



This is a repository copy of *Towards improving user awareness of search engine biases: a participatory design approach*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/202396/>

Version: Published Version

Article:

Paramita, M. orcid.org/0000-0002-9414-1853, Kasinidou, M., Kleanthous, S. et al. (3 more authors) (2024) Towards improving user awareness of search engine biases: a participatory design approach. *Journal of the Association for Information Science and Technology*, 75 (5). pp. 581-599. ISSN 2330-1635

<https://doi.org/10.1002/asi.24826>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

RESEARCH ARTICLE

JASIST WILEY

Towards improving user awareness of search engine biases: A participatory design approach

Monica Lestari Paramita¹  | Maria Kasinidou²  | Styliani Kleanthous^{2,3}  |
Paolo Rosso⁴  | Tsvi Kuflik⁵  | Frank Hopfgartner^{1,6} 

¹University of Sheffield, Sheffield, UK

²Open University of Cyprus,
Nicosia, Cyprus

³CYENS CoE, Nicosia, Cyprus

⁴Universitat Politècnica de València,
Valencia, Spain

⁵University of Haifa, Haifa, Israel

⁶Universität Koblenz, Koblenz, Germany

Correspondence

Monica Lestari Paramita, University of
Sheffield, Sheffield, UK.

Email: m.paramita@sheffield.ac.uk

Funding information

Horizon 2020 Framework Programme,
Grant/Award Number: 810105

Abstract

Bias in news search engines has been shown to influence users' perceptions of a news topic and contribute to the polarisation of society. As a result, there is a need for news search engines that increase user awareness of biases in the search results. While technical approaches have been developed to mitigate biases in search, very few studies have investigated *user* preferences in interface designs for potentially raising their awareness of biases in news search engines. In this study, we utilized a participatory design methodology to develop eight prototypes with different features that could potentially be used to raise user awareness of biases in news search engines. We conducted three user studies, involving 132 participants with Computer Science backgrounds, to evaluate these prototypes. Our findings indicate the importance of news search engines that (a) inform users of possible biases in the results (*bias visualization approach*) and (b) allow users to access alternative search results (*results-reranking approach*). Our study provides further insights into the strengths and possible risks of each approach, which are important for future research on designing interfaces for raising user awareness of biases in news search engines.

1 | INTRODUCTION

Media biases (e.g., political, framing, or coverage biases) have been shown to influence how a topic is represented in news articles (Hamborg et al., 2019). These biases may further be exacerbated when search engines are used to retrieve news articles, due to the lack of transparency in the search algorithms (Kordzadeh & Ghasemaghahi, 2022). Search engines have been shown to provide different results to different users (Paramita et al., 2021; Urman et al., 2022), prioritize results from certain sites (Introna & Nissenbaum, 2000; Makhortykh et al., 2020; Nechushtai & Lewis, 2019), or produce biased results that are discriminatory to the society (Noble, 2018). Unaware of these biases, search engine users

may often perceive their search results to be objective (Gillespie, 2014) and trustworthy, especially those in the top ranking (Pan et al., 2007). These biases have been shown to manipulate user understanding of unknown topics (Novin & Meyers, 2017), sway the decisions of undecided voters (Epstein & Robertson, 2015), and contribute to the ideological polarisation of news readers (Beam, 2014; Spohr, 2017).

Given users' reliance on search engines to retrieve news (Jürgens & Stark, 2022), there is a need for news search engines that raise user awareness of potential biases in their search engine results. Various studies in mitigating biases in search engines have proposed altering ranks of results (*results-reranking*) to incorporate

This is an open access article under the terms of the [Creative Commons Attribution](#) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Journal of the Association for Information Science and Technology* published by Wiley Periodicals LLC on behalf of Association for Information Science and Technology.

diversity in the results (e.g., Celis et al., 2018; Gao & Shah, 2019) or present results with different perspectives (e.g., Draws et al., 2020). Other studies opt to use visualizations (*bias visualization*) to increase user awareness of biases in search results (e.g., An et al., 2012; Papadakos & Konstantakis, 2020). Despite these developments, very few studies have incorporated a *user-in-the-loop* approach in understanding the user perspective when designing interfaces for raising user awareness of biases in search engines, and more importantly, investigating which bias mitigation approach is preferred by users. Evaluations of previous designs are also very limited, often restricted to fewer than 20 people (Hamborg et al., 2017; Park et al., 2011). The study aims to investigate different interface designs for communicating potential biases in news search engines using a participatory design methodology. This study answers two research questions:

RQ1. What design approaches would users of news search engines prefer to raise their bias awareness?

RQ2. What aspects of the designs are found to be the most valued by news search engine users?

We conducted three user studies to gather eight designs of potential search engine interfaces, which captured different features for bias awareness across two approaches: (a) to inform users of biases (*bias visualization*) and (b) to modify the ranking of retrieved items (*results-reranking*). After developing them into prototypes, we conducted three studies to evaluate each prototype, assess how each approach influences users' search tasks, and identify users' preferred approach in a news search engine. We view this work as a starting point towards a more in-depth, thorough research on developing approaches for raising user awareness of biases in news search engines.

2 | PREVIOUS WORK

Algorithmic biases have been shown to be a highly intricate issue that influences the trustworthiness of search engines (Noble, 2018). Biases in search engines can be influenced by many aspects, from how the data were created, indexed, ranked, and used by the users (Baeza-Yates, 2018). News aggregators, which collate news from different sites, have been shown to contain coverage biases in their inclusion (and exclusion) of specific media sites, or the ranking algorithms adopted (Bui, 2010). Editorial slant has been found to provide a biased coverage of a political campaign

(Druckman & Parkin, 2005), which may influence voting decisions (Epstein & Robertson, 2015). Outlets with different political slants may affect news readers' understanding of the topic (Mokhberian et al., 2020). Political coverage of politicians in news articles has also been shown to contain gender bias (Leavy, 2019). In the long term, these biases can influence public opinion and introduce political polarisation (Beam, 2014; Spohr, 2017), and may reinforce existing inequalities, such as racism (Noble, 2018), without user awareness (Gillespie, 2014).

Previous approaches for mitigating biases in search engines can be grouped into two categories. The first, *bias visualization*, aims to use visualizations to increase user awareness of possible biases in the results. Papadakos and Konstantakis (2020) created the bias goggles model, which allows users to explore bias characteristics of web domains (e.g., political sites) using user-defined concepts (e.g., political parties). Other studies utilized highlighting the use of slanted language in the results (Spinde et al., 2020) and commenting on parts of the text that are causing bias in a description of events (Hamborg, 2020; Spinde et al., 2020). Other studies also aimed to increase user awareness by presenting the bias level of specific media sources (An et al., 2012; Kevin et al., 2018) and informing users of their political leaning based on articles they read over time (Munson & Resnick, 2013). The second, *results-reranking*, aims to rerank retrieved items in order to reduce or remove biases in the results. Draws et al. (2021) identified that ranking biases strongly influenced users' attitudes towards a topic and therefore proposed reranking results to expose users to different perspectives on contentious issues. Exposing viewpoints to the users through news framing has also been adopted to propose a more balanced overview of news topics (Park et al., 2011). Other studies developed ranking algorithms to optimize fairness and relevance in the results, by ensuring that minority views/groups are represented in the results (Celis et al., 2018; Gao & Shah, 2019). These approaches allow results-reranking to be done automatically without any user input. Other studies focused on developing different designs for presenting search results to the users. Instead of providing the traditional list of search engine results (such as one adopted by Google News; see Figure 1a), All Sides¹ (Figure 1b) aims to produce a balanced news consumption by providing selected articles from news outlets with different political affiliations (i.e., left, center, and right). Ground News² further aggregated similar topics together to allow users to see the reporting from different perspectives in one view. Hamborg et al. (2017) proposed a matrix-based design to

¹<http://www.allsides.com/>

²<http://ground.news/>

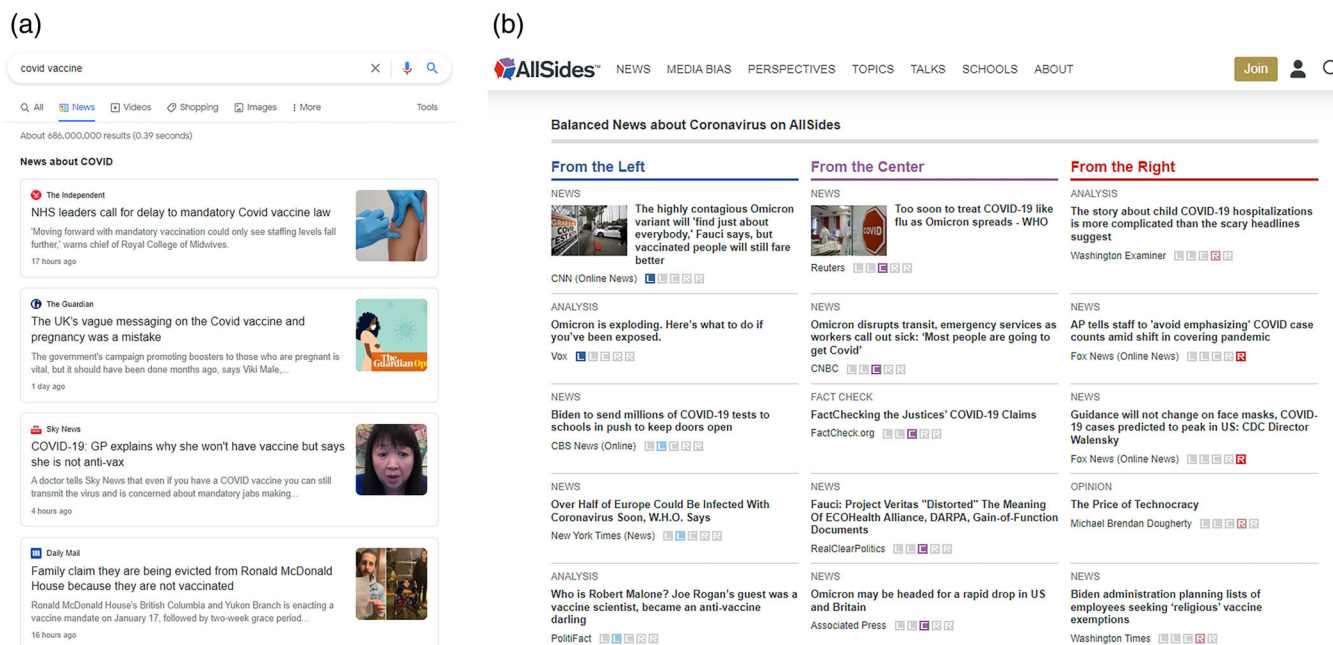


FIGURE 1 Presentation of results in news search engines

present search results for an international news topic involving multiple countries; this design allows users to retrieve news reported from different countries' perspectives (i.e., different publishers). Previous research often focused on one type of bias, such as political bias (An et al., 2012), or highlighted specific parts of the document that contain biased information (Spinde et al., 2020). Hence, these approaches cannot be utilized for informing users of *multiple* types of biases, which are often found in news search engine results. Moreover, there is a lack of understanding of which bias mitigation approaches users prefer, and how these approaches influence users in their search tasks. Our study focuses on gaining a better understanding of these areas.

3 | METHODS AND MATERIALS

We adopt a participatory design approach (Hussain et al., 2012) to involve end-users in designing and evaluating interactive prototypes developed for improving bias awareness in news search engines. We carried out our study in three stages. In Stage 1, we used a participatory approach to gather designs for raising user awareness of biases in news search engine results, which we later developed as prototypes (section 3.1). In Stage 2, we invited users to evaluate these prototypes and the underlying approaches, that is, bias visualization and results-reranking approach (section 3.2). Finally, in Stage 3, we analyzed the results (section 3.3). These methods are summarized in Figure 2.

3.1 | Stage 1: Design

Stage 1 (design workshop) aims to gather design ideas for potential interfaces that will raise bias awareness in search engines. Aiming for input from a culturally diverse group of users of search engines, with comparable lifestyles, we organized three (online) user studies with participants located in countries with different cultural traditions to each other, namely Israel, Italy, and Cyprus.

Methods: We started the online design workshops by giving a brief introductory presentation on bias in information retrieval (IR) to provide sufficient context, followed by an overview of the impact of bias on search engine users and some examples of political and gender bias. We then asked participants to imagine that they were using a news search engine to look for news related to "COVID-19." Participants worked in a group of 2–4 members to complete two activities: (a) to identify a list of biases that, in their opinion, should be highlighted by news search engines³ and (b) to create a mock-up search engine design (low-fidelity prototype) to inform users of these biases. We asked participants to suggest designs taking the users' perspective. It was, therefore, possible that participants suggested designs could be difficult—if not impossible—to implement, given the complex nature, and the limitations of available algorithmic methods for measuring biases. However, given the participatory

³Activity 1 was only intended to help participants identify the types of biases that should be included in the designs (Activity 2) and is not further analyzed in this study.

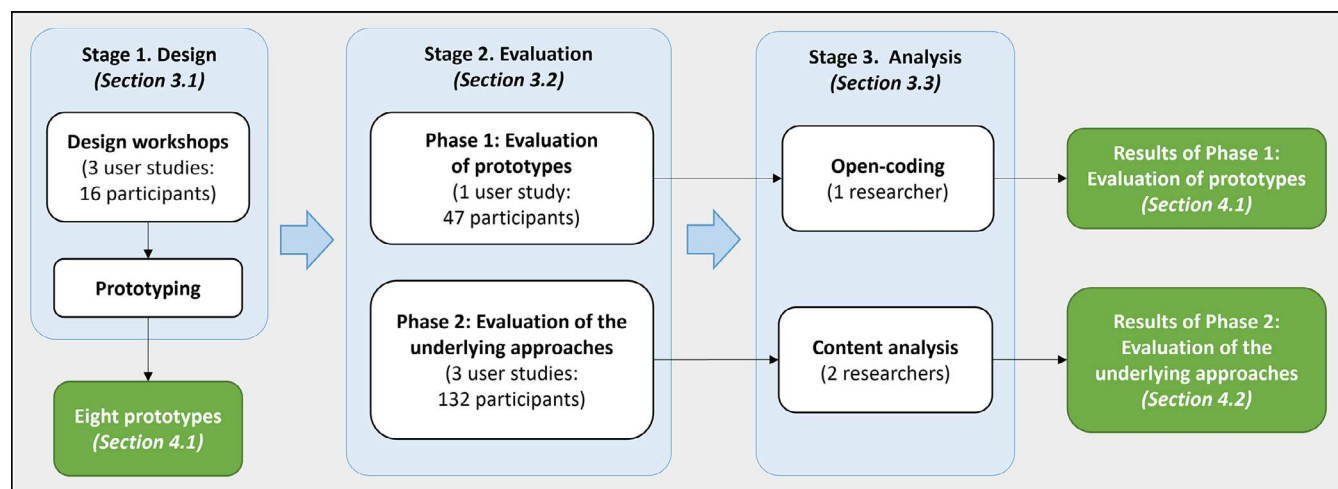


FIGURE 2 Methods

approach, we opted against intervening to avoid influencing the participants' design choices.

Participants were asked to produce a sketch of their proposed search interface in Google Slides. They were allowed to re-use available search engine interfaces (e.g., using screenshots), and to use any data/graphs (e.g., icons, diagrams) in creating their designs. They were also asked to provide a textual description of the features. An example of participants' designs is shown in Figure 3a, which was then developed into prototype V2 (see Section 4.1.2). In some cases, participants proposed a results-reranking approach (i.e., without any additional features in the search results) and described how reranking works in the text description (see Figure 3b), which later on was developed into prototype R4 (see Section 4.1.8).

After all three user studies were completed, two of the authors collated all the proposed designs, removed similar designs, and selected eight distinctive interfaces to develop in the study. Four design interfaces (System V1–V4) aim to inform users of biases in the results (i.e., *bias visualization approach*), and the remaining four interfaces (System R1–R4) aim to modify the ranking of retrieved items (i.e., the *results-reranking approach*).

Participants: A total of 18 participants took part in the design workshops, all of them were students at The University of Haifa, the University of Trento, or the University of Cyprus. Most participants (13) identified themselves as males, four were females, and one preferred not to say. Sixteen were between 18 and 30, and two between 31 and 50. Five were enrolled in a Bachelor's Degree, and the rest were postgraduate students (Masters and PhD). Most students came from a Computer Science background, and one studied Business Administration.

Prototyping: The selected designs were then developed into interactive prototypes using Proto.io, which

allowed multiple screens to be created and linked to one another to interactively simulate how the finished products will function. Given the role of Google as the world's most commonly used Web search engine,⁴ all the eight prototypes were modeled after Google's basic graphical user interface, with additional features added to the main page (e.g., Section 4.1.5) and/or results page (e.g., Section 4.1.1) as informed by the participants' designs. Users interact with the prototype by submitting their query on the main page, clicking "News" to view the news articles, and viewing the result pages. General pandemic-related queries were used in the study to make the news search experience more relatable for users across different countries. No algorithms were applied to measure biases or actually rerank the results. Instead, the results were manually created to demonstrate the functionality of each prototype.

3.2 | Stage 2: Evaluations

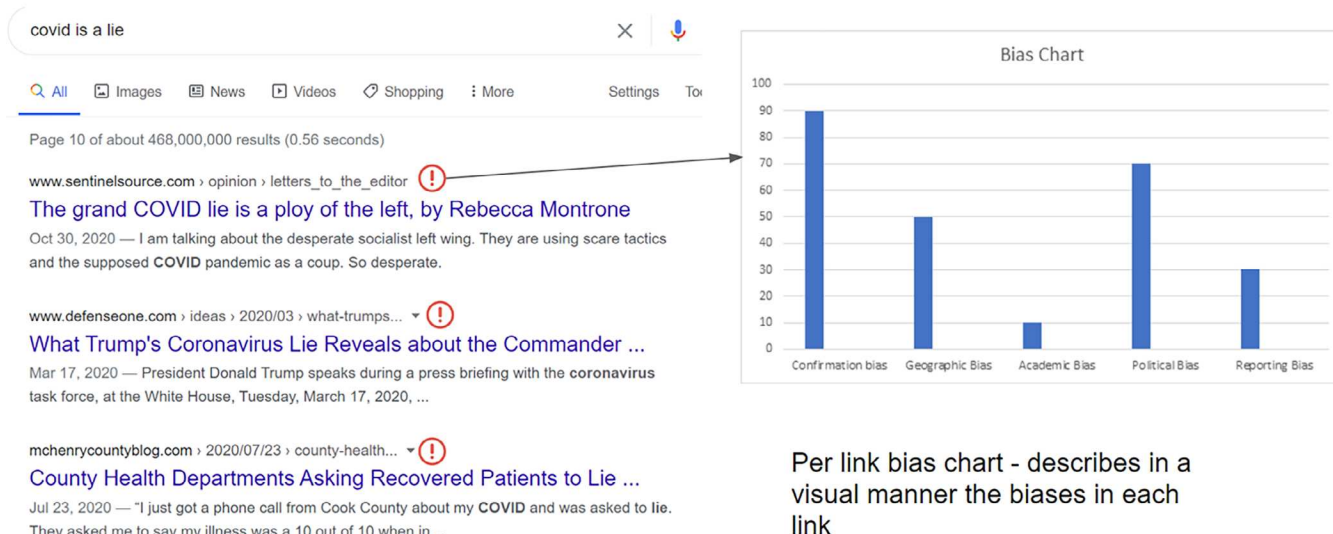
In Stage 2, we gathered users' evaluation of prototype interfaces designed in Stage 1. We divided the evaluation into two phases to gather more detailed feedback. Phase 1 evaluated the strengths and weaknesses of each prototype design. Phase 2 evaluated the underlying approaches.

3.2.1 | Phase 1: Evaluation of prototypes

Methods. We provided participants with online lecture recordings about search engine bias and a demo introducing the eight prototypes (Systems V1–V4 and Systems R1–R4). We then provided them with the link to the prototypes and guidelines on how to use them (e.g., what

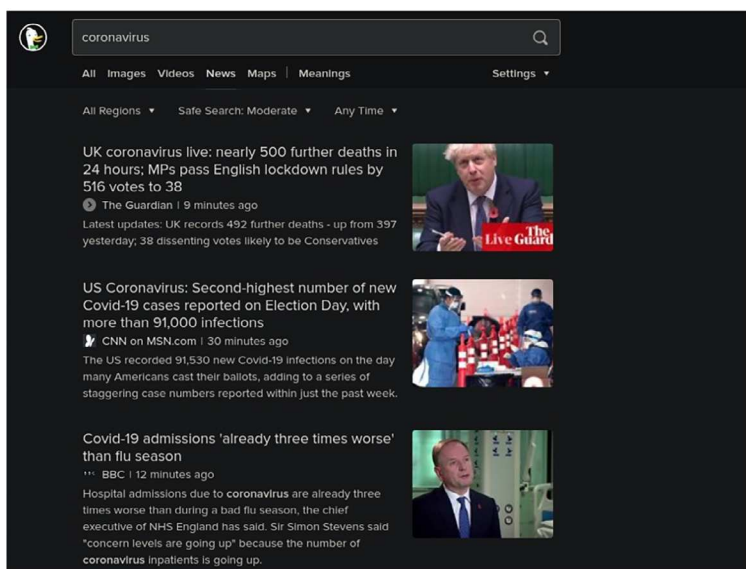
⁴<https://gs.statcounter.com/search-engine-market-share>

(a) Bias visualisation approach



Per link bias chart - describes in a visual manner the biases in each link

(b) Results-reranking approach



When searching for something the top 3 results will be most relevant to your search but the next 3 will try to be opposite.

FIGURE 3 Examples of designs created by participants in Stage 1

query to use and how to access the features). We used counter-balancing to present the prototypes to the participants to reduce the order effects. Participants were asked to engage with the prototypes and provide feedback on the features they liked and disliked for each system.⁵ We reminded the participants that (a) the prototypes were not working systems, (b) information about the biases presented might not necessarily relate to the content, and (c) the evaluation should be focused on the design and features, rather than the accuracy of the information itself.

Participants. The study was run as a part of the Information Retrieval module at the University of

Sheffield. A total of 47 MSc students participated in the study. All participants used search engines daily, with 74% participants rating their understanding of how search engines work to be advanced (5 or above on a 7-point Likert Scale). More information is shown in Table S1, Supporting Information.

3.2.2 | Phase 2: Evaluation of the underlying approaches

Methods. First, we delivered a short lecture to discuss a brief background of bias in information retrieval. Second, we described the two underlying approaches (i.e., bias

⁵The questionnaire is shown in Figure S1.

Q1)	How would bias visualisation approach influence your information seeking task? (<i>open-ended question</i>)
Q2)	How would you rank the bias visualisation prototypes (Rank1=best, Rank 4=worst)?
Q3)	How would reranking search results influence your information seeking task? (<i>open-ended question</i>)
Q4)	How would you rank the results-reranking prototypes (Rank1=best, Rank 4=worst)?
Q5)	Which approach would you like to see implemented by Web search engines? (<i>bias visualisation/results-reranking/others</i>)

FIGURE 4 Questionnaire (Phase 2)

visualization and results-reranking approaches) and demonstrated the eight prototypes. Finally, participants were asked to familiarize themselves with the prototypes before answering the evaluation questions (Figure 4). To obtain a broader range of feedback, we conducted three user studies that involved participants with different academic backgrounds. There were slight differences in how the activities were carried out (individual vs. group work) and how the contents were delivered (Study 1 used a mixture of recorded and synchronous online sessions, while Studies 2 and 3 used all synchronous online sessions). However, the same material and questionnaire were used across three studies.⁶ Studies 1 and 2 were run by the same researcher. Study 3 was run by a different researcher, who was present in Study 2, and therefore was aware of how the previous sessions were run. All sessions were run online due to COVID-19 restrictions.

Participants. A total of 85 responses from 132 participants were gathered across three studies:

- *Study 1:* Included 60 postgraduate students from the University of Sheffield. The evaluation was run during the Information Retrieval module. Participants worked in groups, resulting in a total of 26 group responses.
- *Study 2:* Included 16 participants (ranging from postgraduate students to practitioners and academics). This session was run as part of a Winter School on bias and transparency. Participants worked in groups, resulting in a total of three group responses.
- *Study 3:* Included 56 undergraduate students from the Universitat Politècnica de València. The evaluation session was run as part of their Natural Language and Information Retrieval module. Participants worked individually, resulting in a total of 56 individual responses.

We asked participants to complete a pre-questionnaire on their demographics and academic background.⁷ All participants were frequent search engine users (87%–91% used search engines daily). Participants specified having a moderate to advanced understanding of how search engines work (mean = 5.07, 5.21, and 4.89 for Studies 1, 2, and 3, respectively; 1 = *no understanding* and

7 = *excellent understanding*).⁸ The differences between their understanding levels are not statistically significant. None of the participants in the evaluation contributed to the design of the prototypes.

3.3 | Stage 3: Analysis

We analyzed participants' answers to Evaluation Phase 1 (i.e., identifying the strengths and weaknesses of the prototypes) using open-coding, which is often utilized as a first approach in thematic analysis to identify interesting concepts. Different from a thematic analysis which requires multiple annotators, open-coding does not require the use of multiple annotators (Kelly, 2009). Given the small amount of data gathered in this phase, this approach was sufficient for identifying interesting aspects noted across participants.

For Evaluation Phase 2 (i.e., evaluation of the underlying approaches), a content analysis was used to analyze the rich qualitative comments (Thomas, 2006). In cases where comments were not written in English,⁹ Google Translate was used to translate the comment into English prior to carrying out the analysis, which in most cases was sufficient to understand the aspect discussed in the comment. In cases where the translation was poor and difficult to understand, the original comments were manually translated by one of the researchers who spoke the language as a native speaker. Two researchers independently read the comments to familiarize themselves with the data, identified significant aspects of the responses, and generated initial categories. Both researchers then compared these categories, discussed any disagreements, and made the necessary amendments to reach a consensus on the finalized categories. Researchers then independently re-coded the comments using the finalized categories, allowing each comment to be recoded to multiple categories. The responses were then compared again and any disagreements were discussed and resolved. Finally, the categories (subthemes) were sorted into

⁸The distributions of scores across the three studies are shown in Figure S2.

⁹While we had encouraged participants to answer in English, some participants in Study 3 preferred to articulate their thoughts in their native language (i.e., Spanish).

⁶More details on the sessions are described in Table S2.

⁷A detailed overview of the participants' background is provided in Table S3.

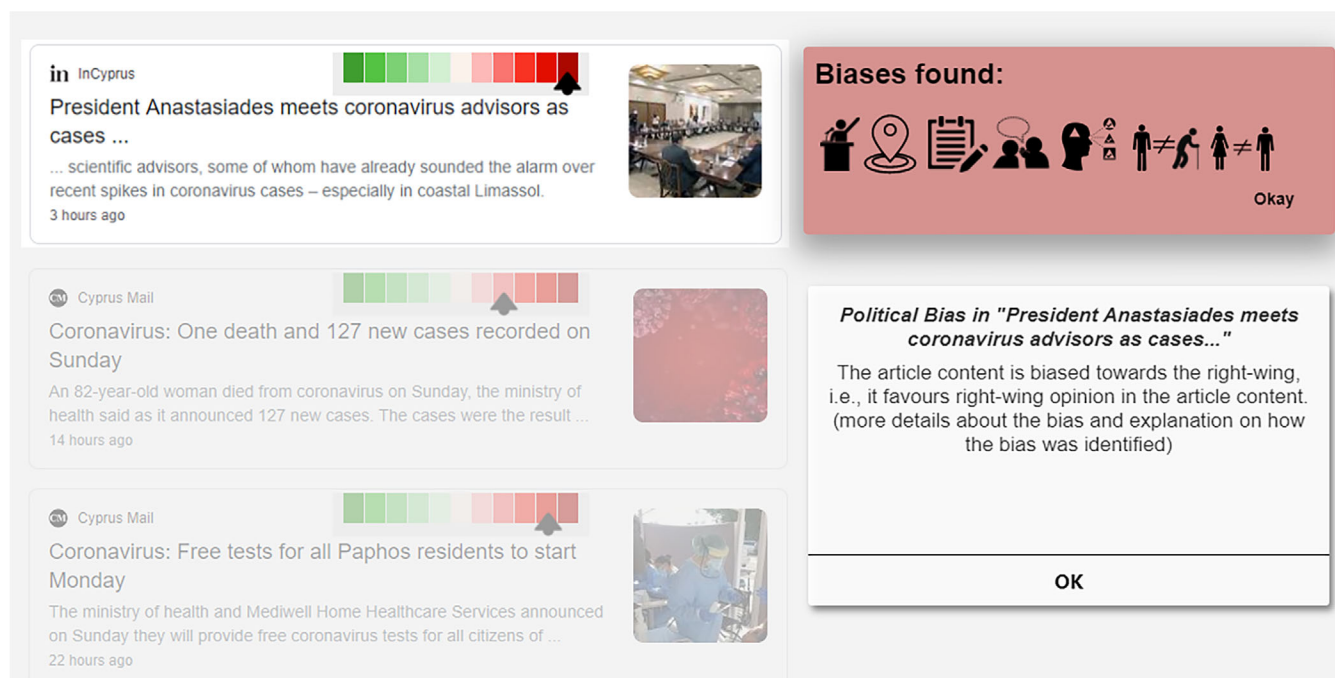


FIGURE 5 System V1

themes. Descriptive statistics were used to present these data to understand the magnitude of each category. Due to the small number of responses in Study 2, we did not perform any comparison between the studies.

4 | RESULTS

The proposed interface designs gathered in Stage 1 resulted in eight prototypes¹⁰ that can be categorized into *two underlying approaches*:

1. **Bias visualization** aims to improve user awareness of potential biases by visualizing them in the search results. Four different prototypes were designed: Visualization 1 to Visualization 4 (V1–V4).
2. **Results-reranking** aims to potentially rerank the results to allow users to retrieve alternative/diverse results. Four prototypes were developed to demonstrate this approach: Rerank 1 to Rerank 4 (R1–R4). Two of them (R1 and R4) *automatically* rerank the results without users' input, while the other two (R2 and R3) allow users to *manually* influence the results-reranking process.

We remind the readers that we only use the interface to demonstrate the approach and that we did not use any

algorithms to actually measure the biases or perform reranking of the results. To aid with readability, we describe the design of each prototype and the prototype evaluation (Phase 1) together in section 4.1. The evaluation of the underlying approaches (Phase 2) is described in section 4.2.

4.1 | Phase 1: Evaluation of prototypes

4.1.1 | Visualization 1 (V1)

Design. V1 displays biases for each result in a form of a “bias meter” (see Figure 5), ranging from green (no bias detected) to red (high level of bias detected). When the meter is clicked, a pop-up window shows a number of icons that display the types of biases found in the document. When hovering over an icon with the mouse, a message bubble appears presenting the type of bias presented by this particular icon (e.g., “political bias”). When the icon is clicked, a new window shows up to present more information about the biased aspect, for example, “the article is identified to be biased towards the right-wing.”

Strengths and weaknesses. Participants liked how V1 clearly displays the biases in the results (participant 11–p11, p17, p47), further highlighting that “*it shows how serious is the bias of the result*” (p13). Participants particularly liked the bias meter as it “*is very clear to see and understand*” (p43) and it “*indicates the amount of bias involved on each article*” (p34). Participants were concerned about the

¹⁰A larger screenshot of each system is shown in Figures S3–S10.

transparency and possible subjectivity of the methods used to identify these biases (p5, p7). Another participant pointed out that it was unclear how the different types of biases are aggregated into the bias meter and that the icons did not indicate the amount of bias involved in each article (p34). Some participants also pointed out that the amount of information provided may be confusing (p45) and time-consuming for novice users (p13).

4.1.2 | Visualization 2 (V2)

Design. V2 adopts a more minimalistic approach. Instead of a bias meter, an icon is shown on each result (see Figure 6). When clicked, a histogram opens to show the types of biases found in the result and the severity rate. When users hover over the bar, more information about the bias (e.g., definition) is displayed.

Strengths and weaknesses. Participants indicated that the histogram provided a clear insight into the amount of bias found in the article (p16, p18, p34, p36, p46, p47) and the different types of biases (p14, p38, p45). More specifically participants mentioned that *“the use of the bar chart gives a much clearer indication as to the amount of bias involved”* (p34), although others mentioned that the information was too detailed (p19, p20). Others pointed out that the degree of biases shown was not very intuitive (p14) and it *“(was) not clear what the charts measure”* (p45). Participants did not like the bias icon (red exclamation mark) (p11, p23, p26, p36, p38, p46) and would prefer something less ominous. Participants also disliked not being able to see the overall bias in the results (p13, p36, p44).

4.1.3 | Visualization 3 (V3)

Design. V3 informs users of related aspects that are not discussed in the article content. As shown as an example in Figure 7, when a user hovers over an article “COVID-19 vaccine: First person receives Pfizer jab in UK,” a notification appears to inform users that “Many companies are developing vaccine.” This allows search engine users to understand other viewpoints/aspects relating to the topic that might otherwise be unknown to them. When users click on the question mark icon, a new window shows up and displays more information about the article, for example, “This article is biased toward ‘Pfizer’ company. Alternative companies exist in related topics.”

Strengths and weaknesses. Participants liked the straightforward and simple design (p7, p22, p24), noting that *“[user] can quickly look through the content”* (p7) to see the biases in the article (p34). However, others mentioned that the information is too concise (p47) and does not provide the severity of the biases (p13, p22, p34). One participant noted the limitation of V3 is providing information on multiple biases, noting that *“it may become complicated to present if a listing has many types of biases”* (p44).


4.1.4 | Visualization 4 (V4)

Design. V4 omits biased articles from the results (as shown in Figure 8, third-ranked results). A similar approach has been adopted by Twitter with regard to tweets containing misleading information (Roth & Pickles, 2020). Users still have the option of seeing the



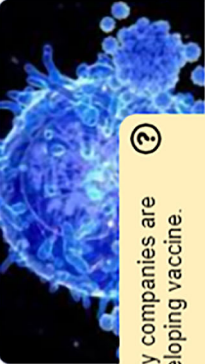
FIGURE 6 System V2

Top stories



Covid-19 vaccine: First person receives Pfizer jab in UK


1 hour ago



Many companies are developing vaccine.

"Mix-and-match" coronavirus vaccines to be tested

6 hours ago



First Britons receive Covid-19 vaccine, a landmark moment in the pandemic

2 hours ago


→

View all

FIGURE 7 System V3

There are few results that were omitted because it contains biased information. For more information about the identified biases click on the question mark.






President Anastasiades meets coronavirus advisors as cases ...

... scientific advisors, some of whom have already sounded the alarm over recent spikes in coronavirus cases – especially in coastal Limassol.

3 hours ago



Coronavirus: One death and 127 new cases recorded on Sunday

An 82-year-old woman died from coronavirus on Sunday, the ministry of health said as it announced 127 new cases. The cases were the result ...

14 hours ago

This search result was omitted because it contains biased information...

[Click here if you want to see the article](#)

FIGURE 8 System V4

information if they want to by clicking on the article. Users can also click the question mark icon to get more information about the biases identified.

Strengths and weaknesses. Participants liked the feature of hiding the biased articles (p14, p17, p22, p38), and others also found it simple and easy to use (p8, p13,

p44). Some participants also discussed that V4 “saves time, cause it hides some bias information” (p5). A few participants were concerned that users might miss important relevant information if some results were hidden (p5, p17, p46), but liked that users had the option to see the hidden articles (p36, 45, p47). Participant 35 mentioned that the biggest strength of V4 is that “[it] informs the user there are biased results and it allows them to make a decision towards exploring both biased and unbiased results.” Other weaknesses noted were that there was no information on the detected biases to help users make an informed decision on whether they want to see the articles or not (p44). Furthermore, a few participants pointed out that there was very limited control for the users with regard to specifying the types of biases they were interested in the results (p13, p35, p44).

4.1.5 | Rerank 1 (R1)

Design. R1 allows users to access a set of results (different from the original results) by incorporating a new search button. As shown in Figure 9, the button “I’m feeling unbiased” (on the right) can be used to *automatically* retrieve results that are identified to be unbiased. The search results (not shown here) look similar to figure in Section 4.1.8 but instead contain the notification: “The results are reranked and you are seeing only the most unbiased results.”

Strengths and weaknesses. Participants liked that this system reranked the results to show the most unbiased results at the top (p5, p7, p20, p34). Participants liked the clear layout (p1) and simple interface (p8) and that it was easy to use (p13, p33, p43, p44). However, others did not like that any biased information is hidden without the ability to customize the results (p2, p17, p36), specifying that “[the] user has no control within the interface over switching off/on unbiased results” (p35) and that

“it’s ‘all-in or all-out,’ I cannot see what sort of biases there are/were” (p44). Another noted that it could be difficult to define what unbiased results are (p34).

4.1.6 | Rerank 2 (R2)

Design. R2 allows users to *manually* define specific aspects that they would like to see in the results (Figure 10). In the prototype, only four aspects are able to be modified: geographical bias, gender bias, age bias and political bias. For example, if users want to see news that are politically biased to the “right-wing,” they are able to modify the value of “political affiliation” accordingly. This approach does not reduce or remove those biased contents, but provides users with the control and the awareness that the results they see are biased to the aspects that they formerly specified. This approach also allows users to easily view results from other aspects using a few clicks.

Strengths and weaknesses. Participants liked that the filters allow users to customize their results using different aspects (p7, p10, p36) based on their preferences (p13, p17, p24, p34, 45). As noted by participant 34, “the ability to filter the search allows for a lot of personalisation and allows the user to influence their search.” Other participants, however, pointed out the risk of polarising users “if people choose filters that suit only their preferences” (p45). Others commented that the filters were not comprehensive enough (p14, p17, p43).

4.1.7 | Rerank 3 (R3)

Design. Similar to R2, R3 also requires user input to rerank the results. This can be accessed by clicking “Customise my search,” in which a pop-up window will open that lists the different types of biases

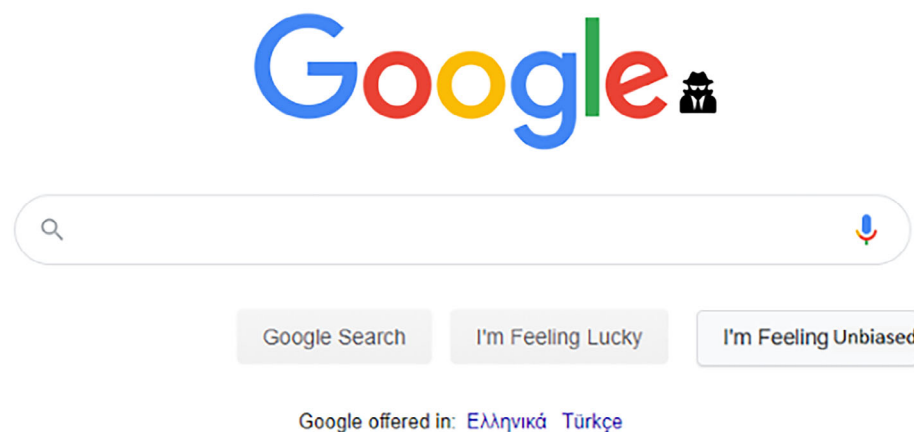


FIGURE 9 System R1

FIGURE 10 System R2

Country ▾ Gender ▾ Age ▾ Political affiliation ▾ ?

in InCyprus

President Anastasiades meets coronavirus advisors as cases ...

... scientific advisors, some of whom have already sounded the alarm over recent spikes in coronavirus cases – especially in coastal Limassol.

3 hours ago

CM Cyprus Mail

Coronavirus: One death and 127 new cases recorded on Sunday

An 82-year-old woman died from coronavirus on Sunday, the ministry of health said as it announced 127 new cases. The cases were the result ...

14 hours ago

default
Left-wing
Centre
Right-wing

About 1,540,000,000 results (0.76 seconds) Customise my search

in InCyprus

President Anastasiades meets coronavirus advisors as cases ...

... scientific advisors, some of whom have already sounded the alarm over recent spikes in coronavirus cases – especially in coastal Limassol.

3 hours ago

CM Cyprus Mail

Coronavirus: One death and 127 new cases recorded on Sunday

An 82-year-old woman died from coronavirus on Sunday, the ministry of health said as it announced 127 new cases. The cases were the result ...

14 hours ago

CM Cyprus Mail

Coronavirus: Free tests for all Paphos residents to start Monday

The ministry of health and Mediwell Home Healthcare Services announced on Sunday they will provide free coronavirus tests for all citizens of ...

22 hours ago

Bias Customization ?

0 100 Confirmation bias ?

0 100 Information bias ?

0 100 Political bias ?

0 100 Gender bias ?

0 100 Racial bias ?

0 100 Age bias ?

0 100 Geographical bias ?

0 100 Business/Capitalistic bias ?

0 100 Reporting bias ?

OK

FIGURE 11 System R3

identified (see Figure 11). The types of biases used in this prototype were derived from the participant's design and previous studies, for example, Baeza-Yates (2018). Users can *manually* modify the value for each

bias. A low value means users would like to have results with no/low biases of that type. Or, alternatively, users can increase the bias level to include biased content in the results.

Strengths and weaknesses. Participants liked that R3 allowed them to customize the search (p7, p10, p20, p21, p35) and get more personalized results (p9, p32). While some liked the ability to fine-tune results selection in R3 (p5, p44, p45), for example, using the percentile system (p14), others highlighted that this introduced difficulties in understanding what the scale represents and how to quantify the different biases (p14, p36, p45). One participant (p44) further specified, “*it’s not clear to me how the various selection interact with each other, e.g., can I really filter for results that have 20% informational bias, 40% age bias and 10% racial bias?*”

4.1.8 | Rerank 4 (R4)

Design. R4 proposes an *automatic* reranking approach to include articles from alternative viewpoints in the search results. For example, when using the query “vaccine COVID-19,” the search results are reranked to ensure that articles presenting alternative viewpoints exist in the top results. This includes articles about how vaccines are able to save people (see rank 1 in Figure 12), and also contradictory articles on how vaccines are not enough to solve the pandemic (see rank 2). Similar to R1, this approach does not require any input from the users.

Strengths and weaknesses: Participants found the introduction of different viewpoints in R4 to be useful (p7, p17, p45, 46) as it “*may help the user expand on their knowledge*” (p35). Others, however, noted that these new results may compromise other aspects, such as relevance (p33) and timeliness (p14, p30). Others were also concerned that seeing alternative viewpoints may confuse users instead and did not find the ability to see all viewpoints to be particularly useful (p2, p17). One participant noted that in R4 “*it is hard to decipher when the re-ranking is happening and why*” (p34). A lack of control given to the users has also been noted as a weakness of R4 (p35, p44, p45), although as a result, many highlighted that the system was clear and easy to use (p43, p44).

4.2 | Phase 2: Evaluation of the underlying approaches

4.2.1 | Bias visualization approach

Influence to information seeking tasks. Three themes and nine subthemes (Table 1) were identified from participants’ responses to Q1 (Figure 4).

Awareness: Participants pointed out that this approach allows users to **obtain information**

The results are re-ranked and you are now able to see results from all the viewpoints

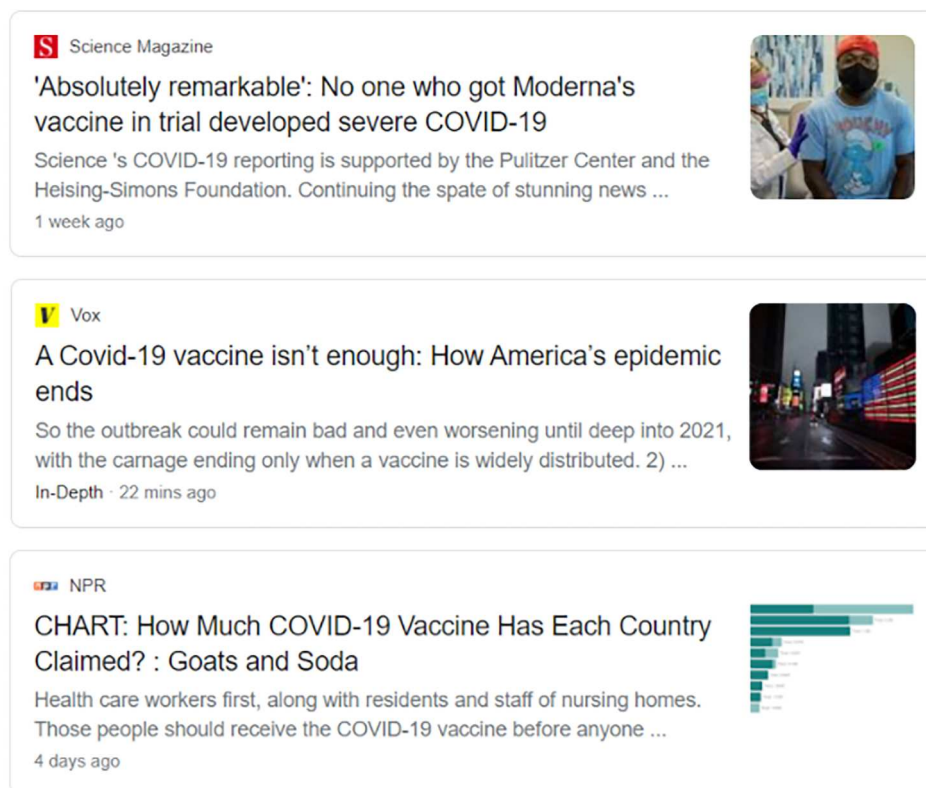


FIGURE 12 System R4

selectively, as discussed by 38 out of 85 responses. Participants pointed out that users can “choose whether to read the returned results” (study 1 response 12 – s1r12) and “read the article from a more critical perspective” (s3r44). Some participants also argued that sometimes users might want to read biased articles and mentioned examples such as coronavirus (s3r36) or politics (s3r38, s3r55). It comes as no surprise that one of the most discussed themes was **increase users’ awareness** (28 responses), noting that this approach can “make people notice and be aware of the bias present in their search results” (s3r1), and “raise awareness about the social and ideological bias that surrounds us” (s3r48). On the other hand, six responses mentioned that this approach may **reduce critical thinking** instead, as users would rely on the visualizations and not use their own judgments (s1r21).

Trustworthiness: Some participants were concerned about the possible **algorithmic bias** (14 responses) in the visualization itself, mentioning that any identified biases in the results may be inaccurate (s3r41), or biased (intentionally or not) (s3r50, s3r53), and may further mislead the users instead (s3r44).

Usability: Five subthemes were related to the usability of the approach. Eight responses mentioned that this approach allows users to **retrieve better results**, further making reference to the ability to retrieve more comprehensive (s1r9) and unbiased (s1r15) results that are “closer to what [users] want” (s1r4). However, others

argued that this approach can lead to **losing relevant information** instead, noting that the accuracy of the results is not guaranteed (s1r4) and that “[it can] filter out some truly good information” (s1r13). Eight responses mentioned that this approach was **easy to understand**, noting that the visualizations such as the use of icons and the bar charts (s3r45) make it easier to interpret (s1r19) and that “it is clear and more convenient to see the specific degree of different [biases]” (s1r14). However, eight responses disagreed and considered that this approach was **difficult to use**, leading to a reduced search efficiency (s1r24) since users would spend time checking which article to read (s1r12, s2r2). Some participants also were concerned that users with lower digital literacy (e.g., older users) might be facing difficulties in understanding and using this approach (s1r7). Six responses highlighted that this approach would **save time** for the users in identifying biases.

Preferred prototypes. Figure 13 shows that participants mostly preferred V1 (Figure 5), elaborating that it was easy to understand (s3r1) and participants liked the use of icons (s1r26), colors (s3r29, s3r37) and the bias bar (s3r54) to present the biases. Participants also ranked V2 (Figure 6) highly as the histogram allowed users to easily interpret the data (s3r13, s3r18) and “see what is the major bias in the article” (s2r23). Others preferred V2 over V1 because it further “quantifies the level of any types of biases” (s3r31). Some responses that ranked V3 highly

TABLE 1 Themes found in Q1 (influence of bias visualization approach)

Theme	Subtheme	Description (bias visualization approach ...)	#
Awareness	Obtain information selectively	Provides users with the knowledge to decide on which information to consume	38
	Increase user awareness	Makes users aware of the existence of biases in the search results	28
	Reduce critical thinking	May decrease people's critical thinking skills (e.g., relying on visualization)	6
Trustworthiness	Algorithmic bias	Can also be biased and promote biased information to users	14
Usability	Retrieve better results	Can achieve better, more comprehensive, or more unbiased results	8
	Lose relevant information	Can cause relevant information to be filtered out	8
	Easy to understand	Is clear, convenient and easy to understand	8
	Difficult to use	Is complicated to use and may reduce search efficiency	8
	Save time	Saves users' time in identifying biases and/or unbiased content	6
Miscellaneous	Prototypes evaluation	Evaluated each prototype individually and did not discuss the bias visualization approach	17
	Others/unclear	[falls outside of the established themes]	17

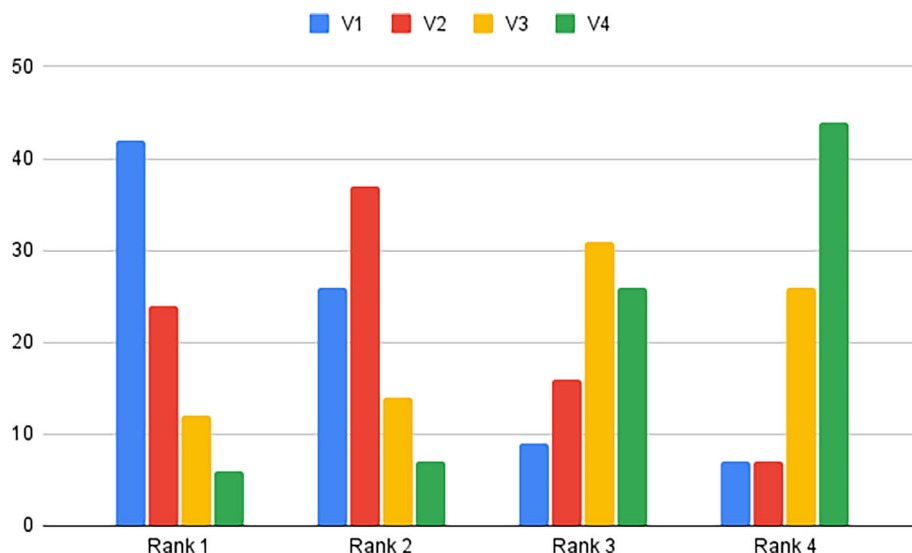


FIGURE 13 Ranks of bias visualization prototypes (rank 1 = best, rank 4 = worst)

said that it was easier to use (s2r1) and understand (s3r11). However, others discussed that V3 (Figure 7) “is less useful [...] it just shows the text, not graphs or icons that are easier to understand” (s3r48). The majority of responses preferred V4 (Figure 8) the least, as the feature of hiding biased information is seen as a form of news censorship that can be dangerous (s3r47).

4.2.2 | Results-reranking approach

Influence of results-reranking approach to information seeking tasks. Three themes and ten subthemes were found in participants' responses (Table 2).

Usability: Participants mentioned that results-reranking allowed them to **retrieve better results** (31 out of 85 responses). One response noted that this approach gives “results that users are interested in according to the user’s individual needs” (s1r22). Some participants further mentioned that they could get more suitable results in a higher ranking (s3r37, s3r41). Most of the comments on this theme were directly related to the second most popular subtheme, **customize search results**. The main idea, shared by the vast majority of the responses, can be summarized by a response which mentioned that “[t]his method provides searchers with different directions to choose and adjust their preferences [...] which can better eliminate bias and help them find the information they want more accurately” (s1r14). Others, however, noted that this approach may **lose relevant results** (9 responses) if the customization removes results that may be relevant to what the users want (s1r12, s3r18, s3r49). Eight responses mentioned that this approach might be **difficult to use**, noting that it would take more time and effort to refine and go through the results (s1r19, s3r9),

and required more patience from users (s1r14). Others mentioned that the customization is only usable for “people who really understand the meaning of bias” (s3r45), but may be too complicated for amateur users (s3r45). One response, however, mentioned that this approach could **save time** on specific searches (s3r46).

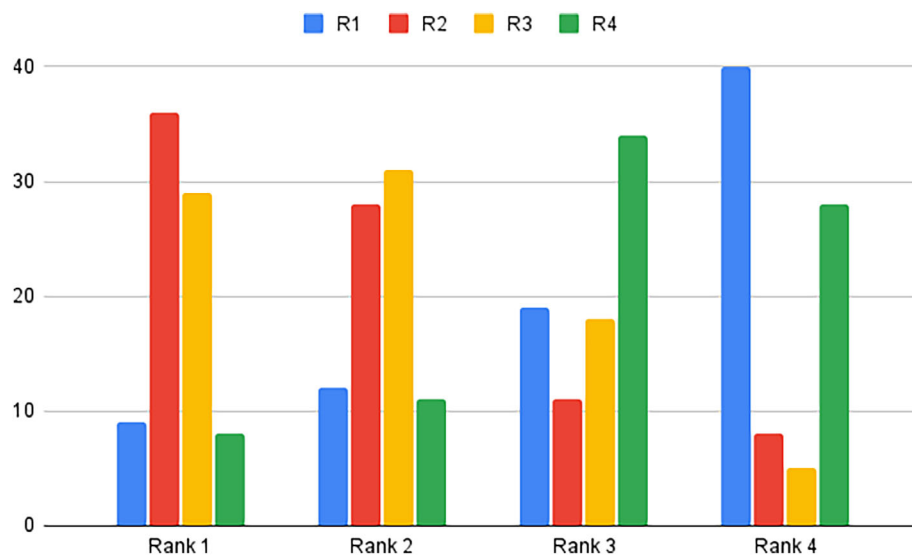
Awareness: Thirteen responses highlighted the risk that the customization feature may **worsen polarisation**. Participants highlighted concerns such as losing divergence of results when specific preferences are set (s1r4, s3r2), seeing only information that fits with their ideas (s3r31), strengthening the filter bubble (s1r16), and cause more polarisation issues for users (s3r44). On the other hand, 12 responses noted that this approach can help **educate the users**, highlighting that the customization allows users to be “more conscious of the type of results they get and possible biases” (s1r16) and “leads to self-criticism and a state of awareness of what [the user] is searching.” Participants further noted that it “will make all the people smarter and more objective” (s3r1) and, eventually, may achieve a society that is “less polarised in terms of opinions” (s3r12). Eight responses further mentioned that this approach can **show alternative viewpoints** to the users, allowing them to “access information that normally [...] would not access” (s1r12) and that it “can help [users] to see that there are people with other points of view” (s3r47).

Trustworthiness: Participants discussed that this approach will **reduce bias** in the results (12 responses), further mentioning that “reorder[ing] search results [can] reduce unfairness and bias” (s1r9) and that “it can help consumers quickly filter out biased messages” (s1r13). Similar to the bias visualization approach, the risk of **algorithmic bias** was also highlighted in this approach (11 responses). Responses mentioned the

TABLE 2 Themes found in Q3 (influence of results-reranking approach)

Theme	Subtheme	Description (results-reranking approach ...)	#
Usability	Retrieve better results	Returns personalized results and/or better search results that fit users' needs	31
	Customize search results	Gives users the ability to customize/filter the results	25
	Lose relevant results	Can lead to missing some relevant information	9
	Difficult to use	Is more complicated and required more effort from the user	8
	Save time	Saves time spent in the search tasks	1
Awareness	Worsen polarisation	May produce results that are affected by users' own biases, leading to a more polarised society	13
	Educate users	Makes people smarter and more objective can achieve a less polarised society	12
	Show alternative viewpoints	Provides users with results containing different points of view	8
Trustworthiness	Reduce bias	Reduces unfairness and bias in the results	12
	Algorithmic bias	Could also be biased and promote biased information to users	11
Miscellaneous	Prototypes evaluation	Evaluated each prototype individually and did not discuss the results-reranking approach	17
	Unclear/no responses	[falls outside the established themes]	10

FIGURE 14 Ranks of results-reranking prototypes (rank 1 = best, rank 4 = worst)



lack of trust in the reranking algorithm (s3r26, s3r34), due to the lack of transparency of the algorithm (s3r38, s3r50), the complexity of identifying biases (s3r53), and the possible manipulation of the algorithms (s2r3, s3r5, s3r48).

Preferred prototypes. As shown in Figure 14, participants highly ranked the *manual* reranking prototypes, R2 and R3 (Figures 10 and 11), further noting that both gave a higher level of control to customize the results (s1r22, s3r50) and select the levels of bias they want

(s3r12, s3r16). They also agreed that R2 has a good balance between functionality, complexity, and convenience (s1r24, s1r26). While participants also liked the advanced customization in R3 (s3r30, s3r39), one pointed out that “average user[s] may have difficulties to do the bias customization” (s3r20). R1 (Figure 9) and R4 (Figure 12) which incorporated *automatic* reranking were the least preferred systems, due to the limited user control (s3r5, s3r39) and that it “omitted some information without justification” (s3r20). One participant also noted that the

TABLE 3 Themes found in Q5 (reasons for the preferred approaches)

Theme	Subtheme	Description	#	Bias vis.	Rerank	Both
Awareness	Increase user awareness	Make users aware of the existence of biases in the results	31	16	2	13
	Obtain information selectively	Provide users with the knowledge to decide on which information to consume or trust	27	12	4	11
	Show alternative viewpoints	Provide users with results containing different points of view	6	1	2	3
Usability	Customize search results	Give users the ability to customize/filter the results	18	2	10	6
	Retrieve better results	Return higher quality search results that better fit users' needs	13	1	5	7
	Better efficiency	It is more efficient, more intuitive, easier to use, and more convenient	10	2	4	4
	Lower quality of results	May negatively alter search quality	2	1	1	0
	Difficult to use	It is more complicated or difficult to use	2	0	2	0
Trustworthiness	Reduce bias	Reduce unfairness and bias in the results	7	0	1	6
Miscellaneous	Other/unclear	[falls outside of the established themes]	18	7	4	7

different viewpoints provided in R4 might confuse users instead (s3r36). While R1 was seen as a straightforward option to view the most unbiased news (s3r1, s3r11, s3r54), others were concerned about the accuracy of the ranking algorithms (s3r33, s3r39, s3r47).

4.2.3 | Preferred approaches for addressing biases in search results

When asked which approach they preferred to have (Q5 in Figure 4), the vast majority of responses indicated that *both approaches* were needed (35 out of 85). Twenty-seven responses preferred the *bias visualization* approach, twenty-two preferred the *results-reranking* approach, and one response did not choose any of these options and was excluded. In the responses elaborating on these choices, 10 subthemes were found (see Table 3). As expected, no new themes emerged in this question. However, this analysis further explores which themes were identified to be the most important in the preferred approach for Web search engines. We further show the distribution of these themes across the three chosen approaches: bias visualization, results-reranking, or both.

Participants who preferred the *bias visualization* approach liked that this approach **increased users' awareness** and provided them with the knowledge to **obtain information selectively**. Meanwhile, those who preferred the *results-reranking* approach liked how it gave them the control to **customize search results** and

retrieve better results. These four subthemes were also the most commonly discussed amongst responses that preferred *both approaches*. Participants argued that this combination would give users the option to be notified of biases (s2r1), and filter and reorder the results (s3r50, s3r53) to get more suitable results for their needs and objectives (s1r8, s1r21).

4.3 | Limitations

No study comes without limitations and this work is not an exception. The eight designs we investigated in this study were proposed by a small number of participants and might not provide an exhaustive range of interface designs for raising user awareness of biases. Nevertheless, these designs captured a wide variety of features that allowed us to gain valuable insights into the usefulness of various aspects and approaches for dealing with biases in search engines; all of which are important for developing future work in this area.

In addition, our study focused on how search engine interfaces should be designed to improve users' awareness. Specific methods, on *how* these biases should be measured are beyond the scope of this work. We acknowledge that this lack of semantic information might have influenced the participants' perceptions of the different bias awareness interface designs, but this is a commonly accepted disadvantage of prototype-based user design and evaluation.

Finally, these prototype evaluations were carried out by participants (mostly students in Computer Science) with a higher digital literacy compared to the general population. Previous studies have shown that users' expertise (prior knowledge) may influence their search behavior and performance in completing search tasks in interactive IR systems (Liu & Belkin, 2014; Scott et al., 2013). It is, therefore, possible that users with lower digital literacy may interpret these interfaces differently, or have different preferences of the designs and underlying approaches. Immediate future work is required to further investigate how search engine users with various levels of expertise (e.g., the general public) perceive these designs and biases in search engines more generally.

5 | DISCUSSION

Often seen as a gatekeeper (Bui, 2010; Diakopoulos, 2015; Wallace, 2018), search engines influence the results users see and trust. We acknowledge that at the moment there is no way to develop a bias-free search engine. Therefore, search engines should be designed to provide more transparency, while managing the risk of overwhelming users through a complex representation of information (Diakopoulos & Koliska, 2017). Our study contributed to this area by utilizing a *human-in-the-loop* approach for assessing different interface designs that could be integrated into a search engine to improve user awareness of biases in the search results.

Overall, the results show that a combination of bias visualization and results-reranking approaches should be implemented in search engines. Nonetheless, for a successful implementation of both approaches, participants have continuously highlighted the importance of transparent and trustworthy algorithms for measuring and identifying bias accurately, to avoid further misleading the users and perpetuating biases in society (Kordzadeh & Ghasemaghahi, 2022; Noble, 2018; Novin & Meyers, 2017). Participants indicated that, similar to Horne et al. (2019), visualization features (e.g., V1) provided them with a brief explanation of the severity of biases in the results. However, they expressed their concerns over the transparency of the approach and whether the potential algorithm for calculating this bias could be itself biased. It is further evident from participants' comments, that *automatic* reranking, was not perceived to be trustworthy due to the lack of transparency of the algorithm and the possibility of compromising the quality of the results or user satisfaction (Gao & Shah, 2020). Participants prefer the *manual* reranking approach instead to allow customizing search results based on their preferences, similar to the liberal approach in recommendation systems proposed by Helberger (2019). However, there is also a risk that

some users may use this feature to render invisible articles that provide different perspectives and therefore, polarising them more in their news consumption and encouraging more narrow-minded individuals (Pariser, 2011).

Finally, despite efforts in designing interfaces to promote user awareness of biases in search engines, users' selectivity and recommendation from other information access systems (e.g., social media) have been shown to play a stronger part in limiting users' exposure to diverse content (Bakshy et al., 2015). Search engines should, therefore, further investigate designs and features that improve users' digital literacy skills to help users be more critical in their information access.

6 | CONCLUSIONS

Using a participatory design methodology, our research investigated the users' perspective on different interface designs and approaches that should be utilized in news search engines to improve user awareness of biases. Eight designs across two underlying approaches (bias visualization and results-reranking) were created and evaluated. Specific design features to visualize biases (such as the bias-meter in V1, or the histogram in V2) were identified to be more useful than textual description (V3) or hiding biased information (V4). Designs that debias results automatically (R1 and R4) were least preferred due to the lack of transparency. Instead, participants preferred manual reranking systems (R2 and R3) because they provide users with a higher level of control in customizing their results. However, others were concerned that this feature comes with the risk of strengthening users' filter bubble and promoting polarisation.

We have also gathered valuable insights into how each underlying approach influenced users in their search tasks. Findings from this study suggest that *bias visualization* approach plays an important role in raising user awareness of existing biases, and as a result, allows users to be more critical in obtaining information from the Web. *Results reranking* approach, on the other hand, allows users to customize their results to retrieve search results that better fit their preferences or needs. Our findings further highlighted the importance to utilize *both* bias visualization and results-reranking approaches in search engines to help users mitigate biases in search results. Participants further asserted the importance of reliable and transparent methods for both approaches, in order to reduce any subjectivities in the biased information presented to the users.

The rich insights gathered in this study are important for sharpening further discussions and research in designing bias-aware user interfaces. Immediate future

work will investigate whether search engine users with different levels of expertise (e.g., the general public) have different perceptions of the designs of interfaces for raising user awareness in search engines.

ACKNOWLEDGMENTS

We would like to thank the reviewers for their valuable comments on this work. We would also like to thank all participants in this study. This project is funded by the European Union's Horizon 2020 research and innovation programme under grant agreements No. 810105 (CyCAT).

ORCID


Monica Lestari Paramita  <https://orcid.org/0000-0002-9414-1853>

Maria Kasinidou  <https://orcid.org/0000-0002-2438-9095>

Styliani Kleanthous  <https://orcid.org/0000-0003-1594-1340>

Paolo Rosso  <https://orcid.org/0000-0002-8922-1242>

Tsvi Kuflik  <https://orcid.org/0000-0003-0096-4240>

Frank Hopfgartner  <https://orcid.org/0000-0003-0380-6088>

REFERENCES

- An, J., Cha, M., Gummadi, K., Crowcroft, J., & Quercia, D. (2012). Visualizing media bias through Twitter. In *The potential of social media tools and data for journalists in the news media industry* (ICWSM Workshop Technical Report WS-12-01) (Vol. 6, pp. 2–5). Association for the Advancement of Artificial Intelligence.
- Baeza-Yates, R. (2018). Bias on the web. *Communications of the ACM*, 61(6), 54–61.
- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130–1132. <https://doi.org/10.1126/science.aaa1160>
- Beam, M. A. (2014). Automating the news: How personalized news recommender system design choices impact news reception. *Communication Research*, 41(8), 1019–1041.
- Bui, C. (2010). How online gatekeepers guard our view—news portals' inclusion and ranking of media and events. *Global Media Journal*, 9, 16.
- Celis, L. E., Kapoor, S., Salehi, F., & Vishnoi, N. K. (2018). *An algorithmic framework to control bias in bandit-based personalization*. CoRR.
- Diakopoulos, N. (2015). Algorithmic accountability. *Digital Journalism*, 3(3), 398–415. <https://doi.org/10.1080/21670811.2014.976411>
- Diakopoulos, N., & Koliska, M. (2017). Algorithmic transparency in the news media. *Digital Journalism*, 5(7), 809–828. <https://doi.org/10.1080/21670811.2016.1208053>
- Draws, T., Tintarev, N., Gadiraju, U., Bozzon, A., & Timmermans, B. (2020). Assessing viewpoint diversity in search results using ranking fairness metrics. *ACM SIGKDD Explorations Newsletter*, 23(1), 50–58.
- Draws, T., Tintarev, N., Gadiraju, U., Bozzon, A., & Timmermans, B. (2021). This is not what we ordered: Exploring why biased search result rankings affect user attitudes on debated topics. In *Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval* (Vol. 21, pp. 295–305). Association for Computing Machinery. <https://doi.org/10.1145/3404835.3462851>
- Druckman, J. N., & Parkin, M. (2005). The impact of media bias: How editorial slant affects voters. *Journal of Politics*, 67(4), 1030–1049. <https://doi.org/10.1111/j.1468-2508.2005.00349.x>
- Epstein, R., & Robertson, R. E. (2015). The search engine manipulation effect (seme) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences of the United States of America*, 112(33), E4512–E4521.
- Gao, R., & Shah, C. (2019). How fair can we go: Detecting the boundaries of fairness optimization in information retrieval. In *ICTIR '19: Proceedings of the 2019 ACM SIGIR international conference on theory of information retrieval* (pp. 229–236). Association for Computing Machinery.
- Gao, R., & Shah, C. (2020). Toward creating a fairer ranking in search engine results. *Information Processing & Management*, 57(1), 102–138. <https://doi.org/10.1016/j.ipm.2019.102138>
- Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media technologies* (pp. 167–194). MIT Press. <https://doi.org/10.7551/mitpress/9780262525374.003.0009>
- Hamborg, F. (2020). Media bias, the social sciences, and nlp: Automating frame analyses to identify bias by word choice and labeling. In *Proceedings of the 58th annual meeting of the Association for Computational Linguistics: Student research workshop* (pp. 79–87). Association for Computational Linguistics.
- Hamborg, F., Donnay, K., & Gipp, B. (2019). Automated identification of media bias in news articles: An interdisciplinary literature review. *International Journal on Digital Libraries*, 20(4), 391–415. <https://doi.org/10.1007/s00799-018-0261-y>
- Hamborg, F., Meuschke, N., & Gipp, B. (2017). Matrix-based news aggregation: Exploring different news perspectives. In *2017 ACM/IEEE joint conference on digital libraries (JCDL)* (pp. 1–10). IEEE. <https://doi.org/10.1109/JCDL.2017.7991561>
- Helberger, N. (2019). On the democratic role of news recommenders. *Digital Journalism*, 7(8), 993–1012. <https://doi.org/10.1080/21670811.2019.1623700>
- Horne, B. D., Nevo, D., O'Donovan, J., Cho, J. H., & Adali, S. (2019). Rating reliability and bias in news articles: Does AI assistance help everyone? In *Proceedings of the international AAAI conference on web and social media* (Vol. 13, pp. 247–256). Association for the Advancement of Artificial Intelligence.
- Hussain, S., Sanders, E. B. N., & Steinert, M. (2012). Participatory design with marginalized people in developing countries: Challenges and opportunities experienced in a field study in Cambodia. *International Journal of Design*, 6, 2.
- Introna, L. D., & Nissenbaum, H. (2000). Shaping the Web: Why the politics of search engines matters. *Information Society*, 16(3), 169–185. <https://doi.org/10.1080/01972240050133634>
- Jürgens, P., & Stark, B. (2022). Mapping exposure diversity: The divergent effects of algorithmic curation on news consumption. *Journal of Communication*, 72(3), 322–344. <https://doi.org/10.1093/joc/jqac009>
- Kelly, D. (2009). Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends in*

- Information Retrieval*, 3(1–2), 1–224. <https://doi.org/10.1561/1500000012>
- Kevin, V., Högden, B., Schwenger, C., Şahan, A., Madan, N., Aggarwal, P., Bangaru, A., Muradov, F., & Aker, A. (2018). Information nutrition labels: A plugin for online news evaluation. In *Proceedings of the first workshop on fact extraction and verification (FEVER)* (pp. 28–33). Association for Computational Linguistics.
- Kordzadeh, N., & Ghasemaghahi, M. (2022). Algorithmic bias: Review, synthesis, and future research directions. *European Journal of Information Systems*, 31(3), 388–409. <https://doi.org/10.1080/0960085X.2021.1927212>
- Leavy, S. (2019). Uncovering gender bias in newspaper coverage of Irish politicians using machine learning. *Digital Scholarship in the Humanities*, 34(1), 48–63. <https://doi.org/10.1093/lc/fqy005>
- Liu, J., & Belkin, N. J. (2014). Multi-aspect information use task performance: The roles of topic knowledge, task structure, and task stage. *Proceedings of the American Society for Information Science and Technology*, 51(1), 1–10. <https://doi.org/10.1002/meet.2014.14505101031>
- Makhortykh, M., Urman, A., & Roberto, U. (2020). How search engines disseminate information about COVID-19 and why they should do better. *The Harvard Kennedy School (HKS) Misinformation Review*, 1, 1–12. <https://doi.org/10.37016/mr-2020-017>
- Mokherberian, N., Abeliuk, A., Cummings, P., & Lerman, K. (2020). Moral framing and ideological bias of news. In S. Aref, K. Bontcheva, M. Braghieri, F. Dignum, F. Giannotti, F. Grisolia, & D. Pedreschi (Eds.), *Social informatics* (pp. 206–219). Springer. https://doi.org/10.1007/978-3-030-60975-7_16
- Munson, S., & Resnick, P. (2013). Encouraging reading of diverse political viewpoints with a browser widget. In *Seventh international AAAI conference on weblogs and social media*. Association for the Advancement of Artificial Intelligence.
- 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>
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.
- Novin, A., & Meyers, E. (2017). Making sense of conflicting science information: Exploring bias in the search engine result page. In *CHIIR '17: Proceedings of the 2017 conference on conference human information interaction and retrieval* (pp. 175–184). Association for Computing Machinery.
- Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google we trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication*, 12(3), 801–823. <https://doi.org/10.1111/j.1083-6101.2007.00351.x>
- Papadakis, P., & Konstantakis, G. (2020). bias goggles: Graph-based computation of the bias of web domains through the eyes of users. In *European conference on information retrieval ECIR 2020: Advances in information retrieval* (pp. 790–804). Springer.
- Paramita, M. L., Orphanou, K., Christoforou, E., Otterbacher, J., & Hopfgartner, F. (2021). Do you see what I see? Images of the COVID-19 pandemic through the lens of Google. *Information Processing & Management*, 58(5), 102654. <https://doi.org/10.1016/j.ipm.2021.102654>
- Pariser, E. (2011). *The filter bubble: What the internet is hiding from you?* Penguin Press.
- Park, S., Ko, M., Kim, J., Choi, H. J., & Song, J. (2011). News-Cube2.0: An exploratory design of a social news website for media bias mitigation. In *2nd international workshop on social recommender systems* (p. 5). International Workshop on Social Recommender Systems.
- Roth, Y., & Pickles, N. (2020). *Updating our approach to misleading information*. Retrieved from: https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information.html
- Scott, D., Hopfgartner, F., Guo, J., & Gurrin, C. (2013). Evaluating novice and expert users on handheld video retrieval systems. In *Advances in multimedia modeling, 19th international conference, MMM 2013, Huangshan, China, January 7–9, 2013, proceedings, part II* (Vol. 7733, pp. 69–78). Springer. https://doi.org/10.1007/978-3-642-35728-2_7
- Spinde, T., Hamburg, F., Donnay, K., Becerra, A., & Gipp, B. (2020). Enabling news consumers to view and understand biased news coverage: A study on the perception and visualization of media bias. In *JCDL '20: Proceedings of the ACM/IEEE joint conference on digital libraries* (pp. 389–392). Association for Computing Machinery.
- Spohr, D. (2017). Fake news and ideological polarization: Filter bubbles and selective exposure on social media. *Business Information Review*, 34, 150–160. <https://doi.org/10.1177/0266382117722446>
- Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, 27(2), 237–246. <https://doi.org/10.1177/1098214005283748>
- Urman, A., Makhortykh, M., & Ulloa, R. (2022). The matter of chance: Auditing web search results related to the 2020 U.S. presidential primary elections across six search engines. *Social Science Computer Review*, 40(5), 1323–1339. <https://doi.org/10.1177/08944393211006863>
- Wallace, J. (2018). Modelling contemporary gatekeeping. *Digital Journalism*, 6(3), 274–293. <https://doi.org/10.1080/21670811.2017.1343648>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Paramita, M. L., Kasinidou, M., Kleanthous, S., Rosso, P., Kuflik, T., & Hopfgartner, F. (2023). Towards improving user awareness of search engine biases: A participatory design approach. *Journal of the Association for Information Science and Technology*, 1–19. <https://doi.org/10.1002/asi.24826>