Bachelor thesis
Computing Science

Radboud University

Search personalization on Google.nl

Author:
Roland Leferink
S4557972

First supervisor/assessor:
Dr. ir. Eelco Herder
eelcoherder@cs.ru.nl

Second assessor:
Dr. Harrie Oosterhuis
harrie.oosterhuis@ru.nl

March 19, 2021

Abstract

With more and more people gaining access to the internet, search engines like Google gain a more important role as the gatekeepers to information. With this role of gatekeeper also comes the power to decide what a user can see when searching for something, will you show a user both perspectives on a certain topic, or will you only show the perspective that you like?

In this thesis we will try to uncover if internet users in the Netherlands are caught in a so-called filter bubble, an online echo chamber, that only caters to a person's interest, created by Google. This could be done by selectively choosing which users get which results when searching for the same subject. More specifically, we will test if certain controversial topics are more likely to result in a more curated search result than more normal, non-controversial topics.

Our research shows that, while there are statistically significant differences between controversial and non-controversial search queries in terms of how unique the search results are, there isn't a clear reason as to why this is the case.

Contents

1 Introduction **3**

2 Preliminaries **5**
- 2.1 Google Search ranking 5
- 2.2 PageRank 5
- 2.3 Contextualization 6

3 Related Work **8**
- 3.1 What did you see? Personalization, regionalization and the question of the filter bubble in Google's search engine 8
- 3.2 Location, Location, Location: The Impact of Geolocation on Web Search Personalization 9
- 3.3 Measuring Political Personalization of Google News Search . 10

4 Research **12**
- 4.1 Questionnaire 12
  - 4.1.1 Why a questionnaire 12
  - 4.1.2 Why these questions 12
  - 4.1.3 Distributing questionnaire 14
- 4.2 Data processing 14
  - 4.2.1 Python script 14
- 4.3 Website standings 15
- 4.4 Comparing search results 15
  - 4.4.1 Levenshtein distance functioning 15
  - 4.4.2 Why Levenshtein distance 15
  - 4.4.3 Similarity 16
  - 4.4.4 Why similarity 16

5 Results **17**
- 5.1 Survey demographics 17
  - 5.1.1 Personal information 17
  - 5.1.2 Controversial survey 19
  - 5.1.3 Non-controversial survey 21
- 5.2 Websites 21

5.2.1 General 21
5.2.2 Controversial websites 22
5.2.3 Non-controversial websites 24
5.3 Statistical significance 24
5.3.1 T-test 25
5.3.2 Cause 25
5.3.3 Interference by Google 26

6 Conclusions **27**

7 Future Work **28**
7.1 Search results between search engines 28
7.2 More diverse test population 28

A Appendix **32**
A.1 Controversial Questionnaire 32
A.2 Non-controversial Questionnaire 35

Chapter 1

Introduction

In his 2011 book, The Filter Bubble: What the Internet Is Hiding from You, [10] Eli Parser came up with the term "filter bubble". He describes it as: "A personal ecosystem of information that's been catered by algorithms to who they think you are." In other words, when two people are using the same questions, they might receive different answers based on assumptions made by an algorithm. While most of the personalized results have good intentions, like making sure you don't see the website of a bakery on the other side of the planet when searching for local bakeries, or only receiving Chinese websites when searching in Dutch, these algorithms could prevent you from seeing relevant information because of their assumptions.

In this thesis we will investigate if the algorithms that Google uses actually lead to differing search results for the same search query, or if they actually create a filter bubble. The first reason for investigating this, is because Google is by far the most used search engine with a market share of around 90%[1]. The second reason is that, if users aren't receiving relevant information, because they think Google is impartial and shows everybody the same results but instead results based on a profile built by their algorithms, users should be made aware of this.

Even though a difference in search results on Google has already been studied in the past, this hasn't been done on `Google.nl` and the focus is usually on different search results for political search queries, like the names of politicians and political parties [7] or news articles [9]. For this thesis we will focus on `Google.nl`, since this kind of research is usually very dependent on language and location, and we will specifically focus on both controversial and non-controversial queries to see if there is a difference between these two.

As we mentioned in the previous paragraph studies like this have already

[1]https://gs.statcounter.com/search-engine-market-share

been done before in the past, but never for the Netherlands, and the search queries in previous studies were never split between controversial and non controversial topics to see if there is a noticeable difference.

Our research will be done by comparing the search results of around 200 people, split into two groups of roughly 100: one for each set of queries. We can then use statistical analysis to determine if the search results differ too much for it to be a coincidence, which could mean that Google filters possible relevant search results from certain users. By using the exact same search query for every person in the group, and registering when and where they search the mentioned queries, we can filter out the expected differing search results caused by location and time which leaves us with search results that could only differ because of personalization based on information by Google.

In chapter 2, we will give some necessary background information on how Google search works. In chapter 3, we will compare our work to related studies and highlight a couple of key differences and similarities. In chapter 4 we will explain the methods used to obtain our data, and why we made certain choices with regards to using this data. Chapter 5 will be used to present the results of our research. In chapter 6 we will give a conclusion based on our results. Finally, in chapter 7 we will give some suggestions for future research.

Chapter 2

Preliminaries

2.1 Google Search ranking

While Google is very confident in their explanation of how their search algorithm works[1], its precise working isn't exactly known. While previous research has reported significant personalization based on location, and Google has been somewhat open about this, using searching for the term "football" in the US and the UK as an example of how they personalize your search results, it isn't very clear how much other factors like a user's browser history contribute to it [15, 7, 5]. While Google does admit that they use browser history in their personalization, the examples that they give for it are incredibly specific, e.g. If you search for Barcelona after recently searching Barcelona against Arsenal, it uses that to give you information about the football, club and not the city. Besides this the only other thing Google has published about their search algorithm, is that they don't use it to deduce personal and sensitive info from it, like race, religion, or political affiliation.

2.2 PageRank

One of the more well-known and oldest algorithms that Google uses to decide which web page ends up where on the search results page is PageRank. This algorithm was developed by Google founders Larry Page and Sergey Brin, when they were still PhD students at Stanford, and is still in use today. While it's not the most important search algorithm anymore, and thus isn't used as one of the main ranking factors to decide which web page goes where on the search results page, those are Links, Content[2] and RankBrain[3], it does give a bit of an insight into what Google uses to rank search results.

[1] https://www.google.com/search/howsearchworks/

[2] https://web.archive.org/web/20160325232656/http://www.bloomberg.com/news/articles/2015-10-26/google-turning-its-lucrative-web-search-over-to-ai-machines

[3] https://daizy.media/googles-zoekalgoritme-mythes-bevestigd/

In short, PageRank works by ranking web pages based on how often other web pages link to that web page, and the quality of those web pages. The output of PageRank is a probability distribution that represents the probability that someone randomly clicking on links ends up at a particular page in a set of web pages.

A simple explanation of how PageRank calculates the output distribution is as follows: Initially all web pages n in the set N get the same probability $1/N$ as their PageRank value $PR(n)$. Then for the first iteration each page n divides its $PR(n)$ equally over its outbound links $L(n)$. This process would then be repeated for a certain number of times to eventually get the final PageRank for every web page. So, in general the following formula can be used to calculate the PageRank value for a page u:

$$PR(u) = \sum_{v \in B_u} \frac{PR(v)}{L(v)}$$

Here $PR(u)$ is the PageRank value of u after the iteration, v is a page that links to u, B_u is the set of v thus the set that contains all the pages that have a link to u, $PR(v)$ the previous (or initial if it's the first iteration) PageRank value of page v, $L(v)$ the total amount of outbound links on page v.

While in reality the algorithm is a bit more complex, with also a damping factor that represents the imaginary user that eventually stops clicking on links, and some extra rules for web pages that don't have any outgoing links to make it more fair for pages that do, these aren't very relevant in this case to the overall use of PageRank and what it means for search results.

Finally, after computing the PageRank values for a set of web pages Google uses these values in combination with other ranking factors to determine which pages should be near the top of the search results, and which should be lower.

2.3 Contextualization

On December 9$^{\text{th}}$, 2019 Google rolled out their biggest update in the last five years in over 70 languages including Dutch, which is the language we used for our search queries. This update applies a new way of contextualization called Bidirectional Encoder Representations from Transformers or BERT for short[4]. BERT was developed by Google as a new technique for natural

[4]https://blog.google/products/search/search-language-understanding-bert/

language processing to better understand search queries and thus to return better search results [3].

Language representation models come in two different forms: context-free and contextual. These contextual models can then either be unidirectional or bidirectional. Context-free models don't take into account the context of where a word is placed. For example, the word "bill" has the same representation in "duck bill" and "restaurant bill". A contextual representation would also look at the other words in the sentence to give a more accurate representation of the word "bill". This can be done either unidirectional where only one "direction" of the sentence is used for context, or bidirectional where both "directions" would be used. For example in the sentence "I touched the bill of my pet duck" a unidirectional contextual model would represent "bill" based on "I touched the" but not "of my pet duck" or vice versa. However, BERT would represent "bill" based on both the previous and the next context because of it being bidirectional. This helps returning better search results because it can make more use of context in the query than before.[5]

In a blog posted earlier that year[6] Google gave some good examples on how the use of BERT improved search results. One of these examples was the query "Parking on a hill with no curb". Before the implementation of BERT Google's systems would focus too much on the word "curb" and not on the "no" in front of it and would thus return results on how to park on a hill that has a curb.

[5] https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html
[6] https://blog.google/products/search/search-language-understanding-bert/

Chapter 3

Related Work

Personalization of search results and by extension the filter bubble has been the subject of other research in the past. While this research is closely related to this one they also differ in some key aspects. These differences are usually most noticeable in which type of personalization they're interested in. Some are focused on personalization based on location [7, 15], personalization based on search history [9, 12], personalization in general [2, 4, 5], or personalization based on other characteristics [1, 6, 8]. For some of these papers I'll go more in depth and explain both the similarities and the differences between their research and ours, especially since at first glance they might look almost identical since they're all just focused on personalization of search results.

3.1 What did you see? Personalization, regionalization and the question of the filter bubble in Google's search engine

One of these related studies was done by the Technical University of Kaiserslautern in 2016. In their paper [8] Tobias Krafft, Michael Gamer, and Katharina Zweig analyze Google search results to investigate the degree of personalization of search results, the proportion of regionalization, and the risk of algorithm based filter bubble formation or reinforcement by the leader in the search engine market.

While there are similarities, both analyzed Google search results for instance, there are some key differences. For their study Krafft, Gamer, and Zweig developed a browser plugin that would automatically search a predefined set of German politicians and political parties every 4 hours if the browser was open. The study had a couple of main focus points: investigate the degree of personalization of search results, how much regionalization affected those search results, and the risks of the formation of reinforcement of an algorithm-based filter bubble.

One big difference between this study and ours is, that for this study the researchers wanted to analyze the changes of search results over time - which is the reason why the plugin searched Google every 4 hours - while we are only interested in one search result per person.

Some other noticeable differences between this research and ours are that for this study, we only looked at the "organic" search results from Google, this means no ads or featured stories but only the results returned by the search query, while Krafft, Gamer, and Zweig also included top stories in their search results, and looked at the results from Google News. They only focused on results from Germany while in this study the focus is on results from the Netherlands. The last difference, which is an extension from the previous, is that they focused on websites in German while our research is focused on websites in Dutch.

In the end, Krafft, Gamer, and Zweig came to the conclusion that at least for search results from the same country, and using the same language, an algorithm-based filter bubble isn't present due to the sufficient overlap in search results. They also noted that users that received non-German websites in their search results, which is probably due to search language settings, had some overlap with each other, which means that there are signs that a filter bubble could exist. They mentioned that this is unavoidable however, due to the difference in languages.

Finally, they mention a chicken-and-egg problem with investigating the existence of a filter bubble on Google due to their market share. Because Google is the most popular search engine in Germany, and a sizeable amount of visits to a website are achieved through Google, it's not known how much of a websites traffic is only due to Google, and thus the results of their study do not represent a popularity measure independent of Google's algorithm.

3.2 Location, Location, Location: The Impact of Geolocation on Web Search Personalization

This paper[7] was written by researchers from Brown University and Northeastern University in the US about how important location is for the personalization on Google Search results. To measure this, they divided both their queries and locations into 3 categories: local, politicians, and controversial for the queries, and county, state, and national for locations. The queries are pretty self-explanatory, but the locations are a bit more complicated. For each location category, they chose the centroid of a set of location divisions within that category to measure the personalization on that level. For the national level they chose the centroids of 22 random states, for the state level they chose the centroids of 22 counties within the state, and for the county level they chose the centroids of 15 voting districts in the county. By doing this they could check how much influence distance had on differing search results.

This research is quite similar to ours with only a few major differences. The main ones being the slightly different query categories and the main focus of the study. While they are interested in personalization of Google search results, they're mostly interested in the effect of location on it, not personalization in general. Furthermore, this research was done in the US while we're doing ours in the Netherlands. This final difference is especially important due to the different political systems, the US has a two-party system while the Netherlands has a multi-party system. This difference in political systems combined with a much more polarized political landscape in the US compared to the Netherlands, makes it difficult to draw useful conclusions from this research for the Netherlands.

The main conclusion of their paper is that just location isn't very influential on the personalization of general search queries like politicians and controversial topics but that it does have a great effect on queries that search for locations, with a greater effect when comparing search results from farther away.

3.3 Measuring Political Personalization of Google News Search

In this paper[9] researchers from three American universities studied the effects of search history on the personalization of search results on Google News. They did this by creating 3 different profiles: An anti-immigration profile, a pro-immigration profile, and a control profile. They did this by using fresh installs of the Firefox browser, and using the anti-immigration profile to access links tweeted by an anti-immigration Twitter account, the pro-immigration profile to access links tweeted by a pro-immigration Twitter account, and finally the control profile didn't access any links at all.

After training all profiles, they searched for 10 policy issues and used 5 different search terms for each policy, for example for the policy issue Veterans they used the search terms 'Support our Veterans', 'Veteran affairs', 'Veterans', 'Veteran benefits', and 'PTSD'. Since they had control over all the profiles that they built, they configured Google to return 100 search results for each query. After collecting all the search results, they compared the three profiles in two ways: the first was by calculating the intersections of the sets of search results for every query, and using those to see the differences between the lists, and the second was by computing the Damerau-Levenshtein distance to see how much deletions, insertions, and substitutions are needed to transform the search results of one profile into the search results of another profile.

This research is almost identical to our research in regards to the two comparison methods that they use: the difference measure they use is the opposite of the similarity measure that we use, and their edit distance is calculated in the exact same way.

There are still quite a decent amount of differences between the two studies. While our study uses the search results of real people these researchers trained three different "profiles" to get different search results by letting them read news articles which hyperlinks were posted by two Twitter accounts: one with a pro-immigration and another with an anti-immigration stance. After they trained their three profiles, they used them to search for the topics they selected on Google News, while we had people use regular Google search. Furthermore, because the researchers used Google News they only received results from newspapers, which meant that every website they found had a certain political bias, which they estimated with the use of `mediabiasfactcheck.com`, which provided the researchers with the political bias of 1540 different domains on a 5-point scale, which the researchers gave different values to, and consisted of left with -100, left-center with -50, center with 0, right-center with 50 and right with 100. This bias was then used to calculate the political bias of the search terms where the difference between the pro and anti immigration profile was higher than 5%, which resulted in a total of 9 search terms. To calculate this, they simply added up the different scores of all domains found in both the search results, and in the list of `mediabiasfactcheck.com`. From this, they found that search terms that received the most personalization tend to get results that reinforce the opinion of the person searching for that term.

In the end the researchers concluded that there is significant personalization based purely on browsing history, and that these personalized results tend to reinforce the opinions that Google saw in the browsing history.

Chapter 4

Research

As mentioned in the introduction, this research set out to gather the search results of nearly 200 people split into two equal groups to investigate the personalization of Google search results. This turned out to be 181 people with 92 in the controversial group, and 89 in the non-controversial group. This chapter will delve into the details of this research.

4.1 Questionnaire

The foundation of this research is based on two questionnaires consisting of 3 general questions, and either 5 controversial search queries, or 5 non-controversial search queries. Both of these questionnaires can be found in the Appendix A.

4.1.1 Why a questionnaire

There were two primary reasons why we chose a questionnaire was to collect the necessary data: it was easier to both process and collect the data with a questionnaire than with other methods like browser extensions. By using a Google Form all the data gets automatically imported to an Excel sheet which are easy to process with a Python script. Regarding the collection of data compared to a browser extension a questionnaire is more user friendly, since it doesn't require installing anything, and a questionnaire is easier to make than a browser extension.

4.1.2 Why these questions

The general questions used in this questionnaire asked about the age, general location, and which political party they voted for in the last general election. These questions were used to both give context for certain search results, and to see if different demographics had differing amounts of personalization.

Initially the questionnaires were combined into one big questionnaire but due to the size of the answers they couldn't be submitted so the questionnaire was split into two, with one having the controversial search queries, and the other the non-controversial search queries. We use both controversial and non-controversial queries to check if the degree of personalization is different between controversial and non-controversial topics. These queries are listed in table 4.1 below.

Controversial Query	**Non-Controversial Query**
Abortus tot hoeveel weken?	Brood bakken recept
Oorzaken klimaatverandering	Honden namen
Zwarte piet of roetveegpiet?	Wat is het grootste bot in het menselijk lichaam?
Gevaren vaccinaties	Hoeveel van een komkommer is water?
Gevolgen illegale immigratie	Hoeveel mensen wonen er in Nederland?

Table 4.1: The Controversial and Non-Controversial queries

The controversial queries were chosen to be about topics where there are two opposing opinions and be general enough that almost everybody in the Netherlands would have an opinion on it. With these restrictions the topics of abortion, climate change, Black Pete, vaccinations, and illegal immigration were chosen.

These topics had to be phrased in such a way to be as neutral as possible as to not give preference to either advocates, or opponents of the subject in question. This is because suggestive search queries would influence the search results we would get, e.g. phrasing the query about Black Pete as: "Why is Black Pete racist?" already implies that Black Pete is racist and would thus lead to more results against Black Pete than the used phrasing of "Black Pete or Soot Pete?". By keeping the phrasing as neutral as possible the probability that differing search results are caused by personalization by Google are increased.

The non-controversial queries were chosen to be about at least one of three things: topics that are either factual, topics where there aren't vastly different opinions between two sides, or not location based. With these restrictions the topics of bread recipes, dog names, largest bone in the human body, percentage of water in a cucumber, and the population of the Netherlands were chosen.

For these queries the phrasing isn't as important as the controversial topics, since these queries were either about facts, i.e. "What is the largest bone in the human body?" or about topics that don't have vastly different opinions, i.e. "Dog names". Thus, the probability that someone would get vastly different results just because of the phrasing of the question was slim to none.

Our hypothesis is twofold: First we hypothesize that search results are personalized, and secondly that personalization has more influence on controversial queries than non-controversial queries. This higher influence for controversial queries would result in more unique results for these controversial queries. This is our hypothesis because topics that have two opposing opinions would be easier to personalize than factual topics.

4.1.3 Distributing questionnaire

The questionnaires were distributed through Facebook and Surveyswap.io[1], a website for exchanging surveys. These were chosen since Facebook and Surveyswap had the highest amount of people willing to fill in the questionnaires. Because there are two questionnaires a webtool[2] was used to randomly redirect to one of the two to ensure that both, constantly had roughly an equal amount of responses.

4.2 Data processing

Because the responses to the surveys were copy and pasted Google search results, it needed to be processed to retrieve the websites from the answer fields.

4.2.1 Python script

To extract the addresses of the websites a Python script was made to filter them from the text and put them in a list. Because there wasn't a general form for the addresses, the filtering was a bit more complicated than first anticipated. Not all websites found in the search results used "http" or "https" in their prefix, which meant we couldn't filter based on that. Since there were also multiple top-level domains, just filtering everything that didn't end in ".com" or ".nl" wasn't an option either. This meant that the best possible way of finding all the websites was to split every response into a list with each word being a single entry in that list. Then finding for "words" that contained text followed by a '.' followed by text again. This meant that every website was filtered out, but this also meant that some websites were counted twice, and some entries weren't websites due to typos in the description on the search results page. Unfortunately, these had to be filtered out manually since there wasn't any way to differentiate between typo's in the description, websites having their address in their

[1]https://surveyswap.io/
[2]https://allocate.monster/

description, and websites without “http” or “https”. Luckily these were exceptions and not the norm, so this wasn’t as time consuming as manually filtering everything.

4.3 Website standings

After getting the results we tried to give all websites either a left, right, or neutral rating. These ratings were based on the entire website, and not the specific page that was linked in the results, since the specific web pages weren’t included in the copied text, and using ratings for specific web pages could lead to conflicting ratings for the same website. After giving each website a rating, they were compared to the political party a user voted for. If the ratings of the websites and the political party’s stances were very similar for a lot of people, it could suggest the existence of a filter bubble, while a lack of similarity could suggest the opposite.

4.4 Comparing search results

Finally, to measure the amount of personalization of a user’s search results, all users were compared to each other with the Levenshtein distance and a similarity function.

4.4.1 Levenshtein distance functioning

The Levenshtein distance works as follows: when given two lists it compares the two by seeing how much single-character edits are needed to transform one of the lists into the other. Single-character edits are deletions, insertions and substitutions. Because all lists have the same length only substitutions are used for the transformations. If two lists are similar, it will have a low Levenshtein distance since it doesn’t need a lot of single-character edits to transform one into the other, and if two lists are completely different it needs a lot of single-character edits which gives a high Levenshtein distance.

For example, if we have two lists $A = [1, 2, 3, 4]$ and $B = [1, 3, 2, 4]$ the Levenshtein distance between A and B is 2 out of a maximum of 4 since we must substitute the value at $B[1]$ with 2 and the value at $B[2]$ with 3.

4.4.2 Why Levenshtein distance

The Levenshtein distance was chosen for comparing the lists of search results because it is an easy way to see how different two lists are. This is done by just using the Levenshtein function from the TextDistance library. Because only substitutions are used in this instance, the resulting Levenshtein distance is the amount of websites that are in the exact same position in both lists.

4.4.3 Similarity

The similarity function used is a very simple one, it checks how many elements appear in both of the lists no matter the position of those elements.

For example, if we have two lists $A = [1, 2, 3, 4]$ and $B = [3, 2, 1, 5]$ the similarity between A and B is 3 out of a maximum of 4 since 1, 2 & 3 are in both lists.

4.4.4 Why similarity

We use a combination of Levenshtein and similarity because just one of the two wouldn't accurately represent the similarity of 2 lists. For example, while the lists $[1, 2, 3, 4]$ and $[2, 1, 4, 3]$ are very similar the resulting Levenshtein distance would be 4 because even though the lists contain the same elements, they aren't in the same positions. If you'd only look at the Levenshtein distance one would think that they're completely different because Levenshtein doesn't take into account the other elements of the list when looking at a specific element. By adding the similarity function, we'd see that both lists contain the exact same elements and are thus quite similar.

Chapter 5

Results

As mentioned in the previous chapter, this research consisted of two surveys. One with controversial search queries, and one with non-controversial search queries.

5.1 Survey demographics

This section will give a brief overview of the demographics of the respondents.

5.1.1 Personal information

At the start of both surveys respondents were asked to give 3 pieces of personal information: their age range, the province that they live in, and which political party they voted for in the general election of 2017.

These 3 categories were chosen to see if any of these categories had a large influence on the degree of personalization. Age was chosen mainly as a control category because if search results are heavily personalized it wouldn't be done based on someone's age. Location was chosen because Google insinuates that most of their personalization is based on location, e.g. searching for football in the US and the UK give vastly different results. Finally political affiliation was chosen, since filtering and personalizing search results to only show what the user wants to see based on their political affiliation would be one of the most dangerous forms of personalization, because that way people will only see results agreeing with their opinion, and thus enter a filter bubble where other opinions aren't heard.

Age	**All Respondents (% of total)**	**Used Respondents (% of total)**
<18	5 (5.4%)	0 (0%)
18 - 24	52 (56.6%)	28 (56%)
25 - 34	20 (21.7%)	16 (32%)
35 - 44	6 (6.5%)	1 (2%)
45 - 54	5 (5.4%)	3 (6%)
>55	4 (4.3%)	2 (4%)

Table 5.1: Ages of respondents in the controversial survey

Province	**All Respondents (% of total)**	**Used Respondents (% of total)**
North Holland	19 (20.7%)	9 (18%)
South Holland	12 (13%)	5 (10%)
Zeeland	2 (2.2%)	1 (2%)
North Brabant	13 (14.1%)	5 (10%)
Limburg	4 (4.3 %)	3 (6%)
Friesland	1 (1.1%)	0 (0%)
Gelderland	25 (27.2%)	19 (38%)
Drenthe	0 (0%)	0 (0%)
Overijssel	4 (4.3%)	2 (4%)
Flevoland	0 (0%)	0 (0%)
Utrecht	6 (6.5%)	4 (8%)
Groningen	3 (3.3%)	2 (4%)
Abroad	3 (3.3%)	0 (0%)

Table 5.2: Location of respondents in the controversial survey

Political Party	All Respondents (% of total)	Used Respondents (% of total)
VVD	18 (19.6%)	9 (18%)
PVV	1 (1.1%)	1 (2%)
CDA	5 (5.4%)	2 (4%)
D66	23 (25%)	13 (26%)
GL	15 (16.3%)	10 (20%)
SP	2 (2.2%)	1 (2%)
PvdA	2 (2.2%)	1 (2%)
CU	2 (2.2%)	2 (4%)
PvdD	5 (5.4%)	4 (8%)
50PLUS	0 (0%)	0 (0%)
SGP	0 (0%)	0 (0%)
DENK	0 (0%)	0 (0%)
FvD	4 (4.3%)	2 (4%)
None	15 (16.3%)	5 (10%)

Table 5.3: Political affiliation of respondents in the controversial survey

5.1.2 Controversial survey

The controversial survey was filled in a total of 92 times, but due to incorrectly filled in surveys, either by using the same search results for all queries or not answering the questions at all, nearly half of the surveys weren't usable and, in the end, only 50 of the 92 surveys were usable.

The first thing we notice is that in each category a few options are over-represented with just 1 to 3 options accounting for more than half of the respondents in all categories. This is due to the fact that this survey was mainly spread through Facebook contacts and because most of these people are young, progressive students, this group is over-represented in this survey. For the specific categories this means the following: the age group of 18 - 24 accounts for more than half of the respondents. The provinces of Gelderland, North Holland, South Holland and North Brabant, all of which have popular universities, account for more than 75% of the total amount of responses. Finally, VVD, D66, and GroenLinks, all quite popular parties for students, account for more than 60% of the amount of responses.

Because of this, options with less respondents might suggest either large differences or small differences, while the opposite might be true due to the small sample size.

Age	**All Respondents (% of total)**	**Used Respondents (% of total)**
<18	2 (2.2%)	1 (1.6%)
18 - 24	50 (56.2%)	37 (59.7%)
25 - 34	26 (29.2%)	15 (24.2%)
35 - 44	1 (1.1%)	1 (1.6%)
45 - 54	3 (3.4%)	2 (3.2%)
>55	7 (7.9%)	6 (9.7%)

Table 5.4: Ages of respondents in the non-controversial survey

Province	**All Respondents (% of total)**	**Used Respondents (% of total)**
North Holland	14 (15.7%)	7 (11.3%)
South Holland	17 (19.1%)	13 (21%)
Zeeland	1 (1.1%)	0 (0%)
North Brabant	10 (11.2%)	4 (6.5%)
Limburg	2 (2.2%)	1 (1.6%)
Friesland	4 (4.5%)	4 (6.5%)
Gelderland	28 (31.5%)	22 (35.5%)
Drenthe	0 (0%)	0 (0%)
Overijssel	4 (4.5%)	4 (6.5%)
Flevoland	0 (0%)	0 (0%)
Utrecht	2 (2.2%)	2 (3.2%)
Groningen	5 (5.6%)	5 (8.1%)
Abroad	2 (2.2%)	0 (0%)

Table 5.5: Location of respondents in the non-controversial survey

Political Party	All Respondents (% of total)	Used Respondents (% of total)
VVD	14 (15.7%)	11 (17.7%)
PVV	0 (0%)	0 (0%)
CDA	2 (2.2%)	2 (3.2%)
D66	29 (32.6%)	20 (32.3%)
GL	19 (21.3%)	14 (22.6%)
SP	0 (0%)	0 (0%)
PvdA	4 (4.5%)	3 (4.8%)
CU	2 (2.2%)	1 (1.6%)
PvdD	2 (2.2%)	0 (0%)
50PLUS	0 (0%)	0 (0%)
SGP	0 (0%)	0 (0%)
DENK	1 (1.1%)	0 (0%)
FvD	6 (6.7%)	4 (6.5%)
None	10 (11.5%)	7 (11.3%)

Table 5.6: Political affiliation of respondents in the non-controversial survey

5.1.3 Non-controversial survey

The non-controversial survey was filled in a total of 89 times. Here there was also a problem with wrongly filled in surveys which meant that 27 surveys weren't usable meaning that in the end 62 of the 89 surveys were used.

Just as in the controversial survey, because of the way the surveys were spread the young, progressive student is also over-represented in this survey.

5.2 Websites

This section will give a brief overview of the websites found in the Google search results.

5.2.1 General

In total 131 distinct websites were found with a total of 5370 links. Of these 5370 links 2412 were from the controversial survey, and 2958 were from the non-controversial survey. These 5370 links were divided over the 560 different lists of search results pages, with each search page having between 8 and 12 results. Due to this, before doing the Levenshtein distance comparison, each list of websites got scaled down to just the first 8 results of the search results page to better compare the different lists with each other. This resulted in 113 distinct websites and 4480 links to these websites.

As mentioned in the previous chapter, a quantitative comparison was done by using the Levenshtein distance, and the similarity between lists. This resulted in an average Levenshtein distance of 4.01 for the websites found in the controversial lists, and 4.28 for the websites found in the non-controversial lists, and an average similarity of 6.14 for controversial lists, and 6.07 for non-controversial lists. Because these comparisons were only used on the trimmed lists containing 8 websites, the values for these comparisons range from 0 to 8 for both the Levenshtein distance, and the similarity. For similarity a higher value meant that the 2 compared lists were more similar, while a higher value for the Levenshtein distance meant that the 2 compared lists had less elements on the same position.

Levenshtein distance

As mentioned in section 4.4 a quantitative comparison was done by computing the Levenshtein distance for both the controversial and non-controversial lists. For the controversial lists the average Levenshtein distance turned out to be 4.01, and for the non-controversial lists the average was 4.28.

This means that on average, less than half of the search results were the same website in the same position, when compared to other search results for that same query.

Similarity

In addition to the Levenshtein distance a quantitative comparison for the similarity was done for both lists as well. For the controversial lists the average similarity was 6.14 and for the non-controversial lists it was 6.07.

From this we can conclude that, on average, less than 2 websites of the search results were unique, when compared to the search results of another person for that same query.

5.2.2 Controversial websites

In this section we will take a look at some interesting results found in the controversial survey. Before conducting our research, our hypothesis was that, especially for controversial topics, a noticeable difference would be seen between the search results of different individuals. For example, we thought that for the subject of abortion there would be a clear distinction between people that vote for more pro-life parties, and people that vote for more pro-choice parties. Which would result in a fairly high Levenshtein value and a fairly low similarity value. What we actually saw was quite the opposite, especially for the topic of abortion: here the average Levenshtein distance was just 2.7 and the average similarity was 7.1 which means that, out of 2 lists of 8 websites, more than 7 of them were in both lists, with more than 5 of those in the exact same position. This suggests that for the topic of abortion most people regardless of their personal opinion about abortion would get the same search results.

This is also confirmed by the raw data: out of the total 12 websites that were found in the search results of the query: "Abortus tot hoeveel weken?" 6 of them were found in every query. While in general websites that portrayed a more positive opinion on abortion were more frequent, there was no defining attribute that was the cause of either seeing or not seeing a website that was pro-life.

Another category that was surprising were the search results of the question: "Gevaren Vaccinaties", here 7 out of 17 websites were in nearly all search results, 4 of them being for vaccination and 3 of them against. Of all 499 links, which also includes the links that weren't in one of the first 8 positions, found from the vaccination query, 231 of those were to websites in favour of vaccination, while 172 of them were to websites either against or skeptical about vaccination. While having 34.3% more links to websites in favour of vaccination shows a clear advantage for pro-vaccination websites overall, since more than two thirds of users only look at the first 5 results of a search results page[11], and the first 5 results have a relatively balanced amount of websites that are either for or against vaccination.

This means that whether someone is for or against vaccinations, they always saw websites that agreed and disagreed with them. And just as with the abortion query there isn't a clear distinct attribute that indicates whether someone will have mostly pro-vaccination or anti-vaccination websites.

The last surprise for the controversial websites was that websites, with a clear leaning to either the left or the right, didn't just show up in the search results of users with the same political orientation. This is evident with the websites `joop.bnnvara.nl` which is a left leaning opinion website[13], and `jalta.nl` which is at the opposite end of the spectrum[14]. If a filter bubble is present, one would assume that mostly right leaning users would see `jalta.nl` in their search results, and mostly left leaning users would see `joop.bnnvara.nl`, but this isn't the case. Of the 39 users that had `jalta.nl` in their search results only 26% of users voted for a right leaning party, this was less than the amount of users that saw the website and voted for either a left leaning or a central party, 33% and 28% respectively. And of the 11 users that had `joop.bnnvara.nl` in their search results only 18%

was left leaning with right leaning users being 45% and central users also accounting for 18%. For both these two websites the other users either didn't remember or hadn't voted in the previous election.

5.2.3 Non-controversial websites

In this section we will take a look at some interesting results found in the non-controversial survey. As said in the previous section our hypothesis was that a noticeable difference would be seen between search results of different individuals for the controversial topics but not as much for the non-controversial topics.

The first part of our hypothesis was confirmed because roughly 25% of search results were personalized when compared to other search results. Furthermore the average similarity of the non-controversial lists was higher than the average similarity of the controversial lists which would suggest that the second part of our hypothesis was also true. This higher similarity meant that it was quite a surprise that the Levenshtein distance for non-controversial queries was higher than the Levenshtein distance for controversial queries. This meant that while the non-controversial queries are less diverse when it comes to the different websites that they show in the search results, they are more diverse when it comes to the order in which these websites appear. This slightly higher similarity could be explained by the fact that in the 310 search results pages of the non-controversial queries only 60 different websites were found, while the controversial queries contained 61 different websites with 250 search results pages.

For example the search query: "Hoeveel van een komkommer is water?" had only 10 different websites in the search results with 4 of those 10 being present in all results, with one of those being present twice in almost all results, 2 of them were present in all but 1 list each. This resulted in very high similarity of 7.09 but it still had a fairly high Levenshtein distance of 3.81 especially when compared to the results for the controversial query about abortion which had an almost identical similarity of 7.10 but had a Levenshtein distance of only 2.74.

5.3 Statistical significance

To check if the results that we found in our research are statistically significant we performed an independent two-tailed t-test to calculate whether the differences we measured between normal and controversial queries are actually statistically significant and not that they just happen to be different.

5.3.1 T-test

To find this out we used an independent two-tailed t-test on the Levenshtein distances for all search results. First the reason we do an independent, and not a dependent t-test is because the two groups of test subjects are unrelated to each other, since people only filled in 1 of the 2 questionnaires. Second, the reason why we do a two-tailed instead of a one-tailed test is because we aren't just interested in if normal search queries have more unique search results than controversial search queries, but also if normal search queries have less unique search results than controversial search queries because both of these outcomes will suggest that there is a difference in the search results between normal and controversial search queries.

We performed the t-test by comparing all search results of normal search queries with each other, and all search results of controversial search queries with each other, and calculating the Levenshtein distance between them which we then used to calculate both the mean M and the standard deviation SD. This led to 9455 comparisons for normal search queries and 6125 comparisons for controversial search queries, and the following results from the t-test:
For the 9455 normal search results comparisons ($M = 4.28$, $SD = 1.68$) compared to the 6125 controversial search results comparisons ($M = 4.01$, $SD = 1.69$) the normal search results displayed a significantly higher degree of uniqueness, $t(13048.84) = 9.71, p = .000$.

These results suggest that whether you search for a controversial or a non-controversial query does indeed have an effect on the uniqueness of your search results. With normal search queries requiring 4.28 substitutions to transform the results from one person into the results of someone else compared to controversial search queries which require only 4.01 substitutions.

5.3.2 Cause

While these results prove that there is a statistical difference between normal and controversial search queries with normal search queries resulting in more unique search results, they don't explain why this difference exist in the first place. If a filter bubble does indeed exist on `Google.nl` one would assume that controversial search queries would have a higher degree of uniqueness than normal search queries because someone who is left leaning would get wildly different search results than someone who is more right leaning on subjects where these two people have vastly differing opinions. So, you could conclude from this that there isn't a filter bubble present on `Google.nl` due to the fact that more normal search queries had more unique search results than controversial search queries. However, what might be happening is that Google isn't making relatively small filter bubbles for left and right leaning people but one big bubble that contains everybody.

5.3.3 Interference by Google

In November of 2019 the Wallstreet Journal published an article regarding the practices at Google related to search results[1]. In it the writers alleged that Google tries to manipulate search results in multiple ways. For example they compared the organic search results (the results that aren't placed on the page due to ads, newsreel or other snippets) of Google, Bing, and DuckDuckGo for the search term of abortion for 17 days between July and August of 2019, and found that in 93% of the total amount of search results on the first page of Google were from Planned Parenthood, a nonprofit abortion rights organization based in the US, while for Bing and DuckDuckGo these percentages were only 14% and 16% respectively. While a spokeswoman for Google said that it doesn't make use of any ranking implementations to promote Planned Parenthood, this does seem like such a big difference that it can't be a coincidence.

A way that Google influences search results without directly tinkering with the organic search results is by changing what users see on a search results page. Over the years Google has added more and more content besides the actual search results like ads, a Google maps to point to the location of relevant search results, or news reels to name a few. Because these additions to the web page aren't part of the "real" search results and most people click on the first link that appears on their screen a company that pays more to be featured in an ad will receive more traffic than a company that doesn't.

These and other tactics mentioned in the article by the Wallstreet Journal suggest that Google tries to both directly, with the use of content moderation and blacklists, and indirectly, with the previously mentioned tactics, influence which content users see not by deciding what certain groups of users get to see but what all users get to see and thus creating one big filter bubble for all users instead of multiple smaller ones for particular groups.

[1] How Google interferes with its search algorithms and changes your results

Chapter 6

Conclusions

Without access to the algorithms that Google uses to give users their search results it is very difficult, if not impossible, to give a definitive answer to the question: "Does Google create filter bubbles in their results?". And while Google will always deny that they're moderating their search results, which in turn creates filter bubbles, there is evidence as mentioned in the previous chapters that suggests the opposite.

We have shown in section 5.3 that the personalization we found in search results is indeed statistically significant. And while the more unique search results turned out to belong to the normal search queries, which wasn't what we had hypothesized it does show that as a user you don't have complete control over what will be served to you by Google when using their search engine.

However, we didn't find a clear reason as to why the non-controversial search results ended up being more diverse than the controversial ones. As we described in sections 5.3.2 and 5.3.3 this could be due to Google filtering more search results for controversial topics in general, but since that wasn't in the scope of this research we can't make any substantial claims about this.

All in all more research should be done in the future to determine how much Google actually personalizes someone's search results and if it's, like Google claims, mostly to filter out low quality web results, or, as others frame it, to subconsciously influence people by cherry picking what they're seeing.

Chapter 7

Future Work

In this chapter we will go over certain topics that fell outside the scope of this paper but are related to it. In section 7.1 we will discuss comparing search results between search engines to see if this problem is specific to Google or if it's more widespread. In section 7.2 we will discuss getting a more diverse and representative test population to improve both the quantitative, and the qualitative conclusions we can get from this.

7.1 Search results between search engines

Just as the journalists at the Wallstreet Journal[1] did as mentioned in section 5.3.3, checking the differences for certain search queries between different search engines, to see just how much impact using a different search engine could have, could be an interesting topic for a paper to further investigate the existence of filter bubbles on search engines.

7.2 More diverse test population

As mentioned in sections 5.1.2 and 5.1.3, each of the three categories had a disproportionate representation for certain groups in their specific category. This made the results obtained for the smaller groups less useful, especially for groups that had only 2 or 3 respondents.

For example, the controversial questionnaire was only filled in by 5 people in the age range 45 - 54, and only 3 of those were used. This group ended up with a similarity of 7.86 and a Levenshtein Distance of 1.06. Now, because of the small pool of test subjects in this age range, there is no definitive conclusion to make about why this result varies quite significantly from the average. It could be a coincidence, or it could also be the norm for how

[1]How Google interferes with its search algorithms and changes your results

similar the search results are for people between the age of 45 - 54, but it can also be due to the other 2 attributes. As it turns out, 2 out of the 3 data subjects in this age range live in Gelderland, and voted for D66 which means that it's unlikely that these very similar search results are solely due to age, but mostly because of the similarities in all 3 categories between the data subjects.

A similar study with a more diverse group of data subjects could prevent these kinds of outliers in the data by distributing the population more evenly over the different categories, and thus preventing certain categories being dominated by a very specific group.

Bibliography

[1] Judit Bar-Ilan, Kevin Keenoy, Eti Yaari, and Mark Levene. User rankings of search engine results. *Journal of the American Society for Information Science & Technology*, 58(9):1254 – 1266, 2007.

[2] Abhinandan S Das, Mayur Datar, Ashutosh Garg, and Shyam Rajaram. Google news personalization: scalable online collaborative filtering. In *Proceedings of the 16th international conference on World Wide Web*, pages 271–280, 2007.

[3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding, 2019.

[4] Zhicheng Dou, Ruihua Song, and Ji-Rong Wen. A large-scale evaluation and analysis of personalized search strategies. In *Proceedings of the 16th international conference on World Wide Web*, pages 581–590, 2007.

[5] Aniko Hannak, Piotr Sapiezynski, Arash Molavi Kakhki, Balachander Krishnamurthy, David Lazer, Alan Mislove, and Christo Wilson. Measuring personalization of web search. In *Proceedings of the 22nd International Conference on World Wide Web*, WWW '13, page 527–538, New York, NY, USA, 2013. Association for Computing Machinery.

[6] Anikó Hannák, Piotr Sapieżyński, Arash Molavi Khaki, David Lazer, Alan Mislove, and Christo Wilson. Measuring personalization of web search, 2017.

[7] Chloe Kliman-Silver, Aniko Hannak, David Lazer, Christo Wilson, and Alan Mislove. Location, location, location: The impact of geolocation on web search personalization. In *Proceedings of the 2015 Internet Measurement Conference*, IMC '15, page 121–127, New York, NY, USA, 2015. Association for Computing Machinery.

[8] Tobias D. Krafft, Michael Gamer, and Katharina A. Zweig. What did you see? personalization, regionalization and the question of the filter bubble in google's search engine, 2018.

[9] Huyen Le, Raven Maragh, Brian Ekdale, Andrew High, Timothy Havens, and Zubair Shafiq. Measuring political personalization of google news search. In *The World Wide Web Conference*, pages 2957–2963, 2019.

[10] Eli Pariser. *The filter bubble: What the Internet is hiding from you.* Penguin UK, 2011.

[11] Advanced Web Ranking. Advanced web ranking, 2019.

[12] Bin Tan, Xuehua Shen, and ChengXiang Zhai. Mining long-term search history to improve search accuracy. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 718–723, 2006.

[13] Eric van den Berg. Joop: Jouw online opinie pagina, Oct 2009.

[14] Frank van Kolfschooten. Een ouderwets opiniemagazine voor het smartphonetijdperk, Mar 2015.

[15] Xinyu Xing, Wei Meng, Dan Doozan, Nick Feamster, Wenke Lee, and Alex C. Snoeren. Exposing inconsistent web search results with bobble. In Michalis Faloutsos and Aleksandar Kuzmanovic, editors, *Passive and Active Measurement*, pages 131–140, Cham, 2014. Springer International Publishing.

Appendix A

Appendix

A.1 Controversial Questionnaire

All pages of the controversial questionnaire that was used:

Vragenlijst filterbubbel Nederland

Met deze vragenlijst ga ik voor m'n bachelorscriptie onderzoeken of er in Nederland op Google een filterbubbel bestaat. Een filterbubbel is het resultaat van een zoekmachine die op iemand persoonlijk is afgesteld. Dit is over het algemeen redelijk gewenst maar kan voor problemen zorgen als er door een filterbubbel mensen in een echokamer terecht komen die alleen maar informatie toont die ze willen zien. Met behulp van deze vragenlijst ga ik kijken of er een filterbubbel in Nederland bestaat en zo ja in welke mate. Alvast heel erg bedankt voor het invullen!

Hoe moet u deze vragenlijst invullen

Deze vragenlijst bestaat uit 2 delen: deel 1 is algemene vragen, deel 2 bestaat uit een aantal Google zoekopdrachten. De werkwijze van het tweede deel is daar beschreven.

Next

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms

Vragenlijst filterbubbel Nederland

*Required

Algemene vragen

Wat algemene vragen om rekening te houden met factoren als datum, locatie etc.

Wat is uw leeftijd? *

- ○ <18
- ○ 18 - 24
- ○ 25 - 34
- ○ 35 - 44
- ○ 45 - 54
- ○ >55

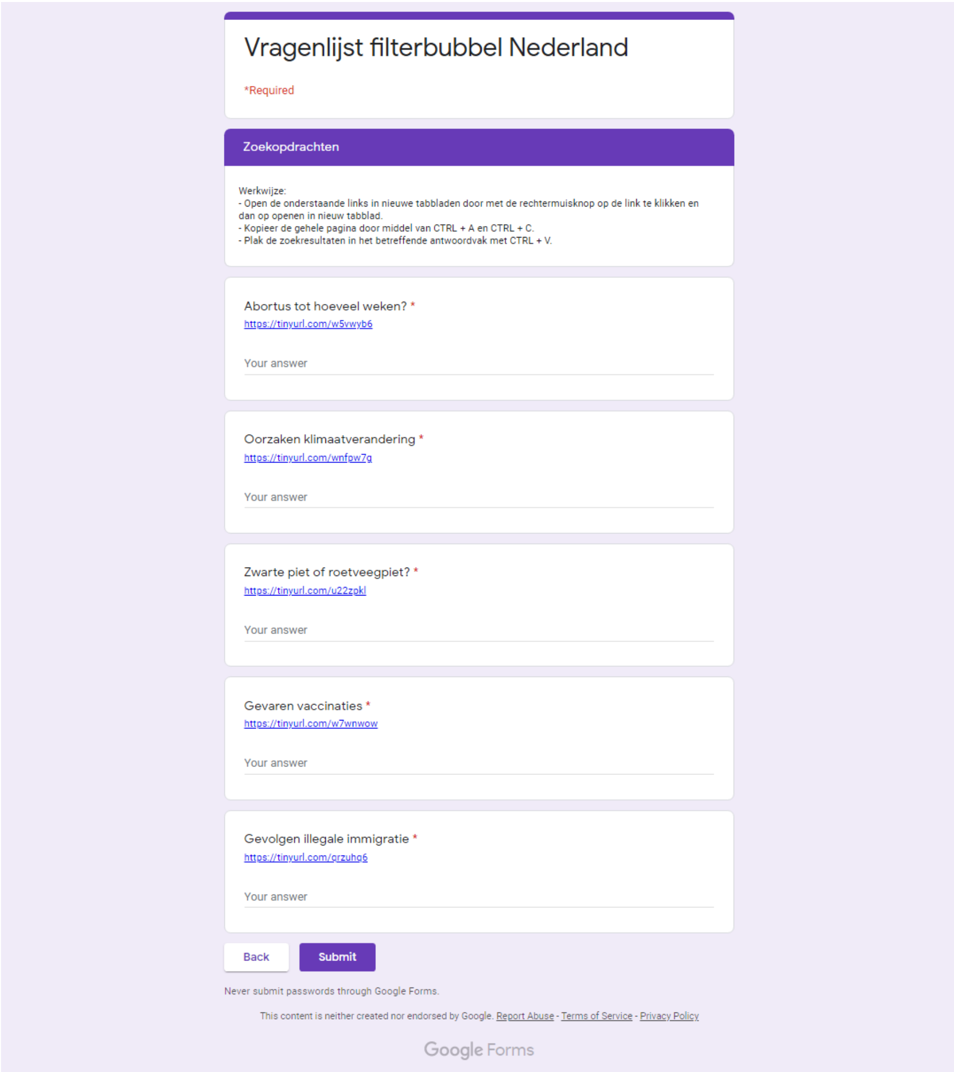
In welke provincie vult u deze vragenlijst in? *

- ○ Noord-Holland
- ○ Zuid-Holland
- ○ Zeeland
- ○ Noord-Brabant
- ○ Limburg
- ○ Friesland
- ○ Gelderland
- ○ Drenthe
- ○ Overijssel
- ○ Flevoland
- ○ Utrecht
- ○ Groningen
- ○ Buiten Nederland

Op welke politieke partij heeft u bij de laatste tweede-kamer verkiezingen gestemd? *

- ○ VVD
- ○ PVV
- ○ CDA
- ○ D66
- ○ GroenLinks
- ○ SP
- ○ PvdA
- ○ ChristenUnie
- ○ PvdD
- ○ 50PLUS
- ○ SGP
- ○ DENK
- ○ FvD
- ○ Other: ____________

Back Next



Vragenlijst filterbubbel Nederland

*Required

Zoekopdrachten

Werkwijze:
- Open de onderstaande links in nieuwe tabbladen door met de rechtermuisknop op de link te klikken en dan op openen in nieuw tabblad.
- Kopieer de gehele pagina door middel van CTRL + A en CTRL + C.
- Plak de zoekresultaten in het betreffende antwoordvak met CTRL + V.

Abortus tot hoeveel weken? *
https://tinyurl.com/w5vwyb6

Your answer

Oorzaken klimaatverandering *
https://tinyurl.com/wnfpw7g

Your answer

Zwarte piet of roetveegpiet? *
https://tinyurl.com/u22zpkl

Your answer

Gevaren vaccinaties *
https://tinyurl.com/w7wnwow

Your answer

Gevolgen illegale immigratie *
https://tinyurl.com/qrzuhq6

Your answer

Back Submit

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms

A.2 Non-controversial Questionnaire

All pages of the non-controversial questionnaire that was used:

Vragenlijst filterbubbel Nederland

Met deze vragenlijst ga ik voor m'n bachelorscriptie onderzoeken of er in Nederland op Google een filterbubbel bestaat. Een filterbubbel is het resultaat van een zoekmachine die op iemand persoonlijk is afgesteld. Dit is over het algemeen redelijk gewenst maar kan voor problemen zorgen als er door een filterbubbel mensen in een echokamer terecht komen die alleen maar informatie toont die ze willen zien. Met behulp van deze vragenlijst ga ik kijken of er een filterbubbel in Nederland bestaat en zo ja in welke mate. Alvast heel erg bedankt voor het invullen!

Hoe moet u deze vragenlijst invullen

Deze vragenlijst bestaat uit 2 delen: deel 1 is algemene vragen, deel 2 bestaat uit een aantal Google zoekopdrachten. De werkwijze van het tweede deel is daar beschreven.

Next

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms

Vragenlijst filterbubbel Nederland

*Required

Algemene vragen

Wat algemene vragen om rekening te houden met factoren als datum, locatie etc.

Wat is uw leeftijd? *

- ◯ <18
- ◯ 18 - 24
- ◯ 25 - 34
- ◯ 35 - 44
- ◯ 45 - 54
- ◯ >55

In welke provincie vult u deze vragenlijst in? *

- ◯ Noord-Holland
- ◯ Zuid-Holland
- ◯ Zeeland
- ◯ Noord-Brabant
- ◯ Limburg
- ◯ Friesland
- ◯ Gelderland
- ◯ Drenthe
- ◯ Overijssel
- ◯ Flevoland
- ◯ Utrecht
- ◯ Groningen
- ◯ Buiten Nederland

Op welke politieke partij heeft u bij de laatste tweede-kamer verkiezingen gestemd? *

- ◯ VVD
- ◯ PVV
- ◯ CDA
- ◯ D66
- ◯ GroenLinks
- ◯ SP
- ◯ PvdA
- ◯ ChristenUnie
- ◯ PvdD
- ◯ 50PLUS
- ◯ SGP
- ◯ DENK
- ◯ FvD
- ◯ Other: ______

Back Next

Vragenlijst filterbubbel Nederland

*Required

Zoekopdrachten

Werkwijze:
- Open de onderstaande links in nieuwe tabbladen door met de rechtermuisknop op de link te klikken en dan op openen in nieuw tabblad.
- Kopieer de gehele pagina door middel van CTRL + A en CTRL + C.
- Plak de zoekresultaten in het betreffende antwoordvak met CTRL + V.

Brood bakken recept *
https://tinyurl.com/s4twjbr

Your answer

Honden namen *
https://tinyurl.com/quwz3jm

Your answer

Wat is het grootste bot in het menselijk lichaam? *
https://tinyurl.com/vdctr9t

Your answer

Hoeveel van een komkommer is water? *
https://tinyurl.com/rm924yl

Your answer

Hoeveel mensen wonen er in Nederland? *
https://tinyurl.com/vo8z7cz

Your answer

Back Submit

Never submit passwords through Google Forms.