Research paper

What does Google recommend when you want to compare insurance offerings? – A study investigating source distribution in Google's top search results

Dirk Lewandowski [1] and Sebastian Sünkler [2]

[1] Hamburg University of Applied Sciences, Germany; dirk.lewandowski@haw-hamburg.de
[2] Hamburg University of Applied Sciences, Germany; sebastian.suenkler@haw-hamburg.de

This is a preprint of

Dirk Lewandowski and Sebastian Sünkler (2019) "What does Google recommend when you want to compare insurance offerings? A study investigating source distribution in Google's top search results", Aslib Journal of Information Management, Vol. 71 No. 3, pp. 310-324

Abstract: Purpose - We describe a new method to improve the analysis of search engine results by considering the provider level as well as the domain level. This approach is tested by conducting a study using queries on the topic of insurance comparisons.

Design/methodology/approach - We conducted an empirical study that analyses the results of search queries aimed at comparing insurance companies. We used a self-developed software system that automatically queries commercial search engines and automatically extracts the content of the returned result pages for further data analysis. The data analysis was carried out using the KNIME Analytics Platform.

Findings - Google's top search results are served by only a few providers that frequently appear in these results. We show that some providers operate several domains on the same topic and that these domains appear for the same queries in the result lists.

Research limitations/implications - We demonstrate the feasibility of this approach and draw conclusions for further investigations from the empirical study. However, our study is a limited use case based on a limited number of search queries.

Originality/value – The proposed method allows large-scale analysis of the composition of the top results from commercial search engines. It allows using valid empirical data to determine what users actually see on the search engine result pages.

Keywords: Web searching, search engines, Google, Search engine results pages, search results concentration, Web scraping, data analysis, results ranking, insurance comparison.

Introduction

On a theoretical level, search engines can be seen as intermediaries between users and the information objects scattered on the web. The task of search engines is the collection (crawling), preparation (indexing), evaluation (ranking) and provision (user interface) of these information objects. In this sense, search engines are usually regarded as neutral intermediaries (Halavais, 2018; Lewandowski, 2017), which rank search results in a way that can be regarded as objective in the sense that each information object in the index is treated according to the same criteria and thus has at least potentially the same chance of being displayed for a query on a particular position. However, search engine optimization (SEO) methods, in particular, have a considerable influence on the ranking of commercial search engines, and SEO can have both positive and negative effects.

The importance of the question about the influence on search results of commercial search engines results from the importance of search engines themselves, and the importance of Google as the dominant search engine, which reaches a market share of more than 90 per cent in most European countries.[1] The importance of commercial search engines like Google results from several factors: Search engines are a vital service of the Internet without which users would not be able to find their way on the Web (Tavani, 2012; Varian, 2006). This is also reflected in the massive use of search engines: Google alone now serves more than 2 trillion search queries per year (Sullivan, 2016a). Besides, search engines and Google, in particular, enjoy a high level of trust among users. This was shown in studies on the credibility of search results (Purcell et al., 2012), on assuming the actual accuracy of search results (Purcell et al., 2012), on trust in the "correct" result ranking (Pan et al., 2007) and on the use of search engine rankings as a criterion for the quality of content (Westerwick, 2013).

While classical information retrieval research asks about the relevance of search results for users, little is known about the composition of search result sets by commercial search engines. Although a search engine bias (usually in the sense of unwanted top results for specific search queries) is often complained about, a large part of the empirical findings on that topic is based on rather small amounts of data or is even only anecdotal. This raises the question of whether these are only isolated cases or whether systematic biases actually exist.

We define search engine bias as "the tendency of a search engine to prefer certain results through the assumptions inherent in its algorithms" (Lewandowski, 2017). This means that every search engine is biased, and that

> "search engine bias does not mean that search results are deliberately manipulated by the search engine vendor but simply that results are ordered in a certain way that is determined by assumptions of what constitutes a good or relevant result in response to queries. It is even at the core of every idea of ranking, based on certain technically mediated assumptions, that certain items are preferred over others." (Lewandowski, 2017, pp. 66–7)

With our research, we address a particular type of queries, defined by that they (1) have high commercial potential and (2) are considered as informational on the side of the user. So, on the one hand, the user is looking for (unbiased) information on the topic. On the other hand, commercial sites are presenting interest-led information and investing in marketing and search engine optimization, respectively, to gain visibility. This may result in a situation where a user seemingly finds relevant information to her query while in fact, that information is highly biased towards the interests of website owners who use search engine optimization techniques. This is especially of interest when a user searches for things like medical treatment, loans, insurances, and the like. These are highly commercialized areas where companies heavily fight for the best positions in Google. This is because there is a high query volume for such queries and that there is much money to make from selling products or closing a contract.

In this paper, we present a method for collecting and analysing search engine data. Based on a list of search queries, search engines are automatically queried, and the displayed results are collected and stored. In the analysis the results are grouped by position and domain; furthermore, the service providers are determined from automatically extracted imprint data. With this method, the distribution of search results within a search engine on a specific topic can be determined automatically. This method goes beyond previous approaches in that not only simple domain comparisons but also real provider comparisons can be carried out. The method is particularly attractive for the analysis of topics that are discussed controversially or topics where there is a high

---

[1] http://gs.statcounter.com/search-engine-market-share/all/europe/

interest in influencing the search results using search engine optimization. We show the feasibility of the method by a first study on the subject of insurance comparisons.

The remainder of this paper is organised as follows: First, we review the literature on user behaviour on search engine results pages, and on search engine bias on a theoretical as well as on an empirical basis. Then, we present the three research questions guiding our study. After giving details on the methods used, we present the results and discuss them. In conclusion, we summarize and give recommendations for future research.

Literature review

User behaviour on the search engine results pages

The following properties can characterize the selection behaviour on the search engine results pages (SERPs):

1. Position effect: users prefer to look at the results listed first; an overwhelming number of clicks are made on these results. This effect has been demonstrated in numerous studies (e.g., (Bar-Ilan et al., 2009; Craswell et al., 2008; Joachims et al., 2005; Keane et al., 2008; Schultheiß et al., 2018; Yue et al., 2010).
2. Focus on the visible area: Users prefer to focus on the immediately visible area of the search results page, i.e. the area that is visible without scrolling (Höchstötter and Lewandowski, 2009). The results are preferably selected from this area (Kelly and Azzopardi, 2015).
3. Design of the search results snippets: Results that take up more space on the search results page are more likely to be perceived (Liu et al., 2015) and selected. Results that are graphically more attractive are perceived more strongly and clicked accordingly more often (attraction bias; Liu et al., 2015).

The selection behaviour on the search engine results pages can be characterized as strongly oriented to the given order and representation. Users behave in their selection with little information literacy (in the sense of making optimal decisions). The most important overarching explanatory models for user behaviour on the search results pages are the principle of least effort (Zipf, 1949) and satisficing (Simon, 1955). Pan et al. (2007) further determine a trust bias in Google.

The relatively limited number of results from which users preferentially choose raises the question of precisely what is displayed within the area from which they choose and how this affects the diversity of the selected results. A study from Yahoo Research (Goel et al., 2010), for which 2.6 billion search queries were evaluated, gives a first impression. It was shown that in web search, about 80 per cent of all clicked results are accounted for by only 10,000 websites (p. 203). This underlines not only the importance of the result position but also the somewhat limited choice of sources in these positions. However, this study did not report the sources and the frequency with which users selected them. Furthermore, the analysis is based on domains and not on providers. This does not consider that a provider can operate several domains that may appear more than once in the same result list.

Methods and influence of search engine optimization

Search engine optimization (SEO) is alongside the booking of keyword ads ("AdWords") the second way to gain visibility in search engines (Moran and Hunt, 2015, p. 10). While booking search ads involves direct costs (pay per click; PPC), the clicks on the organic results, whose position can be improved utilizing search engine optimization techniques, are not associated with direct costs for the content provider. However, this optimization often requires considerable effort, which increases with the number of competing offers on a topic or more precisely: for a search query. The extent to which search queries have become highly competitive can also be seen from the already mentioned substantial investments in SEO.

Optimization for products and services has become standard; even small companies have recognized the importance of visibility in search engines and are investing accordingly. However, the optimization of informative content, including news and public relations materials, is also becoming increasingly important. This raises new questions about search engines as information intermediaries.

The influence of search engine optimization can have a considerable positive influence on the findability and accessibility of content. First of all, information objects are created or improved so that they can be easily found and indexed by search engines. Measures include optimizing HTML code, optimizing navigational structures and adding Sitemaps. Another, but indirect, positive effect of search engine optimization is the improvement of the usability and accessibility of (collections of) information objects. It was demonstrated that search engine optimization methods have a positive effect on the usability (Thurow and Musica, 2009) as well as on the accessibility of websites (Moreno and Martinez, 2013).

However, it is unclear whether search engine optimization also has a positive effect on the relevance (in terms of content quality) of search results. On the one hand, search engine optimization can be viewed positively here if content is prepared in a way that allows the user to find potentially relevant information objects easily (Thurow, 2015). Since current ranking algorithms of search engines are aimed at optimizing user satisfaction (Diaz, 2016), it is fundamentally assumed here that optimization for search engines also helps the users.

On the other hand, search engine optimization can be regarded as the manipulation of irrelevant or less relevant information objects in a way that they pretend to the search engine ranking algorithms that they are particularly relevant results. In this sense, search engine optimization can be considered the intentional creation of a bias in the search results (Jürgens et al., 2014). This is particularly important in the case of informative content, as search engines are also frequently used by their users to check their own views: "In a way, Google becomes a tool to reflect on (one's own and others') biases, a tool for conscious confirmation bias". (Sundin et al., 2017)

This raises the question of what content users see, especially regarding controversially discussed topics.

Search engine bias

Tavani (2012) summarizes the literature on search engine bias, identifying three distinct areas of concern: "(1) search-engine technology is not neutral, but instead has embedded features in its design that favor some values over others; (2) major search engines systematically favor some sites (and some kind of sites) over others in the lists of results they return in response to user search queries; and (3) search algorithms do not use objective criteria in generating their lists of results for search queries." (Tavani, 2012)

(2) Addresses search engine vendors directly manipulating search results, mainly to favour their own offerings, and (3) addresses the question of whether all documents in the index are treated the same. We will focus on (1) where the central question is not only whether it is at all possible to design "bias-free" search engines (see Grimmelmann, 2010, and Lewandowski, 2017)) but what effect algorithmic biases in search engines produce. Results from such investigations will inform users and search engine vendors alike.

Apart from the reasons of why commercial search engines favour some websites over others, empirical evidence shows that search engines by far favour a relatively small subset of sites from the web (see the study from Goel et al.(2010), p. 203) mentioned above.

Research focusing on the contents current commercial search engines show found biases in regards to race (Noble, 2018) and gender (Noble, 2018; Otterbacher et al., 2017), search engines preferably showing confirmatory information to queries regarding conspiracy theories (Ballatore, 2015), search engines promoting hate speech (Bar-Ilan, 2006), and biased health information (White and Horvitz, 2009).

The question, though, remains whether these content biases are a result of search engine algorithms merely reproducing biases inherent in the processed content (i.e., in the underlying database) or whether the algorithms introduce new biases. To our knowledge, the only study addressing this question empirically is the "Cyberchondria" study by White & Horvitz (2009). The authors found that for queries related to medical symptoms, the results produced by the commercial search engine highly over-represent dramatic interpretations, as compared to the content distribution in a web crawl.

To our knowledge, the role of search engine optimization (SEO) techniques in producing biases in search results has not yet been considered. While it is difficult to isolate single factors that lead to biases on the search engine results pages (SERPs), we can assume a vast effect of SEO on search results, as SEO is now a multi-billion dollar industry (Sullivan, 2016b). When investigating search engine bias, the focus has been on internal (algorithmic) reasons from the search engines. However, search results are generated from an interplay between search engine providers, website providers, and users (Röhle, 2010). Therefore, we need to widen our view of search engine bias, taking the possible influences of all these actors into account.

Research questions

The search for comparison offers for insurances of all kinds is a popular use case in the area of web search. Many online platforms allow users to compare every possible type of insurance concerning the providers and the associated conditions and costs. This results in a highly competitive market and competition for customers, which is also reflected in the fact that a lot of search engine optimization is carried out to bring one's own offerings to the top positions in the search engine rankings. However, online platform providers often do not offer complete comparisons without self-interest. In practice, therefore, search results may not only contain "neutral" comparison sites like the German *Stiftung Warentest*, a not-for-profit foundation whose aim is "helping consumers by providing impartial and objective information based on the results of comparative investigations of goods and services" (Stiftung Warentest, n.d.). Rather, free comparison sites may take commissions from vendors, may not consider all vendors in their comparisons, or base their comparisons on outdated prices (see Stiftung Warentest, n.d.). With our research, we aim to map the distribution of different types of sites within the results sets on the chosen topic.

With our research, we aim to investigate whether there is a bias among the top positions in the use case of insurance comparisons. We examined how the domains and providers (i.e., the companies that possibly run more than one domain) are distributed among the top positions for 121 search queries.

In the context of this research, a domain is defined as a top-level domain (TLD) while a provider is defined as a company owning one or more domains.

For our study, we formulated three research questions that we examined using explorative data analysis:

RQ1: How many different domains can be found in the top positions?

RQ2: How many different providers can be found in the top positions?

RQ3: How often do providers with more than one result appear in the Top10 for the various search queries?

Methods

Our study was carried out in six steps (see Figure 1) which show a mix of automatic and manual processes to generate the data sets to address the research questions and for our explorative data analysis. We show that time consuming tasks like collecting search results to a high number of queries, the extraction of contact data of the providers and data transformation processes can be processed automatically with our developed software and the use of the free tool KNIME. However, not all processes can be conducted by such software tools. It is still necessary to do manual data cleansing and also the explorative data analysis needs to be carried out manually.

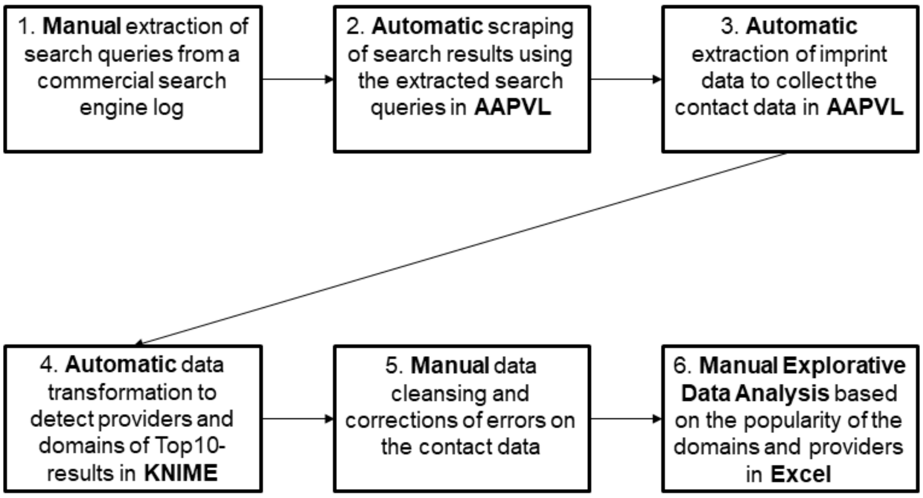


Figure 1: Data collection and analysis workflow

In the first step, thematically relevant search queries were extracted from a German commercial search engine log file consisting of more than 640,000 different search queries (see Lewandowski, 2015). From the log, we extracted a variety of query formulations for the same topic, i.e., queries containing the same word or phrase. The selection was based on pre-defined keywords in the context of insurance comparisons. The queries from the log file were automatically selected by combining the terms "*insurance*" and "*comparison*" (including left as well as right truncation). Examples of resulting queries are "Autoversicherung Vergleich" ("car insurance comparison"), "Berufsunfähigkeitsversicherung Vergleich" ("occupational disability insurance comparison"), "Haftpflichtversicherung im Vergleich" ("liability insurance in comparison"). This procedure identified a total of 121 different search queries. All queries were in German and are here translated for illustrative purposes only.

Next, we sent the queries to Google using the Relevance Assessment Tool's (RAT; Lewandowski and Sünkler, 2013) screen scraping component (step 2). Scraping took place between 8 and 9 May 2018, and a total of 22,138 webpages were saved. A copy of each website was also captured, and the contact details of the website provider were retrieved using the imprint crawler. This is necessary because it can be assumed that a provider may have a number of websites on the same subject. The result of the scraping is a CSV file that contains the Top10 search results for each search query (see the dataset available, Lewandowski and Sünkler, 2019).

In the third step, the search results were processed using software developed within the AAPVL project (Krewinkel et al., 2016), namely modules for crawling websites found through RAT and for extracting information from webpages. This allowed us to identify the imprint page from each website found and extracting company information including addresses. This information about the owners of the websites allows us not only to analyse the frequency by which certain websites (domains) appear in the search results but also to identify the purpose of the vendor (e.g., commercial vs not-for-profit), and to find networks of sites that belong to the same (commercial) entity.

In the fourth step the collected data (Figure 1) were processed in KNIME to transform and aggregate data for identifying the distributions of domains and providers within the Top10-results. We used KNIME mainly for grouping the data for further processing, e.g., aggregating webpages to domains and providers.

In the next step, the domains with the corresponding contact data were extracted from the file and corrections were made to the assignment of the contact data to the domains in a manual data cleansing process. This enabled us to determine how many providers operate more than one domain.

In the last step, an exploratory data analysis was carried out using methods from descriptive statistics based on the popularity of the domains, as well as the popularity of the providers. The data was limited to the top ten positions. The popularity of the domains and the providers were derived from the data set itself, depending on the frequency of the occurrence of the domains and the providers in the top 10 results for the search queries. No external factors like popularity of domains, e.g., measured by the Alexa Rank[2], or popularity of companies by their sales have been considered. The evaluation was based on the three or five most popular domains and providers, respectively.

Results

Description of the dataset

The dataset consists of 22,138 search results, allocated over 121 queries. It consisted of 3,278 different domains. On average, 182.96 results were collected per query, with the highest number of search results for a query being 298. While the scraper was set to capture the maximum number of search results for each query, search engines like Google actually return only a comparatively small amount of search results, contrary to the estimates given on the search engine results pages (SERPs) ("Results 1-10 of XXX"). A manual inspection of the data from the imprint crawler showed that it reaches an accuracy of 82.5 percent. This value is based on street, ZIP and city. In some cases, variants of the company name had to be conflated manually.

Domain level analysis

Distribution of the domains on the top 10 positions

Looking at the distribution of the domains, 116 of the 3,278 different domains can be found in the top 10 positions. When only considering the first rank, we can only find results from ten different domains. When we increase the number of ranks considered to three, we can find results from 37 different domains, and the first five positions have results from 57 different domains (see Figure 2).

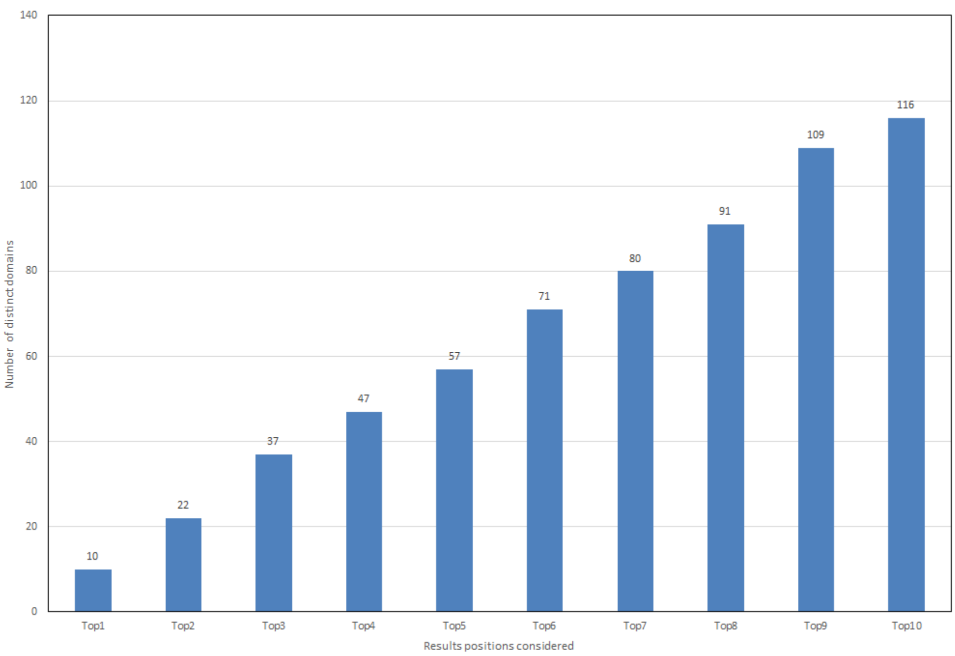


Figure 2: Distribution of the domains among the Top10 search results

[2] https://www.alexa.com/topsites/

Distribution of the five most popular domains in the data set

The popularity of the domains was determined by counting the occurrence of all domains in the Top 10 or Top 5 results, respectively. Table 1 shows the ranking of the top domains based on their share in the Top10 and Top5 search results.

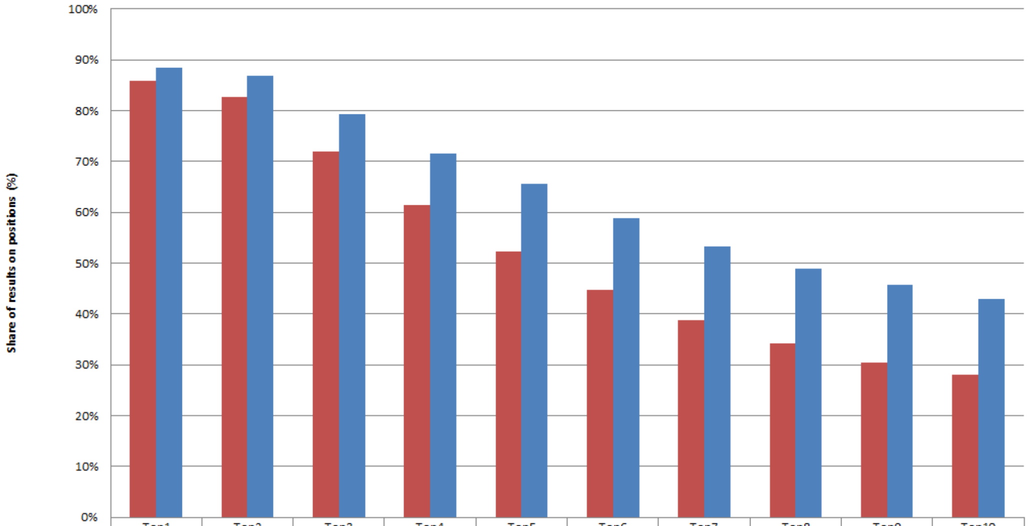
Table 1: The five most popular domains in the Top10 search results

Domain	Number of results in the top 5 (n = 605)	Relative share of results in the top 5	Number of results in the top 10 (n = 1,210)	Relative share of results in the top 10
www.verivox.de	118	19.5%	123	10.2%
www.check24.de	118	19.5%	120	9.9%
www.tarifcheck.de	80	13.2%	90	7.4%
www.financescout24.de	52	8.6%	97	8.%
www.preisvergleich.de	29	4.8%	89	7.4%

A comparison of the five most popular domains in the Top 5 in relation to all other domains shows that 65.6% of all results in the ranking up to position five are covered by search results from the five most popular domains. Considering the first ten position in the ranking shows that 42.9% are still covered by search results that can be assigned to the five most popular domains.

Distribution of the five most popular domains among the top 10 positions

Figure 3 shows the distribution of the three respectively five most popular domains on the positions from one to ten within the search results. This shows that 86% of the search results in the first position are already covered by search results that belong to one of the three most popular domains. The share of the five most popular domains is 88.4%. It becomes clear that with increasing position the share of the most popular domains in the number of search results is steadily decreasing, whereby up to the top 10 ranking 28.1% of the search results are still provided by URLs from the three most popular domains. For the five most popular domains this is as much as 42.9% (see Figure 3).



	Top1	Top2	Top3	Top4	Top5	Top6	Top7	Top8	Top9	Top10
Share of the 3 most popular domains	86.0%	82.6%	71.9%	61.4%	52.2%	44.6%	38.8%	34.2%	30.5%	28.1%
Share of the 5 most popular domains	88.4%	86.8%	79.3%	71.5%	65.6%	58.8%	53.4%	49.0%	45.6%	42.9%

Figure 3: Relative distribution of the domains on the positions 1 – 10

Provider level analysis

Distribution of the different providers to the top 10 results

Besides the evaluation of the most popular domains in the data set, the five most popular providers in the top positions were also considered. In contrast to the evaluation of the domains, the providers were determined from the recorded contact data from the imprints of the websites. This enabled us to determine which providers operate several domains. These analyses aim to determine whether the same providers with different domains appear several times in the top rankings of the search engine. Figure 4 shows the distribution of providers in the top 10 rankings. It turns out that across all search queries only ten different providers are displayed in the first position. Looking at the complete top 10, 93 different providers were identified.

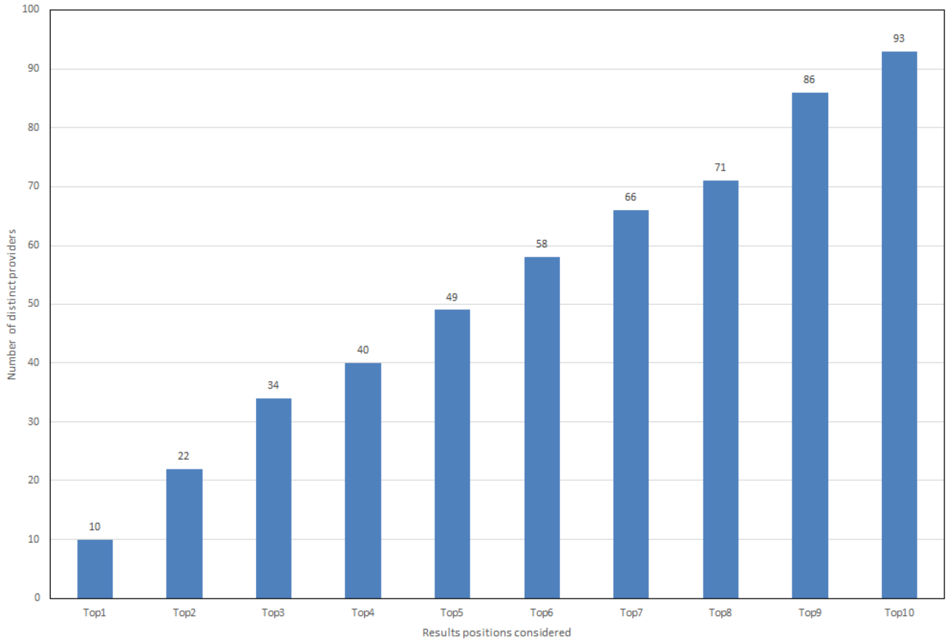


Figure 4: Distribution of different providers in the result positions

Distribution of the five most popular providers in the data set

The definition of the most popular providers in the search results was similar to the definition for the domains. It was determined how often a provider appeared in the Top10 over all searches. Table 2 shows the five most popular providers with their respective shares of the five and ten first positions in the ranking.

Table 2:The five most popular providers in the top 10 search results

Provider	Absolute share of results in the top 5 (n = 605)	Relative share of results in the top 5	Absolute share of results in the top 10 (n = 1,210)	Relative share of results in the top 10
Verivox GmbH	118	19.5%	127	10.1%
CHECK24 GmbH	118	19.5%	120	9.9%
TARIFCHECK24 GmbH	80	13.2%	90	7.4%

Scout24 Holding GmbH	52	8.6%	97	8%
finanzen.de Vermittlungsgesellschaft für Verbraucherverträge AG	41	6.8%	133	11%

The distribution of the five most popular providers in the overall search results up to position five shows that these providers cover 67.6% of these results (see Figure 5). Looking at the top 10, 46.5% of search results are still covered by these five providers.

Distribution of the five most popular providers within the top 10 positions

The next step is to measure the share of the three respectively the five most popular providers in the top positions of the SERPs. The analysis of the individual positions shows that there are no differences in the top position in the ranking. Eighty-six percent of the search results are provided by the three respectively five most popular providers. In terms of distribution, the share of the most popular providers decreases and ends with a share of 31.4% among the top 10 search results for the three most popular providers and 46.8% for the five most popular providers (see Fig. 4).

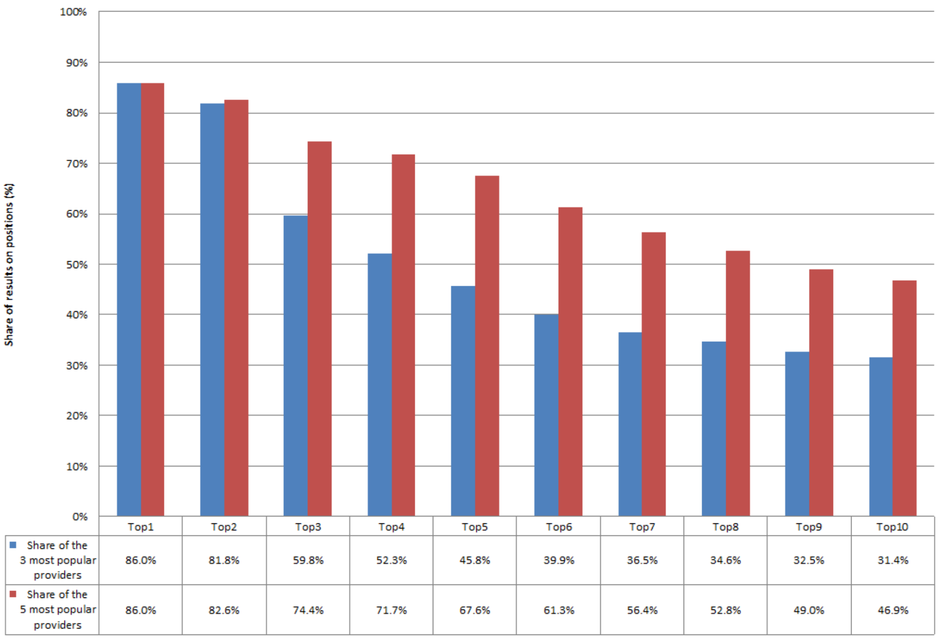


Figure 5: Relative distribution of the most popular providers on the result positions from 1 – 10

Multiple occurrences of a provider in the Top10 results

Besides the analysis of the occurrence of the most popular providers in the top results, the assignment of the domains to the individual pages was used to evaluate whether a provider with more than one domain appeared in the top rankings. The first step was to determine how many domains the provider had in the top ranks and with what frequency. As can be seen from Table 3 , some providers operate different domains (up to 8) and are able to place them in Google's top 10 results.

Table 3: The five most popular providers in the top 10 search results with their primary domains

Provider	Number of domains in the top 10	Number of results in the top 10

finanzen.de Vermittlungsgesellschaft für Verbraucherverträge AG	8	149
Müller & Kollegen UG (haftungsbeschränkt)	5	22
G. Zmuda	4	7
Axel Springer SE	2	53
Verivox GmbH	2	127

In a further evaluation, we investigated whether providers appeared several times with search results for their primary domains. This was done to determine whether providers were represented at least twice with search results from one of their primary domains within a list of results. It was found that Verivox Gmbh with the search results "verivox.de/versicherungen" and "verivox.de/kfz-versicherung-vergleich" was ranked 2 and 3 in the search query "preisvergleich versicherung". Figure 6 shows which providers were found several times in the top 10, and which share they accumulate.

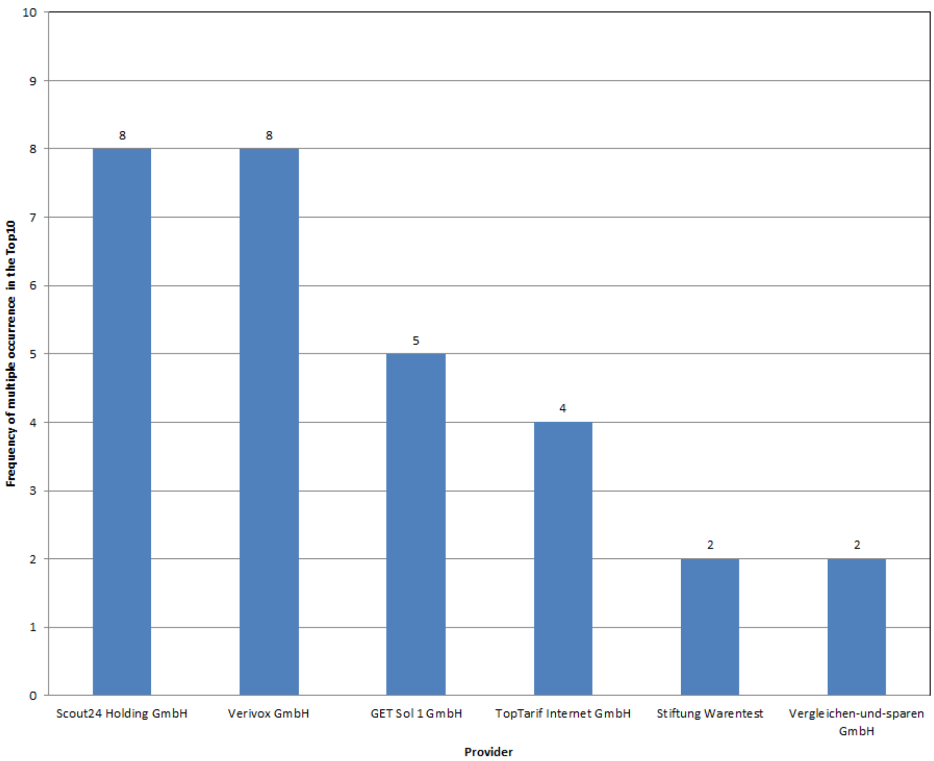


Figure 6: Frequency of multiple occurrences of a provider in the Top10

Discussion

This study shows that a relatively small number of providers often make it to the top positions in Google search. Looking at the overall distribution within the top 10, only 116 different domains (RQ1), which are offered by 93 different providers (RQ2), can be found in the 1,210 search results analysed. The total share of the five most popular sources within the top 10 search results is 43%. Looking at the provider level, this value increases to 46%. For the first five results, the figures are even higher: the five most popular sources have a share of 66% and 67%, respectively, when considering the providers. It turns out that some providers offer several domains (RQ3), which confirms our approach of considering the provider level in addition to the domain level.

Given the typical selection behaviour on SERPs, as we have described it in the literature review, we assume that users mainly click on results from these few sources when searching for insurance comparisons. Few domains/providers dominate the top positions. Other providers are listed, but they tend to be in the lower positions within the top 10 search results. This may indicate on the one hand the high popularity of only a few providers (which, however, may have been caused at least in part by the good positioning in the search engine), on the other hand, it is likely that this positioning was achieved by measures of search engine optimization.

Our study uses a methodology similar to that used in studies measuring domain distribution in search engines' top results (e.g., Höchstötter and Lewandowski, 2009) or the overlap between the top results of different search engines (e.g., Spink et al., 2006). However, we supplement this by reading out imprint information, which enables aggregation at the provider level.

As laid out in the introduction and literature review section, research on search engine bias mostly lacks frameworks for empirical studies, and valid empirical results, as well. Our approach addresses this by making search results concentration for a predefined set of queries measurable. A problem, though, is still a missing baseline for comparing the results of a search engine to a "gold standard", as favoured in computer science research. It is questionable whether such a baseline can at all be established, as it is unclear what degree of concentration should be wished for. Put differently, more diversity might not in all cases be desirable.

However, our approach can also be used to compare the results from different search engines. Then, one could measure the bias of a search engine's results in relation to the other individual search engines, or in relation to the average over all search engines considered.

Research on search engines in general, and search engine bias, in particular, would surely benefit from more empirical findings on what search engines display instead of theoretical assumptions often not grounded in empirical data. Our approach to data collection and analysis could benefit this branch of research in that it allows researchers to easily check the results of one or more search engines for a given set of queries.

Of course, this study is not without limitations. First of all, it should be noted that we were only able to analyse a relatively small sample of search queries. This is due to our data source. Although it contains several hundred thousand search queries, these are only search queries that were submitted at least with a certain frequency. There may be many more, but less popular searches on our topic. Furthermore, the search queries originate from an older data set, so that they may not reflect the current query behaviour.

We used the German version of Google search for our study. Accordingly, the results in other countries may be different. Also, we have taken advantage of the fact that there is an obligation to provide an imprint in Germany, which makes it easier to find out who is offering each website. However, this may be more problematic in other countries.

Personalization is an issue that has not been addressed in this study. Personalization means that different users may see different (top) results for the same search query. In our study, however, we submitted all search queries from the same computer without adapting the user profile. This may have led to the results being adapted to this computer or the results being influenced by context information (e.g., location of the computer). In future studies, this possible effect of personalization and contextualization should at least be reduced by collecting data on the same search queries several times. Ballatore (2015) has made a proposal that could be implemented quite easily by repeating the data collection at several points in time and then aggregating the results found.

In future studies, larger data sets should be created and analysed. This will be possible without problems, as any number of search results can be collected with the automatic data collection demonstrated. Larger data sets, on the one hand, would help investigating a particular topic in more depth, i.e., strengthen the validity of the results on that topic. On the other hand, larger data sets spanning several topics could be constructed, i.e., results would be strengthened through cross-validating findings over several topics.

Furthermore, query frequencies could be considered in future studies. This would not only enable us to determine which providers can be found in the top positions for which search query, but also to calculate the effect on the actual clicks of users with a certain probability. Such data could be collected using the Google AdWords API, for instance, which contains a prediction of the assumed search frequency for the queries.

The methodology described in this paper can be readily applied to a variety of topics. First and foremost, controversial issues such as nuclear power or abortion should be considered here. It is important to see whether (and if, to what degree) search engines present biased information when users wish to inform themselves on such topics and expect neutral or unbiased information. Research on search engine bias towards particular positions on particular topics or on the over-representation of certain aspects in the search results (e.g., Noble, 2018) often lack a valid empirical basis. Using the methods described in this paper, research on search engines' results could be put on a solid empirical basis.

Turning to the users' perspective, a key question is how users who expect a search engine to provide neutral information react when they are confronted with results that are above all commercially motivated.

Conclusion

We were able to show that the extension of search result analysis by adding the provider level makes sense and produces fruitful results. Already in our case study, there are differences in the results on domain level vs on provider level. We will investigate this in further studies on other topics and on larger data sets, as well.

We are confident that this approach will be fruitful especially for investigating highly competitive queries, where search engine optimizers fight for the best positioning in Google. Our approach could be used for a multitude of topics like health, news and loans to reveal potential biases in search results, as often lamented by researchers (e.g., (Diaz, 2008; Introna and Nissenbaum, 2000). We plan to investigate a broad set of topics to show whether search engines' tendencies to prefer certain results or providers over others is a general tendency found in different informational contexts. Especially worthwhile seems investigating health-related queries because a large number of websites are seem-to-be neutral information sources, while in reality, they are financed by large pharmaceuticals.

However, our study cannot answer the question of how precisely to define search engine bias and what an "ideal search result set" should look like. Nonetheless, descriptive studies such as ours can help to understand the presence and, if necessary, the extent of influences on the SERPs and to discuss possible measures of search engine bias on a solid foundation of data.

Acknowledgements

Dirk Lewandowski, being Editor-in-Chief of *Aslib Journal of Information Management*, was uninvolved in the review process, which was handled entirely by Joachim Griesbaum and Isabella Peters, the other guest editors of this special issue.

References

Ballatore, A. (2015), 'Google chemtrails: A methodology to analyze topic representation in search engine results', *First Monday*, Vol. 20 No. 7. Retrieved from, http://www.firstmonday.org/ojs/index.php/fm/article/view/5597/4652

Bar-Ilan, J. (2006), 'Web links and search engine ranking: The case of Google and the query "Jew"', *Journal of the American Society for Information & Techology*, Vol. 57 No. 12, pp. 1581–1589.

Bar-Ilan, J., Keenoy, K., Levene, M. and Yaari, E. (2009), 'Presentation bias is significant in determining user preference for search results-A user study', *Journal of the American Society for Information Science and Technology*, Wiley Online Library, Vol. 60 No. 1, pp. 135–149.

Craswell, N., Zoeter, O., Taylor, M. and Ramsey, B. (2008), 'An experimental comparison of click position-bias models', *Proceedings of the International Conference on Web Search and Web Data Mining*, pp. 87–94.

Diaz, A. (2008), 'Through the Google Goggles : Sociopolitical Bias in Search Engine Design', in Spink, A. and Zimmer, M. (Eds.), *Web Search: Multidisciplinary Perspectives*, Springer, Dordrecht, pp. 11–34.

Diaz, F. (2016), 'Worst Practices for Designing Production Information Access Systems', *ACM SIGIR Forum*, Vol. 50 No. 1, pp. 2–11.
Goel, S., Broder, A., Gabrilovich, E. and Pang, B. (2010), 'Anatomy of the long tail', *Proceedings of the Third ACM International Conference on Web Search and Data Mining - WSDM '10*, ACM Press, New York, New York, USA, p. 201.
Grimmelmann, J. (2010), 'Some Skepticism About Search Neutrality', edited by Szoka, B. and Marcus, A.*The next Digital Decade: Essays on the Future of the Internet*, TechFreedom, Washinton, Vol. 031, pp. 435–460.
Halavais, A. (2018), *Search Engine Society*, Second edition, Polity, Cambridge, UK.
Höchstötter, N. and Lewandowski, D. (2009), 'What users see – Structures in search engine results pages', *Information Sciences*, Vol. 179 No. 12, pp. 1796–1812.
Introna, L.D. and Nissenbaum, H. (2000), 'Shaping the Web: Why the Politics of Search Engines Matters', *The Information Society*, Vol. 16 No. 3, pp. 169–185.
Joachims, T., Granka, L., Pan, B., Hembrooke, H. and Gay, G. (2005), 'Accurately interpreting clickthrough data as implicit feedback', *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '05*, ACM Press, New York, New York, USA, p. 154.
Jürgens, P., Stark, B. and Magin, M. (2014), 'Gefangen in der Filter Bubble? Search Engine Bias und Personalisierungsprozesse bei Suchmaschinen', in Stark, B., Dörr, D. and Aufenanger, S. (Eds.), *Die Googleisierung Der Informationssuche*, De Gruyter, Berlin, Boston, pp. 98–135.
Keane, M.T., O'Brien, M. and Smyth, B. (2008), 'Are people biased in their use of search engines?', *Communications of the ACM*, Vol. 51 No. 2, pp. 49–52.
Kelly, D. and Azzopardi, L. (2015), 'How many results per page?', *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '15*, ACM Press, New York, New York, USA, pp. 183–192.
Krewinkel, A., Sünkler, S., Lewandowski, D., Finck, N., Tolg, B., Kroh, L.W., Schreiber, G.A., et al. (2016), 'Concept for automated computer-aided identification and evaluation of potentially non-compliant food products traded via electronic commerce', *Food Control*, Vol. 61, pp. 204–212.
Lewandowski, D. (2015), 'Evaluating the retrieval effectiveness of web search engines using a representative query sample', *Journal of the Association for Information Science and Technology*, Vol. 66 No. 9, pp. 1763–1775.
Lewandowski, D. (2017), 'Is Google Responsible for Providing Fair and Unbiased Results?', in Taddeo, M. and Floridi, L. (Eds.), *The Responsibilities of Online Service Providers*, Springer, Berlin Heidelberg, pp. 61–77.
Lewandowski, D. and Sünkler, S. (2013), 'Designing search engine retrieval effectiveness tests with RAT', *Information Services & Use*, Vol. 33 No. 1, pp. 53–59.
Lewandowski, D. and Sünkler, S. (2019), 'What does Google recommend when you want to compare insurance offerings? – A method and empirical study considering Google's top search results [Data set]'. Retrieved from, https://zenodo.org/record/2572691#.XKxfvC1XZ24
Liu, Z., Liu, Y., Zhou, K., Zhang, M. and Ma, S. (2015), 'Influence of Vertical Result in Web Search Examination', *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '15*, ACM Press, New York, New York, USA, pp. 193–202.
Moran, M. and Hunt, B. (2015), *Search Engine Marketing, Inc.: Driving Search Traffic to Your Company's Website*, Third edition, IBM Press, Upper Saddle River, NJ.
Moreno, L. and Martinez, P. (2013), 'Overlapping factors in search engine optimization and web accessibility', *Online Information Review*, Vol. 37 No. 4, pp. 564–580.
Noble, S.U. (2018), *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York University Press, New York, USA.
Otterbacher, J., Bates, J. and Clough, P. (2017), 'Competent Men and Warm Women', *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, ACM Press, New York, New York, USA, pp. 6620–6631.
Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G. and Granka, L. (2007), 'In Google We Trust: Users' Decisions on Rank, Position, and Relevance', *Journal of Computer-Mediated Communication*, Vol. 12 No. 3, pp. 801–823.
Purcell, K., Brenner, J. and Raine, L. (2012), *Search Engine Use 2012*, Washington, DC: Pew Research Center. Retrieved from, http://pewinternet.org/~/media/Files/Reports/2012/PIP_Search_Engine_Use_2012.pdf
Röhle, T. (2010), *Der Google-Komplex: Über Macht Im Zeitalter Des Internets*, Transcript, Bielefeld.

Schultheiß, S., Sünkler, S. and Lewandowski, D. (2018), 'We still trust in google, but less than 10 years ago: An eye-tracking study', *Information Research*, Vol. 23 No. 3. Retrieved from, http://www.informationr.net/ir/23-3/paper799.html

Simon, H.A. (1955), 'A Behavioral Model of Rational Choice', *The Quarterly Journal of Economics*, Vol. 69 No. 1, pp. 99–118.

Spink, A., Jansen, B.J., Blakely, C. and Koshman, S. (2006), 'A study of results overlap and uniqueness among major Web search engines', *Information Processing & Management*, Elsevier, Vol. 42 No. 5, pp. 1379–1391.

Stiftung Warentest. (n.d.). 'About us: An introduction to Stiftung Warentest', available at: https://www.test.de/unternehmen/about-us-5017053-0/.

Stiftung Warentest. (n.d.). 'Hausrat-versicherung: Die besten Tarife für Sie', available at: https://www.test.de/Vergleich-Hausratversicherung-4682054-0/.

Sullivan, D. (2016a), 'Google now handles at least 2 trillion searches per year', *Search Engine Land*, available at: http://searchengineland.com/google-now-handles-2-999-trillion-searches-per-year-250247.

Sullivan, L. (2016b), 'Report: Companies Will Spend $65 Billion On SEO In 2016', *Media Post*, available at: http://www.mediapost.com/publications/article/273956/report-companies-will-spend-65-billion-on-seo-in.html (accessed 14 May 2016).

Sundin, O., Haider, J., Andersson, C., Carlsson, H. and Kjellberg, S. (2017), 'The search-ification of everyday life and the mundane-ification of search', *Journal of Documentation*, Emerald Publishing Limited, Vol. 73 No. 2, pp. 224–243.

Tavani, H. (2012), 'Search Engines and Ethics', 27 August, available at: http://plato.stanford.edu/entries/ethics-search/ (accessed 12 August 2015).

Thurow, S. (2015), 'To Optimize Search, Optimize the Searcher.', *Online Searcher*, Vol. 39 No. 4, pp. 44–48.

Thurow, S. and Musica, N. (2009), *When Search Meets Web Usability*, New Riders, Berkeley.

Varian, H. (2006), 'The Economics of Internet Search', *Rivista Di Politica Economica*, No. November-December, pp. 177–191.

Westerwick, A. (2013), 'Effects of Sponsorship, Web Site Design, and Google Ranking on the Credibility of Online Information', *Journal of Computer-Mediated Communication*, Vol. 18 No. 2, pp. 80–97.

White, R.W. and Horvitz, E. (2009), 'Cyberchondria', *ACM Transactions on Information Systems*, Vol. 27 No. 4, p. Article No. 23.

Yue, Y., Patel, R. and Roehrig, H. (2010), 'Beyond position bias', *Proceedings of the 19th International Conference on World Wide Web - WWW '10*, ACM Press, New York, New York, USA, p. 1011.

Zipf, G.K. (1949), *Human Behaviour and the Principle of Least Effort*, *Human Behaviour and the Principle of Least Effort*, available at:https://doi.org/10.2307/2226729.