Personalised Filter Bias with Google and DuckDuckGo: An Exploratory Study

Awais Akbar[1(✉)], Simon Caton[2], and Ralf Bierig[1]

[1] Department of Computer Science, Maynooth University, Kildare, Ireland
`awais.akbar.pk@gmail.com`
[2] School of Computer Science, University College Dublin, Dublin, Ireland

Abstract. Personalisation in search has improved performance, focus, and user experience to a great extent, however, it also arguably polarises informational perspectives. This paper seeks to illustrate an experimental methodology to quantify how three situational user variables affect personalisation across two search engines: Google and DuckDuckGo. We find that the presence of cookies and prior search history markedly affect the first page of search results on both platforms, but that prior (shallow) browsing history has no observable effect. We also find that there is very little in common between the results of both search engines. We argue that these results advocate more consideration of how personalisation fosters filter biases.

Keywords: Personalisation · Filter bias · Simulation experiment · Search engine

1 Introduction

Search engines (and the web in general) are highly aligned to users' perceived interests. While this often delivers "relevant" content, it is arguably informational polarisation and can negate the serendipity of search through the development of "Filter Bubbles". Whilst researchers have begun to investigate search engine performance in relation to user information needs (e.g. [8,33,35]), we argue that methodologies that more directly quantify the contribution of situational user variables to personalised results are needed to better understand potential biases and their implications.

To provide some initial empirical insights, we investigate the extent to which search result personalisation is informed (or influenced) by: 1) user's information stored in *browser cookies*; 2) user's *prior search history*; and 3) user's *prior browsing history*. To derive empirical results, we investigate two common search engines: Google and DuckDuckGo, with three experiments designed to expose any discernible differences in search engine behaviour by analysing the content of Search Engine Results Page(s) (SERP). To investigate these aspects, we leverage a simulation-based controlled experiment, i.e. we instrument an automated search process within an engineered user context. Our methodology controls for noise, specifically the carry-over effect [15], to accurately attribute the differences

to personalization in the returned results. The carry-over effect is a phenomenon that occurs within a browsing session when a user immediately searches for one query after another. In this case, the search for the first query may influence the results received by the immediate search for the second query. This strategy has been documented on Google by Hannak et al. [15].

Our motivation for a simulation-based study is that it provides significant control over key variables (cookie information, prior browsing/search history, the ordering and nature of search terms as well as situational context of search, i.e. browser headers etc.) that can provide initial insights for the design of a more expansive user-based study. Thus, the contribution of this paper is a set of empirical results that investigate the impact of browser cookies in general (without being logged into the search engine's ecosystem), the impact of prior user searches, and the impact of users' prior browsing behaviour on SERPs as a juxtaposition over two major search engines: Google and DuckDuckGo using a variety of search terms from multiple categories. In general, we find that both search engines are influenced, albeit differently, by our construed situational search context in a manner that is indicative of personalisation biases.

2   Related Work

While personalisation for the web and for search services has been explored and practised for the past two decades [7,26], its downsides have been equally established [28]. Personalised search engines help people to focus and increase their effectiveness, but they also potentially overexpose their users with information experiences that are highly aligned to their long-standing digital profiles. Pariser [28] coined the term "Filter Bubble" to describe this effect, defining it as "the personal ecosystem of information that's been created by the personalisation algorithms". He argued that Google's personalisation algorithms provide users with information that reinforces their ideas and hides the information that opposes their viewpoints, thus, decreasing the diversity of their views. Due to this interference, users might not see the contrasting viewpoints on a moral or political issue [6]. As a result, they will be trapped in a filter bubble without even knowing what they are missing [28]. This may lead to fewer serendipitous information encounters in the short term, and narrower views, informational blind spots, or radical polarisation in the longer term [4,28]. Awareness of bias in news and media has gained substantial attention on its own [5,19,34] as well as in relation to personalised search and news services [10,12,15,23,32].

While relevance and link structure of online resources are decisive factors in determining the placement of search results, studies indicate that several other factors such as politics [21], economics and social biases [2], etc. play a role in ranking and may lead to biased results [29]. Bias based on geographical location also occurs because popular search engines are in the USA. A study by Vaughan and Thelwall [40] testing three main search engines for national bias discovered that websites based in the USA were much better covered. A new study by Cooper et al. [9] identified significant variations when extracting scientific articles for composing review papers. The study compared results from

the same queries across 12 countries. Some of its geographical locations (based on the IP address) suppressed more than half of its relevant results. Bias can also be caused by search engines showing popular search results first [17] and learning from user click behaviors. Google's auto-complete feature has also been shown to be biased towards more popular searches and sometimes offers some questionable choices[1]. White [42] investigated inherent search engine biases and their effect on information quality. He showed that half of the time, the combined effect of inherent biases and user preferences leads people to incorrect beliefs. Epstein and Robertson [12] investigated the impact of search results on the election outcome and showed that voting preferences of undecided users can change by at least 20% due to biases in search results. At present, search engines retrieve and present biased information to users. Google, for instance, provides personalised search results based on ∼57 different signals including user's search history, location, past click behavior, etc. [28]. Thereby, creating a filter bubble by limiting the search results that we get for a particular topic.

Our work differs in a number of ways from previous research. While the prior work [12,20,21,24,31,32,39,41] aimed to quantify personalisation bias in web search, the studies were rather limited to the political searches only. Furthermore, the authors did not control the noise in search results i.e. the carryover effect [15]. Besides, some other studies [9,18,27] also focused on search bias quantification, nevertheless, only single-user features such as geolocation was considered. For instance, Cooper et al. [9] used a virtual private network (VPN) to conduct the same Google searches in 12 different countries to study the impact of users' geographical location on returned results. The authors find that the user's location appears to be influencing the results returned in response to the searches conducted for systematic reviews. Likewise, Silver et al. [18] conducted a series of searches over a period of 30 days using 240 queries. The results collected from the 59 GPS coordinates in the US revealed that location-based personalisation leads to ∼40–50% change in search results for localised queries and a minimal change in results for more general queries. Similarly, the impact of location on search results personalisation was studied in [27], however, only image search results were considered and the queries were also kept limited to Covid-19. The authors conducted the same search experiments in four different parts of Europe and compared results across the countries. Surprisingly, they only found a ∼46% overlap in search results, which became minimal when the queries were expressed in different languages. Other researchers [14,23] quantified search bias in Google News, thereby, limiting the scope of their research to news outlets.

Our research, compared to the previous works, has a wider scope. It spans a wide category of search terms, includes various user features, and also controls for noise. Comparing it with the work of Hannak et al. [15], we maintain the IP address as a control variable. Hannak et al. also focused on training the browser

---

[1] The Wired article from 2018 reports on this issue https://www.wired.com/story/google-autocomplete-vile-suggestions. Note that Google's auto-complete suggestions can now be reported. Nevertheless, the issue remains relevant.

profiles to represent various demographic properties (e.g. age, gender, ethnicity), however, we focused on training our browser profiles to absorb the history of search and browsing behaviour. Furthermore, we included DuckDuckGo as a second, more neutral counterpart. We chose DuckDuckGo mainly because of two reasons: first, it claims to respect users' privacy and does not track them during web search sessions, and second, it is the most widely used privacy-protecting search engine. About 35 billion queries were searched on DuckDuckGo during the year 2021, with a monthly average of 3 billion searches and a daily average of up to 101 million searches[2]. In addition to the choice of search engines, we also used different search terms to measure personalisation. It is reported in the literature that the magnitude of personalisation varies with search terms [15,21]. Lastly, the personalisation algorithms of Google have a changing nature. It is known that Google continuously updates its data sources and personalisation algorithms over time. For instance, there have been changes in Google's privacy policy in recent years [13,30], which allowed Google to aggregate users' data throughout its services (e.g., Gmail, Search, DoubleClick, Google Analytics, etc.) for content personalisation and targeted advertising. Therefore, using a third-party tracking and analytics network, Google now infers users' browsing history and personalises search results in a more effective way. We believe these changes in privacy policy have also produced a need for a more recent study on the subject as a significant amount of time has already passed since the prior research was conducted.

3   Methodology

Table 1. Search terms from four different categories

Category	Queries
News (6)	Corona virus; stimulus check; boris johnson; public health england; Weather tomorrow; democratic debate
Health (6)	Pandemic; covid-19 antibody testing; face masks; quarantine; Hand sanitizer; immune system
Sports (3)	FIFA 2022, ICC Men's T20 World Cup; kobe bryant;
Science (5)	Brain out; weathering with you; locust, asteroid; What dinosaur has 500 teeth

We investigate three situational aspects of search: cookies, past search, and past browsing history; corresponding to three separate experiments. All experiments use both the Google and the DuckDuckGo search services with the query collection shown in Table 1. Default settings were used in both search engines. All experiments were conducted with an in-house tool built on PhantomJS[3],

a headless-browser framework that allows simulating real user interaction with search engines collecting SERP in real-time. We avoided search engine APIs as they have been suspected of presenting results differently [25]. All the experiments run during the summer of 2020 in Dublin, Ireland.

The ***cookie-tracking experiment*** captures any personalisation bias that is driven by user information collected, stored, and maintained in browser cookies. Search engine providers can use cookies to create a user model even though the user is currently not logged into their ecosystem [15]. To evaluate the impact of cookies, we conducted a series of web searches during which all cookies throughout the search session were either enabled or disabled.

The ***search-history experiment*** investigates the personalisation of search results over time. We conducted a series of web searches once per day, for four consecutive days, once with cookies enabled and once with cookies disabled. Experiments run every day from 12 noon GMT for approximately 5 h.

The ***browsing-history experiment*** reviews the interactive effects of search personalisation. We examined whether Google and DuckDuckGo personalise search results based on users' browsing history outside the usual search activity. First, all news-related queries (see Table 1) were searched. The script then browsed four news domains from four countries (`dw.com` (Germany), `news.com.au` (Australia), `cbc.ca` (Canada), and `scmp.com` (China)) and followed two random links on each portal to simulate a brief episode of shallow browsing. The script then ran all news-related queries again and compared SERPs with the earlier search results.

All experiments used a set of 20 queries covering four topics (news, health, sports, and science, as shown in Table 1). Similar to other researchers (e.g. [15, 16]), we selected queries from Google Trends[4], and WebMD[5] for the health-related topics. Google Trends was chosen as a platform for query collection as it shows the queries that remained popular over a particular period of time and also sorts the queries based on different categories, geographical locations, etc. We chose the queries that were trending in the last year but did not limit query selection to any particular region/city, that is, the selected region was "Worldwide". Between each subsequent search in all experiments, our script waited for 15 min to prevent any "carry-over effects" [15].

4   Results

We measured search result personalisation as the difference in URLs (links to the target page) between two SERPs. We only consider the first SERP from both search engines, based on prior findings [3] which showed that users often limit interactions to the first SERP. If 1 of the 10 search results (web links) differed across the two SERPs, we define the personalised difference to be 10% regardless of any ordering differences.

---

[4] https://trends.google.com/trends/.
[5] https://www.webmd.com/.

As part of the *cookie-tracking experiment* we executed our set of queries in a single session on both search engines once with cookies enabled, and once with cookies being cleared between individual queries. We found that cookie-based personalisation with Google is relatively high (∼37%) in comparison to DuckDuckGo (∼20%). This implies that Google changed on average 3 or 4 results, while DuckDuckGo adapted about 2 results between these two conditions. Results from the *search-history experiment* are more differentiated and therefore depicted in Fig. 1. Here, our query collection was repeatedly submitted over four consecutive days. During this time, personalisation ranges from 28% to 41% (3–4 adapted results) with Google and 9% to 28% (1–3 adapted results) with DuckDuckGo. Specifically, search-history-based personalisation with cookies ranges from 32–35% for Google and 12–28% for DuckDuckGo. Without cookies, personalisation varied from 28–41% (Google) and 9–27% (DuckDuckGo). The upper two lines in Fig. 1 show the differences in personalised results for Google, whereas the lower two lines show the variations in personalisation for DuckDuckGo. Note that the first day is used as a reference, and is therefore 0 for all cases.
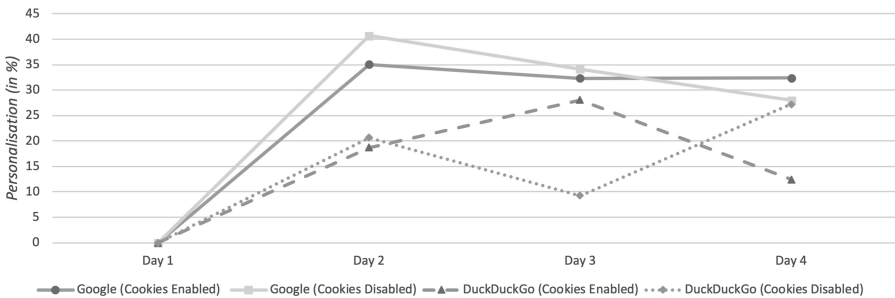

Fig. 1. Search-History Based Personalisation in Google and DuckDuckGo (in %)

The *browsing-history experiment* reviewed the impact of simulated shallow browsing on search result personalisation. Usually, SERPs provide a section that shows the latest news in relation to a submitted query. During this experiment, our script executed only news-related queries both before and after browsing the links on four different news domains. Neither Google's "Top Stories" nor DuckDuckGo's "Recent News" revealed personalised adaptations in response to our simulated browsing behaviour. However, Google returned localised results while DuckDuckGo remained neutral. This suggests that DuckDuckGo does not use IP addresses as a personalisation signal which supports DuckDuckGo's claim of not tracking its users. Nevertheless, our results show some evidence that Duck-DuckGo may use other signals to personalise search results: e.g. search history.

Additionally, we found that very few search results were commonly shared between search services – on average between 2–8%, as shown in Fig. 2. Specifically, Google and DuckDuckGo share only about 2% and 4% of their results in the news and health categories. While for the other two categories (science and sports), this percentage is slightly higher (∼6% with cookies disabled, and ∼8%
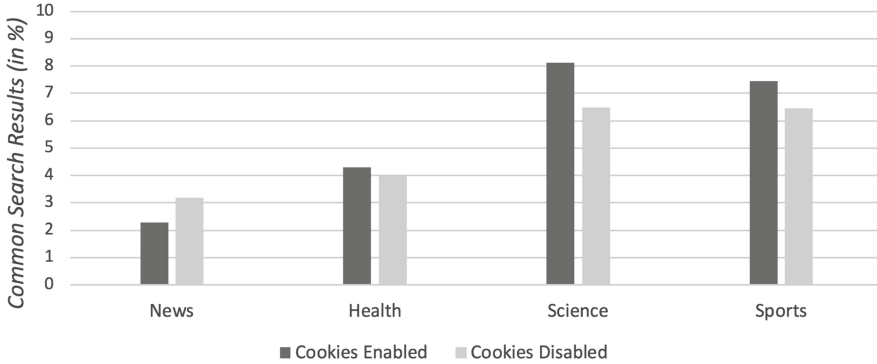
Fig. 2. Overlap between Google and DuckDuckGo results, based on queries in the four topical categories (in %)

with cookies enabled). To the best of our knowledge, this is a finding that has not been investigated previously and it is generally surprising that there is that little in common between Google and DuckDuckGo, even on rather objective queries covering categories such as science or health-related topics. Prior studies, however, focused on measuring the overlap between other search engines – e.g. Google, MSN, Yahoo, and Ask Jeeves [38], and Google and Bing [1], with [38] finding only a minimal overlap of about 1%. In a later study, Spink et al. [37] found less commonality in the first page results of four search engines compared to their previous study. Similarly, Ding and Marchionini [11] studied the distinctiveness in the search results (of InfoSeek, Lycos, and OpenText), and Selberg and Etzioni [36], who conducted a study to measure the overlap in search results (of Galaxy, Infoseek, Lycos, OpenText, Webcrawler and Yahoo) both found that the search engines returned the results that were unique to each other.

We have presented a methodology to evaluate personalisation bias in common search engines for three different situational variables: 1) browser cookies, 2) users' search history, and 3) users' browsing history. Our results show that Google adapts on average about 40% of its first results page, whereas DuckDuckGo adapts about 20%. Even though DuckDuckGo claims that it does not track its users, we found the service appears to perform certain forms of personalisation in response to different situational variables, and that this spans multiple query categories. While our results indicate that users' search history influences SERP variation for both Google and DuckDuckGo, Google search results depicted increased levels of personalisation. This indicates that further research is needed to quantify the effects of search history on search personalisation. Shallow browsing appears to not significantly affect personalised results for both search services. However, as we have only simulated a simple browsing episode, further research is needed to conclusively exclude this parameter as a potential source for personalisation bias.

5   Conclusion and Future Work

In this paper, we explored the potential for personalisation biases in search under different experimental user context settings: information stored in cookies, search history, and browsing history. Our results have shown that personalisation biases exist in both Google and DuckDuckGo, even if a user is not actively logged into the search engine ecosystem. As a result, users are consistently provided with adapted answers for their queries which may alter judgment and decision making [12,22,42]. While personalisation can be a useful measure to help people overcome handling an overabundance of information, we need to be aware of the cost of personalisation. This is less about users settling for "incorrect" answers, but rather the potential for over-exposure to one-sided viewpoints that reinforce beliefs on a potentially critical subject matter – a filter bubble that conveniently allows people to avoid learning alternative and competing views inhibit healthy information society.

Furthermore, all the previous studies in the literature find a small overlap in the first search result page of different search engines for a variety of search terms. There could be many reasons for this little overlap. First, there are constraints on the search engines in the portion of the web they index, owing to disk storage, computational power, and network bandwidth. Different technologies are used by search engines for finding the pages and indexing them. Furthermore, proprietary algorithms are deployed by search engines for determining the results' ranking and their demonstration to users. Hannak et al. [15] consider implicit personalization as a plausible reason. From our study, we form the opinion that the use of different search engines could be beneficial for users. It increases information viewpoint diversity since each search engine share a different perspective on a topic, therefore, the filter bubble effect can be mitigated using different search engines.

This work has derived its empirical findings via a simulation-based approach, a natural extension would be to use these findings to inform the design of a larger-scale user study to both corroborate and extend the findings. Similarly, there would be several additional user context variables that could be further explored. Key examples here are location, and web browser (as well as specific settings). An additional extension of this research involves exploring additional search engines, and also conducting similar experiments on popular news outlets such as the New York Times and the Washington Post for detecting bias in the provision of news stories.

References

1. Agrawal, R., Golshan, B., Papalexakis, E.: Overlap in the web search results of google and Bing. J. Web Sci. **2**, 17–30 (2016). https://doi.org/10.1561/106.00000005
2. Baeza-Yates, R.: Bias on the web. Commun. ACM **61**(6), 54–61 (2018). https://doi.org/10.1145/3209581, https://dl.acm.org/doi/10.1145/3209581
3. Bar-Ilan, J., Keenoy, K., Yaari, E., Levene, M.: User rankings of search engine results. J. Am. Soc. Inf. Sci. Technol. **58**(9), 1254–1266 (2007). https://doi.org/10.1002/asi.20608, http://doi.wiley.com/10.1002/asi.20608

4. Bierig, R., Caton, S.: Special issue on de-personalisation, diversification, filter bubbles and search. Inf. Retr. J. **22**(5), 419–421 (2019). https://doi.org/10.1007/s10791-019-09365-w, http://link.springer.com/10.1007/s10791-019-09365-w

5. Bourgeois, D., Rappaz, J., Aberer, K.: Selection bias in news coverage: learning it, fighting it. In: The Web Conference 2018 - Companion of the World Wide Web Conference, WWW 2018, pp. 535–543. Association for Computing Machinery, Inc., April 2018. https://doi.org/10.1145/3184558.3188724

6. Bozdag, E.: Bias in algorithmic filtering and personalization. Ethics Inf. Technol. **15**(3), 209–227 (2013). https://doi.org/10.1007/s10676-013-9321-6, http://link.springer.com/10.1007/s10676-013-9321-6

7. Brusilovsky, P., Maybury, M.T.: Special issue: from adaptive hypermedia to the adaptive web. Commun. ACM **45**(5), 30–33 (2002). https://doi.org/10.1145/506218.506239

8. Clarke, C.L., et al.: Novelty and diversity in information retrieval evaluation. In: Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR 2008, p. 659. ACM Press, New York, New York, USA (2008). https://doi.org/10.1145/1390334.1390446, http://portal.acm.org/citation.cfm?doid=1390334.1390446

9. Cooper, C., Lorenc, T., Schauberger, U.: What you see depends on where you sit: the effect of geographical location on web-searching for systematic reviews: a case study. Res. Synth. Methods (2021). https://doi.org/10.1002/jrsm.1485

10. Dillahunt, T.R., Brooks, C.A., Gulati, S.: Detecting and visualizing filter bubbles in google and Bing. In: Conference on Human Factors in Computing Systems - Proceedings, vol. 18, pp. 1851–1856. Association for Computing Machinery, New York, New York, USA, April 2015. https://doi.org/10.1145/2702613.2732850, http://dl.acm.org/citation.cfm?doid=2702613.2732850

11. Ding, W., Marchionini, G.: A comparative study of web search service performance. In: ASIS Annual Meeting, pp. 136–42 (1996)

12. Epstein, R., Robertson, R.E.: The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. Proc. Natl. Acad. Sci. U. S. A. **112**(33), E4512–E4521 (2015). https://doi.org/10.1073/pnas.1419828112, https://www.pnas.org/content/112/33/E4512

13. Google: Google's Privacy Policies. https://policies.google.com/privacy/archive?hl=en-US

14. Haim, M., Graefe, A., Brosius, H.B.: Burst of the filter bubble?: Effects of personalization on the diversity of google news. Digit. Journal. **6**(3), 330–343 (2018). https://doi.org/10.1080/21670811.2017.1338145

15. Hannak, A., et al.: Measuring personalization of web search. In: WWW 2013–Proceedings of the 22nd International Conference on World Wide Web, pp. 527–537. ACM Press, New York, New York, USA (2013). https://doi.org/10.1145/2488388.2488435, http://dl.acm.org/citation.cfm?doid=2488388.2488435

16. Hoang, V.T., Spognardi, A., Tiezzi, F., Petrocchi, M., De Nicola, R.: Domain-specific queries and web search personalization: some investigations. In: Electronic Proceedings in Theoretical Computer Science, EPTCS, vol. 188, pp. 51–58. Open Publishing Association, August 2015. https://doi.org/10.4204/EPTCS.188.6

17. Introna, L.D., Nissenbaum, H.: Shaping the web: why the politics of search engines matters. Inf. Soc. **16**(3), 169–185 (2000). https://doi.org/10.1080/01972240050133634

18. Kliman-Silver, C., Hannak, A., Lazer, D., Wilson, C., Mislove, A.: Location, location, location: the impact of geolocation on web search personalization. In: Proceedings of the 2015 Internet Measurement Conference, pp. 121–127. IMC 2015, Association for Computing Machinery, New York, NY, USA (2015). https://doi.org/10.1145/2815675.2815714

19. Knoche, M., Popović, R., Lemmerich, F., Strohmaier, M., Stroh-maier, M.: Identifying biases in politically biased wikis through word embeddings. In: Proceedings of the 30th ACM Conference on Hypertext and Social Media. ACM, New York, NY, USA (2019). https://doi.org/10.1145/3342220.3343658

20. Krafft, T.D., Gamer, M., Anna, K.: What did you see? A study to measure personalization in Google's search. EPJ Data Sci. **8**, 1–23 (2019). https://doi.org/10.1140/epjds/s13688-019-0217-5, https://epjdatascience.springeropen.com/articles/10.1140/epjds/s13688-019-0217-5

21. Kulshrestha, J., et al.: Search bias quantification: investigating political bias in social media and web search. Inf. Retr. J. **22**(1–2), 188–227 (2019). https://doi.org/10.1007/s10791-018-9341-2, https://link.springer.com/article/10.1007/s10791-018-9341-2

22. Lai, C., Luczak-Roesch, M.: You can't see what you can't see: experimental evidence for how much relevant information may be missed due to Google's web search personalisation. In: Weber, I., et al. (eds.) SocInfo 2019. LNCS, vol. 11864, pp. 253–266. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-34971-4_17

23. Le, H., Maragh, R., Ekdale, B., High, A., Havens, T., Shafiq, Z.: Measuring political personalization of google news search. In: The World Wide Web Conference on–WWW 2019, pp. 2957–2963. Association for Computing Machinery (ACM), New York, New York, USA (2019). https://doi.org/10.1145/3308558.3313682, http://dl.acm.org/citation.cfm?doid=3308558.3313682

24. Martinovic, M.: Exploring the effect of search engine personalization on politically biased search results (2018)

25. McCown, F., Nelson, M.L.: Agreeing to disagree: Search engines and their public interfaces. In: Proceedings of the ACM International Conference on Digital Libraries, pp. 309–318. ACM Press, New York, New York, USA (2007). https://doi.org/10.1145/1255175.1255237, http://portal.acm.org/citation.cfm?doid=1255175.1255237

26. Micarelli, A., Gasparetti, F., Sciarrone, F., Gauch, S.: Personalized search on the world wide web. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The Adaptive Web. LNCS, vol. 4321, pp. 195–230. Springer, Berlin, Heidelberg (2007). https://doi.org/10.1007/978-3-540-72079-9_6

27. Paramita, M.L., Orphanou, K., Christoforou, E., Otterbacher, J., Hopfgartner, F.: Do you see what i see? Images of the COVID-19 pandemic through the lens of google. Inf. Process. Manag. **58**(5), 102654 (2021). https://doi.org/10.1016/j.ipm.2021.102654, https://www.sciencedirect.com/science/article/pii/S0306457321001424

28. Pariser, E.: The Filter Bubble: What the Internet Is Hiding from You. The Penguin Group, London (2011)

29. Pitoura, E., et al.: On measuring bias in online information. SIGMOD Rec. **46**(4), 16–21 (2018). https://doi.org/10.1145/3186549.3186553

30. ProPublica.: Google Has Quietly Dropped Ban on Personally Identifiable Web Tracking (2016). http://bit.ly/2eAjC9w

31. Puschmann, C.: Beyond the bubble: assessing the diversity of political search results. Digit. Journal. **7**(6), 824–843 (2019). https://doi.org/10.1080/21670811.2018.1539626

32. Robertson, R.E., Lazer, D., Wilson, C.: Auditing the personalization and composition of politically-related search engine results pages. In: The Web Conference 2018–Proceedings of the World Wide Web Conference, WWW 2018, pp. 955–965. Association for Computing Machinery Inc, New York, New York, USA, April 2018. https://doi.org/10.1145/3178876.3186143, http://dl.acm.org/citation.cfm?doid=3178876.3186143

33. Sakai, T., Kando, N., Macdonald, C., Soboroff, I.: Introduction to the special issue on search intents and diversification. Inf. Retr. **16**(4), 427–428 (2013). https://doi.org/10.1007/s10791-013-9223-6, http://link.springer.com/10.1007/s10791-013-9223-6

34. Sales, A., Balby, L., Veloso, A.: Media bias characterization in Brazilian presidential elections. In: HT 2019 - Proceedings of the 30th ACM Conference on Hypertext and Social Media, pp. 231–240. Association for Computing Machinery, Inc., September 2019. https://doi.org/10.1145/3342220.3343656

35. Santos, R.L.T., Macdonald, C., Ounis, I.: Search result diversification. Found. Trends® Inf. Retr. **9**(1), 1–90 (2015). https://doi.org/10.1561/1500000040

36. Selberg, E., Etzioni, O.: Multi-service search and comparison using the MetaCrawler. In: 4th International Conference on World Wide Web (1995)

37. Spink, A., Jansen, B.J., Wang, C.: Comparison of major web search engine overlap: 2005 and 2007. In: 14th Australasian World Wide Web Conference

38. Spink, A., Jansen, B.J., Blakely, C., Koshman, S.: A study of results overlap and uniqueness among major web search engines. Inf. Process. Manag. **42**(5), 1379–1391 (2006). https://doi.org/10.1016/j.ipm.2005.11.001, https://linkinghub.elsevier.com/retrieve/pii/S0306457305001500

39. Urman, A., Makhortykh, M., Ulloa, R.: The matter of chance: auditing web search results related to the 2020 U.S. presidential primary elections across six search engines. Soc. Sci. Comput. Rev. 08944393211006863. https://doi.org/10.1177/08944393211006863, https://doi.org/10.1177/08944393211006863

40. Vaughan, L., Thelwall, M.: Search engine coverage bias: evidence and possible causes. Inf. Process. Manag. **40**(4), 693–707 (2004). https://doi.org/10.1016/S0306-4573(03)00063-3, https://linkinghub.elsevier.com/retrieve/pii/S03064573000633

41. Weber, I., Garimella, V.R.K., Borra, E.: Mining web query logs to analyze political issues. In: Proceedings of the 4th Annual ACM Web Science Conference, pp. 330–334. WebSci 2012, Association for Computing Machinery, New York, NY, USA (2012). https://doi.org/10.1145/2380718.2380761

42. White, R.W.: Beliefs and biases in web search. In: SIGIR 2013–Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 3–12. ACM Press, New York, New York, USA (2013). https://doi.org/10.1145/2484028.2484053, http://dl.acm.org/citation.cfm?doid=2484028.2484053