

# **To Exist Is to Be Found**

The Impact of a Search Result

# To Exist Is to Be Found

## *The Impact of a Search Result*

Anais Siebers

Cognitive Science, Ruhr-University, Bochum, Germany,

`anais.siebers@rub.de`

2th November 2022

### **Abstract**

Search engines have become an important aspect of everyday life. They are nearly invisible, ubiquitous, and shape people's actions in and knowledge of the world. But because of this invisibility, the daily impact a search result has on the individual and society when searching as well as when offering information is unaware. This paper investigates the impact a search result has by analysing different stages of the search: from entering a search query over retrieving results up to the presentation and selection of a result among others. The focus is on the existential power search engines exert.

**Keywords:** search engines; human-centered computing; social computing; categorisation; Google

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Motivation and Relevance . . . . .	4
<b>2</b>	<b>The Search Query</b>	<b>5</b>
2.1	Opening the Search Engine . . . . .	6
2.2	Knowing How to Write a Search Query . . . . .	8
<b>3</b>	<b>Retrieval of the Search Result</b>	<b>9</b>
3.1	Materialities of the Search . . . . .	11
3.2	Metadata and GoogleAds . . . . .	14
<b>4</b>	<b>The Power of Page 1</b>	<b>15</b>
4.1	Presentation of the Search Results . . . . .	16
4.2	Selecting a Search Result . . . . .	17
<b>5</b>	<b>Being Found</b>	<b>18</b>
<b>6</b>	<b>Discussion</b>	<b>21</b>
	<b>References</b>	<b>23</b>

# 1 Introduction

Everyday people around the world search for information on the internet. Many search for information multiple times a day. Therefore, the impact and influence search engines exert on everyday life are immense, but nevertheless they are nearly invisible to users in daily life.

The central topic of this paper is the material and discursive impact of a search result. To do this, the data infrastructure of the most known public search engine Google<sup>1</sup> will be analysed in regard to social and cultural consequences. The focus will be on the process of entering a search query until the appearance and selection of a specific search result. During the analysis, material as well as immaterial aspects of the search are exhibited, as well as the (in-)visibility of the infrastructure and their consequences. As basis for the analysis, aspects of data infrastructures defined by Bowker and Star (2008) and Dourish (2022) are used. Detailedly, the following topics and questions – to highlight the impact of search engines on the existence in virtual space as well as the existence in real life – are discussed:

**The Search Query** – How, and when, do we open search engines? How are they interwoven in our daily and cultural routines? And how does the formulation of the search query influence our knowledge and world-view?

**Retrieval of the Search Result** – Which algorithms and what kind of infrastructure determine the search result? What materialities can be found during a search process? What influence does Google have on standardisation and normalisation?

**The Power of Page 1** – How are the search results presented and what social and cultural consequences does this imply? What happens when a user selects a search result?

In the end, the impact of being and not being found with search engines as Google is discussed. Not only from the perspective of those who want to be found and offer information, but also considering those who do not yet exactly know what they want to find and only have a rough search query in mind. The epistemic goal of this paper is to outline the existential influence search engines exert on those who want to be

---

<sup>1</sup>Google is chosen as search engine, because it is the world-dominating search engine. Of course, there are other search engines: Especially in Russia and China, proper search engines were developed and are used. Nevertheless, Google is most known and used world-wide.

found and offer something, as well as the influence on public opinion and knowledge of societies, cultures and the individual.

## 1.1 Motivation and Relevance

The necessity and need to order and make information retrievable has been around for a long time. Over the last 300 years, various attempts to create a unified encyclopaedia and “store all knowledge in a single form” were taken (Bowker and Star, 2008, 107). Whereas knowledge was usually saved and distributed in the form of books, it has long transitioned to the internet. As the internet became more popular and the amount of available information exploded, it was increasingly difficult to find the information one searched for (Bowker and Star, 2008, 7). Therefore, systems enabling and facilitating the navigation, retrieval and management of the information were needed (Escandell-Poveda et al., 2022, 49). Thus, the evolution of search engines was only natural. In 1999, 800 million pages were estimated to be reachable in the public web and that this number would increase to 13 billion pages. Nevertheless, in 1999 search engines were only capable to index 16% of all pages (Escandell-Poveda et al., 2022, 54).

Search engines have been around the web for some time now. The first search engine was called Archie (derived from archive) and allowed to query the files from existing File-Transfer Protocol servers by their filename (Escandell-Poveda et al., 2022, 50). Nowadays, the content of the files is indexed – text as well as images and videos – and search engines are very diverse. Search engines have become one of “society’s key infrastructures for knowing and becoming informed” and are even classified as critical infrastructure in Europe (Haider and Sundin, 2019, i, 57). They are distributed across many social practices, having a significant impact on personal as well as cultural knowledge, memories, world view and the daily routines.

Although there are a lot of different search engines (from the search engine in a single document over the file system up to internet search engines), Google is the most prominent search engine. It has worldwide market shares above 90%, roughly 92,000 conducted searches every second worldwide and makes up 66% of the web traffic in the USA (Escandell-Poveda et al., 2022, 49). Google became successful during the commercialisation of the internet (the dot-com boom and bust) and convinced with a simple interface, speed and the sorting of the search result (Haider and Sundin, 2019, 12). However, it is difficult to compare Google’s power because there is no

basis for comparison (Patterson, 2012, 8). As for Facebook and Amazon, there is no transparency and accountability concerning the algorithms and infrastructures used in Google (Haider and Sundin, 2019, 12).

The ubiquity and integration of search engines in everyday life makes it necessary to study the impact of search engines on the individual as well as the society. Since Google is dominating the market, its algorithms, design choices and infrastructure ultimately determine what information one finds and gets to see and what epistemological knowledge one draws. There are two forms to study search engines: One is to study the medium itself – the algorithm, make-up and interface of a search engine – and the other method is to study the societal using the data provided by search engines, for example, in their results – studying trends, dominant voices, etc (Schäfer and Es, 2017, 75). Both methods yield difficulties since the algorithms are kept secret and the personalisation of search results makes it difficult to get “objective” results.

This paper does not really distinguish between the research methods. But the fact that the search results are used to study the societal reveals that search engines like Google have long transitioned from pure information retrieval based on keywords contained in documents to more complex and influential searches influencing the individual’s opinion, knowledge, world-view, having an impact on society.

## 2 The Search Query

Every search starts by opening a search engine and entering a *search query*. Users who want to find the same “thing” often use different queries and therefore get different search results. Unknown to many users, the search results will almost always differ if they enter the same search query. The same applies if one person enters the search query later in time. This complicates re-finding information one has already found and wants to find again. Furthermore, it shows that there are factors influencing the search apart from the search algorithm. In general, three types of search queries are differentiated: *common queries* which are committed on a daily basis like finding the address of a restaurant, *trending queries* which display a sudden spike in interest like searches concerning an election, or *medical queries* which concern medical problems (Vincent and Hecht, 2020, 2).

## 2.1 Opening the Search Engine

Opening a search engine is a choice: not only a choice to search something, but also where to find it. Previously, search was considered a professional task and experts were consulted to retrieve information (Sundin et al., 2017, 240). If a person wanted to call somebody, telephone operators connected the person to the right number, information was acquired in libraries where librarians helped and consulted. Nowadays, search is simplified and every user can search for information and does so every day (Haider and Sundin, 2019, 1). The mode of search has also changed: Whereas people used indices and table of contexts previously, today this level is abstracted away (Sundin et al., 2017, 238). Moreover, earlier in the days, people used specialised search engines for specific searches. Today, however, searching is made possible at every time and everywhere through tables, smartphones and other devices (Sundin et al., 2017, 224). Sundin et al. (2017) call this phenomenon the *mundification* of search.

*Invisibility* is an aspect of infrastructures often highlighted by Bowker and Star (2008). The decision to open Google is unconscious and unreflected most of the time, which shows that Google is nearly invisible in daily life (Sundin et al., 2017, 239). That “googling” has become a synonym for searching for information highlights this invisibility (Andersson, 2017, 1244). Students, for example, instantaneously connect Google to fact-finding and to school – they forget everyday searches as in tasks like looking for clothes online while commuting on the bus (Andersson, 2017, 1251, 1256). This reflects two types of searches that can be distinguished: On the one hand, the in-between, unreflected and unconscious search which is integrated in everyday life place-independent by the ubiquity of devices and also programs allowing to access search engines easily. On the other hand, there is information seeking during empirical work (Sundin et al., 2017, 227). Google has become the “door” to the internet – it is the way people navigate and orientate themselves on the internet by starting every interaction over Google (Andersson, 2017, 1244). As a consequence of the search personalisation by Google, users less and less encounter something that challenges their search behaviour and give them grounds to change their behaviour (Sundin et al., 2017, 239).

The *ubiquity* of search engines has largely contributed to the fact that they are almost invisible and mundane. Mobile devices allow searching for various things as “timetables, opening hours, news, recipes, job information, symptoms of illnesses or simply for distraction”, becoming an external memory resource for people (Sundin

et al., 2017, 238). Google dominates the search engine market and also is a dominant global player (Haider and Sundin, 2019, 1). Its dominance is reinforced by various services such as “Google Search, Google Mail, Google Maps, Google Earth, Google Analytics, [...] Google+, and its share in the smartphone operating system Android” (Mager, 2012, 4). This ubiquity of Google, makes it a fundamental infrastructure which shapes social practices (Haider and Sundin, 2019, 1). Because people trust “search engines as neutral ahistorical tools and as politically and ethically unproblematic” (Haider and Sundin, 2019, 143), results are less critically reflected. The ubiquity of Google as a dominating search engine led to a standardisation of search engines and other services (Bowker and Star, 2008, 37).

Apart from the fact that searching has become mundane, invisible and ubiquitous, there are other aspects which do not only concern the choice of opening and choosing a search engine. Two major aspects have been discussed in the literature about search engines and social media. They also influence the search results the user is presented with when he<sup>2</sup> opens the search engine (Bruns, 2019, 4):

**Echo Chambers** In social media, this is the effect of preferentially connecting with friends and thus excluding outsiders. Transferred to search engines, it is the effect of getting increasingly similar search results, after choosing specific search results.

**Filter Bubbles** The effect of preferentially communicating with friends and thus excluding outsiders in social media. Hence, in search engines this describes the personalisation of search results based on a “profile” which among others entails the search history, the previous choices of search results, the location and used search terms (as outlined in the paragraph about ubiquity, Google’s services allow for a myriad of data points for personalisation) (Kuhn and Hauck, 2012; Mager, 2012, 4).

Whereas most research agree upon the phenomena of echo chambers and filter bubbles, Fletcher and Nielsen (2018, 979, 986) find effects that the use of search engines leads to a confrontation with different results and also information from the opposite political spectrum, because it increases user experience, prevents user annoyance and is a design goal by engineers and normative ideal for scholars. They are called “serendipitous suggestions”. But it has to be remarked that the institute, which found these results, is funded by Google.

---

<sup>2</sup>In the interest of readability, I will go without gendering and only use the male form.



The *personalisation* of search results suggests that, as for the internet, the notion of a singular search experience is no longer useful, because everyone experiences the search very differently (Dourish, 2022, 171) and has his own “personalised universe of information” (Pariser, 2015, n.p. in Bruns, 2019, 2). Furthermore, by personalising the search results, the users are also classified according to different categories which make up their “digital person”.

To sum up, before even entering a search query, the user chooses to search and how to search, which is a mainly unconscious process, deeply interwoven in social practices. Furthermore, a lot of data is already collected and known about him, which influences the search. But this is invisible and often unknown to the user.

## 2.2 Knowing How to Write a Search Query

After opening a search engine, a *search query* has to be formulated and send off. It is important to remark that the act of searching tells the computer / program / service to search for something, but it does not say and determine how this task should be fulfilled (Dourish, 2022, 23). There are always different ways to implement an algorithm, and the realisation establishes a bias and tendency of the search. Dourish (2022, 57) names this the “expressiveness of code”, with program code as a medium between humans and the machine. A search query also is such a medium, and the machine’s understanding of this query depends on the parallelism between human and digital affairs.

When starting to enter a search query in Google, people do not tend to recognise the suggestive power and restrictions imposed by the search engine. But there are certain queries which are not allowed and blacklisted (Haider and Sundin, 2019, 71), implying that certain questions are not allowed, and certain information are prohibited. Furthermore, Google automatically suggests questions and next words which also forms and shapes the query which is send off.

A search query often starts with or contains *keywords* – “a word that you type into a computer so that the computer will find information that contains that word”<sup>3</sup>. It is something that acts like a clue and often contains more information than just the ‘the semantic content’ of the word. It is also coined by society and culture: “the social content” (Schäfer and Es, 2017, 76). This “meta-information” is used by

---

<sup>3</sup>Definition of keyword from the Cambridge Advanced Learner’s Dictionary & Thesaurus © Cambridge University Press, 31<sup>st</sup> October 2022

*PageRank* (Google’s sorting algorithm) to resolve ambiguity. By disambiguating, it works social-epistemologically and reveals social hierarchies (Schäfer and Es, 2017, 87). Concerning the Israeli–Palestinian conflict, for example, two different keywords were around for a while, both telling a different narrative and demonstrating different political opinions. When using the term “apartheid wall” (coined by the Palestinian), Google retrieves images with a massive wall and violent images whereas the Israeli term “security fence” returns a harmlessly looking, partly broken down fence (Schäfer and Es, 2017, 89). By choosing the words used in the query, the returned results reflect very different world views and political positions. Scholars use this effect to try to find out the dominant interpretation by searching with neutral terms and detect which results are returned. But in this case, users are often unknowingly confronted with the echo chamber effect and not a neutral compilation.

### 3 Retrieval of the Search Result

After a user formulated a search query and sent it off, he often receives the answer in under a second. But the search result is retrieved among millions of possible results, using a lot of different algorithms and searching in big data infrastructures. Google works by indexing information on the web and querying it upon request. Thus, it classifies the information about the world, but it does so without knowing which information will be relevant in a specific use case (in the form of a search query). A dilemma highlighted by Lemke in 1995 (Bowker and Star, 2008, 116). The categories created by Google’s indexing are situated artefacts of a specific time period and try to completely cover the whole world, everything there is to know. But there is no real-world classification system that can ever completely cover the world it describes (Bowker and Star, 2008, 287, 11). Furthermore, classifying is a way to make sense of the world and is ultimately linked to the culture behind / shaping the system (Dourish, 2022, 2).

When retrieving information, Google uses its algorithm PageRank which is based on the principle of *wisdom-of-crowds*: by aggregating incoming links to a webpage in a way which allows evaluating the importance of this page on the web (Masterton and Olsson, 2018, 593; Escandell-Poveda et al., 2022, 56). This approach is also used for scientific papers to determine the scientific impact and quality or relevance of the paper. While this is the idealised idea and key algorithm behind Google’s search engine, implementations first make this idea feasible (Dourish, 2022, 2) and

they underlie more practical and physical restrictions, which form the retrieval of a search result to a significant amount.

Google’s ranking algorithm has been through several changes. The most recent and relevant modifications of the algorithm had great consequences on the retrieved search results (Escandell-Poveda et al., 2022, 58f.):

- 2003** The update called “Florida” punished abusive keyword usage which was used to get a higher position in the search results.
- 2004** Google started rating by the relevance of anchor texts and hyperlinks.
- 2010** Indexing became faster, and search results became more current.
- 2012** Keyword overuse was again penalised, as well as the trading and buying of links.
- 2013** Understanding of natural language and voice queries was improved.
- 2015** Mobile search was promoted by prioritising websites with optimisation for mobile devices and *RankBrain* was introduced – a machine learning algorithm to better understand the intention behind search queries.
- 2016** Two updates were supposedly released but never confirmed: giving more relevance to local businesses and the punishment of excessive advertising or content solely created for monetarisation.
- 2019** With BERT (Bidirectional Encoder Representations from Transformers), it was attempted to understand language even better and later that year a websites’ speed and response was included as a factor in the ranking.

Overall, the changes show that there are specific design decisions influencing the ranking of the results, differing significantly at various stages of development. By including artificial intelligence in the retrieval of search results, societal biases are re-enforced (Moreau et al., 2020, 3). The decision to “reward” mobile or fast loading websites, for example, sets standards for web development and also creates a new social hierarchy. But it also defines by default what result relevant for the user is, without considering individual preferences / definitions of relevance.

When returned, the search results are presented in a specific order by the *relevance* of the search result concerning the search query. Relevance is difficult to measure and highly individual. Moreover, there are many different forms of relevance such as topic, system, cognitive, situational or affective relevance. In the context of search engines, relevance has been called the “the invisible hand” governing the system (Haider and Sundin, 2019, 69). Google’s results are at best a scientific opinion as

to what is most relevant for the user (Patterson, 2012, 22). Google’s maxime is user relevance, and it uses search rater tests and experiment tests to increase the relevance ranking relevance for the users. Furthermore, Google published a list of over 200 factors influencing the ranking, which can be summarised by the following categories: “referring domains, organic click-through-rate, domain authority, mobile usability, dwell time, total number of backlinks, content quality and on-page SEO” (Haider and Sundin, 2019, 70, 67). SEO stands for *search engine optimisation* – an aspect which will be discussed later.

Because the ranking algorithm determines who will be displayed first and also defines what is relevant for the user, search engines have a *social responsibility* (Lindsköld, 2018, 778). In the debate about the “right to be forgotten”, Google argued that they only index and return results. But some people might not want to be found due to various reasons. Since Google provides users with information, they can influence and assess which information is found and therefore Google is “conceptualised as being a judge of character in the discourse” of the right to be forgotten, because two main reasons for wanting to be forgotten are identified: being innocent / a victim or being culprit (Lindsköld, 2018, 777).

By ranking the search results, Google does editorial work, but Google does not consider having editorial responsibility – the maxime is to return what people want to find. There is a dilemma, because the users tend to choose a result quickly by impulse by the title and less about the content. In combination with the echo chamber effect, this makes Google’s definition of relevance with regard to their responsibility problematic (Kuhn and Hauck, 2012). Additionally, the amount of data Google uses could be revised: it has been found out that the search could be realised with more privacy without lowering the quality of the search results (Chiou and Tucker, 2014, 19).

### **3.1 Materialities of the Search**

After inspecting the search algorithm and its history of Google, the materialities of the search are of interest now. The algorithm strongly influences the ranking and retrieval of results, but it cannot be forgotten that every digital service / infrastructure also is material and thus has physical limitations and restrains (Dourish, 2022). Infrastructures as databases in their physical as well as virtual form always are trade-offs between expressiveness, efficiency and transparency, etc. which constrains,

enables and forms the creation, storing, transmission and manipulation of information (Dourish, 2022, 6, 123).

In the beginning of the paper, it was argued how search engines and Google are invisible in everyday life. But this is only the case as long as everything works fine, and the search engine is easily reachable. The physical world can make itself felt, because the information is in the end always only encountered in physical and material form (Dourish, 2022, 3). The recent world-wide inaccessibility of What's App on 25<sup>th</sup> October 2022 showed users how much they rely on and use What's App in everyday life. Students feel that “[being] without Google would be the same as being without facts” (Andersson, 2017, 1251). Moreover, the search results themselves are experienced as goods and not only pure information (Patterson, 2012, 11).

Bowker and Star (2008, 13) argue that classifications and standards always impact the material and symbolic. While a search primarily is the use of a digital service, only consisting in the information retrieval, it can take material form if used as, for example, informational basis for political acting as in the example of the Israeli–Palestinian conflict. A standard as the information infrastructure of Google consists of agreed-upon rules and spans over multiple communities of practice, making things work over distance, but it has inertia and is difficult to change (Bowker and Star, 2008, 13f.).

Google's search represents a *boundary object* (Bowker and Star, 2008, 297), users want to find something, companies want to sell something, local businesses or institutions want to be found and others want to offer information. These different communities are represented, fitted and standardised in the search Google offers. As Google works world-wide, different cultural habits also have to be “standardised” in search. Nevertheless, there are frictions, if dissonances and incongruencies between expectations and reality are experienced, because of individual requirements, social and cultural interests or other stakeholders (Haider and Sundin, 2019, 143).

Finally, Google's predominance is also manifested in their database sizes, the server capacities and network infrastructure, which requires a lot of people to keep on working (Patterson, 2012, 6). Alphabet (the company behind Google) had 156,500 employees in 2021<sup>4</sup> with more than 70 offices in 50 countries<sup>5</sup>, keeping Google running. The server size is unknown, but estimated to be at around 2,5 million servers

---

<sup>4</sup>“Anzahl der Mitarbeiter von Alphabet weltweit in den Jahren 2007 bis 2021”, L. Rabe, 8<sup>th</sup> February 2022, de.statista.com, last access: 1<sup>st</sup> November 2022

<sup>5</sup>“Our locations”, <https://about.google/locations/>, last access: 1<sup>st</sup> November 2022



Figure 1: On the top left is a Google server room (“Daten und Sicherheit”, <https://www.google.com/intl/de/about/datacenters/data-security/>, last access: 1<sup>st</sup> November 2022), on the top right are the locations of the Google data centers (“Entdecken Sie die Standorte unserer Rechenzentren”, <https://www.google.com/intl/de/about/datacenters/locations/>, last access: 1<sup>st</sup> November 2022) and on the bottom is a European data center (“Unser Engagement an den europäischen Standorten der Google-Rechenzentren”, <https://sustainability.google/intl/de/progress/projects/eu-data-centers/>, last access: 1<sup>st</sup> November 2022).

in 2016<sup>6</sup> with in 2019 approximately around 10-15 exabytes of data<sup>7</sup> <sup>8</sup> in 23 data centres<sup>9</sup>, used to save the index-information, user-data, etc. The biggest data centre is assumed to be 91044.97 m<sup>2</sup> big <sup>10</sup>.

### 3.2 Metadata and GoogleAds

Previously, it was outlined that Google’s definition of relevance of a search result for the user has the highest impact on the ranking in Google search. But it has been argued that search engines nowadays – although first developed and maintained in the academic realm with interest in information provision and retrieval – are commercialised. As discussed in the sections before, different target groups can strongly influence the search result retrieval and infrastructure make-up. Thus, it is necessary to understand and analyse the target group of the search engine (Mager, 2012, 3):

**1994-1997** The period of *technical entrepreneurs* who developed search engines for information retrieval with roots in academia.

**1997-2001** The period of *portals and vertical integration* during the dot-com boom and bust, which stands out with a shift from search engines to portals and content development targeted at an audience (media corporations).

**2002-ongoing** The period of *syndication and consolidation* where search is dominated by technology companies which promote advertising and target big companies. This allowed companies as Google to buy up competition.

There are claims that Google does not return results which are best for the user, but instead rather server their own interest (Patterson, 2012, 2).

---

<sup>6</sup>“Google Data Center FAQ”, Data Center Knowledge, 17<sup>th</sup> March 2017, <https://www.datacenterknowledge.com/archives/2017/03/16/google-data-center-faq>, last access: 1<sup>st</sup> November 2022

<sup>7</sup>10,000,000-15,000,000 terrabytes, which would equal 20-30 million personal computers with a common memory of 500 gigabytes

<sup>8</sup>“What is the current memory size of Google servers including Gmail, YouTube, Google+?”, Rishi Jaswal, <https://www.quora.com/What-is-the-current-memory-size-of-Google-servers-including-Gmail-YouTube-Google>, 17<sup>th</sup> July 2019

<sup>9</sup>“Entdecken Sie die Standorte unserer Rechenzentren”, <https://www.google.com/intl/de/about/datacenters/locations/>, last access: 1<sup>st</sup> November 2022

<sup>10</sup>“Google Data Center FAQ”, Data Center Knowledge, 17<sup>th</sup> March 2017, <https://www.datacenterknowledge.com/archives/2017/03/16/google-data-center-faq>, last access: 1<sup>st</sup> November 2022

To be able to be found, a website has to be optimised for search engines today. The “existence of SEO demonstrates the power of search engines” (Patterson, 2012, 14), because to exist in “the internet”, one has to appear very high in a search engine, which depends on how well the website conforms to search engines’ guidelines which are not publicly revealed (Moradi, 2019, 55). This also emphasises the standardising power search engines have. Because optimising websites for search engines has become increasingly important, it allowed for a new industry to emerge: *search engine marketing* (SEM). Experts in SEM analyse the small amount of information provided by Google and their own search experience (experiments, trials, surveys) to create websites which rank highly in searches (Haider and Sundin, 2019, 12f.). The idea behind SEM is that people, when searching, often do not only want information but maybe want to buy something related to the keyword like, for example, “t-shirt”. Sellers bid money to place their advertisement next to normal results and every time a user clicks on an advertisement link, the seller has to pay for the click: *pay-per-click* (Turow, 2006, 89f.).

Entering data into a data infrastructure is a difficult task and “no matter how good the scheme, its scope is limited by the fact that data entry is never an easy task” (Bowker and Star, 2008, 107). To summarise, the way websites are implemented, strongly determines how and when they are found. Changes of the information infrastructures like, for example, the emergence of SEM, deeply change the world of knowledge as well as the industry (Bowker and Star, 2008, 109).

## 4 The Power of Page 1

The last two sections dealt with entering the search query, which is performed on an individual and societal level (choosing and reflecting about using search engines as well as which keywords to use), and the retrieval of the search result, which is performed by the search result providers (Google’s algorithms, data infrastructure and the websites which are indexed by Google as well as SEM-companies). Now, the presentation and selection of search results is discussed. Whereas the presentation of search results takes up on the ranking of search results, the selection of search results happens by the user. Although digital data in its essence is “only” bits, thus 0s and 1s, it has been shown that not all of them are equally important as in compression processes (Dourish, 2022, 17). The order of search results strongly influences user choice. Users, even if they want to oppose the ranking performed for



them, have difficulties to understand the make-up of search results and fail to assess the quality of search results (Patterson, 2012, 11), which is further complicated by the personalisation (Lindsköld, 2018, 768). Infrastructures and as such search engines are never transparent and increasing complexity impede workability (Bowker and Star, 2008, 33).

## 4.1 Presentation of the Search Results

When search results are presented, only few results are shown in relation to the millions and billions of results Google claims to have found. Google even cuts off the results after some pages (roughly page 34) and says that it omitted the other results because they are less relevant due to similarity to the information shown (Haider and Sundin, 2019, 49). Consequently, this means that out of two suppliers, only one is shown if too similar.

Partly, the *cut-off* is done for simplification and copability with the huge amount of results. It conceals the complexity behind searching and blinds users of the large amount of results (Andersson, 2017, 1245). Although the number of results is shown, users mainly perceive the first top results. If confronted with many different information sources and different information, people usually tend to become more uncertain because of the variety, but when using Google this effect does not occur, because the amount of results is reduced for copability. At the same time, it limits the user in his choice and suggests that he is unable to evaluate the information for his need on his own (Haider and Sundin, 2019, 9).

Design always anticipates data and action (Dourish, 2022, 91). As indicated before, the search for keywords and the display of the title in Google anticipates keyword stuffing, the auto-completion of queries suggests what to search, etc. During the presentation of search results, users are confronted with various suggestions which are part of the user experience design or user journey. With *knowledge panels* (the boxes with a website's content – often Wikipedia – above the search result) which are directly visible without scrolling (Vincent and Hecht, 2020, 1), it is suggested that this is the “ultimate” answer to the query. Furthermore, similar questions or advertisement and products in the first line, suggest the next action and information searched for. By simplifying the search and strongly leading the user throughout the search, Google has become an authoritative source returning user-specific results. This makes the evaluation of the search result for users more difficult, because

although perceived as epistemological machine, the differentiation between *personal* (served to an individual) and *universal results* (served to all) is impossible (Schäfer and Es, 2017, 76f., 80).

As indicated before, advertisements have become increasingly important – for Google as a source of income and for suppliers to be visible in the web and rank high in Google’s algorithm. There is an ongoing conflict between sellers and search engines, because search engines want to preserve their image as neutral fact-finding machine whereas marketers want to appear as normal search result to increase the click-rate and visibility (Mager, 2012, 9). There have already been trials concerning corruption of search engines because advertisements were not sufficiently marked as advertisement (Haider and Sundin, 2019, 76).

How much Google and other search engines shape user’s knowledge is also highlighted during the presentation of results. The previously introduced knowledge panels are very prominent and filled with content from sources Google prefers. Wikipedia’s definition (an open source, crowdsourced encyclopaedia) appears, for example, for 81-84% of the queries concerning common queries, 67-72% of the queries concerning trending and 16-54% of the medical queries (Vincent and Hecht, 2020, 1). Culture has a common understanding of search engines in that it will direct the user to a page that contains the desired information (Lindsköld, 2018, 769). But Google more and more becomes a search engine that does not direct any more.

## 4.2 Selecting a Search Result

The selection of a search result is ultimately linked to its presentation. Since user action and practices are led by the design and presentation of the result, result presentation exploits human fallibilities when selecting search results (Mager, 2012, 9). The suggestive power discussed before only works because “most users: a) click the top results, b) have the results set to the default of ten, and c) do not venture beyond one page of results” (Schäfer and Es, 2017, 80). Another aspect when choosing a result is that people do not question the ranking mechanism and presentation of results, trusting Google as an authority which critically assesses different possible results (Andersson, 2017, 1254). Furthermore, over 60% of users do not distinguish between organic (not an advertisement) and sponsored results and therefore often choose advertisements (Mager, 2012, 9). This is because people tend to do what they are able to do and satisfice if there is too much information (Bowker and Star, 2008,

24).

Bowker and Star (2008, 10) remark that classification is a “spatial, temporal, or spatio-temporal segmentation of the world” which sorts things in clearly distinguished bins. Personalisation algorithms work very similar, sorting the user in one or many categories on the basis of the data which is available. Search engines offer possibilities to personalise their products because this reveals information about the user, which is used in customisation as well as they track the websites the user visits, the queries, etc. (Turow, 2006, 97). But how can categories as humour or other human skills and traits be captured and evaluated (Bowker and Star, 2008, 30) and be included in the finding of the most relevant result for the user?

Personalisation leads to less privacy, the gathering of data, targeted advertisements and personalised results, which are aspects primarily concerning the individual. But personalisation also allows search engines to manifest and strengthen their predominance in the market and in society. Search engines are shaped by their use and the data, which is – on an individual level – used to personalise the search, is used – on a collective level – to improve the whole system. By suggesting common queries, these queries gain dominance and are thus displayed to even more users, increasing the likelihood to be suggested. This is also called the Matthew Effect, similar to the echo chambers: “the known becomes more known” (Haider and Sundin, 2019, 5). The more searches are executed via a search engine, the more data it has to improve the search becoming better at returning satisfying results (Patterson, 2012, 7) which makes it more attractive to users and more difficult to exchange (Kuhn and Hauck, 2012; Vincent and Hecht, 2020, 7).

## 5 Being Found

After analysing the process of a single search, from opening the search engine, entering a search query, executing the search, retrieving the search result, presenting and selecting it, the existential impact search engines have on those who are and are not found will be discussed.

With the increasing amount of digitalised and global information, it has become increasingly difficult to navigate and find the information needed. Moreover, searching simplified many social practices such as finding out opening hours or the address, that the need for search engines has existed for a long time (Mager, 2012, 12). Search engines are negotiated as any other infrastructure by different players and actors

of different interests. The mainly corporate search engines came into existence in a techno-euphoric climate of privatisation and innovation, coined by capitalist ideology, and are nowadays manifested in a socio-political climate (Mager, 2012, 1). Google – as one example – is a political subject because it influences knowledge acquisition, distribution (Lindsköld, 2018, 768). By personalising as well as indexing, Google automatically classifies information and people and their relation, thereby depicting the world in its information infrastructure. But no classification system is universal or natural (Bowker and Star, 2008, 131).

To be successful on the internet and by consequence even in the non-digital world, one has to “pass” Google, as it is the “door” to the internet (Mager, 2012, 8). If the link of a supplier is not part of the first links shown by Google, it usually will not be considered by the user. The hierarchy of suppliers in Google is also called *googlearchies* (Schäfer and Es, 2017, 77). This concerns local businesses as people search for restaurants, individuals as employers search for employees, big business as employees search for jobs, newspapers as users search for information...

The preferred return of Wikipedia or Amazon links in Google, moreover, impacts the non-digital world as well: When a search for a specific varnish is committed, Amazon will probably offer the searched product, but the local specialised shop for varnish, which one is not aware of, might not come up, because it is less optimised for Google and does not pay to be advertised. If this often happens, these local business have to close. The same applies for information providers such as newspapers. A small new local editorial might not be found, although it addresses a big interested audience, because it does not have the resources to compete with and invest as much in SEO as an established newspaper and moreover cannot increase its ranking by clicks, because it is never found. Since the mechanisms behind the ranking are not revealed, companies make profit out of identifying them and selling this knowledge (Moradi, 2019, 55) further distorting the view on the available information. In medium research, it is researched for whom and how Google works: “To which degree does the engine serve a handful of dominant websites such as Google properties themselves in a ‘preferred placement’ critique, or websites receiving the most attention through links and clicks?” (Schäfer and Es, 2017, 77).

Marx argued that capitalism as well as feudalism are distinct periods of time with links concerning the material means of production: “The handmill gives you society with the feudal lord, the steam-mill society with the industrial capitalist”, and recently late capitalism is described by using the words neofeudal as public goods

increasingly become private (Dourish, 2022, 196) as is the case with search engines. Visibility in search engines is achieved by understanding and conforming to Google, which in return leads to more revenues (Turow, 2006, 167).

Although it has previously outlined that people trust Google, there are changes in society and people start to feel distrustful since there is little openness in the search process which hinders the sharing of ideas across society which inhibits a healthy public sphere. Being found and not being found also is important for individuals, not only suppliers of information and goods: Nowadays, employers often search for the people they plan to hire (Turow, 2006, 187). Thus, certain information should not be found, and other information should be found. But if one is not findable at all, people are more sceptical – although this concerns business more, it also applies to individuals. Searchability has a big impact on the role and use of information in a company (Dourish, 2022, 96).

Search engines aspire to make all information about the world retrievable and thus usable. As outlined before, the description is never complete, and it can happen that the world is changed according to the system's description, which is called *convergence* (Bowker and Star, 2008, 49). Such an example is a website which perfectly addresses the audience, contains all necessary information and does not need to work on e.g. mobile phones, but, because of Google's new algorithm, has to be adapted to be found in future as well. If Google defines the make-up of a good website or what should be searched for, it is – from a pragmatist standpoint – “real in [its] consequence” because there is no way around Google, although it might not be the optimal or desired thing (Haider and Sundin, 2019, 140).

By showing only a limited amount of search results, Google automatically classifies information as relevant and irrelevant, important and unimportant, useful and not useful, interesting and boring, trending, ... However, Google is neither transparent nor accountable for their results, and it is questionable if such decisions should be up to a private corporate search engine which has proven to promote the Anglo-Saxon culture in their algorithms (Lindsköld, 2018, 779, 720).

To summarise, being found requires the supplier to conform with the rules and methods imposed by the search engine which is used to find the supplier (Haider and Sundin, 2019, 2). Additionally, if excluded from search results, it can have existential consequences: individuals do or do not get jobs, business cannot keep up with their competition, information providers do not reach their audience and societal as well as political decisions are formed by information political agents search for every day,

influencing the public opinion even in military clashes. The exclusion and denial of access to Google can lead to societal exclusion (Patterson, 2012, 3). In 1980, Derrida – figuratively – warned that the myriad of standards, infrastructure and classification of the modern world will blind people of the “other” category and its importance of the whole (Bowker and Star, 2008, 301). When searching, only a tiny fraction of all possible results is shown and people are blinded of all the other information, which has severe consequences.

## 6 Discussion

To put it in a nutshell, when opening a search engine, not only a search engine is opened, but a commercial tool which is biased and strongly influences the result presented. Therefore, the choice of search engine – similar maybe to the choice of a newspaper – determines what section of the world of information one will be confronted with. By entering a search query, the result is anticipated in a non-objective way (see Israeli–Palestinian conflict) and personalisation is used to resolve ambiguity as well as understand the user’s intention behind the query to retrieve optimal results. Depending on the keywords used, different results will be returned and by using de-personalised search engines, social trends and hierarchies can be retrieved from the search results. The retrieval of the search result is intimately connected to the indexing of websites and information. Websites have to adapt to standards imposed by Google to be ranked high and thus findable. A new industry emerged, commercialising how well suppliers can be found. The presentation of search results strongly suggests certain information to users and by breaking down and concealing the complexity of the search, leaving the user with little choice and possibilities to evaluate the result. Users, on the other hand, enforce the existing pattern because they do not know how to and cannot influence the parameters about the search. Hence, they strengthen the search engine and the presented world view as well as results by using it.

The design of a search engine – in the case of the present analysis Google – strongly impacts what is found, in parts independent of what is searched for. By commercialising the search and because of capitalistic tendencies in society, being found has become a competition, leading to a driving out of smaller suppliers and strengthening big suppliers. Attempts by users to avoid known “traps” fail, because almost all levels of the search are concealed and kept secret and searching has become

nearly invisible – the user is treated as immature. Returning the best result for the user is not necessary, a result which satisfies the user and stands out to other search engines, is sufficient. But it has to be considered and kept in mind how much the results shape society; having material consequences for, for example, local shops, forming political as well as cultural opinions. They are provided by companies in a capitalistic world for free, which means that they have own commercial interests dominating the search.

Since there is a need to find information and make them retrievable, search engines are tightly interwoven with daily life and in everyday practices around the world. A life without cannot be imagined. Thus, being found has strong existential power: What is the use of selling the best products, if nobody finds the store? What is the use of spending a lot of money in SEM, if only Europeans find a website in Chinese? What is the use of providing objective information, if the website is never visited? What is the use of being found among others, if the user cannot assess and cherish the quality of the result? What is the use of finding, if the quality and relevance cannot be assessed? Why searching and finding, if one does not know if one found what one wanted to find?

## References

- Andersson, C. (2017). “Google is not fun”: an investigation of how Swedish teenagers frame online searching. *Journal of Documentation*, 73(6):1244–1260.
- Bowker, G. C. and Star, S. L. (2008). *Sorting things out: classification and its consequences*. Inside technology. MIT Press, Cambridge, Mass., 1. paperback ed., 8. print edition.
- Bruns, A. (2019). Filter bubble. *Internet Policy Review*, 8(4).
- Chiou, L. and Tucker, C. (2014). Search Engines and Data Retention: Implications for Privacy and Antitrust. *SSRN Electronic Journal*.
- Dourish, P. (2022). *Stuff of Bits: an essay on the materialities of information*. MIT Press, S.I. OCLC: 1346213354.
- Escandell-Poveda, R., Iglesias-García, M., and Papí-Gálvez, N. (2022). From Memex to Google: The origin and evolution of search engines. *Informes DigiDoc - EPI*.
- Fletcher, R. and Nielsen, R. K. (2018). Automated Serendipity: The effect of using search engines on news repertoire balance and diversity. *Digital Journalism*, 6(8):976–989.
- Haider, J. and Sundin, O. (2019). *Invisible search and online search engines: the ubiquity of search in everyday life*. Routledge, Taylor & Francis Group, Milton Park, Abingdon, Oxon ; New York.
- Kuhn, J. and Hauck, M. (2012). Eli Pariser und die „Filter Bubble“ – Ausweitung der Komfortzone. *Süddeutsche Zeitung*.
- Lindsköld, L. (2018). Google as a political subject: the right to be forgotten debate 2014-2016. *Online Information Review*, 42(6):768–783.
- Mager, A. (2012). ALGORITHMIC IDEOLOGY: How capitalist society shapes search engines. *Information, Communication & Society*, 15(5):769–787.
- Masterton, G. and Olsson, E. J. (2018). From Impact to Importance: The Current State of the Wisdom-of-Crowds Justification of Link-Based Ranking Algorithms. *Philosophy & Technology*, 31(4):593–609.



- Moradi, M. (2019). Crowdsourcing for search engines: perspectives and challenges. *International Journal of Crowd Science*, 3(1):49–62.
- Moreau, J. T., Baillet, S., and Dudley, R. W. (2020). Biased intelligence: on the subjectivity of digital objectivity. *BMJ Health & Care Informatics*, 27(3):e100146.
- Patterson, M. R. (2012). Google and Search Engine Market Power. *SSRN Electronic Journal*.
- Schäfer, M. T. and Es, van, K., editors (2017). *The Datafied Society. Studying Culture through Data*. Amsterdam University Press.
- 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*, 73(2):224–243.
- Turow, J. (2006). *Niche envy: marketing discrimination in the digital age*. MIT Press, Cambridge, Mass. OCLC: ocm62738704.
- Vincent, N. and Hecht, B. (2020). A Deeper Investigation of the Importance of Wikipedia Links to the Success of Search Engines. arXiv:2004.10265 [cs].

## Competing Interests

This paper was written in the context of the course “Data Infrastructures: Social and Cultural Perspectives” at the Ruhr-University Bochum. The paper is rewarded with credit points and evaluated with a grade.

## Statement of Authorship

I hereby certify under oath that the paper I am submitting is entirely my own original work except where otherwise indicated. I have not used any auxiliary means other than those listed in the bibliography or identified in the text and any use of the works of any other author, in any form, is properly acknowledged at their point of use with indication of the source.

**Anais Siebers**

(firstname, surname)

**2.11.2022, Siegburg**

(date and place)

*Anais Siebers*

(signature)