JOURNAL OF DIGITAL SOCIAL RESEARCH

WWW.JDSR.IO

ISSN: 2003-1998

VOL. 4, NO. 1, 2022, 128–151

EXPLORING SIRI'S CONTENT DIVERSITY USING A CROWDSOURCED AUDIT

Tim Glaesener[a]

ABSTRACT

This study aims to describe the content diversity of Siri's search results in the polarized context of US politics. To do so, a crowdsourced audit was conducted. A diverse sample of 170 US-based Siri users between the ages of 18-64 performed five identical queries about politically controversial issues. The data were analyzed using the concept of algorithmic bias. The results suggest that Siri's search algorithm produces a long tail distribution of search results: Forty-two percent of the participants received the six most frequent answers, while 22% of the users received unique answers. These statistics indicate that Siri's search algorithm causes moderate concentration and low fragmentation. The age and, surprisingly, the political orientation of users, do not seem to be driving either concentration or fragmentation. However, the users' gender and location appear to cause low concentration.

Keywords: Siri; content diversity; crowdsourced audit; voice assistants; US politics; politically controversial issues; algorithmic bias; search algorithms; search results; long tail distribution; concentration; fragmentation.

[a] School of Arts and Communication, Malmö University, Sweden.

1 INTRODUCTION

The algorithms of search engines are powerful[1], as they act as gatekeepers to information (Diakopoulos et al., 2018, p. 321). In their function as gatekeepers, they help users to navigate through the web by deciding what information is most relevant to them (Gillespie, 2014, p. 167). By ranking results, these algorithms strongly impact users' attention (Trielli & Diakopoulos, 2019, p. 1), since users tend to click on top results more often than on lower ones (Agichtein et al., 2006, p. 3).

While it is largely unknown how algorithms actually work (Pasquale, 2015, p. 3), it is known that they function in context (Kitchin, 2017, p. 25). For example, algorithms depend on personal data to create situated outcomes (Sandvig et al., 2014, p. 10). As a consequence, personalized ranking algorithms of search engines can provide two users with different results to the same query (Bozdag, 2013, p. 209; Gillespie, 2014, p. 188).

Due to their power to provide users with different information, the algorithms of search engines can be linked to a larger discussion around the role of audience fragmentation in democracies. Some degree of fragmentation is welcome, since it represents a diverse media environment (Fletcher & Nielsen, 2017, p. 477), which helps citizens to form their political opinions (Sunstein, 2002, p. 9). In well-functioning democracies, citizens need to share information in order to engage with it and debate it (Fletcher & Nielsen, 2017, p. 477).

This is in line with the concept of deliberate democracy, which is a type of democracy in which citizens and their representatives are expected to justify relevant decisions. The most important characteristic is that both parties are expected to give reasons for their decisions as well as to respond to the reasons given by others (Gutmann & Thompson, 2004, p. 3). According to Habermas, to participate in deliberative democracy, people need first and foremost to be able to take part in discourse (Olson, 2011, p. 140). A high degree of audience fragmentation is problematic because it might people might end up only being exposed only to political opinions they already agree with. This could, in effect, end up shielding people from other viewpoints, perhaps encouraging people to hold more rigid and extreme positions. In turn, this could "hinder consensus-building in society" (Vike-Freiberga et al., 2013, pp. 27-28).

The issue of algorithmic content selection has become increasingly relevant, so it is not surprising that in recent years, many researchers have been exploring it (Zuiderveen Borgesisus et al., 2016, p. 9) - often relying on Pariser's (2011) disputed concept of 'filter bubbles'. The term 'filter bubble' refers to the idea that personalization of content caused by algorithms can lead to selective exposure (Pariser, 2011). Concerns about such personalization often focus on the possible

[1] There are numerous definitions of the concept of algorithm. In this paper, I use Diakopoulos' (2015) definition that considers an algorithm to be a series of steps undertake to solve a problem or accomplish an outcome (p. 400).

negative effects of filter bubbles on democracy (Zuiderveen Borgesisus et al., 2016, p. 4).

Scholars have come to different and, at times, contradictory findings. Most of them have challenged the idea that search algorithms cause fragmentation. For example, the findings of meta-researchers, i.e., scholars who study the research of others, state that filter bubbles do not constitute a serious danger. With regard to news websites, Zuiderveen Borgesisus et al. (2016) conclude that "there is little empirical evidence that warrants any worries about filter bubbles" (p. 1). More recently, Bruns (2019), who has critically reviewed the idea of the filter bubble, points to empirical evidence that the users of search engines and social media tend to be exposed to more diverse information that non-users (p. 1).

However, it is not only meta-research that has challenged the notion that algorithms cause fragmentation; empirical studies have done so as well. For example, after having analyzed the search results of 187 US Google users (p. 148), Robertson et al. (2018) conclude that there is little evidence of the existence of filter bubbles. Similarly, Nechusthai and Lewis (2019) found that "users with different political leanings from different states were recommended very similar news" (p. 298). Their study involved 168 participants who used their personal Google accounts (Nechusthai & Lewis, 2019, p. 298). Also in the US, Feezell et al. (2021) conducted a nationally representative survey of young adults and one of the general population and found that "neither non-algorithmic nor algorithmically determined news contribute to higher levels of partisan polarization" (p. 1).

Outside the US, scholars have come to similar conclusions. For example, Krafft et al. (2019) have analyzed the search results of more than 4,000 German Google users and discovered hardly any evidence of filter bubbles (p. 1). Similarly, Haim et al. (2018) has researched the personalization of the "content and source diversity of Google News" (p. 330) in Germany. The researchers found minor effects of personalization, concluding that their results indicate that the danger of algorithmic filter bubbles might be exaggerated (Haim et al., 2018, p. 330).

In Belgium, the research findings of Curtois et al. (2018) did not support the existence of filter bubbles as far as social and political information were concerned (p. 2006). However, the scholars coded (in other words: categorized) search results before comparing them (Curtois et al., 2018, p. 2010). In Denmark, Bechmann and Nielbo (2018) researched the Facebook News Feed of 1,000 people. By analyzing link sources and content semantics, they found that roughly 10% to 28% of the sample size was exposed to content that did not overlap. The researchers use this finding as the basis for their critique of the existence of filter bubbles. (Bechmann & Nielbo, 2018, p. 999).

In contrast, other researchers have confirmed the existence of filter bubbles. For example, Barker (2018) has determined that art directors and copywriters are exposed to personalized search results on Google Search (p. 85). His qualitative study does not allow for generalizations, however, as it involved only 18 Australian participants (Barker, 2018, p. 85). Cho et al. (2020) carried out a lab experiment that included manipulating "user search/watch history" (Cho et al., 2020, p. 150). The researchers prove that "political self-reinforcement" (Cho et al., 2020, p. 150) and "affective polarization are heightened by political videos – selected by the YouTube recommender algorithm – based on the participants' own search preferences" (Cho et al., 2020, p. 150). Theoretically analyzing "twelve different information filtering scenarios" (Geschke et al., 2019, p. 129), Geschke et al. (2019) found that "even without any social or technological filters, echo chambers emerge as a consequence of cognitive mechanisms, such as confirmation bias" (Geschke et al., 2019, p. 129). They further claim that social and technological filtering enhances echo chambers (Geschke et al., 2019, p. 129). The research objects of these studies that confirm the existence of filter bubbles differ greatly from those that refute it.

While the answers of algorithms of graphical and textual search engines have increasingly been researched, the replies of algorithms of voice assistants such as Apple's Siri, Amazon Alexa and Google Assistant have not yet been studied extensively. However, I argue that especially these search algorithms have the potential to shape users' attention. When voice assistants are used to retrieve information from the internet, they typically provide users with only one result to their queries (Dambanemuya & Diakopoulos, 2020, p. 1). In contrast, textual and graphical search engines offer a plurality of results (Natale & Cooke, 2021, p. 1003). In other words, graphical and textual search engines allow users to make a choice among the several results. In contrast, in the case of voice assistants, the choices have been pre-selected for the users, who are presented with only one result. Due to their unique affordances, it seems that the search algorithms of voice assistants are even more powerful gatekeepers than those of other web interfaces.

1.1 Siri in the USA

To address this research gap, I conducted a crowdsourced audit of the search algorithm of the voice assistant Siri[2] in the context of US politics. I have chosen to focus on the USA since voice assistants were used by a critical mass of 51% of the population in 2020 (Valishery, 2021). In addition, there is a high degree of political polarization in the US. (Jurkowitz et al., 2020).

Siri was chosen as a research object for two reasons. Firstly, data suggest that it remains more popular than Google Assistant in the USA (Wagner, 2018; Kinsella, 2020). Although Amazon Alexa is even more popular (Dellinger, 2019), I decided not to focus on this device since smart speakers are communal media (Boothby, 2018), which are typically used in households by more than one user and I focus on content diversity on an individual level. Secondly, while the search

[2] In this study, I use the term 'search algorithm' for the sake of simplicity. I acknowledge, though, that in fact there is not a single search algorithm but only a system of algorithms (Natale & Cooke, 2021, p. 1009). Similarly, I speak of Siri even though this voice assistant is not a real entity but rather a complex system (Seaver, 2013, p. 416).

algorithms of neither Siri nor Google Assistant have been researched sufficiently in relation to content diversity, the search algorithm of Google's textual and graphical search engine has already been investigated (e. g. Haim et al., 2018; Nechusthai & Lewis, 2019). Some of these findings might provide some insight into the workings of Google's voice-based search engine. For this reason, I maintain that there is more need to investigate Siri's search algorithm.

The purpose of this study is to *explore* and *describe* whether Siri answers politically controversial questions of US-based users differently. To do so, I seek to answer the following research question: To what extent does Siri's search algorithm provide 18 to 64-year-old US-based users with different results when asked the same politically controversial questions?

Throughout this study, I refer to the previously mentioned 'different results' as 'unique' and 'personalized' search results. Both terms are used interchangeably to indicate that only one individual has received a certain reply.

Inspired by Fletcher et al. (2020, p. 181), in this study, I refer to the case when users receive non-overlapping content as 'fragmentation' and the case when they receive the overlapping as 'concentration'.

My selection of a diverse group of Siri users lowers the generalizability of this study's results. However, I believe that content diversity, which I define in this study as differences in content as do Krafft et al. (2019, p. 330), can best be studied by using a heterogeneous data set. This view was inspired by Bechmann and Nielbo (2018), who warn scholars about relying on homogenous data sets to investigate differences and similarities in content (p. 994).

1.2 Algorithmic bias

To analyze Siri's search results, I draw on Friedman and Nissenbaum's (1996) concept of 'algorithmic bias'. They argue that computer systems are systematically biased in three ways: pre-existing biases which exist in society (e. g. personal values of software developers) and affect the design process; technical biases which result from the influence of the technology on how computers work; and emergent biases which manifest themselves during the use of a particular software after its release (e. g. through feedback loops) (Friedman & Nissenbaum, 1996, pp. 334–335). As a consequence of these biases, search engines can end up providing users with an "increasingly distorted and limited view of the web" (Introna & Nissenbaum, 2000, p. 54).

While the concept of algorithmic bias seems to be suitable to analyze the outcome of Siri's search algorithm, in one aspect it needs to be adjusted to fit to this study. Friedman and Nissenbaum (1996) distinguish between three types of biases: pre-existing, technical, and emergent (pp. 334-335). The technical bias of Siri's search algorithm has already been mentioned: as is the case with all voice assistants, Siri typically offers users only one answer to their queries (Natale & Cooke, 2021, p. 1003).

For this reason, each search result of Siri can be considered technical biased. To identify, however, whether the search results are subject to pre-existing or emergent biases goes beyond the scope of this study. I only aim to identify whether a bias can be observed, but not if is a pre-existing or emergent one. For this reason, in this study, I do not differentiate between the two types of biases, but only refer to a bias.

2 METHODS

2.1 Participants

The participants included 170 US-based Siri users[3] who were between 18 and 64 years old and of different genders, locations, and political leanings. Ethical approval was obtained from Malmö University's Ethic Council before I began recruiting Siri users. This study was advertised exclusively on the crowdsourcing platform Amazon Mechanical Turk on April 21, 2021. The assignment was visible to all workers on the platform. However, to participate, users needed to be based in the US. Apart from this criterion, I did not use any filters such as age or location because I wanted to reach a heterogeneous group of Siri users. Participants were self-selected and compensated 2 USD for their time. Informed consent was obtained from all participants. Within roughly eight hours, 170 workers had completed the task.

However, for the final data analysis, I filtered out 36 users who occasionally or systematically did not report back Siri's replies as was intended – whether intentionally or not is unknown. Some participants answered the questions themselves, others summarized or interpreted the search results or filled the answer boxes with placeholder text. Even though this quality control increases this study's internal validity to some extent, it could not be done rigorously: I was able to identify whether participants made other queries than the ones they were supposed to and reported back answers to those searches. The quality control reduced the data set.

As table 1 shows, the final sample size consisted of 134 users. Each user was asked to report Siri's answers to five questions. However, not everyone reported back the answers to all five questions sufficiently. This explains why the total number of replies is 631, rather than 670 (this would have been the case, had all 134 users reported all five answers).

Despite the reduction of the size of the data set, the sample remained heterogeneous (see Table A1). However, the vast majority of participants belonged to the age of group 25–34. In addition, more males than females and more liberals than conservatives participated in the audit.

[3] This sample size was inspired by other studies which focus on content diversity of search results using a crowdsourced audit, (e. g. Nechusthai & Lewis, 2019; Cho et al., 2020).

Table 1. Data set

Sample size	134
Answers to the question "Should there be stricter gun laws?"	122
Answers to the question Should immigration be limited?	129
Answers to the question "Should the death penalty be abolished?"	130
Answers to the question "Should taxes be lowered? "	123
Answers to the question "Should abortion be legal?"	127
Total replies	631

2.2 Material and procedure

This study used a crowdsourced audit as proposed by Sandvig et al. (2014, p. 15). Every participant individually completed the assignment "Ask Siri Five Questions" on Amazon Mechanical Turk. After having read the assignment's description, participants signed a consent form containing the description of the assignment, which data would be collected and how it would be handled. Upon agreeing to the terms of this study, I collected basic demographic information (this included two multiple choice questions regarding the participants' gender and political leaning as well as two short answer questions regarding the participants' age and location). Next, participants were asked to use their smartphone to ask Siri and note down the first search results to each of the following politically controversial question (Chams, 2020; Najle & Jones, 2019):

- Should there be stricter gun laws?
- Should immigration be limited?
- Should the death penalty be abolished?
- Should taxes be lowered?
- Should abortion be legal?

Finally, the participants were thanked for the completing the assignment and paid through Amazon Payments.

Before the data could be analyzed, it had to be reduced. To do so, I replaced the actual answers that the participants had reported by numbers. For example, the first answer to the question whether guns laws should be stricter became "1", the second was turned into "2" and so on. Identical answers were labeled with the same number. This method reduced the amount of data significantly.

To measure the diversity of search results, I used univariate and bivariate descriptive statistics. The former focuses on a single variable and can be used to report a sample's distribution. To do so, summary measures such as frequency counts can be used. The latter can be used to analyze two variables (e. g. political orientation and frequency of results) (Blaikie & Priest, 2019, p. 205). In the context of this study, using these two descriptive statistics was the logical choice, because I was only concerned with the frequencies of Siri's search results across a diverse set of users.

Inspired by Webster and Ksiazek (2012, p. 43) and Trielli and Diakopoulos (2019, p. 7), I use the Gini coefficient to calculate the concentration of Siri's replies per query. A Gini coefficient can range from 0 (most equal) to 1 (most unequal). In the field of media and communications studies, defining whether the level of concentration is low, medium, or high is challenging due to a lack of established normative standards. While scholars who study audience concentration use descriptions such as 'high degrees of fragmentation', they do not explain where they see the boundaries between low, medium, and high degrees of concentration (e. g. Webster & Ksiazek, 2012; Fletcher & Nielsen, 2017; Trielli & Diakopoulos, 2019). Because of this, I have chosen to borrow definitions from the field of economics. In the context of this study, a value of 0-0.3 indicates an equal frequency distribution of many search results and therefore low concentration. A Gini coefficient of 0.31-0.7 indicates moderate concentration and a value of 0.71-1 suggests a high degree of concentration, in which few search results dominate (Lambert, 2001, p. 31).

As numerous formulas can be used to calculate the Gini coefficient (Yitzhaki & Schechtman, 2013, p. 11), in the following I present the one that I used to increase this study's reliability (Lehn et al., 2000, p. 61):

$$G = \sum_{i=0}^{n} (H_i + H_{i-1}) \times q_i - 1$$

Once I had calculated the first version of the Gini coefficient for each query's result, I calculated a second and final version of the Gini coefficient that indicates the degree of concentration, independent of the sample size (Lehn et al., 2000, p. 61). To do so, I used the following formula:

$$G* = \left(\frac{n}{n-1}\right) \times G$$

I chose these formulas as they are widely used and less complex than other ones (e. g. Dorfman, 1979; Milanovic, 1997).

To calculate fragmentation, I drew inspiration from Bechmann and Nielbo (2018, p. 994). In this study, I define fragmentation as a lack of overlapping content. This was the case whenever a Siri's search algorithm exposed only one user to a specific search result.

As with concentration, when it comes to fragmentation, there is a lack of established normative standards. Based on my measurement of the degree of concentration among Siri's answers, I decided that a fragmentation of 30% or lower was low, a value of 31%-70% was moderate and one of 71% or higher was high.

For the second step of the analysis, I utilized bivariate descriptive statistics. Using this method of analysis, I analyzed the frequency of search results (dependent variable) in relation to political orientation, location, gender, and age of users (independent variable).

3 RESULTS

The five queries led to a variety of search results. There are similarities across the searches, though (see Table A2). Firstly, Siri's search algorithm provided multiple answers to each question. The number of replies range from 23 to 46. Three questions led to more than 40 search results. The question whether taxes should be lowered received the lowest number of answers: 23.

Secondly, for each query there are one or two frequent replies. These answers reached 39 to 47 users (first question: 47; second question: 45; third question: 46; fourth question: 41, 45; fifth question: 39). The question whether taxes should be lowered is the only one that led to two frequent replies.

Thirdly, most of the answers to each question were personalized ones. The number of these unique replies range from 19 to 33. The question whether taxes should be lowered has the lowest value: 19.

Analyzing the data using univariate descriptive statistics reveals the sizes of the frequent and personalized answers in relation to all the answers (see Table 2). The few top results reached 35% to 70% of the participants, depending on the search. However, it needs to be kept in mind that questions about taxes, the answers of which were heard by 70% of the participants, had two common replies, while the other questions had one.

Fifteen to twenty-five percent of the participants were given unique replies. Four times these answers reached 20% or more of the users. The query about taxes stands out by having the lowest value: personalized replies were heard by 15% of the participants.

Table 2. Percentages of most frequent and personalized search results

Question	Should there be stricter gun laws?	Should immigration be limited?	Should the death penalty be abolished?	Should taxes be lowered?	Should abortion be illegal?	Total
Most frequent answers	39%	35%	35%	70%	42%	44%
Personalized answers	22%	25%	25%	15%	20%	22%

When the units of analysis are arranged horizontally according to the number of its total users, it becomes obvious that Siri's search algorithm has produced a long tail distribution of search results (see Figure 1). Siri's replies to each of the queries are unevenly distributed. A pattern is obvious across the five queries: The few top search results are at the head of the tail. Then, after a sharp decline, the replies that belong to the second category appear along the tail. The long part of the tail consists of the many personalized answers.

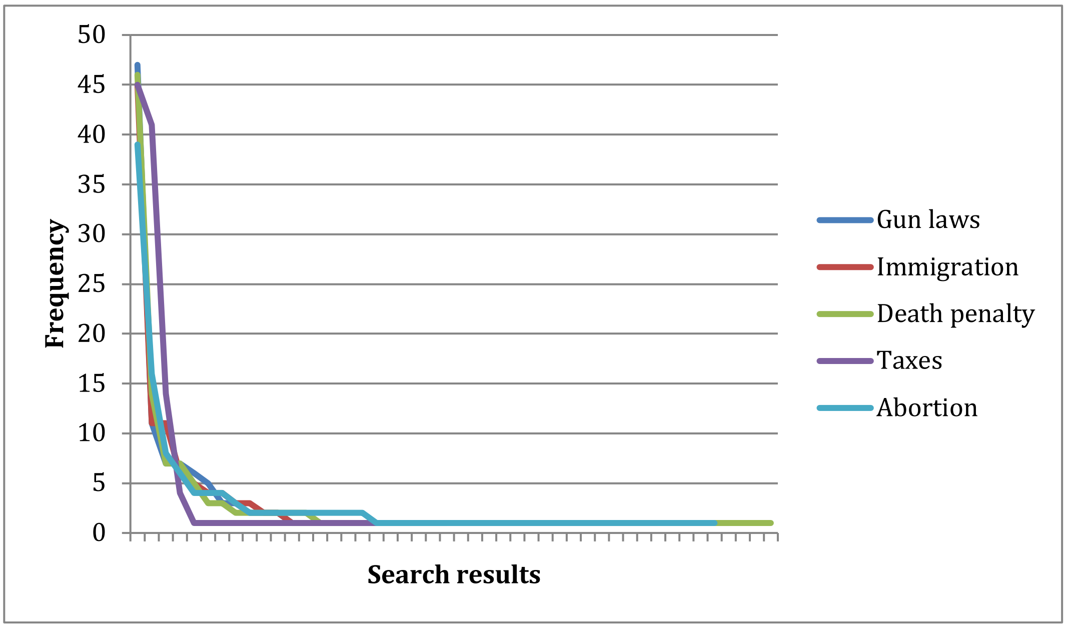

Figure 1. Long tail distribution of search results

To measure the concentration among the search results, I use the Gini index, which showed that the concentration among the search results is mostly moderate. As Table 3 shows, the Gini coefficients of the queries range from 0.588 to 0.758. Four values are below the threshold of 0.71, indicating moderate concentration. The Gini coefficient of the replies to the question about taxes, however, is above 0.71. This value suggests high concentration.

Table 3. Gini coefficients of queries

Query	Gini coefficient
Should there be stricter gun laws?	0.630
Should immigration be limited?	0.611
Should the death penalty be abolished?	0.599
Should taxes be lowered?	0.758
Should abortion be illegal?	0.588

To measure the fragmentation among the search results, I examined the percentage of non-overlapping content as previously explained. To do so, I turned to the unique answers. Table 2 already provided an overview of the percentages of these replies for each query. They range from 15% to 25%. In total, 22% of the users received a personalized answer.

According to previously defined criteria, the fragmentation of search results is low. For every query, and therefore in total as well, the percentage of non-overlapping content is below 31%. Even though the data suggest that Siri's search algorithm causes only low fragmentation, it needs to be emphasized *that* is caused fragmentation throughout the searches at all.

So far the analysis has shown that Siri's search algorithm produced a long tail of replies and consequently caused moderate concentration and low fragmentation. This finding raises the question whether the identified concentration or fragmentation was driven by the political orientation, gender, age, or location of users.

It appears that Siri does not provide one group of politically like-minded users disproportionally frequently with top results or unique answers. The most frequent answers to each of the five queries were given to liberals, conservatives, people with other political views, and users who have no political leaning. Unique answers also reached a politically diverse audience.

The data suggest Siri's search algorithm is biased to some extent towards the gender of users. Siri's search algorithm provided female and male participants equally with unique search results. However, it caused concentration based on gender in relation to the most common answers: Siri provided men with the most frequent replies disproportionally often (see Table 4). Seventy-eight to eighty percent of the users who received five of the six replies were male, although they only made up 57% of the sample. The other reply represents an outlier: 76% of the participants who were provided with this answer were female, although women made up only 43% of the sample.

Table 4. Gender of users who received the most common search results

Question	Participants	Gender
Should there be stricter gun laws?	47	Female: 10 (21%) Male: 37 (79%)
Should immigration be limited?	45	Female: 9 (20%) Male 36 (80%)
Should the death penalty be abolished?	46	Female: 10 (22%) Male: 36 (78%)
Should taxes be lowered?	45	Female: 10 (22%) Male: 35 (78%)
	41	Female: 31 (76%) Male: 8 (19%) Other: 2 (5%)
Should abortion be illegal?	39	Female: 8 (21%) Male: 31 (79%)

It seems that Siri's search algorithm is biased towards location. The data indicate that Siri's search algorithm does not cause fragmentation, as unique answers were given to a diverse set of participants, but concentration based on the location of users. Siri appears to provide users from Pennsylvania more often with the top answers than others (see Table 5). The percentages of Pennsylvanians who were exposed to the most frequent answers (13% to 16%) is in five out of six cases roughly three times its percentages in the sample (5%) (see Table A1). However, due to the small number of users from different states who received the most frequent answers (see Table A3) (Tversky and Kahneman (1971) caution against relying on small numbers) I consider this finding to be highly tentative.

Table 5. Percentage of users from Pennsylvania who received the most common search results

Question	Should there be stricter gun laws?	Should immigration be limited?	Should the death penalty be abolished?	Should taxes be lowered?	Should abortion be illegal?
Percentage of users based in PA	15%	13%	15%	16%	13%

The data suggest that Siri's search algorithm is not biased towards age. It seems that Siri's top search results do not reach one age group disproportionally frequently. In addition, Siri provided users of different ages with personalized answers. The results indicate that Siri's search algorithm does neither cause concentration nor fragmentation based on the age of users.

4 DISCUSSION

This study sought to discover to what extent Siri's search algorithm provides 18 to 64-year-old US-based users with different answers to the same politically controversial questions. The main finding is that Siri's search algorithm seems to produce a long tail distribution of search results. Across the audit's five questions, it provided 42% of the users, disproportionally many men, with six answers. At the same time, Siri presented 22% of the participants with personalized search results. In the context of this study, these data indicate that Siri's search algorithm causes moderate concentration and low fragmentation. However, the thresholds for concentration and fragmentation were set rather arbitrarily due to a lack of established ones. If different thresholds had been chosen, different degrees of concentration and fragmentation would have been noticed.

4.1 Evaluation of findings

The finding that Siri's search algorithm causes both moderate concentration and low fragmentation is important in three ways. Firstly, the finding that the voice assistant has caused fragmentation, even though to a relatively low degree, is relevant in itself. Several scholars have claimed that democracies benefit from the presentation of a variety of voices and viewpoints of public life in diverse media content (e. g. McQuail, 1992; Benson, 2013; Anderson, 2016). Diversity is important when it comes to political information (Diakopoulos et al., 2018, p. 337–338). However, personalized search results can undermine the diversity of political dialogue (Gillespie, 2014, p. 188). If users received strongly biased answers to the audit's politically controversial questions which correspond to their own world view, this would undermine information diversity. However, due to this study's focus, the content of the search results was not analyzed. It is therefore unknown to what extent ideological diversity or topical diversity among the results exists.

Secondly, the concentration of search results, even though it was only mostly moderate, is an interesting finding. Scholars who are concerned with fragmentation claim that a unified body of relevant information should be accessible to the entire society (e. g. Gitlin, 1998; Sunstein, 2002). I mostly agree – especially when it comes to politically controversial questions. However, it is problematic that Siri's search algorithm decides how to answer these questions. Even if this study had analyzed the content of the six answers that 42% of the users heard, it would remain unclear why they were selected. The question arises whether the top search results managed to become "algorithmically recognizable" (Gillespie, 2017, p. 63); in other words, were these replies optimized to be found by Siri's search algorithm? Or were these results "privileged sources" (Diakopoulos et al., 2018, p. 331) selected by Siri? If so, their dominance might be based on some bias of the voice assistant (Trielli & Diakopoulos, 2019, p. 12). Due to the lack of transparency, these questions remain unanswered; in fact, they could not have been answered even if this study had had a different focus. Since Siri's search algorithm gave some answers significantly more often than others, it is clear, however, that Siri promotes some search results over others, just as other algorithms do (see Gillespie, 2015, p. 1). By doing so, it "shapes the things people encounter" (Beer, 2009, p. 1000). Because these *things* are answers to politically controversial questions, Siri's search algorithm actively influences the political information environment.

The finding that male participants more often received the most frequent results worrying, though not surprising. Gender biases in algorithms have long been known, and well documented, (e. g. see Bolukbasi et al., 2016; Fabris et al., 2020). This finding can be linked to the general issue of 'fairness' in searches (Ekstrand et al., 2019) as well as to 'group fairness' in particular (Fabris et al., 2020, p. 5). One might wonder why men received the frequent replies more often than women did. Although a more in-depth answer to this question would go beyond the scope of this study, Friedman and Nissenbaum's (1996) concept of algorithm bias suggests two possible explanations. The identified gender bias could either be a pre-existing bias resulting from widespread social values and, hence, of the programmers themselves. Another explanation could be the or an emergent bias resulting from feedback loops caused by the users' behavior (Friedman & Nissenbaum, 1996, pp. 334-335). The former seems to be more likely, since biases in algorithms tend to reflect biases in society (Bolukbasi et al., 2016, p. 8).

4.2 Relation to previous research

The third way in which the finding that Siri's search causes both moderate concentration and low fragmentation is important becomes obvious when relating it to the outcome of other research.

At first sight, the findings of this study challenge some of the research that was discussed in the introduction. This study's data have shown that Siri's search algorithm provides roughly 22% of users with unique search results. This result suggests that Siri's search algorithm causes low fragmentation. In contrast, the empirical research of Haim et al. (2018), Curtois et al. (2018), Robertson et al. (2018), Bechmann and Nielbo (2018), Nechusthai and Lewis (2019) and Krafft et al. (2019) claims that there is little evidence of filter bubbles. Despite their different geographical foci and methodological approaches, these scholars found no evidence of fragmentation. This study's results seem to contradict them.

However, upon closer examination, the data of this study and the results of some of the previously mentioned academic works are similar. For example, Krafft et al. (2019) found that, on average, around two to four results out of ten were personalized, although this depended on the search term (p. 1). Similarly, Bechmann and Nielbo (2018) reported that roughly 10% to 28% (depending on the methodological approach) of users were presented with different content (p. 990).

While the data of this study as well as the studies reviewed here are comparable, the scholars have interpreted their results differently. One reason might be that they focused exclusively on fragmentation. Because of this, their perspective tends to be quite narrow: for them fragmentation, often in the form of filter bubbles, either exists or not. This study offers a different, more nuanced, view by focusing on *both* fragmentation and concentration. Its finding proves that fragmentation and concentration can coexist when it comes to search results given by Siri's search algorithm. This wider approach can provide a clearer understanding of content diversity and therefore provide new insights into the frequency distribution of search results.

The finding that female users received less often the most frequent results adds to the existing literature on gender bias. So far, studies have found out that search results reinforce gender stereotypes (e. g. Kay et al., 2015; Otterbacher et al., 2017; Bolukbasi et al., 2016; Fabirs, et al., 2020). Taking another perspective, this study has indicated that algorithms of search engines, in this case Siri, can also be biased towards gender by providing female and male users with different results.

4.3 Limitations

Some limitations exist. First and foremost, the findings of this study are not representative. I tested Siri's search algorithm using only 170 Amazon participants. None of these Siri users were found by using probability sampling. As a result, I cannot generalize from this sample to the larger population of US-based Siri users. However, it was never the aim of this research to make claims about a population. Instead, this study has sought to provide a mere indication to what extent Siri provides 18 to 64-year-old US-based users with different answers to the same questions.

Secondly, a high internal validity of the data could not be guaranteed. The study's participants made the queries themselves. As I was not present during the data collection, I cannot assure that the Siri users reported back the replies to the queries they were supposed to make.

Thirdly, this study's results are necessarily constrained by its context. The five queries that were used in the audit have resulted in concentration and fragmentation alike. It is unknown whether other, perhaps less controversial, searches (or, for that matter, different participants) would yield answers that have a different frequency distribution. Similarly, the findings here are US-based. Search algorithms might work differently in different countries; in fact, there is evidence that they produce localized outcomes (Kitchin, 2017, p. 25). It is possible that users who are based outside the USA would receive different answers if they asked Siri the same questions.

Fourthly, this study's focus is highly limited. I decided to research only the extent to which Siri's search results differ. The contents of the search results, including ideological or topical differences, have been disregarded. As a consequence, the information that Siri provided users with was not analyzed. In addition, due to the focus of this study and especially because of its theoretical framework I excluded the perspective of the users themselves. Even though scholars such as Gillespie (2014) and Bucher (2017) are increasingly focusing on users in order to understand algorithms more holistically, I decided not to do this. This limitation is not problematic, however. My research question only focuses on the outcomes of Siri's search results and the concept of algorithmic bias enables me to analyze them.

Fifthly, the results might not be replicable. As algorithms are constantly tested (e. g. through a/b testing), developed further and changed, their outcomes are likely to vary over time (Gillespie, 2014, p. 178; Diakopoulos et al., 2018, p. 322). For this reason, the results of the audit, which was conducted on April 21, 2021, are time-based. If the same users had performed the same searches two weeks later, they might have received different replies.

4.4 Future research needed

Due to this study's limitations, further research is needed. The small number of participants and the potentially low internal validity of the data were identified as the main limitations. Therefore, research with representative sample sizes and a more closely controlled environment (e. g. by being in the same place with participants or being virtually connected via a program such as Zoom) would have to be conducted to support or challenge the finding that Siri's search algorithm causes both concentration, especially among male users, as well as fragmentation. Such studies could use different search terms. However, they should not focus on a different location than the US because Siri's algorithm is likely to work differently in another country.

If future studies are able to disprove that Siri's search algorithm produces a long tail of search results, more diverse research would be needed. To avoid focusing only on the role of technology, the role of Siri's users would need to be considered. This approach would be in line with the growing number of scholars who see algorithms as the entanglement between the technical and social (e. g. Seaver, 2013; Bozdag, 2013; Just & Latzer, 2017). For example, inspired by Bucher (2017), one could ask how users feel about receiving unique search results or the same as other users. Or: how do users react to the replies they receive? In addition to a crowdsourced audit, interviews with users could be conducted. By doing so, not only the frequency distribution of results, but also their impact, could be studied.

FUNDING STATEMENT AND ACKNOWLEDGMENTS

The author would like to thank the two anonymous reviewers for their valuable comments and suggestions for improving the quality of this paper.

REFERENCES

Agichtein, E., Brill, E., Dumais, S., & Ragno, R. (2006). Learning user interaction models for predicting web search result preferences. *SIGIR '06: Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 3–10. https://doi.org/10.1145/1148170.1148175

Anderson, C. (2016). Assembling publics, assembling routines, assembling values: Journalistic self-conception and the crisis in journalism. In J. Alexander, E. Breese, & M. Luengo (Eds.), *The Crisis of Journalism Reconsidered: Democratic Culture, Professional Codes, Digital Future* (pp. 153-169). Cambridge University Press. https://doi.org/10.1017/CBO9781316050774.010

Barker, R. (2018). Trapped in the Filter Bubble? Exploring the Influence of Google Search on the Creative Process. *Journal of Interactive Advertising*, *18*(2), 85–95. https://doi.org/10.1080/15252019.2018.1487810

Bechmann, A., & Nielbo, K. L. (2018). Are We Exposed to the Same "News" in the News Feed? An empirical analysis of filter bubbles as information similarity for Danish Facebook users. *Digital Journalism*, 6(8), 990–1002. https://doi.org/10.1080/21670811.2018.1510741

Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society*, 11(6), 985–1002. https://doi.org/10.1177/1461444809336551

Benson, R. (2013). *Shaping immigration news: A French-American comparison*. Cambridge University Press. https://doi.org/10.1017/CBO9781139034326

Blaikie, N., & Priest, J. (2019). *Designing Social Research*. Polity Press.

Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. *Advances in Neural Information Processing Systems 29*, 4349–4357.

Boothby, H. (2018, October 15). *Radio Production Autumn 2018* [Lecture recording]. Canvas@Malmö University. https://mau.se/canvas/

Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, *15*(3), 209–227. https://doi.org/10.1007/s10676-013-9321-6

Bruns, A. (2019). Filter bubble. *Internet Policy Review*, *8*(4), 1–14. https://doi.org/10.14763/2019.4.1426

Bucher, T. (2017). The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, *20*(1), 30–44. https://doi.org/10.1080/1369118x.2016.1154086

Chams, Z. (2020, October 15). *Republicans vs Democrats: Where do the two main US political parties stand on key issues?*. ABC News. https://www.abc.net.au/news/2020-10-15/us-election-political-parties-explained-democrats-vs-republicans/12708296

Cho, J., Ahmed, S., Hilbert, M., Liu, B., & Luu, J. (2020). Do Search Algorithms Endanger Democracy? An Experimental Investigation of Algorithm Effects on Political Polarization. *Journal of Broadcasting & Electronic Media*, *64*(2), 150–172. https://doi.org/10.1080/08838151.2020.1757365

Curtois, C., Slechten, L., & Coenen, L. (2018). Challenging Google Search filter bubbles in social and political information: Disconforming evidence from a digital methods case study. *Telematics Informatics*, 35(7), 2006–2015. https://doi.org/10.1016/j.tele.2018.07.004

Dambanemuya, H. J., & Diakopoulos, N. (2020, March 20-21). *"Alexa, what is going on with the impeachment?" Evaluating smart speakers for news quality* [Paper presentation]. Computation and Journalism Symposium, Boston, MA, USA. https://cj2020.northeastern.edu/research-papers/

Dellinger, A. J. (2019, April 30). *Survey says that Siri and Google Assistant are the most used voice assistants*. Digitaltrends.

https://www.digitaltrends.com/home/siri-google-asistant-most-used-voice-assistants-alexa/

Diakopoulos, N. (2015). Algorithmic Accountability. *Digital Journalism*, 3(3), 398–415. https://doi.org/10.1080/21670811.2014.976411

Diakopoulos, N., Trielli, D., Stark, J., & Mussenden, S. (2018). I vote for – How search informs our choice of candidate. In M. Moore & T. Tambini (Eds.), *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple* (pp. 320–341). Oxford University Press.

Dorfman, R. (1979). A Formula for the Gini Coefficient. *The Review of Economics and Statistics*, 61(1), 146–149. https://doi.org/10.2307/1924845

Ekstrand, M. D., Burke, R., & Diaz, F. (2019). Fairness and discrimination in retrieval and recommendation. *RecSys '19: Proceedings of the 13th ACM Conference on Recommender Systems*, 576-577. https://doi.org/10.1145/3298689.3346964

Fabris, A., Purpura, A., Silvello, G., & Susto, G. A. (2020). Gender stereotype reinforcement: Measuring the gender bias conveyed by ranking algorithms. *Information Processing & Management*, *57*(6), 102377. https://doi.org/10.1016/j.ipm.2020.102377

Feezell, J. T., Wagner, J. K., & Conroy, M. (2021). Exploring the effects of algorithm-driven news sources on political behavior and polarization. *Computers in Human Behavior*, 116, 1–11. https://doi.org/10.1016/j.chb.2020.106626

Fletcher, R., & Nielsen, R. K. (2017). Are News Audiences Increasingly Fragmented? A Cross-National Comparative Analysis of Cross-Platform News Audience Fragmentation and Duplication. *Journal of Communication*, 67(4), 476–498. https://doi.org/10.1111/jcom.12315

Fletcher, R., Cornia, A., & Nielsen, R. K. (2020). How Polarized Are Online and Offline News Audiences? A Comparative Analysis of Twelve Countries. *The International Journal of Press/Politics,* 25(2), 169–196. https://doi.org/10.1177/1940161219892768

Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transactions on Information Systems*, 14(3), 330–347. https://doi.org/10.1145/230538.230561

Geschke, D., Lorenz, J., & Holtz, P. (2019). The triple-filter bubble: Using agent- based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers. *British Journal of Social Psychology*, 58(1), 129–149. https://doi.org/10.1111/bjso.12286

Gillespie, T. (2014). The relevance of algorithms. In: T. Gillespie. T., P. Boczkowski & K. Foot (Eds.), *Media Technologies: Essays on Communication, Materiality, and Society* (pp. 167–194). MIT Press. https://doi.org/10.7551/mitpress/9780262525374.003.0009

Gillespie, T. (2015). Platforms Intervene. *Social Media & Society*, 1–2. https://doi.org/10.1177/2056305115580479

Gillespie, T. (2017). Algorithmically recognizable: Santorum's Google problem, and Google's Santorum problem. *Information, Communication & Society*, 20(1), 63–90. https://doi.org/10.1080/1369118X.2016.1199721

Gitlin, T. (1998). Public sphere or public sphericules? In J. Curran & T. Liebes, (Eds.). *Media, ritual and identity* (pp. 168–174). Routledge. https://doi.org/10.4324/9780203019122

Gutmann, A., & Thompson, D. F. (2004). *Why Deliberative Democracy?*. Princeton University Press.

Haim, M., Graefe, A., & Brosius, H. (2018). Burst of the Filter Bubble?. *Digital Journalism*, 6(3), 330–343. https://doi.org/10.1080/21670811.2017.1338145

Introna, L. D., & Nissenbaum, H. (2000). Shaping the web: Why the politics of search engines matters. *The Information Society*, 16(3), 169–185. https://doi.org/10.1080/01972240050133634

Jurkowitz, M., Mitchell, A., Shearer, E., & Walker, M. (2020). *U.S. Media Polarization and the 2020 Election: A Nation Divided: Deep partisan divisions exist in the news sources Americans trust, distrust and rely on*. Pew Research Center. https://www.journalism.org/2020/01/24/u-s- media-polarization-and-the-2020-election-a-nation-divided/

Just, N., & Latzer, M. (2017). Governance by algorithms: reality construction by algorithmic selection on the Internet. *Media, Culture & Society*, 39(2), 238–258. https://doi.org/10.1177/0163443716643157

Kay, M., Matuszek, C., & Munson, S. A. (2015). Unequal representation and gender stereotypes in image search results for occupations. *CHI '15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3819–3828. https://doi.org/10.1145/2702123.2702520

Kinsella, B. (2020, November 5). *Voice Assistant Use on Smartphones Rise, Siri Maintains Top Spot for Total Users in the U.S.*. Voicebot. https://voicebot.ai/2020/11/05/voice-assistant-use-on-smartphones-rise-siri-maintains-top-spot-for-total-users-in-the-u-s/

Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society*, 20(1), 14–29. https://doi.org/10.1080/1369118X.2016.1154087

Krafft, T. D., Gamer, M., & Zweig, K. A. (2019). What did you see? Personalization, regionalization and the question of the filter bubble in Google's search engine. EPJ Data Science, 8(38), 1–23. https://doi.org/10.1140/epjds/s13688-019-0217-5

Lambert, P. J. (2001). *The Distribution and Redistribution of Income.* Manchester University Press.

Lehn J., Müller-Gronbach T., & Rettig S. (2000). Konzentrationsmaße. In J. Lehn, T. Müller-Gronbach & S. Rettig (Eds.), *Einführung in die Deskriptive Statistik. Teubner Studienbücher Mathematik* (pp. 50-64). Vieweg & Teubner Verlag. https://doi.org/10.1007/978-3-322-80099-2_3

McQuail, D. (1992). *Media performance: Mass communication and the public interest*. Sage.

Milanovic, B. (1997). A simple way to calculate the Gini coefficient, and some implications. *Economics Letters*, 56(1), 45-49. https://doi.org/10.1016/S0165-1765(97)00101-8

Najle, M., & Jones, R. P. (2019, February 19). *American Democracy in Crisis: The Fate of Pluralism in a Divided Nation*. PRRI. https://www.prri.org/research/american-democracy-in-crisis-the-fate-of-pluralism-in-a-divided-nation/

Natale, S., & Cooke, H. (2021). Browsing with Alexa: Interrogating the impact of voice assistants as web interfaces. *Media, Culture and Society*, 43(6), 1000–1016. https://doi.org/10.1177/0163443720983295

Nechusthai, 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

Olson, K. (2011). Deliberative democracy. In B. Fultner (Ed.), *Jürgen Habermas: Key Concepts* (pp. 140–155). Acumen Publishing. https://doi.org/10.1017/UPO9781844654741.008

Otterbacher, J., Bates, J., & Clough, P. D. (2017). Competent Men and Warm Women: Gender Stereotypes and Backlash in Image Search Results. *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 6620–6631. https://doi.org/10.1145/3025453.3025727

Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Press.

Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms that Control Money and Information*. Harvard University Press.

Robertson, R. E., Jiang, S., Joseph, K., Friedland, L., Lazer, D., & Wilson, C. (2018). Auditing Partisan Audience Bias within Google Search. *Proceedings of the ACM on Human-Computer Interaction*, 2, 1–22. https://doi.org/10.1145/3274417

Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014, May 22). *Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms*. [Paper presentation]. 64th Annual Meeting of the International Communication Association. Seattle, WA, USA.

Seaver, N. (2013). Knowing Algorithms. *Media in Transition*, 8, 412–422.

Sunstein, C. R. (2002). *Republic.com*. Princeton University Press.

Trielli, D., & Diakopoulos, N. (2019). Search as News Curator: The Role of Google in Shaping Attention to News Information. *CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3290605.3300683

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, 76(2), 105–110. https://doi.org/10.1037/h0031322

Valishery, L. S. (2021, January 22). *Share of voice assistant users in the U.S. 2020 by device*. Statista. https://www.statista.com/statistics/1171363/share-of-voice-assistant-users-in-the-us-by-device/

Vike-Freiberga, V., Däubler-Gmelin, H., Hammersley, B., & Pessoa Maduro, L. M. P. (2013). *A free and pluralistic media to sustain European democracy*. High Level Group on Media Freedom and Pluralism. http://ec.europa.eu/digital-agenda/sites/digital-agenda/files/HLG%20Final%20Report.pdf

Wagner, P. (2018, June 29). *Siri Remains The Most Used Mobile Voice Assistant*. Statista. https://www.statista.com/chart/14505/market-share-of-voice-assistants-in-the-us/

Webster, G. J., & Ksiazek, T. B. (2012). The Dynamics of Audience Fragmentation: Public Attention in an Age of Digital Media. *Journal of Communication*, 62(1), 39–56. https://doi.org/10.1111/j.1460-2466.2011.01616.x

Yitzhaki S., & Schechtman E. (2013). More Than a Dozen Alternative Ways of Spelling Gini. In S. Yitzhaki & E. Schechtman (Eds.), *The Gini Methodology:* A Primer on a Statistical Methodology (pp. 11–31). Springer. https://doi.org/10.1007/978-1-4614-4720-7_2

Zuiderveen Borgesius, F. J., Trilling, D., Möller, J., Bodó, B., de Vreese, C. H., & Helberger, N. (2016). Should we worry about filter bubbles? Internet Policy Review, 5(1), 1–16. https://doi.org/10.14763/2016.1.401

APPENDIX

Table A1. Characteristics of the sample (N = 134)

Age	18-24: 8 25-34: 87 35-44: 14 45-54: 16 55-64: 8
Gender	Female: 57 Male: 74 Other: 3
State	Alabama: 3 Arkansas: 1 Arizona: 2 California: 9 Connecticut: 8 Delaware: 1 Florida: 8 Georgia: 3 Idaho: 1 Illinois: 8 Indiana: 2 Kansas: 1 Kentucky: 2 Michigan: 5 Minnesota: 2 Mississippi: 3 Nevada: 3 New Jersey: 8 New York: 26 Pennsylvania: 7 Rhode Island: 2 South Carolina: 1 Tennessee: 2 Texas: 7 Utah: 2 Virginia: 4 Washington: 3 Wisconsin: 1 West Virginia: 1 Only indicating USA: 8
Political inclination	Liberal: 61 Conservative: 43 Other: 4 No political leaning: 26

Table A2. Number and frequencies of search results

	Questions				
Number of replies	Should there be stricter gun laws?	Should immigration be limited?	Should the death penalty be abolished?	Should taxes be lowered?	Should abortion be illegal?
1	11	1	46	4	1
2	1	1	14	1	1
3	1	1	1	1	1
4	1	1	1	1	1
5	5	1	1	1	1
6	1	1	1	1	6
7	1	1	1	1	2
8	1	2	1	1	16
9	7	4	1	1	1
10	2	1	3	1	2
11	1	1	1	1	2
12	1	1	1	45	1
13	1	1	2	1	1
14	1	1	1	1	1
15	1	1	2	41	2
16	1	1	2	1	1
17	2	4	1	1	2
18	1	1	1	1	1
19	1	1	1	1	1
20	1	1	1	1	1
21	1	1	1	1	4
22	1	6	1	14	1
23	1	1	1	1	1
24	7	1	1		1
25	2	1	1		1
26	47	11	1		2
27	1	1	5		4
28	3	1	1		39
29	1	1	1		1
30	1	2	2		1
31	1	1	1		8
32	1	3	1		1
33	1	1	7		4
34	1	45	1		2
35	3	5	2		1
36	6	11	1		1
37	1	1	1		1
38	1	1	1		3
39		1	1		2
40		3	1		1
41		1	1		1
42		1	1		2
43		1	3		
44			7		
45			1		
46			2		

Table A3. Location of users who received the most common search results

Question	**Should there be stricter gun laws?**	**Should immigration be limited?**	**Should the death penalty be abolished?**	**Should taxes be lowered?**		**Should abortion be illegal?**
Participants	47	45	46	45	41	39
Location	AR: 1	AR: 1	AR: 1	AZ: 1	AL: 1	AR: 1
	AZ: 1	AZ: 1	CT: 1	CT: 1	AR: 1	AZ: 1
	CT: 1	CT: 1	FL: 4	FL: 3	CA: 2	CT: 1
	FL: 5	FL: 3	GA: 3	GA: 3	CT: 7	FL: 3
	GA: 3	GA: 3	ID: 1	ID: 1	FL: 1	GA: 3
	ID: 1	ID: 1	IN: 1	IN: 1	IL: 7	ID: 1
	IN: 1	IN: 1	KS: 1	KS: 1	MI: 2	KS: 1
	KS: 1	KS: 1	KY: 1	KY: 1	MS: 1	KY: 1
	KY: 1	KY: 1	MI: 1	MI: 1	NV: 1	MI: 1
	MN: 2	MI: 1	MN: 1	MN: 2	NJ: 3	MN: 1
	MS: 1	MN: 2	MS: 2	MS: 2	NY: 6	MS: 1
	NV: 1	MS: 2	NJ: 3	NJ: 2	TX: 1	NJ: 1
	NJ: 2	NJ: 2	NY: 8	NY: 8	VA: 2	NY: 8
	NY: 7	NY: 8	PA: 7	PA: 7	WA: 2	PA: 5
	PA: 7	PA: 6	RI: 2	RI: 2	USA: 4	RI: 2
	RI: 2	RI: 2	SC: 1	SC: 1		SC: 1
	SC: 1	SC: 1	TN: 1	TN: 2		TN: 2
	TN: 2	TN: 2	TX: 1	TX: 2		TX: 2
	TX: 3	TX: 3	UT: 2	UT 2		UT: 1
	UT: 2	UT: 1	VA: 1	VA: 1		WA: 1
	VA: 2	VA: 1	WA: 1	WA: 1		
	WA: 1	WA: 1				