INTERNET POLICY REVIEW
Journal on internet regulation

Volume 12 **Issue** 4

RESEARCH ARTICLE

OPEN ACCESS

PEER REVIEWED

A complicated picture: Media diversity in the case of Google's video search during the pandemic

Qun Wang *Fordham University*

DOI: https://doi.org/10.14763/2023.4.1728

Published: 30 November 2023
Received: 4 September 2022 **Accepted:** 18 January 2023

Funding: This study received funding support from Tow Center for Digital Journalism and Fordham University.
Competing Interests: The author has declared that no competing interests exist that have influenced the text.

Citation: Wang, Q. (2023). A complicated picture: Media diversity in the case of Google's video search during the pandemic. *Internet Policy Review*, *12*(4). https://doi.org/10.14763/2023.4.1728

Keywords: Platform governance, Google, Media diversity, Pandemic, Journalism

Abstract: With the rise of online video content, especially the growth of online video consumption during the early days of the COVID-19 pandemic, this study focuses on Google's video search as an important but understudied platform to explore Google's role in three media diversity areas — format-type diversity, source diversity, and structural-social diversity—in the online video landscape. Findings of this study reveal a complicated picture: on the one hand, Google can expose users to video content alternative to YouTube video, which involves a variety of media and non-media sources; on the other hand, the lack of diversity can be found in the studied areas. In addition, Google-source partnership and the limited presentation of minority sources at the time of a global crisis, this study argues, has profound implications for media diversity. While normative expectations of digital platforms' role in media diversity depend on different understandings of democracy models, media diversity is an important thread of the ongoing debate regarding platform governance. This study sheds light on one video platform, as a case updating our understanding of media diversity in the digital age.

Introduction

Media diversity is a framework that has been used to assess the media's role in serving a democratic society and to guide policymaking in order to ensure such a role. In the digital era, media diversity has gained increasing attention in understanding digital platforms' influence and governance. When diversity is considered, however, there have been ongoing debates about whether these digital platforms facilitate information access or produce filter bubbles or echo chambers, or whether they benefit or harm media diversity. These debates occur not only because media diversity has various dimensions as both a normative framework and a measuring tool, but also because people have different understandings about the models of democracy, which leads to distinct expectations of the role and function of digital platforms (Helberger, 2019).

Meanwhile, the effort of understanding the relationship between media diversity and digital platforms faces challenges since some concepts and metrics that were once used to assess media diversity in the pre-internet era are not adequate or relevant to study new developments unique to technologies and practices in the digital environment, from the role of algorithms to the use of various tools online. One of such new phenomena in the digital age is the rapid development of online video content as a major information genre. This study pays particular attention to Google's video search service against the backdrop of the rise of online video content in the past decade. SEO observers found that Google has given more priority to video content in its search results (The STAT Team, 2017). Such priority from the world's largest search engine and the most visited website worldwide would have important implications for media diversity. While existing studies about online video tend to focus on YouTube, Google's video hosting and sharing platform, few studies examined Google's video search, a platform with different focus and service from YouTube, and even fewer studied Google's video search results through the lens of media diversity. This study aims to make contributions to this understudied area using the COVID-19 pandemic — an event that has had an impact on almost everyone at a time when online video consumption increased greatly — as a unique time to study media diversity in the case of Google's video search.

This study provides the examination of Google's video search results through three media diversity areas: format-type diversity, source diversity, and structural-social diversity. The following sections first explain the development of online video, a brief history of Google Videos and why studying Google's video search, and key issues in media diversity studies in the pre-internet era and the digital era.

The development of online video

Video content is an important information type to serve the public's critical information needs, especially in times of crisis. For example, studies found surges in television news watching during times of national crisis, such as the 9/11 attacks and the 1990-91 Persian Gulf crisis (Althaus, 2002). With the development of broadband networks and the deepening of the rich media era, video content has been expanding online globally. In the US, the Pew Internet & American Life Project's first major report on online video in 2007 reported that 57% of online adults used the internet to watch or download video (Pew, 2007). That number grew to 78% in 2013 (Pew, 2013). In 2020, the digital video penetration reached nearly 84% in the US (Ceci, 2021). In Europe, online video subscription revenue reached €9.7 billion in 2020, up from 12 million in 2010 (Informitv, 2021). Outside these two most built-out areas, online video has increased worldwide in terms of viewing time and subscription (Bloom, 2022).

The demand for video content further increased during the COVID-19 pandemic, especially in early 2020. Studies have found that TV news viewership, streaming, and online watching all grew greatly in the early days of the pandemic (Adgate, 2020). Globally, online streaming increased over 12% in March 2020 alone (Beech, 2020). In the United States, users spent an average of eight hours a day streaming video content online during the stay-at-home time (Sadlier, 2020). Many of these hours were spent on COVID-related online news watching (Weissbrot, 2020).

In addition to traditional video content producers, digital platforms have become key players in the online video landscape. For example, Google acquired YouTube for US$1.6 billion in 2006; and in 2014, Mark Zuckerberg envisioned a Facebook that would be mostly video (Miners, 2014). These platforms' priority on video also spurred the investment in online video content in the media industry (Kalogeropoulos et al., 2016). A survey from 2016 showed that 79% of CEOs, editors, and digital leaders of the surveyed media organisations across 118 companies in different parts of the world said they planned to invest more in online video (Kalogeropoulos et al., 2016).

Google Videos and why study Google's video search

Google Videos is a search engine for video content, which can be found on Google's regular search page or at video.google.com. On the global online video market that was worth over US$7 billion as of 2021, Google is a leading actor (Skyquest, 2022). Google started its video-related service as early as January 2005,

when Google Video (Beta) was introduced to the public (Lenssen, 2005). In its early days, Google Video was primarily a video hosting service when Google invited users to upload video content to Google Video. At first, Google Video mainly indexed TV programmes and TV shows such as "Who Wants to Be a Millionaire", "Friends", "Saturday Night Live", and more. It did allow users to search, but only returned textual content with snapshots without a playback feature. At that time, other tech companies, such as Yahoo, MySpace, and Microsoft, also invested in the online video market, and Google Video had over 10% of the market share by the time Google acquired YouTube in 2006, while YouTube's market share was 4 times greater than Google Video (Baker, 2006).

Since the acquisition, Google Video's role shifted into a video search engine, which Google claimed was "the most comprehensive [video search engine] on the Web, containing millions of videos indexed and available for viewing" (Chitu, 2007, para. 3). The video search engine, according to Google, focused on its "ability to let people search videos from across the web, regardless of where those videos are hosted", although observers found that websites with .edu domain names appeared to be prioritised among Google Video search results at that time (Kincaid, 2011, para. 6). Today, Google's decision to keep both YouTube and Google Videos indicates the two platforms serve different purposes and are both important to Google as the entry point of the web used by billions of users worldwide.

While there is a large body of research that studies Google's various services, very few scholarly studies provide empirical research of Google's video search. Currently, video studies related to Google often focus on YouTube, but the video content returned by Google search and YouTube are different in many ways. For example, some search engine and SEO professionals pointed out that there are much more video searches on Google than YouTube; and video producers on YouTube also seek traffic from Google (Jarboe, 2019). YouTube search results and Google search results could be very different even with the same search queries (Enge, 2017). The composition of the sources differs between the two platforms as well (Krebs et al., 2021). Google's top video results were also found to focus more on informational videos while YouTube is more entertainment focused (DiSilvestro, 2017). YouTube users were also found to be more likely to use the platform for videos about an interest or hobby (Bump, 2023). In addition, it is costly to maintain a dedicated YouTube channel. A study from Pew Research Center found that about half of the audited local TV outlets did not have YouTube channels; and for those who did, many of their YouTube channels were inactive (Mitchell et al., 2014). This study also found that most media outlets in their sample hosted video content by

themselves and posted the videos on their own websites as opposed to having them on YouTube. These observed distinctions between Google's video search and YouTube and the history of Google Videos discussed above remind us that although owned by the same company, Google Videos and YouTube are two different platforms with different coverage and purposes, and more research efforts are needed to understand Google's video search as an understudied platform.

Media diversity

Media diversity in the pre-internet era

Media diversity is a framework adopted from the pre-internet era. Napoli (1999) pointed out, "diversity is a concept with multiple dimensions, means of assessment, and underlying assumptions" (p. 8). In democratic systems, while there is a general consensus on the normative value of the concept, the ways scholars and policymakers study and assess media diversity vary largely. From an analytical perspective, media diversity studies investigate a wide variety of dimensions and subdimensions. Traditionally, there is a distinction between the supply end and the demand end when media diversity is assessed. On the supply end, content providers and content itself are examined. One component of supply diversity is media ownership diversity that examines the number of parties that own media content and media outlets in the marketplace, the economic structures of these parties, and the diversity in their workplace composition and employment opportunities. This dimension is considered source, actor, or structural diversity that looks into "how the content was made and by whom" (Loecherbach et al., 2020, p. 612). Supply diversity also looks at content diversity that concerns the diversity level of such elements as genre or programme-type of media content, demographic diversity featured in media content, and idea and viewpoint diversity represented in media content.

On the demand end, exposure diversity focuses on media consumption, especially how content diversity is perceived by the audience. According to Napoli (2011), traditional policymaking concerned supply diversity more than exposure diversity due to the concern that whether policymakers have the regulatory authority to intervene audiences' choice and sovereignty. This trend, argued Napoli, may change with some efforts that aim to shape media consumption patterns, for example, linguistic diversity has been addressed to include audiences that access content using languages other than those available online; other factors, such as an individual media outlet's audience reach and audience satisfaction, have also been considered by policymakers.

Media diversity in the digital era

The media environment is evolving; so is the concept of media diversity. Although a framework originated in the pre-internet era, media diversity is still a relevant and important concept when studying the democratic roles and influences of platforms and technologies in the digital age. Scholars, policymakers, and practitioners, however, have different understandings about models of democracy, which results in different expectations of how and to what extent digital platforms should make commitment to diversity in their technologies, practices, and design. For algorithmic recommendation systems, such as search engines, social media sites, and information aggregators, Helberger (2019) argues that the first wave of digital recommenders commonly adopted the liberal model of democracy, in which users' personal autonomy and freedom of expression can outweigh other interests. This giving-users-what-they-want model aims to embrace users' free and autonomous choices but also creates "new concentrations of market or opinion power", which threatens the democratic goal (p. 1000). Compared to the liberal model of democracy, participatory democracy and deliberative democracy require a more proactive role of digital platforms in promoting diversity. Concerns about these models focus on issues such as where the line between informing, educating, and manipulating should be.

In the digital age, new technologies as well as their affordances and global reach have complicated the diversity issue. Sjøvaag (2016) proposed a five-level analytical structure of media diversity — structure, organisation, production, output, and reception — and argued that due to the "global, digital, networked, and privately owned" nature of the large digital platforms, these global "superplayers" have impact on all these five levels of media diversity (p. 11). In the meantime, new elements unique to the digital information environment that may not neatly fall into traditionally defined media diversity dimensions have also been studied in the digital context, such as linguistic diversity that addresses language complexity, tone, sentiment, and the syntactic and semantic characteristics; medium diversity that examines the use and variety of multimedia elements in online media content; and software design diversity that assesses whether the software is sensitive to diversity related issues by design (Carpenter, 2010; Sjøvaag, 2016; Helberger et al., 2018).

Although with complexities, scholars find users welcome diversity in their information diet and research attention to balancing diversity and prediction accuracy of algorithm systems has been growing (Zhou et al., 2010). These dynamics are also a response to biases detected in existing technologies and practices. For example,

studies found gender and racial biases in text-based search results, search auto-completion, and image search, where certain professional occupations, emotions, and technologies are more likely to link to certain gender or racial groups (Makhortykh et al., 2021). Some search engines were found to be biassed toward their own content or to prioritise Western perspectives over non-Western perspectives (Jiang, 2014; Watanabe, 2013).

Diversity areas examined in this study

Since media diversity is such a multifaceted and evolving concept, any individual research project would not be able to cover all aspects discussed above. This study focuses on format-type diversity, source diversity, and structural-social diversity. As will be discussed below, it aims to make contributions to these less studied areas in video platform studies.

Format-type diversity. Format-type diversity is one of the subcomponents to assess content diversity (Napoli, 1999). Since the mid-2000s, the growing adoption of online video content has changed the format of digital content that was once predominantly textual (Tremayne et al., 2007). For traditional video content produced by broadcast media, one can investigate different types of television shows, for example, in terms of situation comedies, made-for-TV movies, and variety programmes, but in the digital context, while marketing and data services identified popular online video content types, such as music video, tutorial, product review (Statista, 2021), little is known about non-marketing-driven genres in terms of their video content type and format. Some studies focused on format-type specific to certain video platforms, such as YouTube-style videos (Beatty, 2016), but as discussed earlier, considerable information producers do not have YouTube channels but host the video content on their own websites. As an effort to address this gap, this study pays particular attention to the format-type diversity in Google's video search results and discusses how it could shape the online video landscape.

Source diversity. In existing studies, source diversity may mean different things in different contexts, for example, Loecherbach and colleagues (2020) used "diversity of entities" in their work where entity is a concept often used in the field of computer science. Named entities could refer to people, places, events, and topics in Google's algorithmic systems (Wang, 2020a). When source diversity is addressed, source could be related to actors, outlets, or sectors in the media marketplace, as in studies about media ownership diversity (e.g. Napoli, 1999); it could also refer to parties or documents cited in journalistic work (e.g. Whitney et al., 1989). To clarify, source in this study refers to the producing parties — whether it is media

outlets, individuals, organisations, or other — of the video content returned by Google's video search.

Several studies found source concentration and geographic concentration in Google's search results. For example, Trielli and Diakopoulos (2019) conducted an algorithm audit of Google search with a focus on Google's top story box. They found that just twenty news sources accounted for more than half of the search results they examined. Among which, legacy media such as CNN, The New York Times, and The Washington Post were the sources that often appear in Google's top story box. Search results of Google News were also found to have come from a small number of national publications, most of which were based in metro areas such as New York City, Washington, D.C. and Los Angeles (Nechushtai & Lewis, 2019). Very few studies paid attention to video search results, but Urman, Makhortykh, and Ulloa (2021) found that sources included in Google's video search results are more diverse than other Western and non-Western search engines.

Structural-social diversity. As Sjøvaag (2016) noted, structural diversity addresses "the broader set of conditions that are beyond the scope of individual media outlets'' that can influence media culture, practice, and law- and policy making (p. 5). These broader issues create the environment in which digital platforms operate and function, and shape output diversity as well as policy choices. One social issue that was salient during the COVID-19 pandemic is that racial and ethnic minority groups were disproportionately affected by COVID-19 due to social determinants of health, such as neighbourhood and physical environment, health and healthcare, occupation and job conditions, income and wealth, and education (Centers for Disease Control and Prevention [CDC], 2020). COVID related disparities are considered "more of a social and economic phenomenon" rooted in structural inequalities (Keating at al., 2020, para. 39).

A Washington Post study found that compared to Whites, African Americans, Hispanics, Asian Americans, and Native American were 37%, 16%, 53%, and 26% more likely to die of COVID respectively. Researchers of the study pointed out that many structural-social issues contributed to such disparities, such as the shortage of COVID testing in minority neighbourhoods, the lack of data from communities of colour, language barriers, housing issues as a result of economic and cultural factors, access to medical care and health insurance, and immigration policies (Keating et al., 2020). Some minority groups face specific racial issues that worsen the challenge. For example, researchers reported that COVID related health disparities that the Asian American community has experienced are largely unknown despite the disproportionately high COVID death rate in this group (Yan et al., 2021).

These researchers pointed out that present-day racism, such as the anti-Asian hate during the COVID-19 pandemic, as well as historical racism, such as the model minority stereotype emerged in the 1960s that depicted Asian Americans as a group that can do better than other minority groups and therefore "unworthy of resources", serve as the profound reasons that lead to the exclusion of the Asian American community from conversations about disparities facing minority groups. "The omission of Asian Americans from discussion of health disparities is itself a form of racism that has serious consequences" argued these researchers (Yan et al., 2021, p. 3547). To what extent Google's video search would help users access this social issue and how it would shape the diversity of its video results? This study will explore these questions.

At the industrial level, cross-sector and cross-company relationships also shape the environmental factors. For example, previous studies found that the relationship between platforms and their competitors had influences on media diversity, particularly, it affected how search engines present their own content and their competitors' content in search results (Jiang, 2014; Urman et al., 2021). Studies found that the changing power dynamic between the tech industry and the news industry has shifted the interrelationship of the two sectors in the past few decades. The inter-industry pushes and pulls have resulted in tensions in some cases and collaborations in other cases across tech and news industries (Wang, 2020b). In the case of Google, such changing relationships have direct effects on output diversity through the inclusion or exclusion of certain sources in its search results. For example, if legal or contractual disputes happened between Google and a source, the involved source may be removed from Google's search results whereas a more agreeable relationship often leads to the inclusion of the source in Google's search results. That's the case when Google stopped hosting the Associated Press' content in 2009-2010 due to contractual disagreement between the two parties over copyright and compensation issues (Wang & Keith, 2021). The AP's content re-entered Google's search results when a new licensing deal was reached (Garber, 2010). These structural-social issues have not been studied on Google's video search platform, an area to which this study aims to make contributions.

Method

The author worked with two students to collect data from Google Videos and data collection was conducted in New York City, US. Data were yielded using 17 sets of keywords related to topics about major events in the early days of the pandemic

after consulting such sources as CNN, Google Trends, and CDC (CNN Editorial Research, 2021; CDC, n.d.). These topics cover general, health, political, economic, public policy, international, and racial/ethnic aspects of the pandemic (Appendix A). We used these keywords to query Google's video search for content published in the early days of the pandemic from 7 March to 23 May 2020, a period of time when most states in the US issued stay at home instructions. For each search query, the data collection focused on the first search engine result page that lists Google's top search results. Studies found that search results on the first search engine result page (SERP) have the highest visibility as searchers have strong selection preference over the top-ranked results (Agichtein et al., 2006; Unkel & Haas, 2017).

We used Python script to download the returned search results and obtained a dataset of 13,084 entries. All personalisation features were turned off during this process to control the potential influence of personalisation. Next, we used Python BeautifulSoup, combined with a manual approach, to extract information relevant to this study, such as date, source, and ranking, for each entry. The collected entries were then coded by two coders for source type, media type, and video page format-type. Consistency was reached through meetings on a weekly basis as well as additional meetings when specific questions were raised. During these meetings, the author and the coders went over the coding sheets. Disagreed items were discussed, cross-checked, and revised based on collective agreement.

For source type, we distinguished media sources from non-media ones. For media type, we adopted categories used in previous studies to identify broadcast media, print media, and online media (Mahone et al., 2019; Mitchell et al., 2020). For non-media sources, we noted whether the source was an individual (such as individual bloggers, posters, etc.), organisation (such as government offices, hospitals, etc.), or other. In addition, as discussed earlier, since previous studies discovered geographical concentration in Google services other than Google's video search, we also paid attention to sources' geographical type to identify whether they are national, local, and international (non-US). For local sources, we identified the state where the given source is located and serves. For format-type, we identified format-type categories based on repeating patterns we discovered during the coding process.

Findings

In this section, I present findings based on the three media diversity areas identified earlier: 1) Format-type diversity, in which I identified three main types of online video pages returned by Google's video search: 2) Source diversity that exam-

ines leading sources and source concentration at both aggregate level and individual search query level; and 3) Structural-social diversity, in which I focused on race/ethnicity-related queries and Google-source partnership.

Format-type diversity

The data analysis reveals three main types of video pages that Google's video search returned to our search queries: the first format-type is articles with videos, the second format-type is YouTube videos, and the third format-type is video show. The rest and a very small portion of our data is in a wide range of variety in terms of the video page format-type, from resource page to informational and educational page.

1. "Articles with videos" format-type. In our data, most Google videos search results are articles with videos (70%, see an example in Figure 1), in which videos were embedded in online articles published on the source's website. For this format-type, the main body of the article is textual, while videos and other elements, such as images and graphs, consist of a multimedia online article. In other words, texts accompanying the videos were the main component of the article. The texts in the article usually provided background and context, explanation of the topic, and additional information, while the video content was complementary to such texts. The prioritisation of this type of video page that requires written content in addition to video content returned more video search results from legacy media, especially broadcast media and print media (35% and 34% respectively).

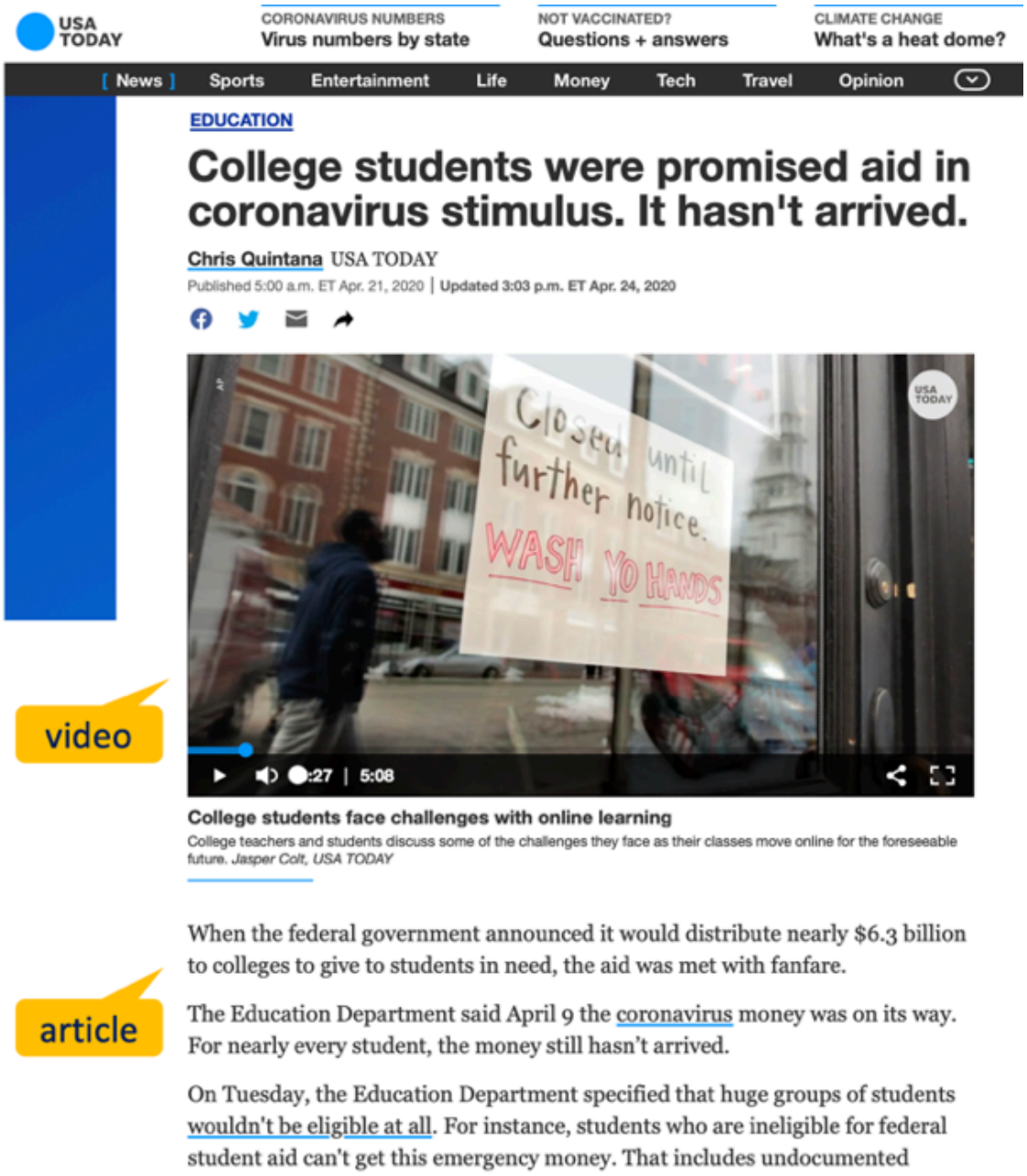

FIGURE 1: An example of articles with videos.

2. "YouTube video" format-type. The second largest category of video page format-type was YouTube videos (12%). These search results directed users to specific YouTube pages, where users can see the video, the description of the video, and information about the YouTube channel. Users can also interact with the video in different ways, such as like, share, and add a comment (see Figure 2 for an example). Unlike articles with videos, there were no online articles accompanying these videos. In our sample, broadcast media had the largest share (53%) in this category. Non-media sources, such as public health authorities, individual producers, and organisations, were the second largest cluster in our YouTube data (33%). Print me-

dia and online media had much smaller shares in this category (9% and 5% respectively).

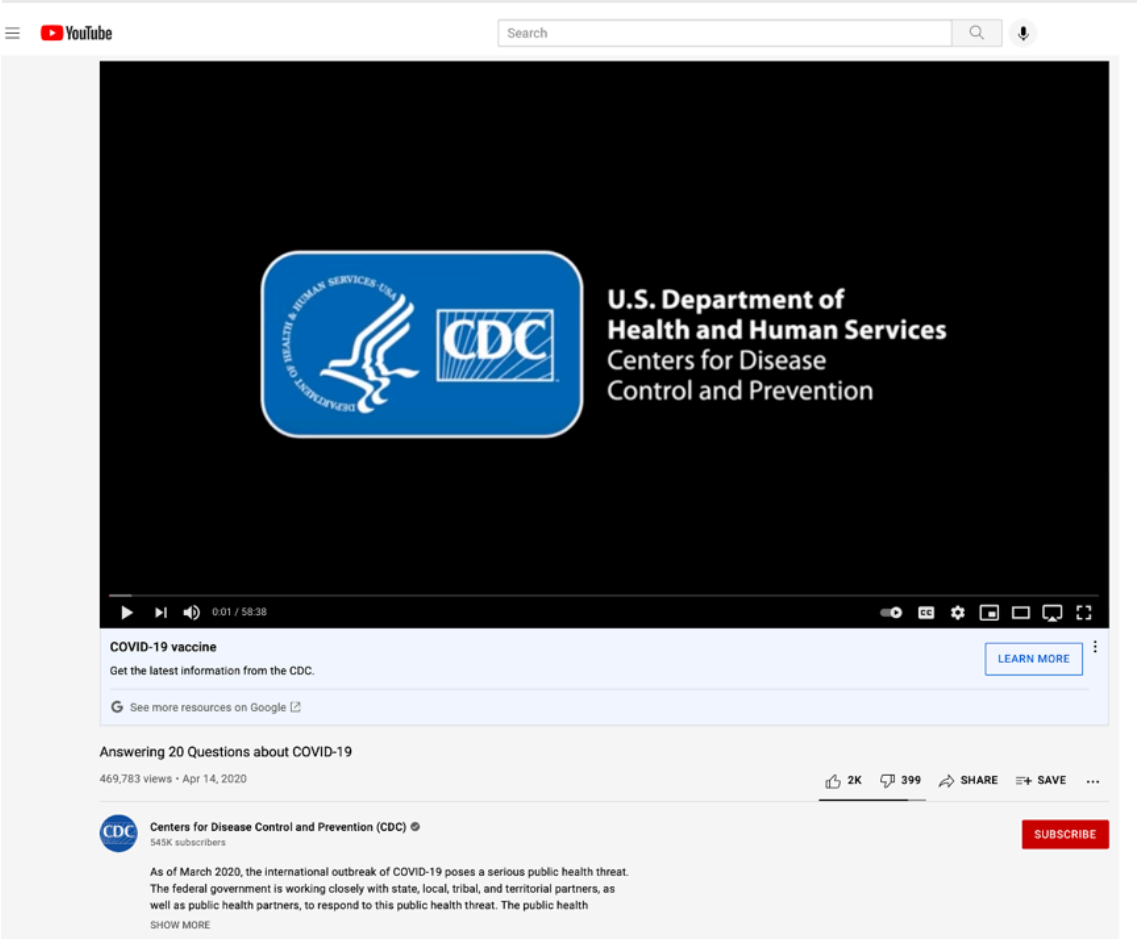

FIGURE 2: An example of a YouTube video.

3. "Video show" format-type. The third video page format-type is video show (6%)[1]. For this format-type, one or multiple video clips were displayed on the source's webpage. This format-type is different from the other two format-types described above in the ways that unlike articles with videos, these webpages feature videos only without any articles accompanying the video; and unlike YouTube videos, the video content of the video show format-type is located on the source's own web page rather than from a YouTube channel. On these web pages, video content was presented in different forms, from a single video clip to video carousel and video list (see Figure 3 for an example). This kind of video shows were primarily from broadcast media.

1. Both "articles with videos" (format-type1) and "video shows" (format-type3) are non-YouTube content.

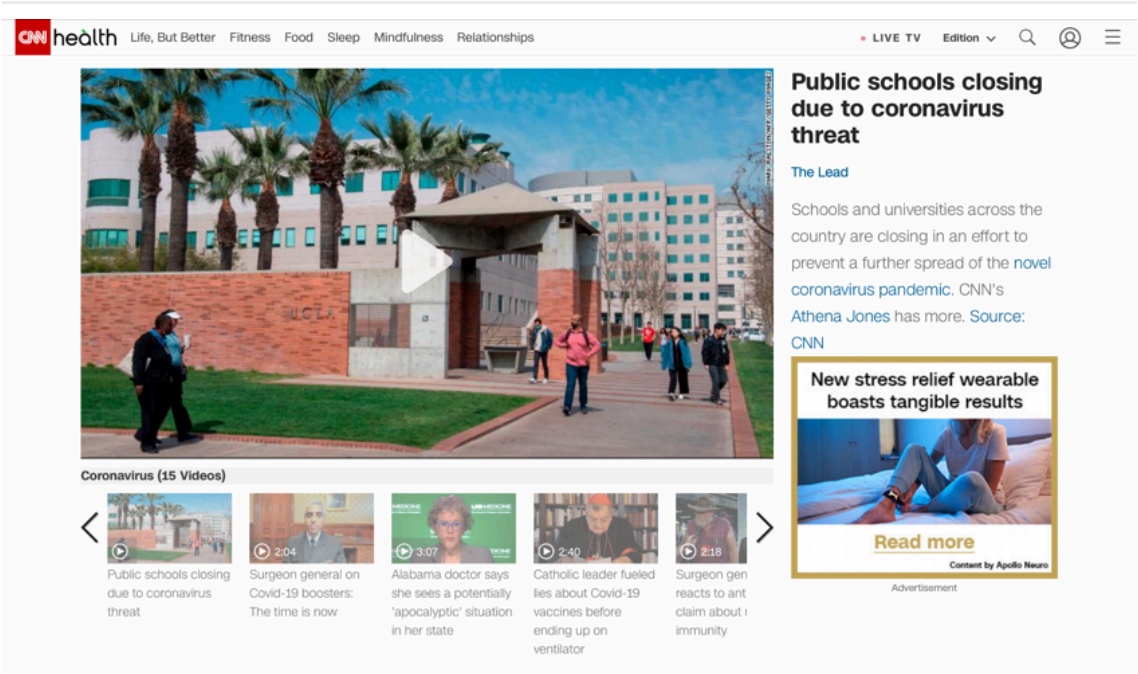

FIGURE 3: An example of a video show.

Findings of this study indicate Google's video search and YouTube are different platforms that prioritise different video content. The format-type diversity examination shows that while YouTube video is the second most popular format-type in our sample, its share is not very large. In fact, Google was found to have the lowest percentage of YouTube content in its top search results when compared to other search engines (Urman et al., 2021). Compared to other major search engines, such as Bing, DuckDuckGo, and Yahoo, that have their top video search results dominated by YouTube content (ranging from 64%-87%, Urman et al., 2021), Google's video search includes more non-YouTube format-types in its top video search results (YouTube content accounted for 12% in our sample). In this sense, Google Videos brings in a certain degree of diversity to the online video landscape in a way that it provides a platform for video format-types alternative to YouTube video to be seen, especially for video producers who are unable to maintain an active YouTube channel, as discussed earlier, and that's the case for many small and local producers. These alternative format-types also involve non-broadcast video producers and help encourage actors who are not traditionally video content producers to enter the online video landscape. For example, print media that traditionally are not video content producers gained a large share in Google's top video search results, especially in the articles with videos format-type category. At the same time, however, these format-types prioritised by Google's video search are very limited and could produce another type of concentration, giving advantages to producing parties who are professionally and financially able to produce both

written and video content. Additionally, given Google's power in the search engine optimisation market, the few format-types that Google prioritises could become templates for video producers who pursue higher search visibility and therefore limit the incentives to explore new format-types for online video content.

Source diversity

In this study, I looked at source diversity at both aggregate level and individual search query level. At the aggregate level, Google data cover both media sources, ranging from broadcast media (36%), print media (28%), and online media (9%), as well as non-media sources, such as public health authorities, government offices, hospitals, and other individuals and organisations, although media sources have a much larger share that took away about three quarters of the entries. In terms of geographical diversity, about 22% of the entries are from local media outlets. The share is lower than national media (35%) but higher than non-US media (9%). These local media outlets cover 46 US states and Washington, D.C.. In the map in Figure 4, the darker the colour the more entries came from local sources in that area. The distribution of these local sources presents a long tail pattern with California, North Carolina, New York, Florida, Pennsylvania, and Texas as the head, representing the majority (53%) of this category. On the tip of the long tail, there were 22 states that each had less than 1% of the entries in this category; and the other 19 states took the remaining entries. In our data, there is no entry from Alaska, North Dakota, South Dakota, and West Virginia.

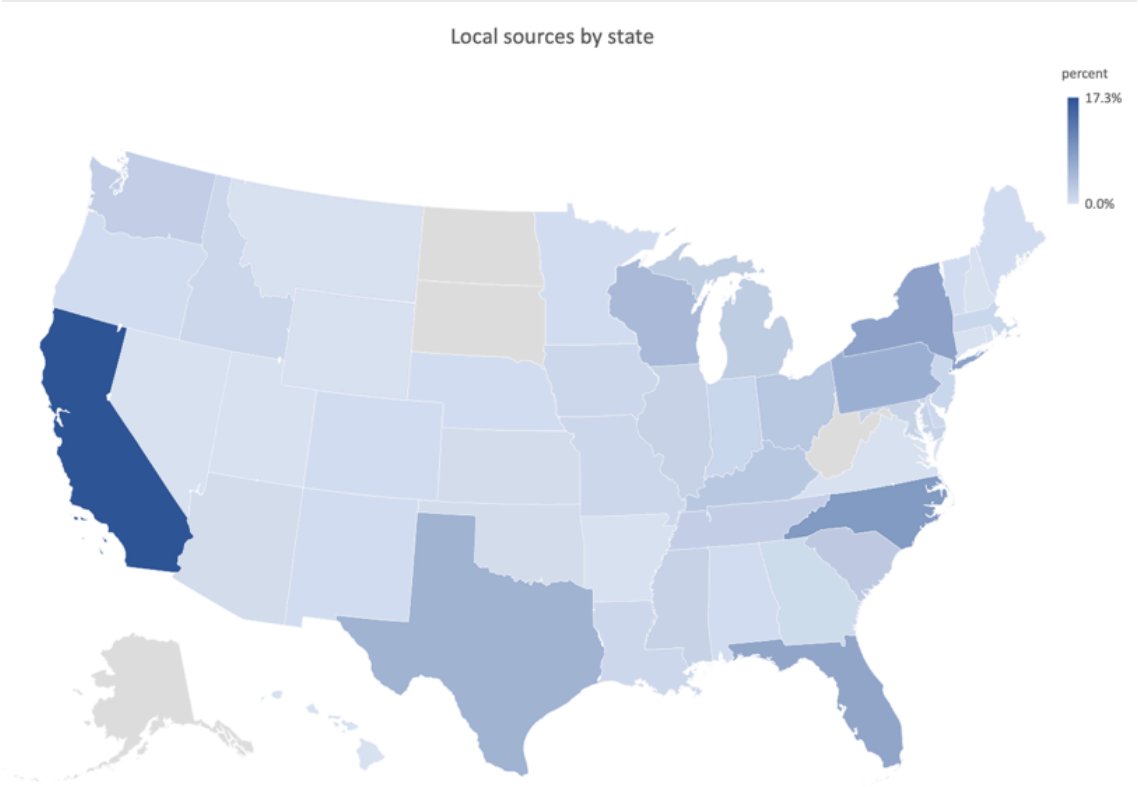

FIGURE 4: Local sources by state.

At individual search query level, the sub-samples that cover political issues (e.g. search queries about election and Trump) related to the COVID-19 pandemic are highly concentrated on a few US national news media sources, such as The Washington Post and USA Today for print media and CNBC, CNN, and CBS News for broadcast media. For their respective media type, these 2-3 leading sources took away about half of the entries in these sub-samples. Local media sources' highest shares were found in two sub-samples, related to search queries about school and unemployment, but source concentration in these sub-samples is lower among local media sources. For example, The News & Observer, a daily newspaper that serves the greater Triangle area based in Raleigh, North Carolina, and The Sacramento Bee, the largest newspaper in Sacramento, California, led in both sub-samples, but combined, they only accounted for less than 17% of these entries. The rest of the data spread out to different local sources on the long tail. Online media had the smallest share across all search queries, but within entries coming from online media, two sub-samples saw more online sources, including the one about COVID stimulus query, where TheHill.com and businessinsider.com combined took away over half of the entries associated with online media; and the Native Ameri-

can sub-sample, where indianz.com (a website that says to "provide you with quality news, information, and entertainment from a Native American perspective") alone accounted for over half of the search results coming from online media sources.

When source diversity is considered, we see a mixed picture. On the one hand, there are a variety of source types and media types in Google's top video search results, many of whom are new actors in the online video landscape as they are not traditionally defined as video content producers, such as newspapers, some online media, and non-media sources. The geographical coverage is wide, but sources are concentrated geographically with a fat head and a long tail. At the individual search query level, source concentration is especially high for national media and online media sources.

Structural-social diversity

As noted earlier, one structural-social issue that was salient during the pandemic is the pandemic's disproportionate impact on racial/ethnic minority groups. To understand the extent to which Google's top video search results reflected this structural-social issue, I examined entries associated with search queries about four racial and ethnic groups in the US — African Americans, Hispanics, Asian Americans, and Native Americans (Standards for maintaining, collecting, and presenting federal data on race and ethnicity, 1997). Since media ownership is an important factor for assessing media diversity (Napoli, 1999), particular attention was paid to sources run by and for racial and ethnic minority groups for each sub-sample. A small portion of such minority sources were found for each sub-sample (8%, 3%, 5%, and 16% for African American, Hispanics, Asian American, and Native American sub-sample respectively). For each sub-sample, the largest shares come from a handful of minority sources that took away over half of the data. These sources include traditional media outlet, such as The Philadelphia Tribune, "the nation's oldest continuously published newspaper reflecting the African-American experience" (Phillytrib, n.d.), but more are social organisations and individuals, such as Salud America!, a Latino-focused research organisation, Asia Society, a non-profit organisation about Asian issues, as well as individual journalists, bloggers, and entertainers.

In 2020, about 1,000 ethnic media outlets were identified in the US (Abernathy, 2020). Studies found that ethnic media play an important role in the media ecosystem as they reach about a quarter of the entire US adult population; and 45% of the surveyed ethnic adults, or 13% of the US adult population, prefer eth-

nic media to mainstream media to get their media information (New California Media, 2005). Given the disproportionate impact of the COVID-19 pandemic on minority groups, the pandemic is a time when ethnic media can be presented more than usual for the public to learn the challenges facing racial and ethnic minority groups. Clearly, the very few minority sources on which Google's top video search results concentrated would not adequately respond to such a need. It's worth noting, however, that many ethnic media are non-English language outlets, which may be less likely to show up in the search results when searching in English. This challenge would require platforms and policymakers to adopt innovative approaches to support these minority sources in the digital age.

Local sources seem to have more weight in Google's top video search results when it comes to queries about racial/ethnic groups. Unlike the source distribution in the overall data, across these race/ethnicity related search queries, local media sources (31%) have a larger share than national media (29%). I identified the top ten states where leading local media sources came from for each racial/ethnic sub-sample (Figure 5-8, left). Then using data from World Population Review, I mapped the top ten states where the given racial/ethnic group has the highest percentage of the state's total population (Figure 5-8, right).

By comparing maps on the left and right, we see both overlaps and mismatches for each racial/ethnic group. For each group, three to four states out of the top ten areas where the given racial/ethnic group has the highest population percentage were included in Google Videos data while most of the areas were not. The mismatches may reflect gaps on both the demand side and the supply side of the information environment. On the demand side, to what extent users would be interested in content outside their own neighbourhoods or their own social group is a factor for commercial digital services to weigh when designing their algorithm and plan their business. On the supply side, the question is how many sources would be able to consistently produce content so that they can be indexed by search engines and other digital platforms. Facebook, for example, complained that it could not find adequate local news for its Today In, a feature launched in 2018 that aimed to promote local content to Facebook users (Holt, 2019). This might especially be the case for online video content considering it requires more resources to produce, publish, and host video content online, a challenge for small, local sources. This observation is also reflected in Google's video search data related to those racial and ethnic minority sources. For example, many entries associated with these sources have videos that do not very well match the content of the source's webpage. In some cases, the same video, such as a video about former US

President Obama's message to the class of 2020, was reused repeatedly for different web pages regardless of the topic and issue addressed in those pages. In other cases, the video is not directly related to the accompanying article, for example, an article that talks about how access to COVID testing and how the option of working from home would impact the African American community was accompanied with a video that explains how soap kills the coronavirus. In still other cases, videos are completely irrelevant to the title, main topic, or content of the webpage. The low relevance of the video content may reflect the challenges many minority media face on the supply end.

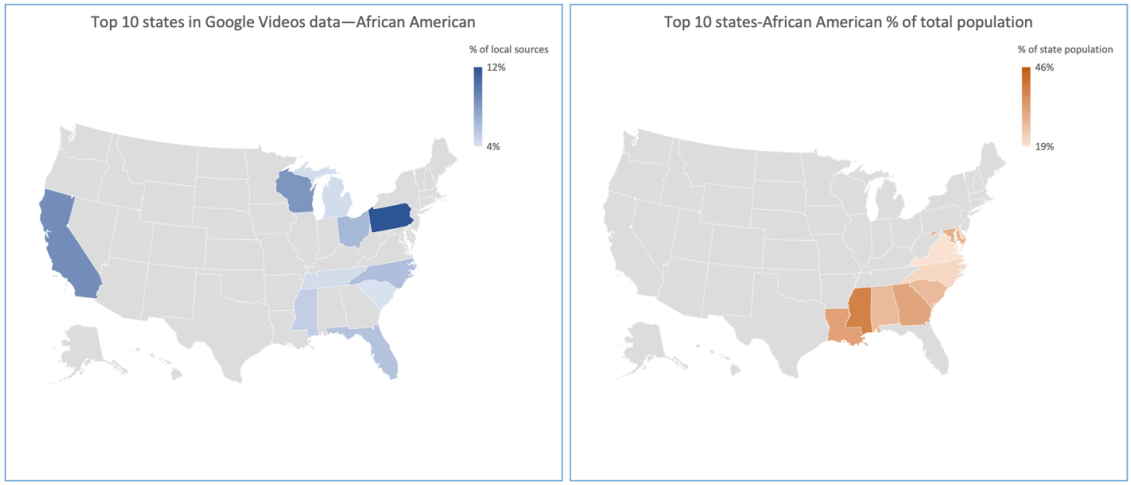

FIGURE 5: Left: African American group — top 10 local sources by state in Google Videos data. Right: African American group — top 10 areas by population %.

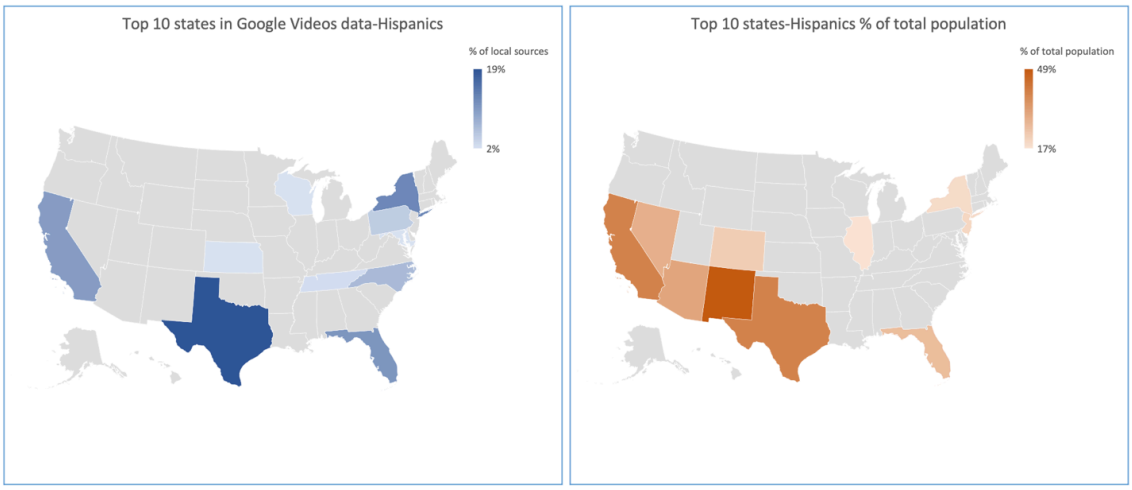

FIGURE 6: Left: Hispanics group — top 10 sources by state in Google Videos data. Right: Hispanics group — top 10 areas by population %.

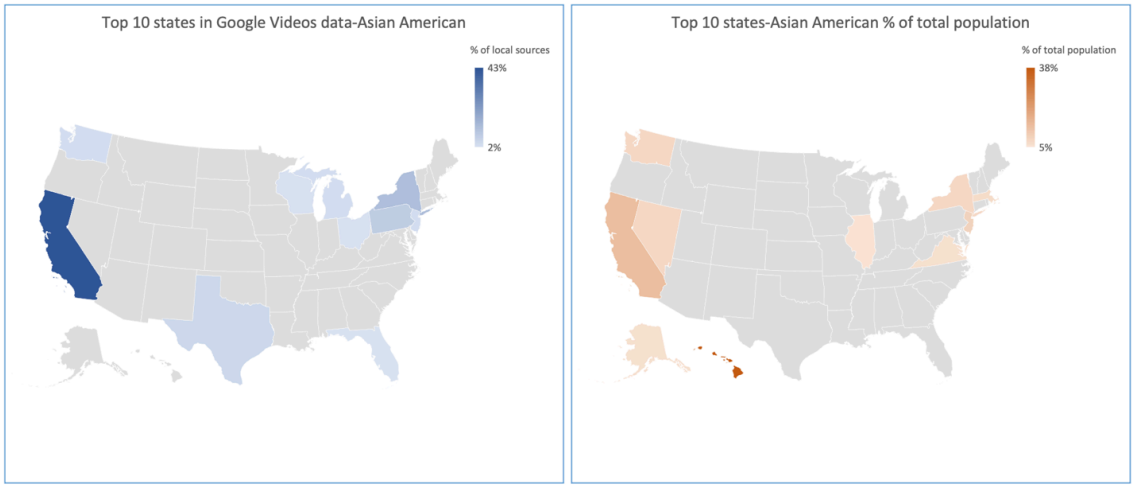

FIGURE 7: Left: Asian American group — top 10 sources by state in Google Videos data. Right: Asian American group — top 10 areas by population %.

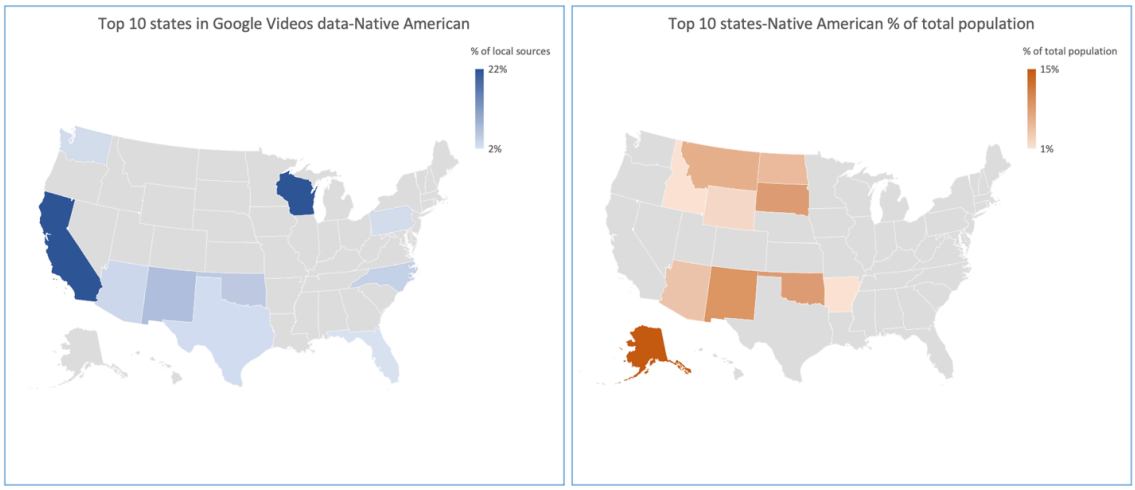

FIGURE 8: Left: Native American group — top 10 local sources by state in Google Videos data. Right: Native American group — top 10 areas by population %.

For industrial relationships, I focused on the leading sources across the entire dataset and examined these sources' partnership relationship with Google. First, I identified the top 10 sources for each search query based on how many times sources appeared in the search results. I then eliminated those that only led in one search query. 31 leading sources are identified, which are top sources for two or more search queries. The examination of the Google-source partnership reveals that most of the identified leading sources have been involved in certain types of partnership with Google at different times (Appendix B). Most of the leading sources are journalistic sources, over a third of which have partnership with YouTube through initiatives such as YouTube innovation funding, the News on YouTube programme, YouTube Player for Publishers, and YouTube Sustainability

Lab (Google News Initiative, n.d.). Another one third of the journalistic sources partnered with Google News Initiative that provides funding, training, and resources for newsrooms (Google News Initiative, n.d.). Over half of the local news sources partnered with Google through its Local News Experiments Project. There are other journalistic sources that were involved in Google's partnership network through a variety of Google products and services, such as Google Ad Manager, Google Analytics, News Consumer Insights, Innovation Challenges funding, and Partnership on AI to Benefit People and Society.

In addition to the journalistic sources, Google has a close relationship with the medical community, including sources on the leading source list. For example, in 2008, Google sponsored the American Medical Association, one of the identified leading sources in our sample, on an AMA Health Care Interoperability and Innovation Challenge project to develop mobile health technology (American Medical Association [AMA], 2018). During the pandemic, Google, through its Ad Grants Crisis Relief program, has developed a partnership with the CDC Foundation, the World Health Organization (both WHO and CDC are on the leading source list), and other public health agencies (Pichai, 2021; Rashidian, 2020).

"MedCram", a 9-year-old YouTube channel run by a Pulmonology doctor, is the only YouTube source on the leading source list. This YouTube channel grew its subscriber number to half million, up from 800 Twitter followers before the pandemic, as a result of its daily update with COVID-related videos (Lewinski, 2020). This source's position in the leading source list, especially considering its smaller subscriber size and short history relative to other sources on the list, suggests that Google's ranking system may reward some high performing users of its own platforms with better visibility in the video search results, such as those that produce native video content and active participants who post on a frequent basis, as noted by algorithm observers (Barnhart, 2021).

While this section does not serve as statistical evidence to prove what partner relationship might be weighed in Google videos search algorithm, given these sources' leading position in the dataset and their dominance among the top video search results, the Google-source partnership identified across the majority of these leading sources is remarkable. When this observation is considered together with previous studies that found Google tends to deprioritise its competitors' content in its video search results (e.g. Schechner et al., 2020; Urman et al., 2021) and aforementioned cases that showed intercompany relationship did play a role in the inclusion and exclusion of sources in Google's search results, this section adds more notes to the effort that explores how industrial relationship at the structural

level could affect output diversity. Google's partnership programs often involve direct funding. In the past two decades, Google has spent billions of dollars in these partnership programs worldwide. For some, details are not disclosed to the public in terms of the terms and conditions, benefits, and short-term and long-term influences. It is also worth pointing out that journalistic sources are the vast majority of these leading sources. Google has recently emphasised on its website that "Google is one of the world's biggest financial supporters of journalism" (Google Supporting news, n.d., para. 2). This section might be a reminder that how such financial relationships would affect the independence of journalism, a key principle of journalism, and therefore media diversity overall, deserves more attention from scholars, industry leaders, and policymakers.

Discussions and future studies

This study, based on data from Google's video search in the early days of the COVID-19 pandemic, reveals a complicated picture in terms of Google's role in format-type diversity, source diversity, and structural-social diversity in the rising online video landscape. In some of these media diversity areas, such as format-type diversity and source diversity, Google's video search brings in some forms of diversity on the one hand and leaves gaps on the other hand. For example, Google's video search results included format-types of online video content alternative to YouTube video, which helps encourage new actors, especially those who are not traditionally video producers, to enter the online video market. On the other hand, format-types that Google's video search prioritised are very limited. In addition, considering Google's power in rule-setting in the digital environment, its prioritisation on articles with videos can play a role in shaping the development of the online video landscape. This dominant format-type in Google's top video search results that involves both written and video content and therefore may require more resources and expertise from content producers. Google's prioritisation can advantage actors who are professionally and financially able to produce such content and/or by setting industry standards for actors on the online video market, which may restrict the incentive to explore new format-types. As far as source diversity is concerned, Google's video search results include a variety of sources that cover media and non-media sources with various geographical types. However, source concentration can be found at both aggregate level and individual search query level. For the highly concentrated search queries, we need to know more about whether such concentration is due to Google's algorithm design or reflects users' preference on certain sources for certain topics. As some scholars pointed out, platforms may have a trade-off between accuracy and diversity (Helberger et al.,

2018), especially in a global crisis like the COVID-19 pandemic when false information spreads widely and with critical consequences.

In the structural-social diversity area, Google seems to prioritise sources in its partnership network. While the statistical significance of partnership relationship in Google's algorithm requires more future studies to investigate, such partner concentration among the leading sources may create an enclosed, centralised space on Google's video search platform. The centralised approach that rewards content and sources from Google's own network could discourage video producers outside Google's partnership network or give them the pressure to join the network. If partnership concentration and centralisation became the underlying logic in algorithm design, independence and transparency related to data collected at the input end and information presented at the output end cannot be guaranteed.

When race and ethnicity related queries are considered, Google's top video search results do include a portion of race and ethnicity minority sources, although the concentration is still notable. From a media diversity's perspective, this study argues that in a time of crisis when racial and ethnic minority groups are disproportionately affected, the presentation of these minority sources could be increased as an opportunity for information seekers to access their content and to learn the challenges facing the minority groups. Local sources have a larger share in race and ethnicity related sub-samples, but only a small portion of the top ten areas with the highest population percentage of the given racial/ethnic group were covered in Google's top local sources. Data of this study, however, cannot explain whether the gap is due to challenges on the supply end or the demand end in the online video content market. There are ongoing debates about how proactive digital platforms should be in their commitment to diversity. For example, should they proactively promote minority content, viewpoints, and sources to users? If so, would such a proactive approach harm user autonomy? Proposals for creating socially sensitive and just algorithms range from technical approaches, such as the adoption of social media signals, to policy intervention, such as the introduction of trusted third party to manage sensitive data (Sheth et al., 2011; Veale & Binns, 2017), but as discussed earlier in this study, these questions are fundamentally normative questions shaped by people's different understandings of democracy. While the goal of this study is not to make normative claims, it argues that technologies bear social responsibilities and issues at the structural-social level matter for media diversity.

From a policymaking point of view, media diversity should be an important thread in the ongoing debate about platform governance. While this study focuses on a

US case, digital platforms, which are largely privately owned entities, often have cross-national businesses and influences and therefore are subject to international regulatory systems. This study as well as many other research efforts that study these platforms provide analyses for global policymakers when addressing platform governance in the digital age. In recent years, platform governance has been moving away from a self-regulation model toward the government regulation model. Multistakeholder governance that involves firms, non-government organisations, international organisations, and states has been practised as well in transnational governance, which forms a "governance triangle" that connects states, firms, and NGOs (Gorwa, 2019). These experiments introduce bottom-up and top-down approaches for platform governance. The former focuses on responsibilities of firms, in terms of their practices, terms, and design. Findings of this study point to areas platform firms can invest in their media diversity commitment, for example, platforms can apply heavier weights to underrepresented data points in algorithm design for better representation of minority groups or they can use different algorithms for different sources and groups. These approaches, however, are facing controversies, such as privacy issues and different voices about digital affirmative action (Daugherty et al., 2018). For the media diversity purpose, platforms can also integrate translation tools in their algorithm for involving non-English sources, or they can adjust their algorithm based on updated or even real-time data in cases like a national or global crisis such as the COVID-19 pandemic. The top-down approach stresses the role of formal regulatory bodies and authorities, such as states and governments, through legislation and traditional regulatory means, such as the "command and control regulation" (Gorwa, 2019). For this approach, law- and policymakers are reminded to avoid the current compartmentalised governance that applies available frameworks to specific issues (such as competition law, privacy law, antitrust law) or specific sectoral areas (such as the banking sector, media sector, or education sector), for a diverse platform ecosystem requires more comprehensive and systematic global platform governance that overlaps many traditional regulatory areas and sectors (Van Dijck, 2021). More conversations between the bottom-up and top-down models are needed to address private actors' commercial incentive and their commitment to public interests.

This study has limitations. Data of this study focus only on Google Videos' first SERP, which may be limited to understanding the long-tail effect. Future studies may want to expand their data collection to cover more search result pages and investigate whether and how information on the long tail would affect diversity. Virtual agents may be used for future studies to further control personalisation. Comparative studies are also needed to provide a baseline for the comparison

across digital platforms and to better understand these platforms' contributions and limitations to media diversity in the digital age.

ACKNOWLEDGMENTS

The author would like to thank the reviewers and editors of the journal for their comments. The author also thanks the students involved in this project: Alex Castro, Brendan McShane, Bridget Conway, and James Hurley.

References

Abernathy, P. M. (2020). *News deserts and ghost newspapers: Will local news survive* [Report]. University of North Carolina at Chapel Hill. https://www.usnewsdeserts.com/reports/news-deserts-and-ghost-newspapers-will-local-news-survive/the-news-landscape-of-the-future-transformed-and-renewed/journalistic-mission-the-challenges-and-opportunities-for-ethnic-media/

Adgate, B. (2020, August 21). Nielsen: How the pandemic changed at home media consumption. *Forbes*. https://www.forbes.com/sites/bradadgate/2020/08/21/nielsen-how-the-pandemic-changed-at-home-media-consumption/?sh=5992c5905a28

Agichtein, E., Brill, E., Dumais, S., & Ragno, R. (2006). Learning user interaction models for predicting web search result preferences. *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 3–10. https://doi.org/10.1145/1148170.1148175

Althaus, S. L. (2002). American news consumption during times of national crisis. *PS: Political Science and Politics*, *35*(3), 517–521. https://doi.org/10.1017/S104909650200077X

American Medical Association. (2018). *AMA, Google launch health care interoperability & innovation challenge* [Press release]. https://www.ama-assn.org/press-center/press-releases/ama-google-launch-health-care-interoperability-innovation-challenge

Asia Society. (n.d.). *Asia Society: Navigating shared futures*. https://asiasociety.org

Baker, L. (2006, October 9). *5 reasons why Google will buy YouTube*. Search Engine Journal. https://www.searchenginejournal.com/5-reasons-why-google-will-buy-youtube/3876/#close

Barnhart, B. (2021, March 26). *Everything you need to know about social media algorithms*. Sproutsocial. https://web.archive.org/web/20210418153831/https://sproutsocial.com/insights/social-media-algorithms/

Beatty, J. (2016). Perceptions of online styles of news video production. *Journal of Visual Literacy*, *35*(2), 126–146. https://doi.org/10.1080/1051144X.2016.1270629

Beech, M. (2020, March 25). COVID-19 pushes up internet use 70% and streaming more than 12%, first figures reveal. *Forbes*. https://www.forbes.com/sites/markbeech/2020/03/25/covid-19-pushes-up-internet-use-70-streaming-more-than-12-first-figures-reveal/?sh=22341b193104

Bloom, D. (2022, May 18). Online video increases worldwide even as U.S. and Europe streaming slows. *Forbes*. https://www.forbes.com/sites/dbloom/2022/05/18/online-video-viewing-growth-flattens-in-us-europe-but-big-screens-roku-rule/?sh=1462833551b3

Bump, P. (2023, April 3). How video consumption is changing in 2023 [Blog]. *HubSpot Blog*. https://blog.hubspot.com/marketing/how-video-consumption-is-changing

Burke, L. V. (2021, September 8). Black Star Network: Roland Martin announces new Black TV network. *The New York Amsterdam News*. https://amsterdamnews.com/news/2021/09/08/black-star-network-roland-martin-announces-new-bla/

Carpenter, S. (2010). A study of content diversity in online citizen journalism and online newspaper articles. *New Media & Society*, *12*(7), 1064–1084. https://doi.org/10.1177/1461444809348772

Ceci, L. (2021, July 12). *Online video usage in the United States—Statistics and facts*. Statista. https://web.archive.org/web/20220207170630/https://www.statista.com/topics/1137/online-video/#dossierKeyfigures

Centers for Disease Control and Prevention. (2020). *COVID-19 racial and ethnic health disparities* [Report]. Centers for Disease Control and Prevention. https://web.archive.org/web/20201210194818/https://www.cdc.gov/coronavirus/2019-ncov/community/health-equity/racial-ethnic-disparities/index.html

Centers for Disease Control and Prevention. (n.d.). *COVID-19 timeline*. https://www.cdc.gov/museum/timeline/covid19.html

Chitu, A. (2007, June 13). Google frames a video search engine [Blog]. *Google Operating System*. http://googlesystem.blogspot.com/2007/06/google-videos-new-frame.html

CNN Editorial Research. (2021). Covid-19 pandemic timeline fast facts. *CNN Health*. https://www.cnn.com/2021/08/09/health/covid-19-pandemic-timeline-fast-facts/index.html

Daugherty, P. R., Wilson, H. J., & Chowdhury, R. (2018, November 21). Using artificial intelligence to promote diversity. *MIT Sloan Management Review*, *60*(2). https://sloanreview.mit.edu/article/using-artificial-intelligence-to-promote-diversity/

DiSilvestro, A. (2017, October 31). *Google videos vs. YouTube: Which is the best video search engine?* Search Engine Watch. https://www.searchenginewatch.com/2017/10/31/google-videos-vs-youtube-which-is-the-best-video-search-engine/?amp=1

Enge, E. (2017, July 19). *Ranking videos on Google and YouTube: Study shows how they differ*. Perficient. https://blogs.perficient.com/2017/07/19/ranking-videos-on-google-and-youtube-study-shows-how-they-differ/

Garber, M. (2010, August 30). The AP and Google reach a licensing renewal agreement—Here's what it might mean for their relationship. *Nieman Lab*. https://www.niemanlab.org/2010/08/the-ap-and-google-reach-a-licensing-renewal-agreement-heres-what-it-might-mean-for-their-relationship/

Google News Initiative. (n.d.). *News on YouTube: YouTube and the news industry*. https://newsinitiative.withgoogle.com/info/youtube

Google Supporting news. (n.d.). *How Google supports journalism and the news industry*. https://blog.google/supportingnews/#overview

Gorwa, R. (2019). The platform governance triangle: Conceptualising the informal regulation of online content. *Internet Policy Review*, *8*(2), 1–22. https://doi.org/10.14763/2019.2.1407

Helberger, N. (2019). On the democratic role of news recommenders. *Digital Journalism*, *7*(8), 993–1012. https://doi.org/10.1080/21670811.2019.1623700

Helberger, N., Karppinen, K., & D'Acunto, L. (2018). Exposure diversity as a design principle for recommender systems. *Information, Communication & Society*, *21*(2), 191–207. https://doi.org/10.1080/1369118X.2016.1271900

Holt, C. (2019, March 18). Facebook can't find enough local news for its local news service [Blog network]. *Engadget*. https://www.engadget.com/2019-03-18-facebook-local-news-availability-today-in-journalism-project.html

Informitv. (2021). *European online video subscription growth* [Report]. https://informitv.com/2021/02/09/european-online-video-subscription-growth/

Jarboe, G. (2019, March 13). *YouTube algorithm: 7 key findings you must know*. Search Engine Journal. https://www.searchenginejournal.com/youtube-algorithm-findings/296291/#close

Jiang, M. (2014). The business and politics of search engines: A comparative study of Baidu and Google's search results of internet events in China. *New Media & Society*, *16*(2), 212–233. https://doi.org/10.1177/1461444813481196

Kalogeropoulos, A., Cherubini, F., & Newman, N. (2016). *The future of online news video* (Digital News Project) [Report]. Reuters Institute for the Study of Journalism. https://ora.ox.ac.uk/objects/uuid:b712713d-5429-4a91-862b-badb41803338

Keating, D., Cha, A. E., & Florit, G. (2020, November 20). 'I just pray God will help me': Racial, ethnic minorities reel from higher covid-19 death rates. *The Washington Post*. https://www.washingtonpost.com/graphics/2020/health/covid-race-mortality-rate/

Kincaid, J. (2011, April 16). Google video prepares to enter the deadpool for good. *TechCrunch*. https://techcrunch.com/2011/04/15/google-video-prepares-to-enter-the-deadpool-for-good/

Krebs, I., Bachmann, P., Siegert, G., Schwab, R., & Willi, R. (2021). Non-journalistic competitors of news media brands on Google and YouTube: From solid competition to a liquid media market. *Journal of Media Business Studies*, *18*(1), 27–44. https://doi.org/10.1080/16522354.2020.1832746

Lenssen, P. (2005, January 25). Google video search live. *Google Blogoscoped*. http://blogoscoped.com/archive/2005-01-25-n90.html

Lewinski, J. S. (2020, February 6). Medical experts employ social media to battle coronavirus. *Forbes*. https://www.forbes.com/sites/johnscottlewinski/2020/02/06/medical-experts-employ-social-media-to-battle-coronavirus/?sh=272506f3776c

Loecherbach, F., Moeller, J., Trilling, D., & Atteveldt, W. (2020). The unified framework of media diversity: A systematic literature review. *Digital Journalism*, *8*(5), 605–642. https://doi.org/10.1080/21670811.2020.1764374

Mahone, J., Wang, Q., Napoli, P., Weber, M., & McCollough, K. (2019). *Who's producing local journalism? Assessing journalistic output across different outlet types* [Report]. DeWitt Wallace Center for Media & Democracy. https://dewitt.sanford.duke.edu/whos-producing-local-journalism-nmrp-report/

Makhortykh, M., Urman, A., & Ulloa, R. (2021). Detecting race and gender bias in visual representation of AI on web search engines. In L. Boratto, S. Faralli, M. Marras, & G. Stilo (Eds.), *Advances in Bias and Fairness in Information Retrieval* (Vol. 1418, pp. 36–50). Springer. https://doi.org/10.1007/978-3-030-78818-6_5

Miners, Z. (2014, November 6). Facebook will be mostly video in 5 years, Zuckerberg says. *PC World*. https://www.pcworld.com/article/2844852/facebook-will-be-mostly-video-in-5-years-zuckerberg-says.html

Mitchell, A., Holcomb, J., Olmstead, K., & Vogt, N. (2014). *News video on the web: A growing, if uncertain, part of news* (State of the News Media, pp. 1–25) [Report]. Pew Research Center. https://www.pewresearch.org/journalism/2014/03/26/developments-in-online-news-video-content/

Mitchell, A., Oliphant, J. B., & Shearer, E. (2020). *About seven-in-ten U.S. adults say they need to take breaks from COVID-19 news* [Report]. Pew Research Center. https://www.pewresearch.org/journalism/2020/04/29/1-americans-are-turning-to-media-government-and-others-for-covid-19-news/

Napoli, P. M. (1999). Deconstructing the diversity principle. *Journal of Communication*, *49*(4), 7–34. https://doi.org/10.1111/j.1460-2466.1999.tb02815.x

Napoli, P. M. (2011). Exposure diversity reconsidered. *Journal of Information Policy*, *1*, 246–259. https://doi.org/10.5325/jinfopoli.1.2011.0246

Nechushtai, E., & Lewis, S. C. (2019). What kind of news gatekeepers do we want machines to be? Filter bubbles, fragmentation, and the normative dimensions of algorithmic recommendations. *Computers in Human Behavior*, *90*, 298–307. https://doi.org/10.1016/j.chb.2018.07.043

New California Media. (2005). *The ethnic media in America: The giant hidden in plain sight* [Final report]. https://legacy.npr.org/documents/2005/jul/ncmfreport.pdf

Pew Research Center. (2007). *Online video proliferates as viewers share what they find online; 57% of online adults watch or download video* [Report]. Pew Research Center. https://www.pewresearch.org/internet/2007/07/25/online-video-proliferates-as-viewers-share-what-they-find-online-57-of-online-adults-watch-or-download-video/

Pew Research Center. (2013). *Online video 2013* [Report]. Pew Research Center. http://www.lifelongfaith.com/uploads/5/1/6/4/5164069/online_video_use_2013_-_pew_research.pdf

Phillytrib. (n.d.). *Our history*. https://www.phillytrib.com/site/about.html

Pichai, S. (2021, January 25). How we're helping get vaccines to more people. *Google The Keyword*. https://blog.google/technology/health/vaccines-how-were-helping/

Rashidian, N. (2020, December 17). Platforms and publishers: The great pandemic funding push. *Columbia Journalism Review*. https://www.cjr.org/tow_center_reports/platforms-publishers-pandemic-funding-news.php

Sadlier, A. (2020, April 14). Americans are streaming 8 hours a day during coronavirus lockdown. *New York Post*. https://nypost.com/2020/04/14/average-american-streaming-content-8-hours-a-day-during-covid-19-according-to-new-research/

Salud America! (n.d.). *About*. https://salud-america.org/about/

Schechner, S., Grind, K., & West, J. (2020, July 14). Searching for video? Google pushes YouTube over rivals. *The Wall Street Journal*. https://www.wsj.com/articles/google-steers-users-to-youtube-over-rivals-11594745232?mod=djemalertNEWS

Sheth, S. K., Bell, J. S., Arora, N., & Kaiser, G. E. (2011). *Towards diversity in recommendations using social networks* (Report CUCS-019-11; Columbia University Computer Science Technical Reports). Department of Computer Science, Columbia University. https://doi.org/10.7916/D81J9JPV

Sjøvaag, H. (2016). Media diversity and the global superplayers: Operationalising pluralism for a digital media market. *Journal of Media Business Studies*, *13*(3), 170–186. https://doi.org/10.1080/16522354.2016.1210435

Skyquest. (2022). *Global online video platforms market drives over 80% of total internet traffic* [Press release]. Skyquest Technology Consulting. https://www.globenewswire.com/en/news-release/2022/08/02/2490661/0/en/Global-Online-Video-Platforms-Market-Drives-over-80-of-Total-Internet-Traffic-Skyquest-Technology.html

Standards for maintaining, collecting, and presenting federal data on race and ethnicity. (October 30, 1997). OMB notice of revisions to the standards for the classification of federal data on race and ethnicity, 62, Fed. Reg., 58788–58790. https://www.govinfo.gov/content/pkg/FR-1997-10-30/pdf/97-28653.pdf

Statista. (2022). *Most popular video content type worldwide in 4th quarter 2022, by weekly usage reach* [dataset]. https://web.archive.org/web/20230609122237/https://www.statista.com/statistics/1254810/top-video-content-type-by-global-reach/

The STAT Team. (2017, November 23). Video carousels nearly doubled in number on the SERP [Blog]. *STAT Blog*. https://getstat.com/blog/video-carousels-nearly-doubled/

Tremayne, M., Weiss, A. S., & Alves, R. C. (2007). From product to service: The diffusion of dynamic content in online newspapers. *Journalism & Mass Communication Quarterly*, *84*(4), 825–839. https://doi.org/10.1177/107769900708400411

Trielli, D., & Diakopoulos, N. (2019). Search as news curator: The role of Google in shaping attention to news information. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3290605.3300683

Unkel, J., & Haas, A. (2017). The effects of credibility cues on the selection of search engine results. *Journal of the Association for Information Science and Technology*, *68*(8), 1850–1862. https://doi.org/10.1002/asi.23820

Urman, A., Makhortykh, M., & Ulloa, R. (2021). Auditing source diversity bias in video search results using virtual agents. *Companion Proceedings of the Web Conference 2021*, 232–236. https://doi.org/10.1145/3442442.3452306

Van Dijck, J. (2021). Seeing the forest for the trees: Visualizing platformization and its governance. *New Media & Society*, *23*(9), 2801–2819. https://doi.org/10.1177/1461444820940293

Veale, M., & Binns, R. (2017). Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society*, *4*(2), 205395171774353. https://doi.org/10.1177/2053951717743530

Wang, Q. (2020a). *Normalization and differentiation in Google News: A multi-method analysis of the world's largest news aggregator* [Doctoral dissertation, Rutgers University]. https://doi.org/doi:10.7282/t3-cn0p-6g66

Wang, Q. (2020b). Differentiation and de-differentiation: The evolving power dynamics between news industry and tech industry. *Journalism & Mass Communication Quarterly*, *97*(2), 509–527. https://doi.org/10.1177/1077699020916809

Wang, Q., & Keith, S. (2021). News aggregators and copyright in the European Union and the United States in the digital age: Evolution, comparisons, and implications. *First Monday*, *26*(9). https://doi.org/10.5210/fm.v26i9.11680

Watanabe, K. (2013). The western perspective in Yahoo! News and Google News: Quantitative analysis of geographic coverage of online news. *International Communication Gazette*, *75*(2), 141–156. https://doi.org/10.1177/1748048512465546

Weissbrot, A. (2020, April 9). Daytime is streaming time: TV viewing habits in the time of COVID-19. *Ad Exchanger*. https://www.adexchanger.com/tv-and-video/daytime-is-streaming-time-tv-viewing-habits-in-the-time-of-covid-19/

Whitney, D. C., Fritzler, M., Jones, S., Mazzarella, S., & Rakow, L. (1989). Geographic and source biases in network television news 1982-1984. *Journal of Broadcasting & Electronic Media*, *33*(2), 159–174. https://doi.org/10.1080/08838158909364070

Yan, B. W., Hwang, A. L., Ng, F., Chu, J. N., Tsoh, J. Y., & Nguyen, T. T. (2021). Death toll of COVID-19 on Asian Americans: Disparities revealed. *Journal of General Internal Medicine*, *36*(11), 3545–3549. https://doi.org/10.1007/s11606-021-07003-0

Zhou, T., Kuscsik, Z., Liu, J. G., Medo, M., Wakeling, J. R., & Zhang, Y. C. (2010). Solving the apparent diversity-accuracy dilemma of recommender systems. *Proceedings of the National Academy of Sciences*, *107*(10), 4511–4515. https://doi.org/10.1073/pnas.1000488107

Appendix A

Search keywords

Google Trends shows that there were much more searches about “coronavirus” than “covid” in the early days of the pandemic, so we used “coronavirus” as our primary search keyword. Main racial/ethnic groups were based on Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity by the U.S. Department of the Interior. https://www.doi.gov/pmb/eeo/directives/race-data

covid
coronavirus
coronavirus and symptom
coronavirus and vaccine
coronavirus and election
coronavirus and Trump
coronavirus and mask
coronavirus and school
coronavirus and China
coronavirus and Europe
coronavirus and unemployment
coronavirus and stimulus
coronavirus and race
coronavirus and African American
coronavirus and Hispanic American
coronavirus and Asian American
coronavirus and native American

Appendix B

TABLE 1: Leading sources and partnership

TOP SOURCES	LEADING IN # OF SEARCH QUERIES	PARTNERSHIP WITH GOOGLE
CNBC	17	Google Assistant
THE WASHINGTON POST	16	YouTube Innovation Funding
CBS NEWS	16	YouTube Innovation Funding
CNN	15	YouTube; GNI
USA TODAY	15	Google Discover
BUSINESSINSIDER.COM	15	GNI
THE GUARDIAN	11	YouTube Innovation Funding
FORTUNE	7	
NBC NEWS	6	YouTube Innovation Funding
THE SACRAMENTO BEE	6	GNI; Local News Experiments Project
SPECTRUM NEWS	6	
THE MIAMI HERALD	5	GNI; Local News Experiments Project
THE SOUTH CHINA MORNING POST	5	YouTube Innovation Funding
THEHILL.COM	4	YouTube Innovation Funding
C-SPAN	4	Google Arts & Culture
ABC30 NEWS	4	GNI Innovation Challenges funding
MCKINSEY.COM	3	Partnership on AI
NEWSWEEK	3	GNI
THE NEWS & OBSERVER	3	Google News COVID-19 special section
WORLD HEALTH ORGANIZATION	3	Google Ad Grants Crisis Relief
ABC NEWS	3	YouTube Innovation Funding
MEDCRAM	2	(YouTube influencer)
THE MILWAUKEE JOURNAL SENTINEL	2	GNI Innovation Challenges funding
FORBES	2	GNI; Google Ad Manager
THE CHARLOTTE OBSERVER	2	GNI; Local News Experiments Project
THE STATE	2	GNI; Local News Experiments Project
THE FORT WORTH STAR-TELEGRAM	2	GNI; Local News Experiments Project
AMERICAN MEDICAL ASSOCIATION	2	Google sponsorship
CDC	2	Google Ad Grants Crisis Relief
PBS	2	YouTube Innovation Funding
AL JAZEERA	2	YouTube

Published by

Alexander von Humboldt
INSTITUTE FOR INTERNET
AND SOCIETY

in cooperation with

UNIVERSITY OF TARTU
Johan Skytte Institute of
Political Studies