

Auditing Source Diversity Bias in Video Search Results Using Virtual Agents

Urman, Aleksandra; Makhortykh, Mykola; Ulloa, Roberto

Veröffentlichungsversion / Published Version

Konferenzbeitrag / conference paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Urman, A., Makhortykh, M., & Ulloa, R. (2021). Auditing Source Diversity Bias in Video Search Results Using Virtual Agents. In J. Leskovec, M. Grobelnik, M. Najork, J. Tang, & L. Zia (Eds.), *WWW '21: Companion Proceedings of the Web Conference 2021* (pp. 232-236). New York: Association for Computing Machinery. <https://doi.org/10.1145/3442442.3452306>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see: <https://creativecommons.org/licenses/by/4.0>

Auditing Source Diversity Bias in Video Search Results Using Virtual Agents

Aleksandra Urman
aleksandra.urman@ikmb.unibe.ch
University of Bern
University of Zurich
Switzerland

Mykola Makhortykh
mykola.makhortykh@ikmb.unibe.ch
University of Bern
Switzerland

Roberto Ulloa
roberto.ulloa@gesis.org
GESIS – Leibniz Institute for the
Social Sciences
Germany

ABSTRACT

We audit the presence of domain-level source diversity bias in video search results. Using a virtual agent-based approach, we compare outputs of four Western and one non-Western search engines for English and Russian queries. Our findings highlight that source diversity varies substantially depending on the language with English queries returning more diverse outputs. We also find disproportionately high presence of a single platform, YouTube, in top search outputs for all Western search engines except Google. At the same time, we observe that Youtube’s major competitors such as Vimeo or Dailymotion do not appear in the sampled Google’s video search results. This finding suggests that Google might be downgrading the results from the main competitors of Google-owned Youtube and highlights the necessity for further studies focusing on the presence of own-content bias in Google’s search results.

KEYWORDS

source diversity bias, algorithmic auditing, web search

ACM Reference Format:

Aleksandra Urman, Mykola Makhortykh, and Roberto Ulloa. 2021. Auditing Source Diversity Bias in Video Search Results Using Virtual Agents. In *Companion Proceedings of the Web Conference 2021 (WWW ’21 Companion)*, April 19–23, 2021, Ljubljana, Slovenia. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3442442.3452306>

1 INTRODUCTION

The public tends to put high trust in the information retrieved via search engines [1]. However, search engine outputs are prone to different forms of bias [11, 21] which can result in distorted representation of subjects which users are searching for [12, 25]. One way of uncovering bias in web search outputs is to engage in algorithmic auditing [24]. Though a few studies have audited bias in text and image search (e.g., [21, 22, 26]), to the best of our knowledge no audits of video search results were conducted. However, video search is important in the societal context since people increasingly consume news via online videos [20] and treat video hosting platforms as a preferred environment for news finding [30]. The fact that video information can have powerful influence on

users and even affect their behaviors [10, 17] prompts the need for auditing whether video search is subjected to bias.

To address this gap, we investigate the presence of source diversity bias - that is a systematic prevalence of specific information sources [11] - in video search outputs. By consistently prioritizing the same set of sources independently of search queries, search algorithms can diminish the quality of outputs by making them less representative [25] and negatively affecting the user experience [6]. Source diversity bias is also related to the phenomenon of search concentration, namely the tendency of search engines to prioritize few well-established domains over other sources [19] that, according to media, often results in search companies promoting their own services (e.g., YouTube in the case of Google [33]). By diminishing the diversity in the composition of source domains, companies can consolidate global media monopolies [19] through gaining unfair advantage over their competitors.

To examine whether video search outputs are subjected to source diversity bias, we audit search results coming from five search engines - four Western (Bing, DuckDuckGo, Google, and Yahoo) and one non-Western (Yandex) - in response to English and Russian queries. Including Yandex along with queries in Russian - a language dominating the main markets of Yandex - allowed us to test whether (some of) our observations can be attributed to structural differences between Western and non-Western markets (e.g., almost monopolistic status of Google in the former and its competition with Yandex in post-Soviet countries). We use virtual agent-based auditing approach to prevent search outputs from being affected by search personalization and search randomization. Then, using a selection of metrics, we assess the level of source domain diversity in search outputs and investigate whether there is evidence of certain engines prioritizing specific information sources. Specifically, we examine whether search engines tend to promote platforms associated with their parent companies (e.g., Alphabet for Google or Microsoft for Bing) or downgrade the competitors as was claimed by earlier research [19].

2 RELATED WORK

The problem of auditing systematic skewness of web search outputs is increasingly recognized in the field of information retrieval (IR) [11, 12, 16, 25]. Existing studies primarily look at it from one of the two perspectives: user bias and retrieval bias. User bias concerns skews in user perceptions of search outputs [15]. Retrieval bias relates to a skewed selection of search results [12, 16].

One form of retrieval bias, which the current paper focuses on, is source diversity bias. Originally discussed in the context of search engines’ tendency to prioritize web pages with the highest number

This paper is published under the Creative Commons Attribution 4.0 International (CC-BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.

WWW ’21 Companion, April 19–23, 2021, Ljubljana, Slovenia

© 2021 IW3C2 (International World Wide Web Conference Committee), published under Creative Commons CC-BY 4.0 License.

ACM ISBN 978-1-4503-8313-4/21/04.

<https://doi.org/10.1145/3442442.3452306>

of visitors [16], source diversity bias is currently investigated in the context of prioritization of certain categories of sources in response to particular types of queries (e.g., [23]). A disproportionate visibility of specific types of web resources can diminish overall quality of search results [6] and provide unfair advantage to companies and individuals that own specific search engines [16] - e.g., through own-content bias, - or direct most of the traffic to a handful of well-established sources, a phenomenon also known as search concentration [19].

To date, source diversity bias has been primarily investigated in the context of text search results [7, 31, 32] with the focus exclusively on one engine - Google. At the same time, few comparative studies that were conducted highlight substantial cross-engine differences in search diversity bias levels [18, 23, 34]. In the case of video search results, there is no systematic comparative assessment of source diversity bias. In 2020, the Wall Street Journal [33] and, subsequently, Mozilla [3] have found that Youtube appears among the top-3 featured "carousel" results in text search 94% of the time [3]. These findings highlight search concentration [19] around Youtube in Google's "carousel" results. However, it is unclear, first, whether the distribution of domains is similar in dedicated video search results. And second, how Google's results compare to those of its competitors - a comparison that is necessary to establish whether Youtube's dominance in video search results from the way Google's algorithm works exhibiting own-content bias. We aim to address these gaps with the present study.

3 METHODS

In this study, we have opted for a combination of methods for bias detection in web search outlined by Edelman [8]: 1) comparative analysis of the results provided by multiple search engines across a variety of search queries in two languages; 2) identification of skewness of results towards specific domains and whether the skewness, if observed, can be explained by market structure incentives. We have used video search results from the 4 biggest Western search engines by market share - Google, Yahoo, Bing, DuckDuckGo, - and one major non-Western search engine - Yandex [2]. Since Yandex has the largest presence in post-Soviet states where large shares of populations speak Russian as a first language, we utilized queries in both, English and Russian, to conduct the search and estimate whether there are differences in the observations. There were 62 queries in total - 31 English queries with translations into Russian, - that concerned contemporary events (i.e., the US presidential elections and coronavirus), conspiracy theories (i.e., Flat Earth), and historical events (i.e., Holocaust)¹. Below we outline the details on the data collection and analysis.

3.1 Data collection

To collect the data, we utilized a set of virtual agents - that is software simulating user browsing behavior and recording its outputs.

¹We used the following queries: *coronavirus, bernie sanders, joe biden, pete buttigieg, elizabeth warren, michael bloomberg, donald trump, us elections, syria conflict, ukraine conflict, yemen conflict, holocaust, holodomor, armenian genocide, second world war, first world war, artificial intelligence, big data, virtual reality, vaccination, vaccination benefits, vaccination dangers, george soros, illuminati, new world order, Flat Earth, UFO, Aliens, misinformation, disinformation, fake news*. All queries were entered into search engines in lower case. For the searches in Russian, we used the exact translations of the English queries listed below into Russian verified by two native Russian speakers.

The benefits of this approach, which extends algorithmic auditing methodology introduced by Haim et al. [13], is that it allows controlling for personalization [14] and randomization [23] factors.

For the current study, we built a network of 84 CentOS virtual machines based in the Frankfurt region of Amazon Elastic Compute Cloud (EC2). On each machine, we deployed 2 virtual agents (one in Chrome browser and one in Mozilla Firefox browser), thus providing us with 188 agents overall. Each agent was made of two browser extensions: a tracker and a bot. The tracker collected the HTML and the metadata of all pages visited in the browser and immediately sent it to a storage server. The bot emulated a sequence of browsing actions that consisted of (1) visiting a video search engine page, (2) entering one of the 62 queries, and (3) scrolling down the search result page to load at least 50 images. Before searching for a new query, the browsers were cleaned to prevent the search history affecting the search outputs, and there was a 7-minute break between searches to mitigate potential effects of previous searches on the results.

The study was conducted on February 26, 2020. We equally distributed the agents between five search engines: Google, Bing, Yahoo, Yandex, and DuckDuckGo (DDG)². Because of technical issues (e.g., bot detection mechanisms), some agents did not manage to complete their routine. The overall number of agents per engine which completed the full simulation routine and returned the search results differed by query - sometimes the search engine would detect automation and temporarily ban the agent. This was particularly often the case with Yandex, where for some queries all 34 deployed agents successfully finished the routine while for others (a minority of queries) Yandex only returned the results for 10 agents and banned the rest. The mean number of agents who completed the full routine by engine across all queries is the following: Bing (29), DDG (34), Google (33), Yahoo (31), and Yandex (17).

After the data was collected, we have extracted top-10 individual video links obtained by each agent for each search query and proceeded with the analysis using this data. Our decision to rely on the top-10 results only is motivated by the fact that users tend to pay the most attention to the first few results - i.e., those on the first results page [27]. A comparison of search results by browser has demonstrated that there are no major between-browser differences - a finding in contrast with those observed for text search results [23] - thus for the analysis we have proceeded aggregating the results for both browsers.

3.2 Data analysis

3.2.1 Source diversity. To assess whether there is evidence suggesting that diversity bias - that is, lack of source diversity, - is present in the sampled results on domain level, we have calculated how many distinct source domains are, on average, present in the results for each query. To account for the potential randomization due to so-called "Google Dance" [5] in the results, for each query we aggregated the calculation over the results obtained from each individual autonomous agent. Afterwards, we have calculated mean numbers of distinct sources separately for English and Russian queries to establish whether there are differences in the observations depending

²For all engines, the ".com" version of the image search engine was used (e.g., google.com).

on the language of the search. We have also qualitatively examined the results per query to find out whether there are distinct patterns with regard to query categories.

3.2.2 Search concentration. To establish whether there is evidence of search concentration in the collected video results, we have calculated, first, the share of times different domains appear as the top result for each search query, and second, the proportion of times different domains appear among the top-10 search results per query at all. We suggest that the consistent appearance of a specific domain or few specific domains at the very top of search results and, in general, among top-10 results more frequently than the others would indicate search concentration.

In addition, we have scrutinized the results with regard to own-content bias exhibited by Google according to the media [33]. We aimed to establish whether Google’s results lend evidence of own-content bias either through the promotion of Youtube in the results or the demotion of the results provided by its main competitors. One way to assess that is to compare Google’s results with those obtained through other search engines [8] and see whether there are major differences in the observed frequencies of appearance of different domains between them. Thus, we compared the proportion of times each domain - Youtube and its competitors such as Vimeo, Dailymotion and Rutube, - appears in the results provided by different engines.

4 FINDINGS

4.1 Source diversity

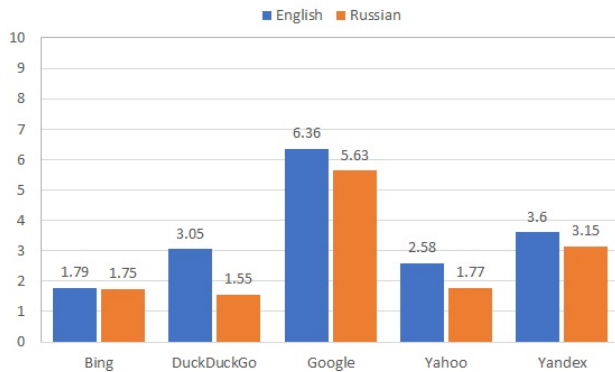


Figure 1: Mean number of distinct domains in top-10 video search results (Y-axis) per query, grouped by query languages (legend) and search engines (X-axis)

As Fig.1 shows, there are differences in the level of source diversity exhibited by the examined search engines. Google has consistently presented more diverse video results in terms of source domains than its competitors. Yandex has taken a second place in terms of domain diversity. This domain diversity hierarchy is similar for both English and Russian queries, albeit for the Russian queries the observed domain diversity is a bit lower on all five search engines than for the English queries.

Qualitative analysis of the domain diversity by query has demonstrated that there are no consistent patterns with regard to the

proportion of distinct sources by query category in our dataset. Hence, we suggest that domain diversity is affected more by the algorithms used by each search engine examined and, probably, the data they are trained on - the latter might explain the observed differences between Russian and English queries, - rather than by the specific topics of the search queries.

4.2 Search concentration

As shown in Fig.2, Youtube has been featured as the top result most frequently in all cases but one - namely, the results for English queries on Yandex where Vimeo surfaced as the first result most often. Remarkably, on Google itself Youtube appeared as the top result less frequently than on other platforms, a finding in contrast with those made by The Wall Street Journal [33] and Mozilla [3] in the context of featured video "carousel" on the first page of Google’s text results. DuckDuckGo, Yahoo, Yandex are the three engines exhibiting sizeable differences in the prominence of Youtube as the top result between English and Russian queries with Youtube being featured as the top result more frequently in response to English queries. We suggest that the findings reported in Fig.2 lend evidence to search concentration bias in video search results on the examined search engines, with the effect in our sample being stronger for English than for Russian queries.

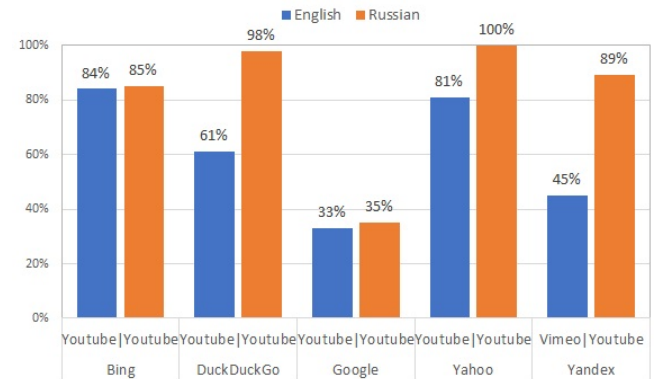


Figure 2: Domain most frequently appearing as the top result by query group and search engine; % of time it appears as the top result.

In Fig.3 we list the domains most frequently featured among top-10 search results on each of the examined search engines for English and Russian queries. As with the top-1 result, the domain most frequently featured in top-10 is Youtube on all engines but Yandex in response to English queries where the most frequent domain is Vimeo. The share of other domains in search results is comparatively marginal - less than 10% - in almost all cases with the exception of Ok.ru for Russian queries on Yandex. Youtube thus emerges as the most prominent domain in search results. Apart from it, there is no other domain that is among the top 5 domains most frequently appearing among the first 10 results on all search engines.

The findings reported in Fig.3 suggest that search engines feature different arrays of domains - with the exception of Youtube - in

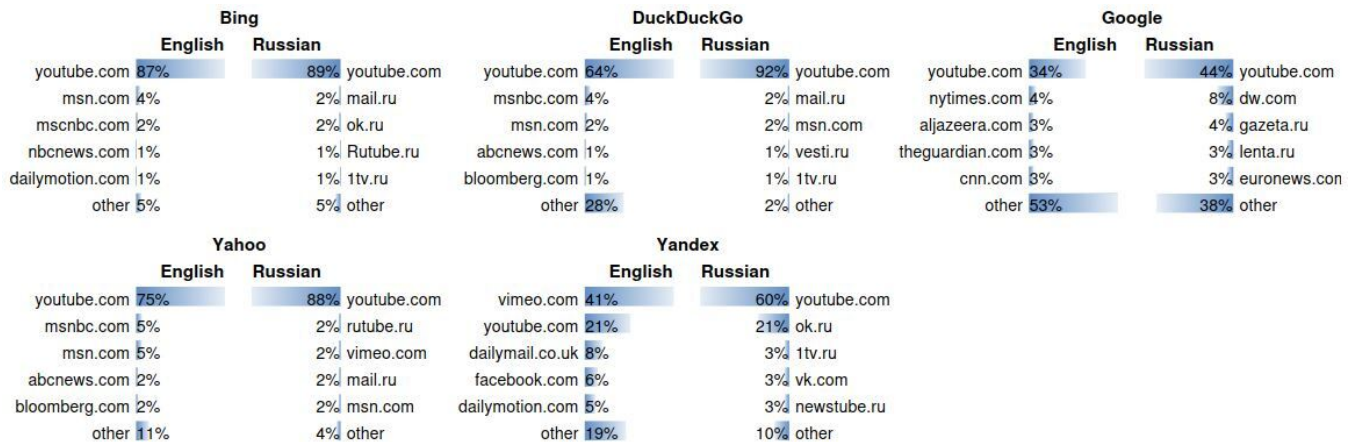


Figure 3: Domains most frequently appearing among top-10 results by query group and search engine and % of time a domain appears among top-10 results.

search results. Google tends to retrieve results from legacy media in both Russian and English more frequently than other search engines, a finding in line with the previous research [7, 31, 32]. Other search engines also include some legacy media, though to a lower extent than Google, as well as social media (e.g., Ok.ru, Facebook.com), and several video portals that are Youtube’s competitors - Dailymotion, Vimeo and Rutube. None of these potential Youtube competitors, however, appear in the top-10 results on Google in our dataset at all despite their presence on other search engines. Vimeo, Dailymotion and Rutube all appear at least once among the top-10 results on all search engines except Google. This finding suggests that Google might downgrade Youtube’s direct competitors, however an analysis based on a broader spectrum of queries is necessary to estimate the scope and persistence of this result.

5 CONCLUSIONS AND FUTURE WORK

Top 10 outputs of video search for most of search engines except Google show limited source diversity. By relying on average on 2-3 unique sources to retrieve top results for English queries, search engines create a situation in which users’ information choices are shaped by a few content providers. This raises concerns about search engines facilitating consolidation of power on the information markets.

The only exception among the five search engines examined is Google, where the degree of source diversity is almost twice as high. This effect can be attributed to Google putting substantial effort into diversifying search results in response to earlier criticisms. With 6 unique domains per 10 top results, Google follows its declared principle of having no more than 2 results coming from the same domain in the top results [4]. The finding also suggests that low diversity on other search engines is likely attributed to the absence of diversification mechanisms which Google implements.

The importance of integrating diversification measurements is highlighted by high degree of source diversity bias. The top results for all search engines (except Yandex in English and Google) are dominated by one platform, namely YouTube. Its systematic prevalence reinforces the platform’s almost monopolistic status. It

is problematic considering that the platform is already used as a major news source among certain shares of the population [20], despite earlier audits demonstrating that its algorithms might lead to user radicalization [29] and, affected by users’ viewing history, aggressively promote pseudoscientific content to users who have watched pseudoscientific videos before [28]. Search concentration around Youtube can only help cement its monopoly and the associated effects.

Google fares better than other search engines in terms of domain-level source diversity and, at a first glance, does not exhibit own-content bias since Youtube is less prominent in its results than on other search engines. However, in our sample, Google is the only search engine that did not provide any results from Youtube’s major competitors. It is thus plausible that Google indeed, as media reports suggested [33], might be downgrading the results coming from the major competitors of Youtube thus exhibiting own-content bias manifested not in promoting Youtube but in lowering the prominence of its competitors in search results. However, it could also be that the obtained results with regards to Google and the absence of Youtube’s competitors in its outputs are specific to the topics addressed by our sample of queries and is absent in other contexts. Hence, to make a robust conclusion about the presence or absence of own-content bias in Google’s video search results, further studies encompassing broader sets of queries are necessary. We suggest that our observations highlight the need for such studies.

Further, the present analysis is based on a snapshot experiment on a limited selection of queries. We believe that our findings warrant subsequent longitudinal audits of video search results to assess the persistence of our observations and, potentially, the changes that occur overtime. Such audits are crucial to inform the decisions of policy-makers and regulators. This is especially timely and pressing given the recent anti-trust cases against Google [9] and calls for putting tech giants under scrutiny, among other, in the context of their market power.

REFERENCES

[1] [n.d.]. 2020 Edelman Trust Barometer. <https://www.edelman.com/trustbarometer>

- [2] [n.d.]. Search Engine Market Share Worldwide. <https://gs.statcounter.com/search-engine-market-share>
- [3] [n.d.]. YouTube Dominates Google Video in 2020. <https://moz.com/blog/youtube-dominates-google-video-results-in-2020>
- [4] 2019. Google search update aims to show more diverse results from different domain names. <https://searchengineland.com/google-search-update-aims-to-show-more-diverse-results-from-different-domain-names-317934>
- [5] John Battelle. 2005. *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture*. Portfolio. Google-Books-ID: FR9PAAAAMAAJ.
- [6] Giovanni Luca Ciampaglia, Azadeh Nematzadeh, Filippo Menczer, and Alessandro Flammini. 2018. How algorithmic popularity bias hinders or promotes quality. *Scientific Reports* 8, 1 (Oct. 2018), 15951. <https://doi.org/10.1038/s41598-018-34203-2> Publisher: Nature Publishing Group.
- [7] Nicholas Diakopoulos, Daniel Trielli, Jennifer Stark, and Sean Mussenden. [n.d.]. I Vote For—How Search Informs Our Choice of Candidate. ([n.d.]), 22.
- [8] Benjamin Edelman. [n.d.]. BIAS IN SEARCH RESULTS?: DIAGNOSIS AND RESPONSE. 7 ([n.d.]), 17.
- [9] Gilad Edelman. [n.d.]. Google's Antitrust Cases: A Guide for the Perplexed. *Wired* ([n.d.]). <https://www.wired.com/story/google-antitrust-lawsuits-explainer/>
- [10] Thane M. Erickson, Adam P. McGuire, Gina M. Scarsella, Tara A. Crouch, Jamie A. Lewis, Ashley P. Eisenlohr, and Tasha J. Muresan. 2018. Viral videos and virtue: Moral elevation inductions shift affect and interpersonal goals in daily life. *The Journal of Positive Psychology* 13, 6 (Nov. 2018), 643–654. <https://doi.org/10.1080/17439760.2017.1365163> Publisher: Routledge _eprint: <https://doi.org/10.1080/17439760.2017.1365163>.
- [11] E. Goldman. 2008. Search Engine Bias and the Demise of Search Engine Utopianism. In *Web Search: Multidisciplinary Perspectives*, Amanda Spink and Michael Zimmer (Eds.). Springer, Berlin, Heidelberg, 121–133. https://doi.org/10.1007/978-3-540-75829-7_8
- [12] James Grimmelmann. 2010. Some Skepticism About Search Neutrality. *Faculty Scholarship* (Jan. 2010). https://digitalcommons.law.umaryland.edu/fac_pubs/1417
- [13] Mario Haim, Florian Arendt, and Sebastian Scherr. 2017. Abyss or Shelter? On the Relevance of Web Search Engines' Search Results When People Google for Suicide. *Health Communication* 32, 2 (Feb. 2017), 253–258. <https://doi.org/10.1080/10410236.2015.1113484>
- [14] Aniko Hannak, Piotr Sapiezynski, Arash Molavi Kakhki, Balachander Krishnamurthy, David Lazer, Alan Mislove, and Christo Wilson. 2013. Measuring personalization of web search. In *Proceedings of the 22nd international conference on World Wide Web - WWW '13*. ACM Press, Rio de Janeiro, Brazil, 527–538. <https://doi.org/10.1145/2488388.2488435>
- [15] Samuel Ieong, Nina Mishra, Eldar Sadikov, and Li Zhang. 2012. Domain bias in web search. In *Proceedings of the fifth ACM international conference on Web search and data mining (WSDM '12)*. Association for Computing Machinery, New York, NY, USA, 413–422. <https://doi.org/10.1145/2124295.2124345>
- [16] Lukas Introna and Helen Nissenbaum. 2000. Shaping the Web: Why the Politics of Search Engines Matters. *The Information Society* 16, 3 (July 2000), 169–185. <https://doi.org/10.1080/01972240050133634> Publisher: Routledge _eprint: <https://doi.org/10.1080/01972240050133634>.
- [17] Michiko Izawa. [n.d.]. WHAT MAKES VIRAL VIDEOS VIRAL?: ROLES OF EMOTION, IMPRESSION, UTILITY, AND SOCIAL TIES IN ONLINE SHARING BEHAVIOR. ([n.d.]), 56.
- [18] Min Jiang. 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 (March 2014), 212–233. <https://doi.org/10.1177/1461444813481196> Publisher: SAGE Publications.
- [19] Min Jiang. 2014. Search Concentration, Bias, and Parochialism: A Comparative Study of Google, Baidu, and Jike's Search Results From China. *Journal of Communication* 64, 6 (2014), 1088–1110. <https://doi.org/10.1111/jcom.12126> _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/jcom.12126>.
- [20] Antonis Kalogeropoulos. 2018. Online News Video Consumption. *Digital Journalism* 6, 5 (May 2018), 651–665. <https://doi.org/10.1080/21670811.2017.1320197> Publisher: Routledge _eprint: <https://doi.org/10.1080/21670811.2017.1320197>.
- [21] Matthew Kay, Cynthia Matuszek, and Sean A. Munson. 2015. Unequal Representation and Gender Stereotypes in Image Search Results for Occupations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3819–3828. <https://doi.org/10.1145/2702123.2702520>
- [22] Juhi Kulshrestha, Motahareh Eslami, Johnatan Messias, Muhammad Bilal Zafar, Saptarshi Ghosh, Krishna P. Gummedi, and Karrie Karahalios. 2019. Search bias quantification: investigating political bias in social media and web search. *Information Retrieval Journal* 22, 1 (April 2019), 188–227. <https://doi.org/10.1007/s10791-018-9341-2>
- [23] Mykola Makhortykh, Aleksandra Urman, and Roberto Ulloa. 2020. How search engines disseminate information about COVID-19 and why they should do better. *Harvard Kennedy School Misinformation Review* 1, COVID-19 and Misinformation (May 2020). <https://doi.org/10.37016/mr-2020-017>
- [24] Brent Mittelstadt. 2016. Automation, Algorithms, and Politics| Auditing for Transparency in Content Personalization Systems. *International Journal of Communication* 10, 0 (Oct. 2016), 12. <https://ijoc.org/index.php/ijoc/article/view/6267> Number: 0.
- [25] Abbe Mowshowitz and Akira Kawaguchi. 2002. Assessing bias in search engines. *Information Processing & Management* 38, 1 (Jan. 2002), 141–156. [https://doi.org/10.1016/S0306-4573\(01\)00020-6](https://doi.org/10.1016/S0306-4573(01)00020-6)
- [26] Jahna Otterbacher, Jo Bates, and Paul Clough. 2017. Competent Men and Warm Women: Gender Stereotypes and Backlash in Image Search Results. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 6620–6631. <https://doi.org/10.1145/3025453.3025727>
- [27] Bing Pan, Helene Hembrooke, Thorsten Joachims, Lori Lorigo, Geri Gay, and Laura Granka. 2007. In Google We Trust: Users' Decisions on Rank, Position, and Relevance. *Journal of Computer-Mediated Communication* 12, 3 (2007), 801–823. <https://doi.org/10.1111/j.1083-6101.2007.00351.x> _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1083-6101.2007.00351.x>.
- [28] Kostantinos Papadamou, Savvas Zannettou, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini, and Michael Srivivasos. 2020. "It is just a flu": Assessing the Effect of Watch History on YouTube's Pseudoscientific Video Recommendations. *arXiv:2010.11638 [cs]* (Nov. 2020). <http://arxiv.org/abs/2010.11638> arXiv: 2010.11638.
- [29] Manoel Horta Ribeiro, Raphael Ottoni, Robert West, Virgilio A. F. Almeida, and Wagner Meira. 2020. Auditing radicalization pathways on YouTube. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20)*. Association for Computing Machinery, New York, NY, USA, 131–141. <https://doi.org/10.1145/3351095.3372879>
- [30] Galen Stocking, Patrick van Kessel, Michael Barthel, Katerina Eva Matsa, and Maya Khuzam. 2020. Many Americans Get News on YouTube, Where News Organizations and Independent Producers Thrive Side by Side. <https://www.journalism.org/2020/09/28/many-americans-get-news-on-youtube-where-news-organizations-and-independent-producers-thrive-side-by-side/>
- [31] Daniel Trielli and Nicholas Diakopoulos. 2019. Search as News Curator: The Role of Google in Shaping Attention to News Information. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, Glasgow, Scotland Uk, 1–15. <https://doi.org/10.1145/3290605.3300683>
- [32] Julian Unkel and Mario Haim. 2019. Googling Politics: Parties, Sources, and Issue Ownerships on Google in the 2017 German Federal Election Campaign. *Social Science Computer Review* (Dec. 2019), 0894439319881634. <https://doi.org/10.1177/0894439319881634>
- [33] Sam Schechner West, Kirsten Grind and John. 2020. Searching for Video? Google Pushes YouTube Over Rivals. *Wall Street Journal* (July 2020). <https://www.wsj.com/articles/google-steers-users-to-youtube-over-rivals-11594745232>
- [34] Andrei Zavadski and Florian Toepfl. 2019. Querying the Internet as a mnemonic practice: how search engines mediate four types of past events in Russia. *Media, Culture & Society* 41, 1 (Jan. 2019), 21–37. <https://doi.org/10.1177/0163443718764565> Publisher: SAGE Publications Ltd.