities of relevant content - or high-quality content - mean that an engine is unable to deliver reliable results for a query. Representatives of major search engines have signaled awareness of this issue, and as previously noted, Google has unrolled a notice for breaking news events, when rapidly changing news presents a problem for the engine's ability to deliver reliable results. This report recommends expanding the use of such notices to other situations where the quality of results is in guestion, including contested terms or topics that are a battleground for search dominance.
- Increase public understanding of how search engines rank and deliver information: More broadly, search engines could contribute to public education about how engines find, rank, and surface content. User trust in engines like Google remains high, but there is little evidence that users have a strong understanding of the factors that determine results.⁶⁹ Digital and media literacy are often touted as ways to make technology users less vulnerable to manipulation, and this certainly rings true in addressing audiences' susceptibilities to state narratives and conspiratorial content. Establishing discrete areas of literacy is essential to this solution, and should include user understanding of how search engines operate and how both search algorithms and users' own online behavior influence results.

- Link search results across related terms: Small but meaningful changes to spelling or certain abbreviations may yield vastly different search results despite referencing the same information. For example, recent research around the conflict in Ukraine found that different content surfaced in Google News when the capital city of Ukraine was spelled "Kiev" (the Russian-rooted spelling) vs. "Kyiv" (the Ukrainian-rooted spelling). The former was more likely to surface Russian state-backed content (before Google decided not to surface it), although both names refer to the same geographic location. Similarly, the terms "Luhansk People's Republic" and "Donetsk People's Republic" did not return a large number of Russian state media, but their abbreviations "LPR" and "DPR" overwhelminalv returned state-backed content.⁷⁰ Linking highly related terms so that they return similar content can help ensure that more authoritative information surfaces in search results, particularly in times of crisis.
- Collaborate and share information with other search engines: Major search engines have collaborated in the past to improve the performance of their technologies. However, responses to state media content and instances of keyword capture have been mixed. Collaboration across the sector could identify best approaches, policies, and technical solutions. The goal should not necessarily be the formation of a new institution or formal mechanism, but rather more substantive information exchange on how to address common vulnerabilities.
- Determine how audiences are searching for information: There are many ways to harness strategic communication strategies to counter the proliferation of state-backed narratives, but those strategies are largely beyond the scope of this research. This report recommends one general course of action for content creators who cover issues related to or of importance to Beijing, including research organizations, aovernment officials. and activists: developing a deeper understanding of how audiences are searching for their stories and

maintaining awareness of efforts to subvert them. Incorporating audience behavior and narrative challenges into tagging, keyword use, and broader communication strategies can go a long way in countering misleading stories.

Users are also an important part of solutions, as their engagement with content and search engine features contributes to the evolution of these tools. However, this report refrains from giving specific recommendations to users at this stage, beyond the support for tailored digital literacy outlined above. User behavior unfolds within search environments controlled by technology companies and contested by the actors that vie for search dominance, and it is imperative to address the issues within these systems first.

These recommendations and findings are based on research into a specific set of state actors and their capture of strategic key terms. As researchers consider the role of search engines in the digital public sphere, we hope to see more inquiries into the role of state actors in other contested search environments, as well as broader inquiries into the vulnerabilities of search results to manipulation by political actors. Future areas of research should consider the role of language and location - two key factors in the personalization of search engine results. The performance of search within social media platforms is another critical area of study, particularly as more users treat the search bars of platforms as search engines.⁷¹ This report also identified gaps in available information about syndication networks and their role in amplifying state narratives, which warrant further analysis. Finally, we hope to see updated research about user trust in search engines, and the role of search within broader information consumption.

APPENDIX

Search engine primer: Purpose and rank factors

Search Engine	Purpose	Main factors that determine search results	
Google Search	"Google Search puts the world's information at your fingertips, helping people find helpful results for billions of queries every day."	 Meaning of query Relevance of webpages Quality of content Usability of webpages Context and settings 	
Google News	"To help everyone understand the world by connecting people with high-quality news from a variety of perspectives."	 Relevance Interests Location Prominence Authoritativeness Freshness Usability 	
Bing Search	"To connect users with the most relevant search results from the web—providing easy access to quality content produced by web publishers."	 Relevance Quality and credibility User engagement Freshness 	
Bing News	"Find and read up-to-the-minute and in-depth coverage of world events."	LocationPage load time	
YouTube Search	"Sorting through a vast number of videos to find the most relevant and useful results to your search query, and presenting them in a way that helps you find what you're looking for."	 Relevance Engagement Quality YouTube Search prioritizes different factors in areas like music or entertainment (where relevance, freshness, and popularity define results) and areas like news, politics, and health (where authoritativeness of content plays a significant role). 	

SOURCES: Google; Microsoft; Microsoft Bing; YouTube⁷²

REFERENCES

1 "Source: The target of the Ukrainian missile could be Vladimir Putin's plane," RT, July 17, 2014, <u>https://russian.rt.com/article/41332</u>.

2 "2021 Trust Barometer," Edelman, January 2021, <u>https://www.edelman.com/trust/2021-trust-barometer</u>.

3 Ibid.

4 Jessica Brandt and Valerie Wirtschafter, "The surprising performance of Kremlin propaganda on Google News," The Brookings Institution, March 1, 2022, <u>https://www.brookings.edu/techstream/the-surprising-performance-of-kremlin-propaganda-on-google-news/</u>.

5 Elen Aghekyan and Bret Schafer, "Deep in the Data Void: China's COVID-19 Disinformation Dominates Search Engine Results," Alliance for Securing Democracy, October 5, 2021, <u>https://securingdemocracy.gmfus.org/data-void-china-covid-disinformation/</u>.

6 "2021 Trust Barometer," Edelman.

7 Jutta Haider and Olof Sundin, *Invisible Search and Online Search Engines: The Ubiquity of Search in Everyday Life* (London / New York: Routledge, 2019).

8 Ramishah Maruf, "These four words are helping spread vaccine misinformation," CNN, September 20, 2021, <u>https://www.cnn.com/2021/09/19/media/reliable-sources-covid-research/index.html</u>.

9 Matt Schrader, "Friends and Enemies: A Framework for Understanding Chinese Political Interference in Democratic Countries," Alliance for Securing Democracy, April 22, 2020, <u>https://</u> <u>securingdemocracy.gmfus.org/friends-and-enemies-a-framework-for-understanding-chinese-political-interference-in-democratic-countries/</u>.

10 "Xi Jinping sent a letter to congratulate China International Television (China Global Television Network)," Xinhua, December 31, 2016, <u>http://www.xinhuanet.com/politics/2016-12/31/c_1120226953.htm</u>.

11 "刘奇葆在中央媒体驻外机构外宣工作座谈会上强调 为党的十九大胜利召开营造良好国际舆 论环境" [Liu Qibao emphasized the work of media abroad to create a positive image ahead of 19th National Congress], Xinhua, September 21, 2017, <u>http://www.xinhuanet.com/politics/2017-</u>09/21/c_1121704621.htm.

12 Renee DiResta, Carly Miller, Vanessa Molter, John Pomfret, and Glen Tiffert, "Telling China's Story: The Chinese Communist Party's Campaign to Shape Global Narratives," (Stanford, CA: Stanford Internet Observatory Cyber Policy Center and Hoover Institution, July 21, 2020, <u>https://www.hoover.org/research/telling-chinas-story-chinese-communist-partys-campaign-shape-global-narratives</u>.

13 Sarah Cook, "Beijing's Global Megaphone," (Washington, DC: Freedom House, 2020), <u>https://freedomhouse.org/report/special-report/2020/beijings-global-megaphone</u>.

14 "Who We Are," China Plus, February 16, 2017, <u>http://chinaplus.cri.cn/aboutus/aboutcri/62/20170216/393.html</u>.

15 Koh Gui Qing and John Shiffman, "Beijing's covert radio network airs China-friendly news across Washington, and the world," Reuters, November 2, 2015, <u>https://www.reuters.com/investigates/special-report/china-radio/</u>.

16 Keir Simmons, Laura Saravia, and Yuliya Talmazan, "China guilty of genocide, crimes against humanity against Uyghurs, watchdog finds," NBC News, December 9, 2021, <u>https://www.nbcnews.com/news/china/china-guilty-genocide-crimes-humanity-uyghurs-watchdog-finds-rcna8157</u>.

17 "Xinjiang genocide' claim biggest lie of the century: Chinese delegation," Xinhua, March 20, 2021, <u>http://www.xinhuanet.com/english/2021-03/20/c_139822760.htm</u>.

18 "Genocide in Xinjiang is a Lie of Century," Chinese Ministry of Foreign Affairs, February 8, 2021, <u>https://www.mfa.gov.cn/ce/ceun/eng/zgyw/t1852721.htm</u>.

19 Francesca Tripodi, "Searching for Alternative Facts," (New York: Data & Society, May 2018), https://datasociety.net/wp-content/uploads/2018/05/Data_Society_Searching-for-Alternative-Facts. pdf.

20 Elen Aghekyan and Bret Schafer, "Deep in the Data Void."

21 ASD Social Data Search on Hamilton Database: search query 'covid' from January 1, 2021 to December 31, 2021, Alliance for Securing Democracy, <u>https://securingdemocracy.gmfus.org/asd-social-data-search/?q=covid*&product=%28product+eq+%27Hamilton%27+or+product2+eq+%27Hamilton%27%29&start-date=2021-01&end-date=2021-12-31.</u>

22 Our dataset includes observations across 120 days. Data was collected incognito to minimize the potential for personalized search results.

23 Numerous user behavior studies have concluded that users heavily favor the first page of search results, and particularly links "above the fold" of a results page. For an overview of major studies, see Yvonne Kammerer and Peter Gerjets, "How Search Engine Users Evaluate and Select Web Search Results: The Impact of the Search Engine Interface on Credibility Assessments," *Web Search Engine Research*, ed. Dirk Lewandowski (Bingley, U.K.: Emerald Group Publishing Limited, 2012), 251-279.

24 The number of observations provided on the first page varies across search engines and, on occasion, by day. For YouTube, we included the top 10 search results. Bing News returns no results for the term "EVALI virologist."

25 "Hamilton Monitored Accounts on Twitter," Alliance for Securing Democracy, <u>https://securingdemocracy.gmfus.org/hamilton-monitored-accounts-on-twitter/#china</u>.

26 For details about Google, Bing, and YouTube search processes, see Appendix.

Across all search types, state-backed media search results are as follows: Xinjiang debunked (3,158); Xinjiang terrorism (1,921); Adrian Zenz (1,241); Kashgar (1,186); Urumqi (810); Xinjiang (676); and Uyghur (13).

28 Stuart Thompson, "Fed Up With Google, Conspiracy Theorists Turn to DuckDuckGo," *The New York Times*, February 23, 2022, <u>https://www.nytimes.com/2022/02/23/technology/duckduckgo-conspiracy-theories.htm</u>; Daniel Bush and Alex Zaheer, "Bing's Top Search Results Contain an Alarming Amount of Disinformation," Stanford Internet Observatory, December 17, 2019, <u>https://cyber.fsi.stanford.edu/io/news/bing-search-disinformation</u>.

29 As previously noted, on platforms without a clearly defined first page, "top results" are defined as the top 10 search results for a given query.

30 Francesca Tripodi, "Searching for Alternative Facts."

31 Francesca Tripodi, "Don't expect breaking up Google and Facebook to solve our information woes," Nieman Lab, December 2020, <u>https://www.niemanlab.org/2020/12/dont-expect-breaking-up-google-and-facebook-to-solve-our-information-woes/</u>.

32 Ibid.

33 Across all search types, state-backed media search results are as follows: Unit 731 (800); Fort Detrick (718); EVALI coronavirus (241); Huanan seafood market (239); and EVALI virologist (188).

34 For example, see Sundar Pichai, "Coronavirus: How we're helping," The Keyword, Google, March 6, 2020, <u>https://blog.google/inside-google/company-announcements/coronavirus-covid19-response/and "COVID-19 medical misinformation policy," YouTube Help, https://support.google.com/youtube/answer/9891785?hl=en.</u>

35 Michael Golebiewski and danah boyd, "Data Voids: Where Missing Data Can Easily Be Exploited," (New York: Data & Society, November 2019), <u>https://datasociety.net/wp-content/uploads/2019/11/</u> <u>Data-Voids-2.0-Final.pdf</u>. For more information about the literature surrounding data voids, see Part 4.

36 Elen Aghekyan and Bret Schafer, "Deep in the Data Void."

37 CGTN, "How terrifying is the history of U.S. Fort Detrick lab?," YouTube, July 28, 2021, <u>https://www.youtube.com/watch?v=leZV55WvOmM</u>.

38 John Dotson, "Xinhua Infiltrates Western Electronic Media, Part 2: Relationships with News Agencies and Distribution Services," (Washington, DC: China Brief, The Jamestown Foundation, August 17, 2021), <u>https://jamestown.org/program/xinhua-infiltrates-western-electronic-media-part-2-relationships-with-news-agencies-and-distribution-services/</u>.

39 "Xinhua News Agency, ANSA to jointly launch Xinhua Italian Service," Xinhua, March 23, 2019, <u>http://www.xinhuanet.com/english/2019-03/23/c_137917481.htm</u>.

40 Xinhua, "China, Nigeria sign MoU on information exchange," The State Council, The People's Republic of China, August 12, 2017, <u>http://english.www.gov.cn/news/international_exchanges/2017/08/12/content_281475784944910.htm</u>.

41 Raksha Kumar, "How China uses the news media as a weapon in its propaganda war against the West," Reuters Institute for the Study of Journalism at the University of Oxford University, November 2, 2021, <u>https://reutersinstitute.politics.ox.ac.uk/news/how-china-uses-news-media-weapon-its-propaganda-war-against-west</u>.

42 Xinhua, "Witnessing the real Xinjiang, foreign diplomats debunk lies," *Helsinki Times*, April 9, 2021, <u>https://www.helsinkitimes.fi/china-news/19019-witnessing-the-real-xinjiang-foreign-diplomats-debunk-lies.html</u>.

43 Gary Machado, Alexandre Alaphilippe, Roman Adamczyk, and Antoine Grégoire, "Indian Chronicles: deep dive into a 15-year operation targeting the EU and UN to serve Indian interests," EU Disinfo Lab, December 9, 2020, <u>https://www.disinfo.eu/publications/indian-chronicles-deep-dive-into-a-15-year-operation-targeting-the-eu-and-un-to-serve-indian-interests/</u>.

44 Amanda Seitz, Eric Tucker, and Mike Catalini, "How China's TikTok, Facebook influencers push propaganda," The Associated Press, March 30, 2022, <u>https://apnews.com/article/china-tiktok-facebook-influencers-propaganda-81388bca676c560e02a1b493ea9d6760</u>.

45 Fergus Ryan, Ariel Bogle, Nathan Ruser, Albert Zhang, and Daria Impiombato, "Borrowing mouths to speak on Xinjiang," (Barton, Australia: Australian Strategic Policy Institute, December 10, 2021), https://www.aspi.org.au/report/borrowing-mouths-speak-xinjiang.

46 Sarah Cook, "Beijing's Global Megaphone."

47 Michael Golebiewski and danah boyd, "Data Voids," 2.

48 Ibid., 16.

49 Ibid., 30.

50 Eli Pariser, The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think (New York: Penguin Press, 2011).

51 "Junk news websites" are defined as websites that "do not employ professional journalistic standards, report on conspiracy theory, counterfeit professional news brands, and mask partisan commentary as news." Samantha Bradshaw, "Disinformation optimised: gaming search engine algorithms to amplify junk news," *Internet Policy Review* 8, no. 4 (December 31, 2019), <u>https://policyreview.info/articles/analysis/disinformation-optimised-gaming-search-engine-algorithms-amplify-junk-news</u>.

52 Michael Golebiewski and danah boyd, "Data Voids," 25–26.

53 For example, see Daniel Bush and Alex Zaheer, "Bing's Top Search Results Contain an Alarming Amount of Disinformation."

54 Tommy Shane and Pedro Noel, "Data deficits: why we need to monitor the demand and supply of information in real time," First Draft, September 28, 2020, <u>https://firstdraftnews.org/long-form-article/data-deficits/</u>.

55 EC, Rafiq Copeland, Jenny Fan, and Tanay Jaeel, "Filling the Data Void," 2020, Assembly at the Berkman Klein Center, <u>https://datavoids.2020.bkmla.org/</u>.

56 JK Kearns, "A Quick Way to Learn More about Your Search Results," The Keyword, Google, February 1, 2021, <u>https://blog.google/products/search/about-search-results/</u>.

57 Danny Sullivan, "A New Notice in Search for Rapidly Evolving Results," The Keyword, Google, June 25, 2021, <u>https://blog.google/products/search/new-notice-search-rapidly-evolving-results/</u>.

58 Brad Smith, "Digital technology and the war in Ukraine," Microsoft on the Issues, Microsoft, February 28, 2022, <u>https://blogs.microsoft.com/on-the-issues/2022/02/28/ukraine-russia-digital-war-cyberattacks/</u>.

59 Geoff Samek, "Greater transparency for users around news broadcasters," YouTube Official Blog, February 2, 2018, <u>https://blog.youtube/news-and-events/greater-transparency-for-users-around/;</u> Nathaniel Gleicher, "Labeling State-Controlled Media On Facebook," Meta, June 4, 2020, <u>https://about. fb.com/news/2020/06/labeling-state-controlled-media/;</u> "About government and state-affiliated media account labels on Twitter," Help Center, Twitter, <u>https://help.twitter.com/en/rules-and-policies/</u> <u>state-affiliated</u>. 60 JK Kearns, "A quick way to learn more about your search results."

61 Shannon Bond, "Facebook, Google and Twitter limit ads over Russia's invasion of Ukraine," NPR, February 27, 2022, <u>www.npr.org/2022/02/26/1083291122/russia-ukraine-facebook-google-youtube-twitter</u>; Smith, "Digital technology and the war in Ukraine."

62 Brad Smith, "Digital technology and the war in Ukraine;" Kent Walker, "Helping Ukraine," The Keyword, Google, March 4, 2022, <u>https://blog.google/inside-google/company-announcements/helping-ukraine/</u>.

63 "Google to 'derank' Russia Today and Sputnik," BBC News, November 21, 2017, <u>https://www.bbc.</u> <u>com/news/technology-42065644</u>.

64 YouTubeInsider (@YouTubeInsider), Twitter, March 11, 2022, <u>https://twitter.com/YouTubeInsider/status/1502335030168899595?s=20&t=wQTscavci4Gw6coAhzz7PQ</u>.

65 Sinéad McSweeney, "Our ongoing approach to the war in Ukraine," Blog, Twitter, March 16, 2022, <u>https://blog.twitter.com/en_us/topics/company/2022/our-ongoing-approach-to-the-war-in-ukraine</u>.

66 Ibid.

67 "Google to 'derank' Russia Today and Sputnik," BBC News.

68 Gerrit De Vynck, "E.U. Sanctions Demand Google Block Russian State Media from Search Results," *The Washington Post*, March 9, 2022, <u>https://www.washingtonpost.com/technology/2022/03/09/eu-google-sanctions/</u>.

69 Francesca Tripodi, "Searching for Alternative Facts."

70 Jessica Brandt and Valerie Wirtschafter, "The surprising performance of Kremlin propaganda on Google News."

71 Tommy Shane, "People Are Using Facebook and Instagram as Search Engines. During a Pandemic, That's Dangerous," Nieman Lab, August 10, 2020, <u>https://www.niemanlab.org/2020/08/people-are-using-facebook-and-instagram-as-search-engines-during-a-pandemic-thats-dangerous/</u>.

72 "How Search Works," The Keyword, Google, <u>https://blog.google/products/search/how-search-works/;</u> "How Search algorithms work," Google, <u>https://www.google.com/search/howsearchworks/</u> algorithms/; "How news works on Google," Google, <u>https://newsinitiative.withgoogle.com/</u> <u>hownewsworks/mission/;</u> "Surfacing useful and relevant content," Google, <u>https://newsinitiative.</u> <u>withgoogle.com/hownewsworks/approach/surfacing-useful-and-relevant-content/;</u> "How Bing delivers search results," Support, Microsoft, <u>https://support.microsoft.com/en-au/topic/how-bingdelivers-search-results-d18fc815-ac37-4723-bc67-9229ce3eb6a3;</u> "Bing Webmaster Tools," Microsoft Bing, <u>https://www.bing.com/webmasters/help/webmaster-guidelines-30fba23a;</u> "YouTube Search," YouTube, <u>https://www.youtube.com/howyoutubeworks/product-features/search/</u>.

ABOUT THE AUTHORS

Jessica Brandt is policy director for the Artificial Intelligence and Emerging Technology Initiative at the Brookings Institution and a fellow in the Foreign Policy program's Center for Security, Strategy, and Technology. Her research interests focus on foreign interference, disinformation, digital authoritarianism, and the implications of emerging technologies for liberal democracies. Jessica was previously head of policy and research for the Alliance for Securing Democracy and a senior fellow at the German Marshall Fund of the United States, a fellow in the Foreign Policy program at the Brookings Institution, special adviser to the president of the Brookings Institution, and an international and global affairs fellow at the Belfer Center for Science and International Affairs at Harvard University.

Bret Schafer is a senior fellow at the Alliance for Securing Democracy (ASD), a program at the German Marshall Fund of the United States, and he leads ASD's information manipulation team. Bret is the creator and manager of Hamilton 2.0. an online open-source dashboard tracking the outputs of Russian, Chinese, and Iranian state media outlets, diplomats, and government officials. As an expert in computational propaganda, state-backed information operations, and tech regulation, he has spoken at conferences around the globe and advised governments and numerous international organizations. Prior to joining ASD, Bret spent more than 10 years in the television and film industry.

Elen Aghekyan is a junior fellow at the Alliance for Securing Democracy and an M.A. candidate at the Jackson Institute for Global Affairs at Yale University. At Yale, Elen focuses on the intersection of disinformation and democratic resilience and serves as a fellow at the Kerry Initiative, an interdisciplinary program to tackle pressing global challenges through teaching, research, and international dialogue. Elen began her career at the democracy watchdog Freedom House, where she was a research analyst for the organization's annual reports on political rights and civil liberties, with a regional focus on Europe and Eurasia.

Valerie Wirtschafter is a senior data analyst in the Artificial Intelligence and Emerging Technologies Initiative at the Brookings Institution. At Brookings, her research explores the effects of contemporary media, including social media and podcasting, on political behavior around the world using novel methodological tools, including machine learning, natural language processing and survey experiments. In fall 2021, she received her Ph.D. from the Department of Political Science at the University of California, Los Angeles (UCLA), where she developed and taught a course on international politics in the digital age. Prior to UCLA, she worked as a researcher in the Women and Foreign Policy program and Latin American Studies program at the Council on Foreign Relations and as a health and development consultant at Global Health Strategies.

Adya Danaditya is a research intern with the Artificial Intelligence and Emerging Technology Initiative at the Brookings Institution. He also teaches a class on ethics in emerging technologies as an adjunct instructor at Carnegie Mellon University, from which he recently graduated with a Master of Information Technology Strategy degree. Prior to his current appointments, he had six years of experience consulting for established corporations, nonprofits, and startups in Indonesia on their digital transformation efforts.

ACKNOWLEDGMENTS

The authors would like to thank Francesca Tripodi and Jenny Fan for sharing their insights and research, Chris Meserole for his invaluable feedback, Joseph Bodnar for his contributions to editing, and Rachel Slattery for layout. The authors are also grateful to NewsGuard, which provided data to support our analysis.

The Brookings Institution is a nonprofit organization devoted to independent research and policy solutions. Its mission is to conduct high-quality, independent research and, based on that research, to provide innovative, practical recommendations for policymakers and the public. The conclusions and recommendations of any Brookings publication are solely those of its author(s), and do not reflect the views of the Institution, its management, or its other scholars.

BROOKINGS

The Brookings Institution 1775 Massachusetts Ave., NW Washington, D.C. 20036 brookings.edu