Faculdade de Engenharia da Universidade do Porto

Web Search Engines - A study on the evolution of user interfaces

Bruno Edgar Évora Rebelo Oliveira

Mestrado Integrado em Engenharia Informática e Computação

Supervisor: Carla Teixeira Lopes

July 27, 2021

© Bruno Edgar Évora Rebelo Oliveira, 2021

Web Search Engines - A study on the evolution of user interfaces

Bruno Edgar Évora Rebelo Oliveira

Mestrado Integrado em Engenharia Informática e Computação

July 27, 2021

Abstract

Web search engines have marked everyone's life by transforming how one searches and accesses information. They crawl the Web, organize world-wide information and make it universally accessible to every user. Since their beginning, a special attention was given to search engines' user interfaces, especially to search engine results pages (SERP), to assure their usability. Consequently, improvements have emerged with the objective of, while maintaining their simplicity, return information in the most practical and intuitive way. The well-known list of "10 blue links" has evolved into richer interfaces, often personalized to the search query, the user, and other aspects. More than 20 years later, the literature has not adequately portrayed this evolution. Knowing more about the evolution of web search interfaces is useful to future improvements in search engines' interfaces. We present a study on the evolution of Web search engine interfaces since their appearance. To study how search engine interfaces evolved during these two decades, we made a systematic analysis of SERP elements throughout the years. We used the most searched queries by year to extract a representative sample of SERP from the Internet Archive, which has been capturing web interfaces throughout the years, and permits to reach those under a specific URL. Using this dataset, we analyzed how SERP evolved in terms of content, layout, design (e.g., color scheme, text styling, graphics), navigation, and file size. We have registered the appearance of SERP elements and the appliance of user interface design patterns to each of them. We found that the number of elements in SERP has been rising over the years, most of them being launched by Google and later replicated by its competitor, Bing. Nowadays, both search engines present similar interfaces and an approximate number of SERP elements. Most of these elements are applications of design pattern solutions, with broader diversity in Google. We found that interface area increased almost exponentially in both cases, where Google leads consistently by a slight difference. In contrast, Bing is more solid when presenting less heavy pages, regarding the size of the files associated with the source code. This systematic analysis portrays evolution trends in search engine user interfaces and, more generally, web design. We expect this work will trigger more specific studies that can take advantage of the extracted captures, made available to the community as an image and HTML dataset, and the website we provide with complete results.

Keywords: Search engines, Search Engine Results Pages, Web interfaces, Web design, Evolution

Resumo

Os motores de pesquisa Web marcaram a vida de cada um ao transformar a forma como procuramos e acedemos informação. Eles navegam a Web, organizam informação mundial e tornam-na universalmente acessível a todo o utilizador. Desde o início, especial atenção foi dada às interfaces dos motores de pesquisa, em particular às páginas de resultados dos motores de pesquisa (SERP), para se assegurar a sua usabilidade. Consequentemente, melhorias foram surgindo com o objetivo de, mantendo a sua simplicidade, retornar informação da forma mais prática e intuitiva. A bem conhecida lista de "10 links azuis" evoluiu para interfaces mais ricas, muitas vezes personalizadas à pesquisa, ao utilizador, entre outros aspetos. Mais de 20 anos depois, a literatura ainda não retratou adequadamente esta evolução. Saber mais sobre a evolução das interfaces dos motores de pesquisa é útil para melhorias futuras nestas interfaces. Apresentamos um estudo sobre a evolução das interfaces dos motores de pesquisa desde o seu começo. Para estudar como evoluíram estas interfaces em duas décadas, fizemos uma análise sistemática dos elementos das SERP ao longo dos anos. Para tal, usamos as pesquisas mais solicitadas por ano para extrair uma amostra representativa das SERP através do Internet Archive, que tem vindo a capturar interfaces Web ao longo dos anos, e que permite encontrar aquelas que estejam abaixo de determinado URL. Utilizando este dataset, analisámos como as SERP evoluíram em termos de conteúdo, layout, design (e.g., esquema de cores, estilo de texto, grafismo), navegação e tamanho de ficheiros. Registámos o aparecimento de elementos SERP e a aplicação de user interface design patterns a cada um. Descobrimos que o número de elementos em SERP tem crescido ao longo dos anos, muitos dos quais são primeiramente lançados pelo Google e mais tarde replicados pelo concorrente, Bing. Hoje em dia, ambos os motores de pesquisa apresentam interfaces similares e um número aproximado de elementos SERP. Muitos destes elementos são aplicações de soluções de design patterns, com maior diversidade no Google. Mostrámos como a área das interfaces aumentou quase exponencialmente nos dois casos, onde o Google lidera consistentemente por uma curta diferença. No entanto, o Bing é mais eficaz ao apresentar páginas menos pesadas, no que respeita ao peso do código fonte e dos seus ficheiros associados. Esta análise sistemática retrata tendências de evolução nas interfaces dos motores de pesquisa Web e, genericamente, em web design. Esperamos que este trabalho seja veículo para trabalho futuro que possa tirar partido das capturas extraídas, disponibilizadas à comunidade como um dataset de HTML e imagem, e do website com os resultados completos.

Palavras-chave: Motores de pesquisa, Páginas de resultados de motor de pesquisa, Interfaces Web, Design Web, Evolução

Acknowledgements

To Professor Carla Teixeira Lopes for her continuous initiative and availability that amply contributed to elevate this work to its best.

To my closest ones that directly and indirectly walked along during this long journey.

To my departed family, especially my grandfather, for continuing to guide my way.

Contents

List of Figures

List of Tables

Abbreviations

API	Application Programming Interface
DOM	Document Object Model
FTP	File Transfer Protocol
HCI	Human-computer Interaction
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
IR	Information Retrieval
SERP	Search Engine Results Page
URL	Uniform Resource Locator
WWW	World Wide Web

Chapter 1

Introduction

This chapter presents the context for this work, along with the main motivations associated with its realization. It concludes by describing the structure and organization of this document.

1.1 Context and Motivation

Search engines have had an impressive growth in terms of personal and professional use, given the wealth of information available on the Web and which needs to be filtered to satisfy one's information needs [6]. Cisco's Visual Networking Index [13] estimates how much data there is in the world, stating that the annual global IP traffic will reach 4,800 zettabytes (ZB) per year in 2022. In 2021, the indexed Web contains at least 2.42 billion pages [3].

Search engines are composed of two main pages, the home page and the results page. The latter was known as the "list of 10 blue links" and has become sophisticated, with the constant evolution of Search Engine Results Pages (SERP) that evolved to improve user experience while searching. The outcome was the introduction of the most varied elements in just a single page: the general SERP, which is the focus of this work.

Interfaces are important for the success of any web application. Especially in search engines, they must support the user's search process, reduce their experience problems and maximize the understanding of the results presented. These roles, described in detail in Section 2.2, constitute a notable motivation for search engines to improve SERP interfaces and usability.

Users scarcely look for search results that are not the first ones, as presented in Figure 1.1 [29]. In most cases, users may simple rephrase the query if they cannot find promising results at the top of the list [23]. They rarely consult collections of information, such as videos or news, when they are on the first page of results, compared to the respective 'tabs' that lead to separate pages. For those reasons, search engines are returning on the first page, when possible, several aggregations of information from other types of results, usually located in other tabs [26]. This phenomenon of "tab blindness" is presented by Danny Sullivan [68], who states that users simply do not see these tabs. It is a priority to place samples from these collections, other than just the usual retrieved results on the SERP, not for the sake of diversity, but because most users will

not reach them outside the first results page. Another similar behavior, presented by Laura et al. [46] as "banner blindness", is related to advertisement banners, which the user also ignores. Overcoming these user tendencies was one of the incentives for the evolution of the interfaces, and, as a consequence, it is a strong motivation to study this evolution.

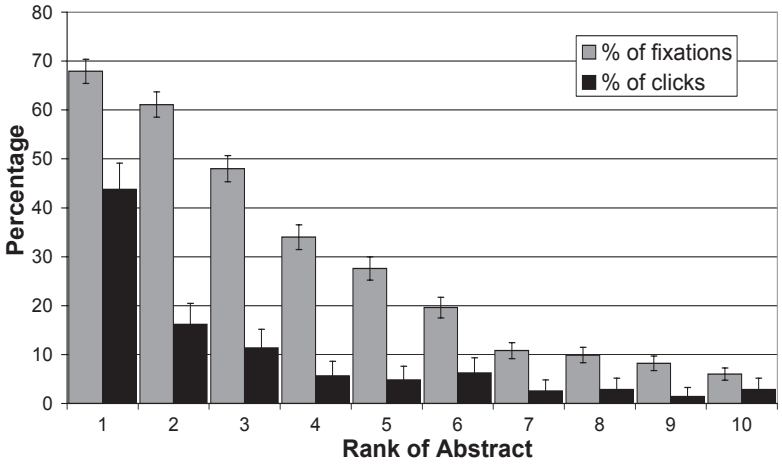


Figure 1.1: Percentage of views/clicks depending on the result's ranking (2017) [29]

Other aspects of the retrieval process received a more significant focus compared to interfaces. Studies addressing Web search engines' interfaces focus and discuss various aspects and SERP elements, but a complete picture of the whole interface, considering all possible results presentations, lacks in the literature and is a powerful tool for understanding what users really see [26]. This author was responsible for the first and only, to the best of our knowledge, research article addressing a complete overview of the elements that the most popular Web search engines use in their results page. Still, it is an outdated work, done when the current SERP features, described in this work, were still nonexistent. Thus, documenting in a more complete and updated way the evolution of search engine interfaces, producing complements for future research purposes, was the primary motivation for developing this work.

1.2 Problem Statement and Goals

This dissertation problem is based on the scarce research regarding the evolution and current state of the interfaces of Web search engines. It is necessary to split the problem into concrete steps, defined in Chapter 3, for the achievement of the following outcomes:

1. **Collect old SERP interfaces of the top search engines on the Web**

   Look for a repository with captures of SERP interfaces during the last twenty years, and define a strategy for the collection of a representative sample of that period.

2. **Study and document the evolution of SERP**

   Analyse overall aspects, like navigation, area and file size, and identify elements that constitute SERP, individually assessing characteristics such as content, layout and design (e.g., color scheme, text styling, graphics).

3. **Compare search engines' SERP**

   Include, when applicable, a comparative perspective of the search engines under study at each stage of the analysis.

4. **Analyze SERP source code to automate analysis**

   Look for identifiers that locate SERP elements in order to scale the evolutionary analysis in an automated way.

Using a systematic analysis, the dissertation's final objective is to analyze and describe the state and evolution to which the interfaces of the Web search engines have been subjected. They increasingly seek to correspond with the user's needs, returning the information in a useful and direct way as possible.

1.3 Contributions

The knowledge aimed in the previous section is essential for further research on Information Retrieval interfaces, and thus, we provide all our research data and results to the community. Our main contributions are listed as follows:

1. **Image and HTML dataset[1] of SERP**

   Dataset with all the 7000+ SERP captures between 2000 and 2020, extracted from Internet Archive. Each capture contains a screenshot, the source code, an its files folder.

2. **List of SERP elements' code identifiers**

   Listing all the code identifiers, used during the elements detection phase, to facilitate future work on analysis automation and coding of these elements.

3. **Website[2] on the study of SERP interfaces**

   Website to provide a personalized experience with complete and extra information resulted from this historical evolution analysis that can not be present in this document.

4. **Full Paper on the evolution of Google SERP**

   Submission of a paper, entitled *The evolution of Search Engine Results Pages*, to the 30th ACM International Conference on Information and Knowledge Management.

[1] Available at: https://doi.org/10.25747/991g-f765
[2] Available at: https://bedgarone.github.io/serpevolution/

1.4 Document Structure

This document is divided into four chapters. The first section of Chapter 2 refers to the state of the art, addressing the role, mechanisms, and history of Web search engines as the problem's context. The following section summarizes the state of the literature regarding the interfaces in Information Retrieval and the composition of Web search engines interfaces. Chapter 3 presents the methodology applied to this study. Chapter 4 exposes in detail the results of the individual SERP elements analysis, which addresses each search engine simultaneously. In an aggregated summary of the various domains, Chapter 5 analyzes results from an overall SERP perspective. Finally, Chapter 6 concludes the work and proposes further research in this subject.

Chapter 2

Background and State of the Art

This chapter explores the literature on Information Retrieval, with a primary emphasis on Web search engines and their interfaces, to better contextualize the problem. Many of the concepts and terminologies described here are crucial to the development of the work and the elaboration of strategies for its realization.

This chapter is organized into three sections. Section 2.1 covers search engines in general, stating the role and importance of Web search engines in society, the mechanisms used for their operation, their history and their popularity. The following section presents fundamental strategies in Information Retrieval interfaces, listing strategic components that underlie the structure of different systems, and analyzing the composition of search engine interfaces. Section 2.3 addresses user interface design patterns, solutions used in SERP elements that are further highlighted in this work. The chapter ends with Section 2.4, which presents related works on the anatomy of SERP, their interfaces, and general interfaces related to information retrieval.

2.1 Web Search Engines

This section contextualizes Web search engines, introducing their role, stating their importance nowadays, addressing their base architecture and mechanisms, resuming their history and concluding with popularity statistics.

2.1.1 Role and Importance

The amount of information present in the daily life of modern society continues to increase. At the beginning of this century, it was estimated that between 3 and 6 Terabytes of original information (stored digitally) were produced each year [35]. This value is growing, with each passing year, while new formats and new data are collected, and for Gregory Smyth [66], human beings cannot be able to memorize all this information, with the needs for tools to access information. In the article, the author says that the best-known tools are the Web search engines, and in fact, the leading owner of the Web search market, Google, already sees its name converted into a verb in the dictionary ("google it"). In 2020, 5.5 billion searches were done on Google per day [76],

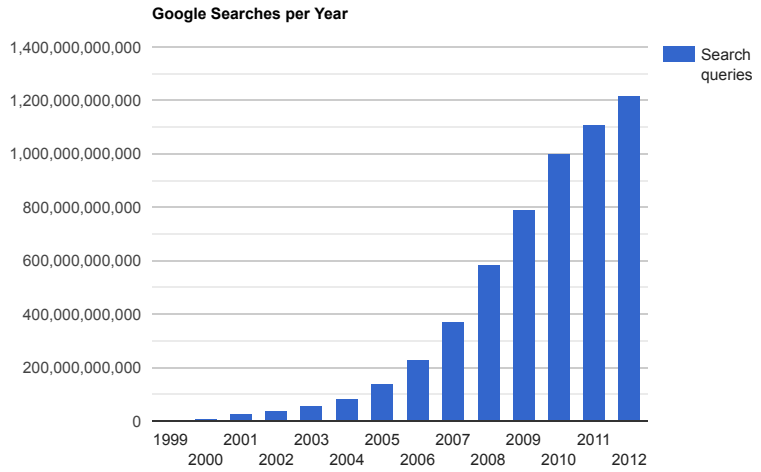


Figure 2.1: Number of Google searches per year (until 2012) [27]

which translates to about 2 trillion searches a year. This evolution of search requests, seen in Figure 2.1, would not happen if these tools were not an asset in returning the desired information. Search engines' role is to filter the massive amount of data available. Without this work, users would have to dive into a sea full of irrelevant information, while with search engines, users can quickly find information deemed relevant on top quality websites [66]. According to Tefko Saracevic [52], the concept of 'relevance' brings together "intents, goals, and motivations of a user, and texts retrieved by a system". In this way, search engines are currently working to better fulfill the role of delivering information to the user that correctly meets their needs.

Search is not always related to obtaining general, cultural, or academic knowledge. Many studies look for information that helps the user to satisfy needs or interests related to his life. Thus, one of the outcomes regarding the information that is returned is that search engines increasingly define the brands, products, and services users will view and potentially consume. Thus, companies must get their content to the results pages presented to the user. For Xing and Lin [77], two methods generally reach this goal. One method is to buy advertising from the search engine, where an advertisement is displayed in a pre-specified region next to the returned results. Another approach is *Search Engine Optimization*, which consists of optimizing the website itself so that it is highly ranked and appears in the first results' positions, constituting organic (unpaid) results. It is an emerging area, mainly because, according to Loren Baker [10], users give more preference to organic results, in contrast to paid results. The presentation of organic and sponsored results in the search engine interfaces will be analyzed throughout this work.

2.1.2 Architecture

For a brief understanding on what is the architecture, as suggested in Figure 2.2, and mechanisms that Web search engines use to deliver results, six steps can be listed: Web crawling, indexing, searching and processing, result matching, result ranking, and presentation of results [6]. The

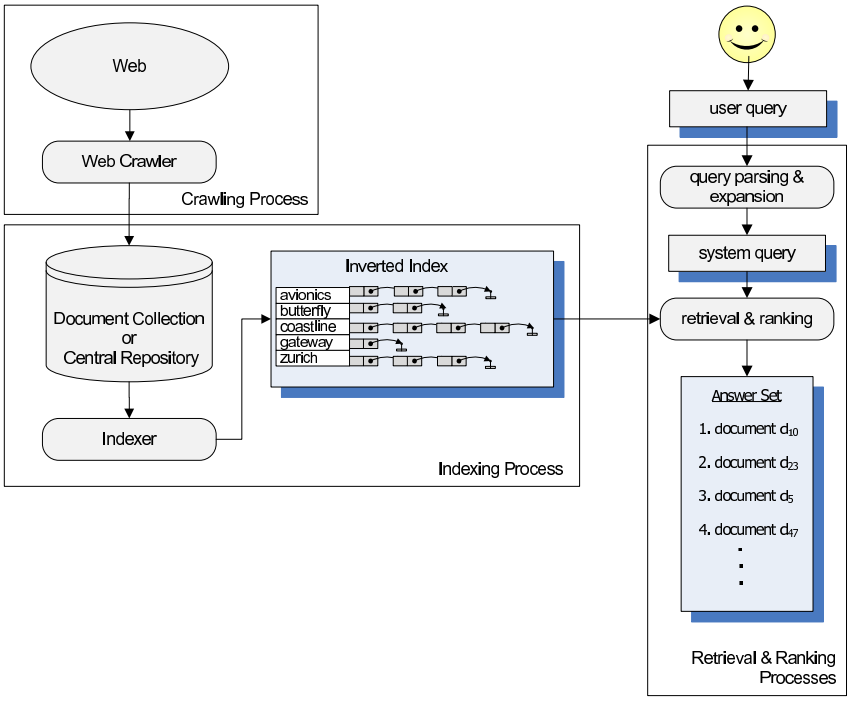


Figure 2.2: High level software architecture of an IR system [9]

first step uses programs called Web crawlers, as the name implies, to browse the information on the Web for a later process of indexing, done by a search engine, of everything they found. They begin with a list of links, which is increased as they are visited, in a recursive search consistent with a series of policies depending on the search engine. This work creates virtual copies of each visited page, stored in a place usually denominated as central repository, ready for a further indexing process that aims, as stated by Baeza-yates and Ribeiro-neto [9], for quick retrieval and ranking of those documents. According to these authors, during indexing, search engines parse the document to generate an inverted index, that compose the most common index structure (inverted indexes) that point distinct words to documents that contain them. By assigning specific keywords to each page, speed and performance in the search for relevant documents are improved. The third step is initiated by the user when he specifies a query reflecting the user's needs [9]. The search engine processes the terms entered and compares them with the pages indexed in the database. Matching the results is associated with this task, in which the query is compared with the similar pages detected. Finally, the result ranking stage defines the order whereby the results will be presented to the user, followed by the results presentation step, which consists of displaying the results along the search engine results page's interface [6].

2.1.3 History

The first Web search engine appeared in the last decade of the 20th century. Until then, search tools did not exist or were scarce. Jon Penland [31] describes the history of search engines, which begins in 1960, at Cornell University, even before the Internet was created, where the SMART Information Retrieval System was developed. The author argues that, despite not being a search engine, this early information retrieval system established concepts such as term weighting, relevance feedback, and many others, all necessary foundations for search engines. Penland includes the WHOIS as a context to this overview, although more related to data retrieval, being created in 1982 as the first search tool in databases on the Internet, for tracking or locating information regarding resources on the Internet. It is still used today, with more limited parameters, to find the registered owner of any domain or website. Searching for content, not for users, Archie was launched, in 1990, being considered the first search engine [62]. When searching for a file name in Archie, the answer consisted of the directory path and the system containing the file's copy [62]. Which, according to Penland [31], was something that was necessary to know by heart in case it was required to do some research on the famous FTP servers. Archie allowed a connection through the command line and queries formatted in email or through a search interface. Figure 2.3 shows Archie's interface.

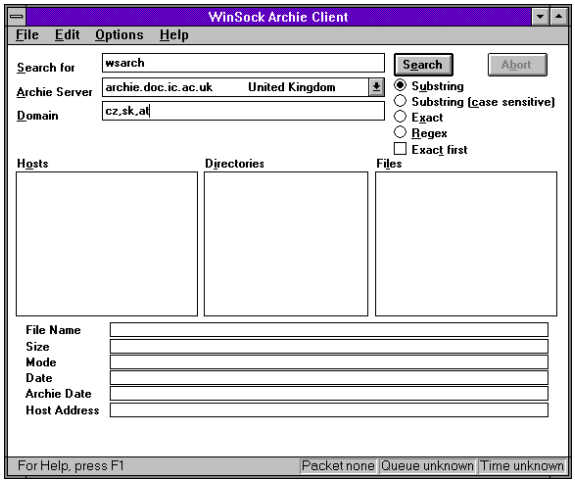


Figure 2.3: Archie's WinSock client interface in 1995 - source: earchiv.cz

Archie required the user to be aware that a file was somewhere on a particular site. That stopped happening in 1991, with the launch of Gopher, which included searching databases and text files; all it took was a search and waiting for the site that contained that information [74]. Closer to the Web than to FTP servers, it was seen as a kind of File Manager, as for enthusiasts it was faster and more organized than the Web itself [31]. Penland's text states the requirement, in the search process, of navigating through various menus and submenus associated with titles and descriptions until the desired file is found. Alternatives to this process emerged, Veronica and Jughead. For the author, Veronica consisted of applying the Archie model to the Gopher protocol, with queries to the Gopher servers. Jughead, also a Gopher tool, had the limitation that

it could only be used on a limited part of Gopher, and, in contrast, it consisted of a powerful tool with compatible advanced search operators, dedicated to a single Gopher server [31]. With the emergence of the Web in 1993, W3Catalog was born, the first Web search engine, with a concept very similar to Archie and Veronica. Seymour, Frantsvog and Kumar [62] describe that it offered a catalog for WWW resources, mirrored the pages, reformatted the contents for individual entries, and provided a front-end for dynamic querying. Aliweb, the second search engine, allowed users to submit file locations for their sites, so for a site to be indexed on Aliweb it was necessary to register on the server [62]. As seen in the previous section, indexing and Web crawlers are essential in the results returned. The first Web crawler was created in 1993, allowing the launch of the first application of a Web crawler, JumpStation [31]. Since then, the evolution has been more and more notorious until the end of the decade, when the most resonant names popular with today's users appeared. AltaVista appeared in 1995 and was one of the most popular search engines until Google arrived [62]. It was the first search engine to use Boolean operators. In the same year, Yahoo! appeard, another search engine with great prominence that later, in 2003, came to buy AltaVista and launch Yahoo! Search one year after, which is still in the ranking of the most used search engines today [12]. In 1996, Ask Jeeves (now known as Ask.com) appeared, with the purpose to answer users' questions in natural language. In 1998, Google was born, a search engine that revolutionized the market, starting and leading innovation in this area. According to Penland, [31] what drove Google to great success (the world's leading search engine) was its patented algorithm, PageRank, which recursively assigned scores to pages based on weighted sums of Page Ranks of pages that linked to them. With the advancement of technology, many other engines have emerged. For the development of this work, it is important to highlight the appearance of another name, Bing, now known as "Microsoft Bing", introduced in the 21st century, in 2009, under the baton of Microsoft. It is currently the second most used search engine [12].

2.1.4 Popularity

Web search engine's popularity can me measured according to the market share. The most recent data are synthesized by Alex Chris [12], who lists the ten most popular search engines today, highlighting the first five in Figure 2.4:

1. **Google** is undoubtedly the leading search engine today. It has a market share of 92.26%, maintaining a domain in all countries on any device.

2. **Microsoft Bing**, despite being runner-up, already has a big difference in market share (2.83%) and is still the default search engine on Windows computers.

3. **Yahoo!** has an average market share of 1% and is, since 2014, the default search engine in Firefox browsers in the United States.

4. **Baidu** is the most popular search engine in China, with a global market share of 0.68% and only available in Chinese.

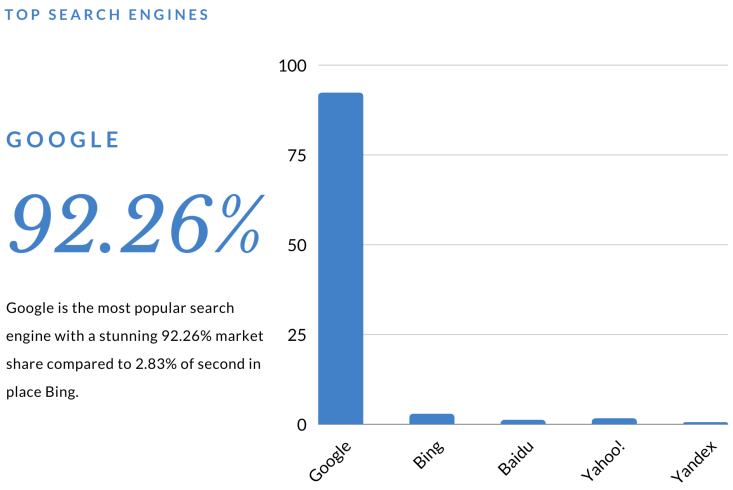


Figure 2.4: Top 5 Web Search Engines in 2021 [12]

5. **Yandex** is another locally popular search engine, in this case in Russia, with a global market share of 0.5% and 65% in that country.

Other generalistic search engines mentioned in the list are, in order, DuckDuckGo, Ask.com, Ecosia, Aol.com, and Internet Archive, which has the "Wayback Machine" tool without which this work would not be possible.

2.2 Interfaces in Information Retrieval

This section is divided into two parts. The first is more focused on information retrieval, related to general search systems, and addressing the strategic components essential to the search process. The second part analyzes the state of the art regarding the composition of Web search engines' interfaces.

2.2.1 Strategic Components

When faced with a new information system, the user does not have complete knowledge to manage and expertize the system tools to achieve his goals. A user interface should help the user to acquire that knowledge intuitively, bringing their needs closer, while facilitating also the user experience. Marti Hearst [24] points that a human-computer interface is a field not as well understood as other information retrieval areas since there are human factors more complicated than computer systems, addressing the difficulty to define or characterize behaviors and motivations in the human case. The author lists three fundamental design principles: offer informative feedback, reduce working memory load and provide alternative interfaces for novice and expert users. Hearst's text clears that determining how much information should be available to the

user is a choice of great importance when designing interfaces in information retrieval. Combining this with the fact that humans have a greater affinity for images and visual information [73] is also desirable. Well suited figures can be more captivating. In this work, we'll look at how search engines applied better graphics to develop new interfaces.

A person relies on a search system to take oriented first steps in a search process. Interaction models can be defined through the different ways in which a user thinks, prepares, and executes this task plan. According to Salton and Shneiderman [51, 64], this process assumes an interaction model that consists of specifying the query, receiving and examining the results, and may end there or, otherwise, reformulate the query and repeat the same process until the results are satisfactory. This is a simple interaction model, which conforming to Hearst [24], assumes that the user's information needs do not change throughout the search, the latter being in a cycle with the refining of queries until the desired result is achieved. O'Day and Jeffries [45] state that search results for an objective tend to generate new information needs. Therefore, search interfaces should allow the user to readjust his objectives and both reformulation and filtering strategies at any time during his search [24]. On the other hand, a user interface must track the user's steps throughout his session, including progress, filtering, reformulation, and intermediate or visited results. Thus, the interface will allow the user to follow different paths, triggered by some unanticipated results, since good usability is applied, being the way to return to the previous status visible and direct.

The search process contains two types of primary activities: searching for information and the consequent analysis/synthesis of results. For Hearst [24], user interfaces must allow these two activities to interrelate. Despite the diversity of systems, types of search, and adaptability to different user experiences, some strategic components in the interfaces' design are identified in most research systems, which are described in detail in Hearst's text, listed as following: starting points, query specification, context, using relevance judgment, and interface support for the search process.

Starting Points

The start of the search process must be intuitive for the user. Therefore, the interfaces should assist the user in the beginning, providing options for a good search from the start. Users introduce small queries in the first phase of the session [28]. They analyze the results and then try to modify these queries in an incremental feedback loop [7].

The definition of the user's idea about the desired information is variable and also varies throughout the search experience. Several ways designed to assist the first steps are possible to identify [24]. Typically, systems require the user to scroll through a list of collections accompanied by the name and a small piece of information to accompany the entry. Others present a wide range of textual collections, with the aim of the user studying an overview of the results' content. These overviews can be of a high level, as when they are presented by categories to structure and facilitate the user's starting point, where he can find references to the topics of interest. Other overviews can be derived automatically, using clustering techniques that orga-

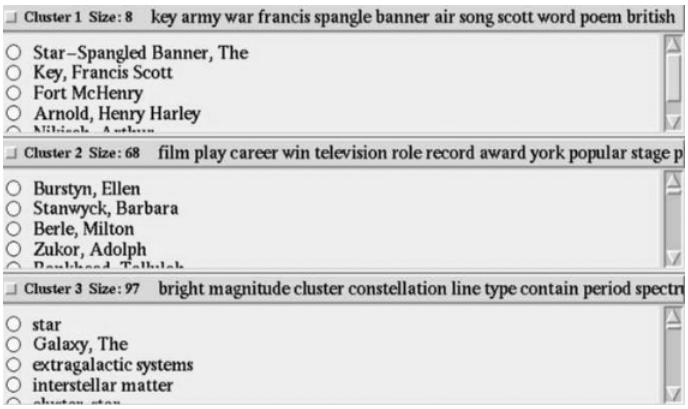


Figure 2.5: Display of Scatter/Gather clustering retrieval results (1992) [14]

nize documents into groups based on their similarity (clusters) and concluding on their themes (centroids). Hearst and Peterson [25] suggest that this method is more effective when integrated with conventional search technologies, applying clustering to the results of a query, organizing them, and directing the user to more relevant sets, shown in Figure 2.5. Finally, in addition to many other presentations, for example, more graphic (piles of books, bookshelves), overviews can also be presented based on patterns in co-citations, applying a variation of co-citation analysis, and trying to match articles or pages regarding the sharing of commonalities [24]. From this approach, it is possible to extract dominant themes and central authors in a field.

Users can be helped to start their search process with an example of interaction with the system, a technique called retrieval by reformulation, which is very important in systems that infer information needs [78]. However, users generally feel that being guided by examples they then have to modify is tedious. A dynamic variation of this aspect is to use interactive dialogues instead of examples. The most well-known and restricted form of dialogue is the so-called wizard, a tool that helps users in tasks regarding efficiency, reducing the user's actions to those that are essential [47].

Query Specification

In the search process, to reach information, the user must specify words or phrases that describe the information you want to find in what is called 'query'. Shneiderman [63] identifies five human-computer interaction types: command language, form filling, menu selection, direct manipulation, and natural language. Unlike modern information access, the systems only supported Boolean queries. According to Hearst [24], these are problematic since the syntax is counterintuitive, suggesting inexperienced users to misinterpret operators like AND and OR. For the author, most users are not familiar with the languages, whose complex syntax is a requirement for the use of other connectors. Web search engines needed to get around this problem, and from there came the search for loose words or phrases without the need to use less intuitive operators. Mechanisms are then applied to combine the terms in conjunction ('all the words') or disjunction ('any of the words'). Another of the problems with Boolean queries is that they

do not provide any early feedback on the results to be returned, so users may experience empty results when they restrict data too much (many joint terms) or wide results when many disjoint terms are specified [24]. Faceted queries, seen in Figure 2.6 have become an approach used in several systems to simplify queries into simpler topics and, in parallel, present a value relative to the number of entries to be given in case of selection of these topics, allowing management of the size of the result set during a query specification.

Figure 2.6: Example of faceted queries (2009) [59]

In other systems, the change ends up replacing command lines with forms, menus, and check-boxes, making the user preferentially perform selections instead of specifications. The graphical way the interface is built allows making explicit many of the tools that users would never explore using a command-line interface. Regarding visual approaches, interface manipulation consists of another alternative, based on three principles [63]: continuous representation of the object of interest, physical actions or button presses instead of complex syntax, and rapid incremental reversible operations whose impact on the object of interest is immediately visible. Interfaces generally give the user extra motivation, as they are easier to use in various contexts, providing more accuracy and speed. Examples given by Hearst [24] are interfaces based on Venn diagrams with rings as terms and interceptions as conjunctions; filter-flow models inspired by the inverted flow of a river depending on the filtering of the documents; block-oriented diagrams, where rows and columns highlight operators; and magic lenses, which, associated with words/categories, go over documents in a pool, filtering and presenting those that are relevant.

Context

After specifying the query, the next step in the search process is the analysis of the results. The way the results are presented and related to the query is associated with the term 'context'. There are several interface techniques to make the result set more understandable, involving the linkage of the results to the requested query, as well as the presentation of descriptive metadata, results overviews, structuring the contents, and relations within the set [24]. The results are typically presented as entries with title, brief description, and metadata, called document surrogates, seen in Figure 2.7. To these it is added the ranking, or proximity to the query, of the result when the search algorithm is ranking-based. Search engines automatically generate these elements that accompany each result, usually extracting the first text lines from each Web page.

Figure 2.7: Document surrogates in a Search Engine Results Page (SERP)

In systems where each document's text is presented in full, it is useful to highlight the terms in the text that correspond to terms present in the query. This mechanism attracts the user's attention to the text's parts that address what they searched for. This represents an improvement in the efficiency of the search [33], and can be done through a colored background behind the word, as seen in Figure 2.8, and, optionally, the use of bold [19].

Some systems contextualize the results not by looking at each document individually but by relating them to each other, summarizing the content of the results concerning the subset of query terms contained by them. The tabular display is another method to show relationships between the returned results, allowing a range of connection forms through coordinated axes, relating metadata, and using colors, shapes, and sizes of icons [24]. Kammerer and Gerjets [32] proposed a grid interface for search engine results, as seen in Figure 2.9.

Less popular approaches have emerged over time, such as the presentation of results in the context of a table of contents, which can be manipulated, expanded, or contracted; assigning hierarchical metadata categories to results; using hyperlinks for organization, among others [24].

Using Relevance Judgments

Between the formulated queries and the user's information need, a semantic gap stands that led information retrieval systems to employ techniques to capture the subjective aspect of rel-

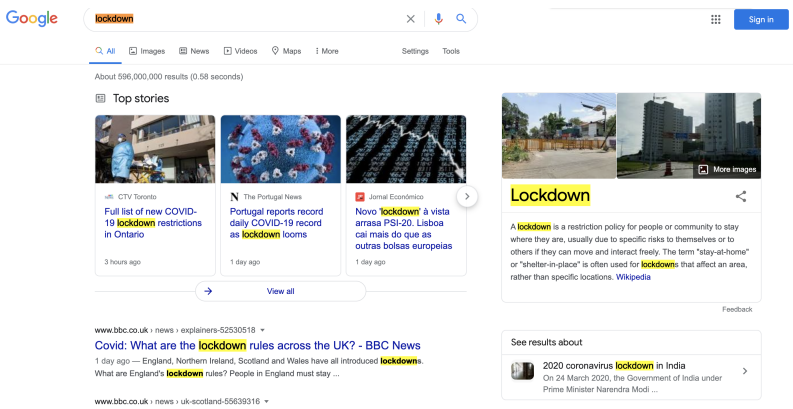


Figure 2.8: Highlighting system (text colored background) applied to Google's SERP

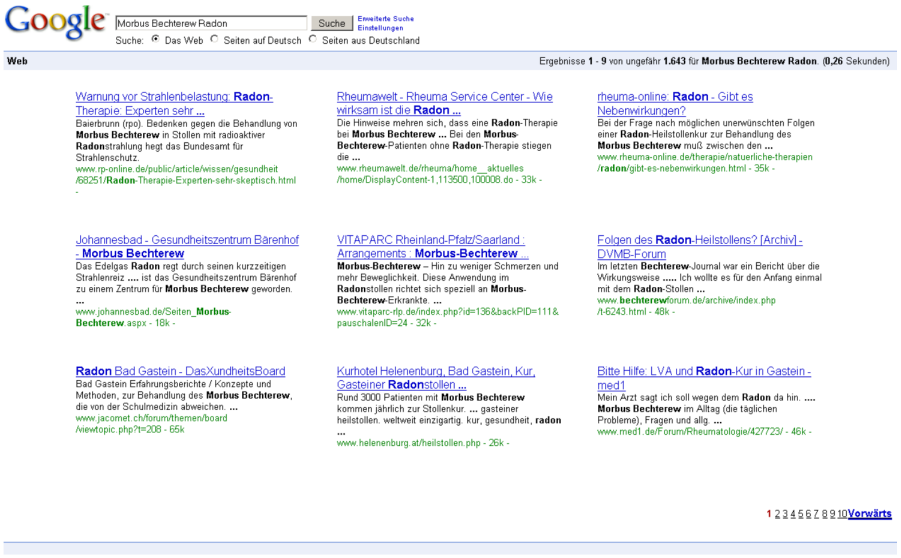


Figure 2.9: Grid layout proposal for SERP [32]

evance and improve the effectiveness of retrieved results [48]. Hearst [24] refers to relevance feedback as a proven effective technique, which takes the user to select a small set of documents that appear to be relevant to the query, and the system ends up deriving from that selection a new set of documents. For the author, in information retrieval systems, the control should allow the manipulation of the type of information returned, not in terms of the query used to modify this information. However, an opinion is left on how challenging it is to design interfaces to allow interaction in this way.

Relevance judgment is a complex task and requires a time-consuming effort. Allegretti et al. [4] point that to successfully destroy the ambiguity inherent in the user's information need, it is necessary to resort to an interactive and iterative process called the relevance feedback cycle. The authors describe that this process lists the user's explicit or implicit indications about relevant objects that the system processes to represent the user's needs better. After presenting the results, for Hearst [24], it is useful to contemplate the indication that a document has already been seen by the user, for example, using less vivid colors such as gray or even smaller sizes and fonts. To reduce this effort, the search engines adopted a terminology of "more like this", an example given by the author that lists some documents with some relation to the searched query and whose similarity may interest the user. This feature calls for a single user click, rather than feedback collection steps, and is still used today in Web search engines, being featured in the elements described in the following section.

Interface Support for the Search Process

Finally, and as a stepping stone to the studies in the next section, the importance of a well-structured interface and screen spaces arranged according to the best layouts is addressed, mainly when this idealization is directed to a complex activity as information retrieval. Hearst [24]

presents several themes related to different systems' interfaces, such as string matching, window management, example systems, and retaining search history, so those interested in any of the themes, in particular, should consult the full text. The text clears that it is essential to know how to limit the amount of information presented at one time on an information retrieval interface, not forgetting the significant space to be reserved for text area, capable of presenting text legibly. Web search results can be visibly very different depending on graphic details such as the font and spacing [40]. Hearst presents the concept of diagrams off monolithic layouts, as the basis for interfaces divided into left and right sides, usually associated with structuring and visualizing information respectively, to which are added components at the top and bottom of each side for search parameters, documents of interest, metadata, among others. Several other interfaces are presented. These must provide guidelines so that the user is helped start directly and towards success in the search, preventing errors, not forgetting to display the path already made and giving tips on further valid paths. There are several studies on the efficiency, advantages, and disadvantages of strategies around the components discussed in this section. Respecting Web search engines, there has been a standardization of which to use in most different engines. The next section looks at what the literature has already cataloged about their current interfaces.

2.2.2 Web Search Engines Interfaces

Search engines have slightly different ways of presenting information, which influences the experience and results selected by users. At this time, Google Search, from now on referred to only as Google, has the largest number of "sophisticated" results on its results pages, as it has always been one step ahead of the competition in implementing these elements [26]. Therefore, because the entire literature is committed to Google's elements, this analysis lists those same elements. As we will see in this work's results, many of these elements will also be present, implemented similarly, in other search engines.

Web Search Engines are divided into two main interfaces: homepage and Search Engine Results Pages (SERP), shown in Figure 2.10, being SERP the target of this study. The purpose of this section is to present the elements identified in the past that compose a SERP.

SERP contain the search query bar and the list of results. The search query consists of one or more terms entered by the user, inserted in an input text box so that the user can resort to

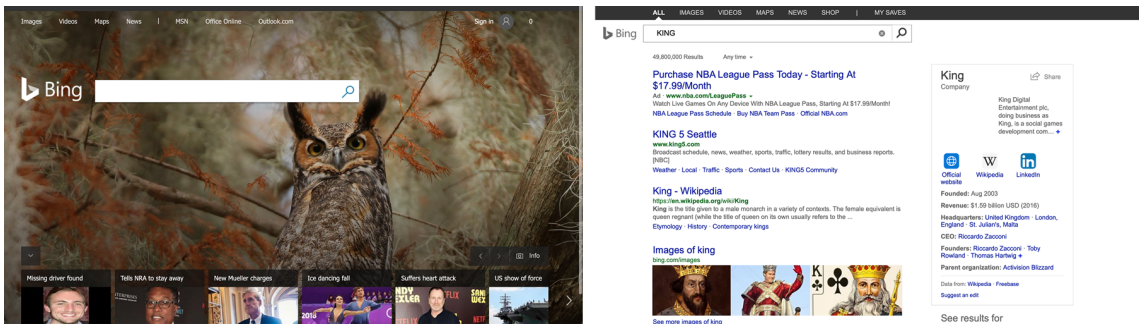


Figure 2.10: Bing homepage (left) and SERP (right), captured by Internet Archive (2017)

when beginning a new search or doing query reformulation. The results list contains the entries selected by the search engine and which are related to the submitted query. Each result points to a Web page identified by title (Web page title), a URL, and a brief description (date of publication, meta description), as suggested in Figure 2.11. Depending on the relevance of the results, as we will see later, other more enriched elements can also be presented, such as images, ratings, or rich snippets, described in this section.

Höchstötter and Lewandowski [26] define two visual areas on the results page: the visible and scrolling areas, also shown in Figure 2.11. The visible area consists of the initial part of the interface as far as the user can view information without interacting with the browser, and immediately after opening a specific page. Information beyond is considered to be in the scrolling area, as the user needs to slide the interface to reach the rest of the information.

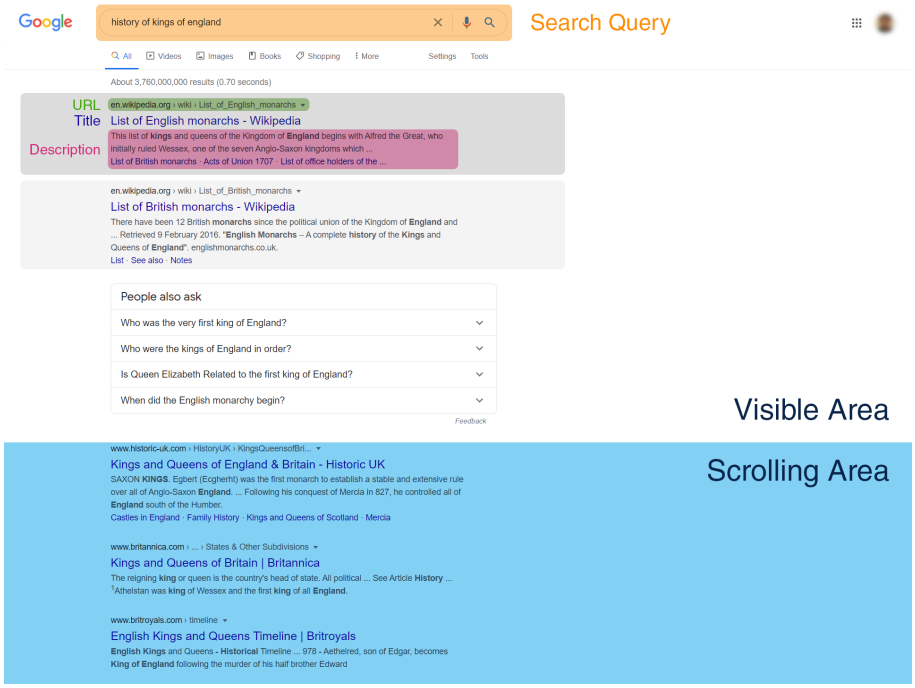


Figure 2.11: SERP components and areas - taken from Google (2020)

Regarding the results list, we can divide its content into three major categories: organic results, sponsored, and SERP features. These elements are tabulated in Tables 2.1, 2.2 and 2.3, and will be described and analysed in greater detail throughout Chapter 4. Introduction dates come from blog posts on *Google Blog*, *SearchEngineWatch* and *SeroundTable*.

Table 2.1: SERP Organic Results

Results	**Description**	**Appearance**
Regular	Usual entry, listing title, url and description	1998 [18]
Enriched	Sitelink, searchbox, breadcrumb or ratings and review	2006 [22]

Table 2.2: SERP Sponsored Results

Results	**Description**	**Appearance**
Textual	Similar to a regular result, with a label 'Ad'	2000 [21]
Shopping	Set of cards listing various information about each product (name, image, store, price, ratings, offers)	2016 [55]

Table 2.3: SERP Features

Feature	**Description**	**Appearance**
Featured Snippets	Answer box responding a question-related query	
Direct Answer Results	Instant answers with short and direct information	
Knowledge Panel	Dynamic panel with contents ranging from text to images, ratings, social profiles, among others	2012 [65]
Local Pack	Main geographical results positioned in a map	2014 [54]
Local Teaser	Similar to the Local Pack, focused on restaurants and hotels	
Image Pack	Set of images related to the query	
Video Pack	Set of videos related to the query	
Top Stories	Blocks of news considered relevant to the query	
Carousel	Line of several cards, with topic and image, presenting related results at the top of the SERP within a same category	2012 [53]
People Also Ask	Set of questions related to the query, commonly asked	2017 [58]
Related Searches	Set of suggestions for related searches	
Twitter Pack	Set of recent tweets related to the query	2016 [56]
Recipe Card	Set of cards linking culinary recipes	2016 [57]

2.3 User Interface Design Patterns

This section is focused on user interface design patterns, addressing their concept, motivations for their use, and listing several well-known examples useful for search engine interfaces.

A fundamental mission of human-computer interaction is to capture past design knowledge and best practices, so that one is not inventing the wheel in every project [61]. Thus, a design pattern is a means of capturing and disseminating this knowledge. Patterns are simultaneously concrete and abstract, which allows them to facilitate solutions to concrete design problems, but being sufficiently abstract to act in different situations, of which we will study, particularly in this work, the search engines interfaces. Patterns capture essential details, leaving aside what might make a problem overly specific, such as technology dependencies, and lose its abstraction capacity. In short, they are an invariant solution that addresses recurrent design problems in a specific context [17].

The use of user interface design patterns has not only become popular among software engineers but also usability engineers and specialists who are dedicated to building usable systems [61]. These have been considered as *lingua franca* for design as they cross cultural and professional barriers [60]. Standing out among the various benefits of using patterns is the intuitive way of documenting knowledge, the content being easily interpreted by designers, developers and other participants, the fact that they are the result of experiences with solid and not artificial knowledge, and, above all, the user-centered perspective within their rationale [61, 60].

Each pattern has three essential elements: a context, a problem and a solution. Therefore, it manages to address the specific context in which the situation occurs, describe the problem and constraints, as well as the intended objectives, and propose a solution with valid design rules for these constraints, applicable in identical situations [60].

Table 2.4 includes several user interface design patterns described by Jeniffer Tidwell [70] and online [72], which are useful for search engine interfaces. Each is accompanied by a short sentence describing the solution and an image example taken from the same source as the description.

2.4 Related work

This section will contain an overview of the literature related to this study. This work is positioned at the intersection of Information Retrieval (IR) and Human-computer Interaction (HCI) areas. It is not difficult to find works in the literature that address this same intersection. However, to the best of our knowledge, it is unlikely to find works specifically close to this one, seeking to collect interfaces from search engines over time and analyze their evolution. Some works address general interfaces in HCI, like the one from M. Hearst [24] that explains many of the essential components of an interface in a research system, contributing a lot to Subsection 2.2.1. There are also works related to the anatomy of SERP and works related to the state of SERP interfaces, mainly of a propositional and evaluative nature. The highlights are now described in the section.

Höchstötter and Lewandowski [26] address the composition of the Search Engine Results Pages (SERP) and count the appearances of the various elements. To the best of our knowledge, this was the first work to analyze the entire structure of the SERP. Besides, the authors examined the retrieved results, their sources, and types (e.g., organic results, advertisements, shortcuts). When authors wrote this paper, advanced features in SERP were not widespread, which was not the case when Moran and Goray [38] studied the anatomy of SERP, defining terminology for SERP elements. The latter is used in numerous articles regarding Nielsen Norman Group's main work related to usability in Web search. Nielsen's work [42] inspired Nunes et al. [44] when studying the interface evaluation of news websites, focusing on interface design, accessibility, and usability.

Nicholson et al. [41] compare the SERP's "editorial precison" by dividing the percentage of organic content (editorial) by the percentage of every content (organic and sponsored). Results show that 67% of a full page was dedicated to organic content.

Morville and Callender [39], in their 'Search Patterns' book, apart from addressing the anatomy of the search process and the related behavior, also list elements and principles of interaction design, describing many user interface design patterns around search websites. This book shows how search blends into everything one does online, making explicit the user behaviors that trigger the search and satisfaction with the retrieved results. Graphics and screenshots play an important role in conveing the authors points.

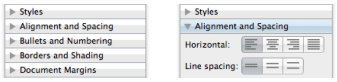
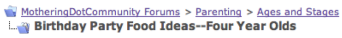
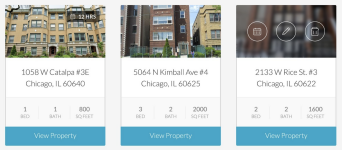
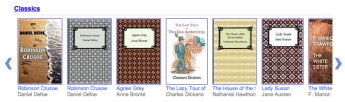
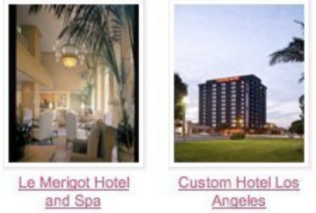
Mostafa Alli [5] proceeded with a propositional work, rather than descriptive, focusing on feature snippets and user query ambiguity. It proposes the creation of visual techniques to improve the user experience. The author claim to be the first to study a visual approach, having noticed that visuality and more appropriate textual snippets and titles would help enhance users' satisfaction with a SERP during an ambiguous search session. In this work, they have generated, with variations, new search snippets and titles with more appropriate feature visuals.

An article by Stella Tomasi [71] tried to study users' behavior based on different SERP interface formats, with a focus on visualization, to find what cues were most useful for users to find the desired information. The work involved the design of three interface prototypes to evaluate each one's strength depending on whether users followed or not their functionalities. Conclusions state that users prefer shorter suggestions than complex and longer ones, mainly website or category titles and keywords, results orientation, and animation. Another similar work by Arguello et al. [8] relates this task complexity to the type of results and how they are displayed in SERP, namely the test of two types of presentation: a page with singly vertical results and a page blending those with other categories of results. The analysis concluded that the user is more satisfied with the blended interface for more straightforward searches more often. The introduction of content from other SERP 'tabs' has been addressed in this work, and it will be interesting to check when its process started. The present work dates from 2012, with no evidence of whether this format already existed or it was an innovative proposal.

There is a work focused on the cognitive attraction of the Knowledge Graph, carried out by Monteiro et al. [37], on a sample of Google SERP interfaces. It consisted of applying an eye-tracking technique and assessed the user's attention throughout the SERP interface, depending on the type of results it presents. According to this article, the smaller the number of SERP features presented throughout results, the greater the attention will be on the Knowledge Graph. Results reveal the attractive role of image, although other cognitive factors may interfere, such as search sort and user's search engine usage expertise. There are other works with the application of a similar technique, evaluating other SERP components, as is the case of Buscher et al. [11] on advertising and related searches.

According to our research, it is a short set of related work, constituting a strong motivation for developing this work that will complement and update this area of the search engine interfaces in the literature.

Table 2.4: Sample of well-known user interface design patterns

Pattern	Description	Example
Accordion	Used to stack modules of similar content, permitting the user to open and close each module independently, freeing up space and allowing the user to access the content only if interested [70].	
Breadcrumbs	They linearly specify hierarchy levels leading to the current page. However, in SERP, Breadcrumbs are associated with content elements, not with the page itself, and may specify different types of hierarchies [70].	
Categorization	Provides the user with categories different from each other, helping suggest what content is to be found [72].	
Cards	Used to display content composed of distinct elements, normally about a single subject, to form one coherent piece of content. Thus, it is usual for cards to be closer to other cards. [72].	
Carousel	Consists of a horizontal strip of simple cards, letting the user to scroll horizontally to view them and encouraging the inspection of the items that are to come [70].	
Grid of Equals	Used to display items in a grid or matrix, each following a common template, linking to respective pages [70].	
Module Tabs	Used when content is groupable and there is no room for everything. Modules of content are divided in small tabbed areas with only one visible at a time, the user being able to click on tabs to reveal other modules [70].	
News Stream	Used to list time-sensitive items chronologically, combining the sources in one place [70].	
Responsive Disclosure	Used when the user is exploring information or proceeding with different actions. The interface is being revealed (disclosured) as the user proceeds [70].	
Thumbnail	Used when the user is provided with an overview of pictures without downloading each in full size [72].	

Chapter 3

Methodology

To study web search engines' SERP user interfaces, we used the methodology summarized at a high level in Figure 3.1.

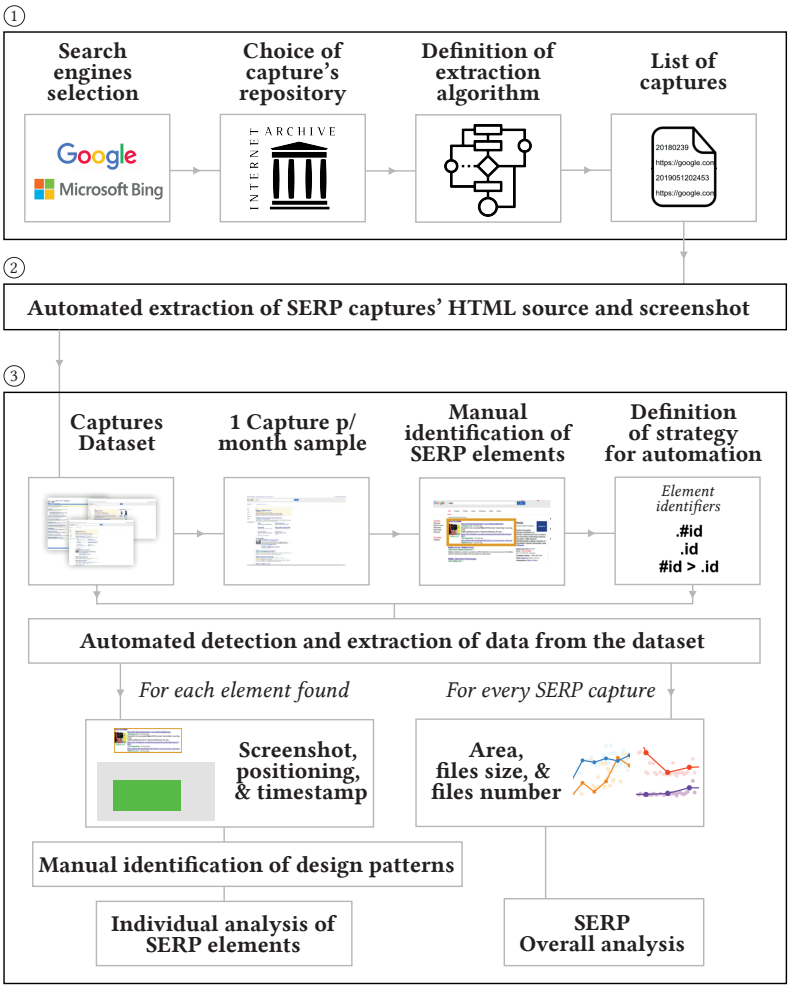


Figure 3.1: Overall methodology for each section

The process starts with the selection of the search engines to be analyzed, and to look for a repository that contains captures of its interfaces; tasks explained in Section 3.1. Thereafter, we created and applied a method to collect a set of these captures, described in Section 3.2. Finally, we studied a sample of the set to understand the global structure of the interfaces and list the target elements to carry out a systematic analysis on the entire set, which is detailed in Section 3.3.

3.1 What SERP have we captured?

There is a noticeable difference between Search Engines' market share [12]. We concluded that the extensive work analyzing search engines whose use is minimal (around 1%) was not justified. For this reason, we decided to focus our analysis on the first two classifieds: Google for its market share hegemony and Microsoft Bing to permit a comparative analysis. These two options are responsible for 95% of the worldwide market share. Although Web search engines are also available on mobile platforms, their interfaces are considerably different from desktop ones. This study will address desktop versions exclusively, from its beginning in 2000 until 2020.

The Internet Archive has been keeping snapshots and the respective HTML version of webpages over time. Its collection contains more than 50 billion webpages [34]. Internet Archive provides a standalone HTTP servlet, the *Wayback CDX Server API*, that allows complex querying, filtering, and analysis of captures. It enables, for example, to obtain the timestamp of captures in a 14 digit format, filtering captures by date (year, month, and day) and URL. While filtering by URL, we can use a wildcard (*) at the end of the URL to specify the latter as a prefix and also receive entries that go beyond the specified URL (e.g., *www.google.com/search?q=cookies**).

Using the API, we found hundreds of thousands of SERP captures during two decades. This large number of SERP and existing resource restrictions led us to devise a method to identify a smaller, yet representative, set of SERP. Initially, we collected a random sample of SERP over time, but results would retrieve captures that referred to non-usual queries, some of them with non-sense characters and most of them being strangely too restrictive. To assure SERP diversity, we have used a set of 129 most searched queries in the last 20 years[1]. These queries contain relevant terms often searched by users and, consequently, trigger features in SERP. Hence, it is highly likely that SERP interfaces derived from these queries are richer and, thus, more relevant for this study than those generated by random searches. We decided to append these queries with the '*' wildcard while submitting them to the API to obtain a larger amount of captures.

Using the most searched queries, we noticed that some years had no captures. In fact, for those years, there are few captures from the generic search domain. Hence, in those years, as when Table 3.1 mentions the *all* method, we collected all the available captures. In Google's case, we also noticed that the two last years had a much larger number of captures (>10 thousand). Therefore, in 2019 and 2020, we restricted the URL submitted to the API to those without queries longer than 37 char, the equivalent of 6 words, with an average length of 5 char according to

[1] Available at: https://bedgarone.github.io/serpevolution/mostsearchedqueries

Norvig [43], and spaces between them. This restriction excludes longer and more specific queries that are probably less useful to the plurality of interfaces. Our dataset has 5.653 captures from Google and 2.267 captures from Bing. Table 3.1's last column consists of ordered lists with the number of captures per year.

Table 3.1: Dataset extraction information

Google	**Method**	**Max. length**	**Screen width**	**Captures**
2000 - 2002	queries	-	800px	200, 3, 23
2003	queries	-	1024px	231
2004 - 2008	all	-	1024px	12, 0, 200, 0, 26
2009	queries	-	1024px	11
2010	all	-	1024px	78
2011	queries	-	1024px	7
2012 - 2015	queries	-	1366px	57, 975, 30, 89
2016 - 2018	queries	-	1366px	172, 192, 548
2019	queries	37 char	1366px	171
2020	queries	37 char	1920px	2628
Bing				
2010 - 2011	queries	-	1024px	71, 5
2012 - 2015	all	-	1366px	30, 38, 41 , 62
2016 - 2019	queries	-	1366px	9, 138, 143, 132
2020	queries	-	1920px	1598

3.2 How have we captured SERP?

We used Python and Selenium Webdriver, for browser automation, to visit each capture online, check if the capture is valid, save the HTML version, and generate a full screenshot. This is shown diagrammatically in Figure 3.2.

Not all the captures listed by the API were considered valid, even though being labeled with an OK status code. Some are inexistent, with a contradictory message of URL not captured, and some are defective (e.g., showing incomplete interfaces without search results). To automatically assess the validity of each capture, the program tries to find an organic result, the element that cannot lack in a Search Engine Results Page (SERP). Captures from other SERP tabs other than the general first page, identified with the substring "tbm" in the URL, were also discarded for being outside of this work's scope. A timeout exception is raised after 6 seconds, which means the program will skip that capture, which was the time that was empirically considered sufficient for a valid capture to be fully loaded. Before downloading the page, we still have to remove graphical elements from Internet Archive, such as its information and donation bars, like in Figure 3.3. Some captures present other distracting banners, overlapping parts of the interface, such as the ones related to cookie consent. The ones we were able to identify have been incorporated in the removal process.

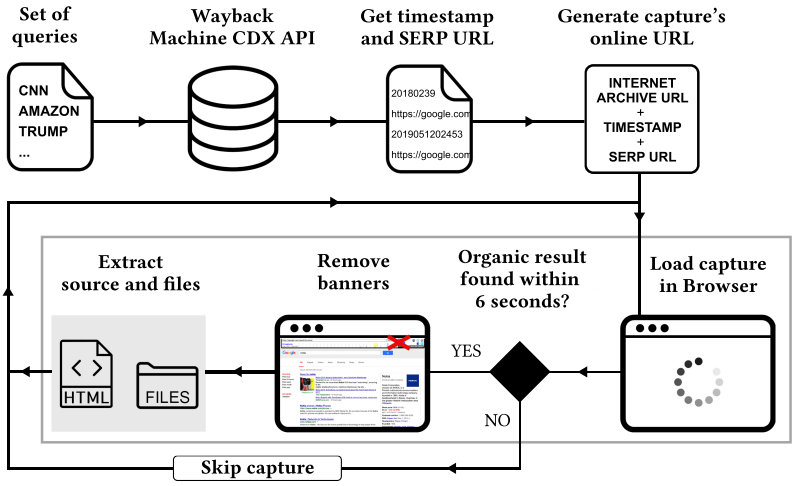


Figure 3.2: Extracting captures procedure

The existing methods can only access the webpage source code. Thus, to download the source code and the respective files, we used an external library *pyautogui* to simulate the keyboard keys to trigger the browser's save function.

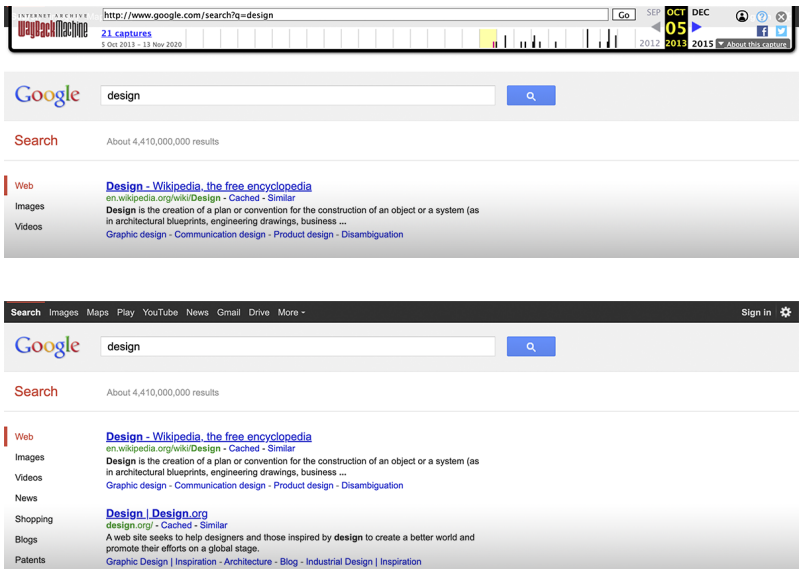


Figure 3.3: Cropped capture with (top) and without (bottom) distracting banners

The process concludes with the generation of full-height screenshots of every HTML version, opened in another browser instance, in headless mode. Because SERP are responsive to the screen's size, we produced screenshots considering the most popular screen size at the time of the capture. We only considered the width measure because SERP height is highly variable. According to screen size usage throughout the years [69], seen in Figure 3.4, we used the browser widths shown in Table 3.1.

The dataset contains all the extracted captures. Each capture is represented by a screenshot,

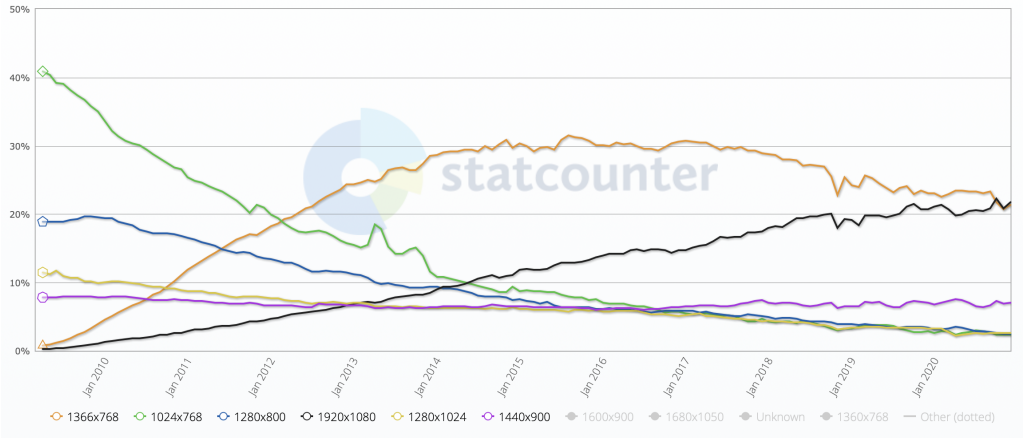


Figure 3.4: Evolution of screen size usage (since 2009) [1]

an HTML file, and a files' folder. We concatenate the initial of the search engine (G or B) with the capture's timestamp for file naming. The filename ends with a sequential integer "-N" if the timestamp is repeated. For example, 'G20070330145203-1' identifies a second capture from Google by March 30, 2007. The first is identified by 'G20070330145203'.

3.3 How have we analyzed SERP?

The analysis process included two main stages, as shown in Figure 3.5. In the first stage, we have extracted from the main dataset a sample of captures to identify SERP elements. We selected the capture with the most features for each month by manually looking at the screenshots of that month's captures. For each element, we analysed its source code looking for identifiers that could locate the element in a later automated process. Since elements' code was frequently changing, it was not possible to list every single identifier that elements might have had, but all the identifiers encountered were logged and are listed in Appendix A. Element identifiers consist of HTML classes, ids, tags or a combination of these using CSS selectors (e.g., featured snippets: *'#knocube'* or *'#res .hp-xpdbox'*).

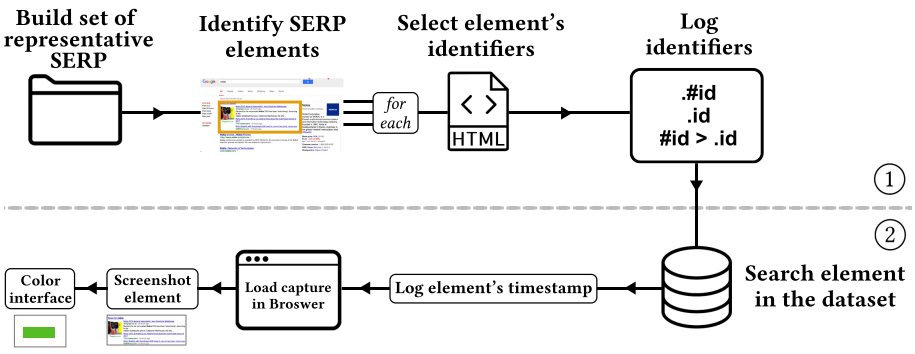


Figure 3.5: Detection and analysis of elements procedure

A brief example on how to associate each SERP element with a code identifier is demonstrated now, through the "King's College Cambridge" search on Bing.

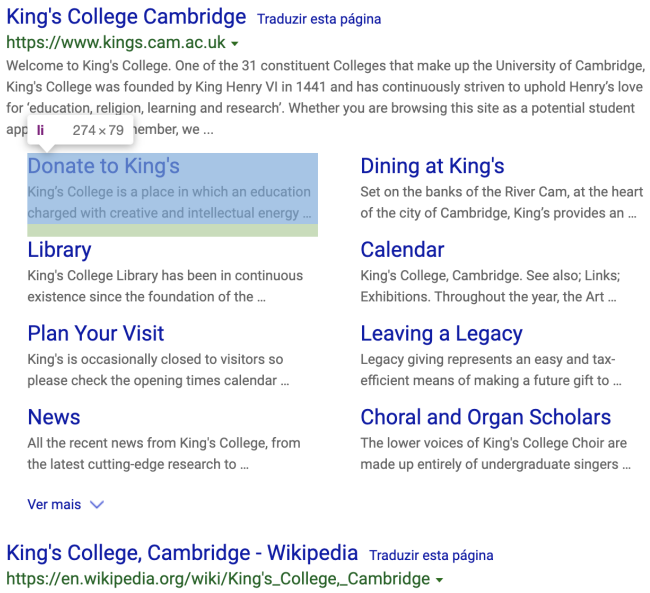


Figure 3.6: Sitelink of an Enriched Result - taken from Bing

Inspecting a Bing SERP's source code, it was possible to extract the code of one sitelink, highlighted in Figure 3.6, embedded in the first organic result, an enriched result. From the code of this element we can detect the class that characterizes it *deeplink_title*, which, after checked for similarity in other captures that have sitelinks, will be labeling this element during the captures' scan.

```
<h3 class="deeplink_title">
    <a href="https://www.kings.cam.ac.uk/members-and-friends/support"
        h="ID=SERP,5333.1">Donate to King's</a>
</h3>
```

Listing 3.1: Sitelink HTML code

The second stage refers to the automated process, that scrapped throughout the entire dataset, finding elements with their identifiers triggers. When an element is found, the capture's timestamp is stored in a log file that keeps a record of the time of the element's appearance. Another function receives the coordinates of the element's upper-left corner, and its width and height, and locates the element, generating and saving an image of it in the element's folder. Due to the large scale of the dataset, we restricted the image generation to 15 images per month. However, it is possible to change the approach and skip images generation, just to store all the elements' timestamps in a file. In this case, no limit of captures per month is needed as the computation permits a full dataset scan in an acceptable time.

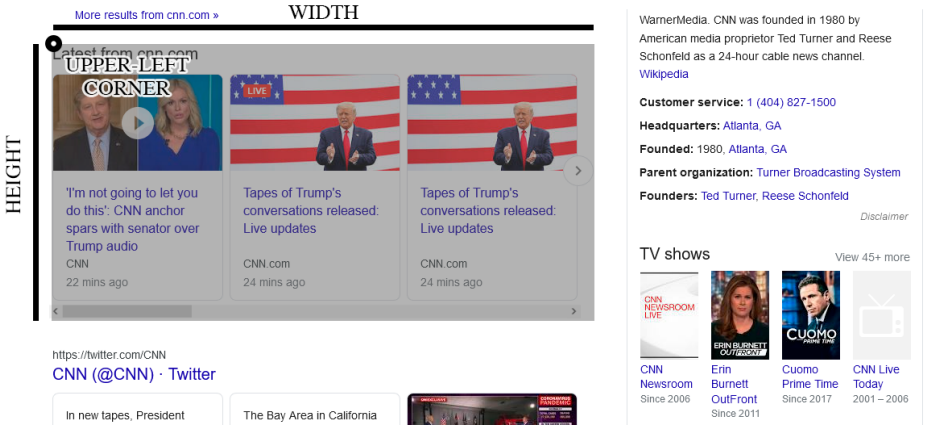


Figure 3.7: Capturing a specific element using its coordinates and size

Another automation was also applied when each element is found, to coloring interfaces according to the following element categories, described individually in Section 5.1: **Welcome & identity** , **Search statistics**, **Navigation & user inputs**, **Organic Results**, **Sponsored Results**, **Features**.

Following a similar procedure, the list of identifiers allowed to automatically detect and color targeted areas of the webpage. We used Python, Selenium Webdriver, and BeautifulSoup to scrape every single HTML capture in colored phases to identify and generate transparency-colored images. Due to the size of the dataset, we imposed a limit of 15 elements/captures per month, as we did in the extraction of elements' images. The end result of each category is an image that overlays all the individual images from single captures. The high level of transparency will enhance the most common areas at the end while leaving the others almost unnoticeable. The overlaying process uses the upper-left corner as the reference for image alignment. However, the *navigation & user inputs* category includes elements in and next to the footer, but these common areas were not evident in the result due to how variable the height of the captures can be. In this case, to reproduce the overlapping of every bottom of the interface, the algorithm had to consider a height value of N, cropping the results from the initial pixel (original height minus N) to the lower-left corner (original height), maintaining the original width. Thus, the end result displayed in Figure 5.2 is trimmed at the middle and should be seen as two separate results, as Figure 3.8 suggests.

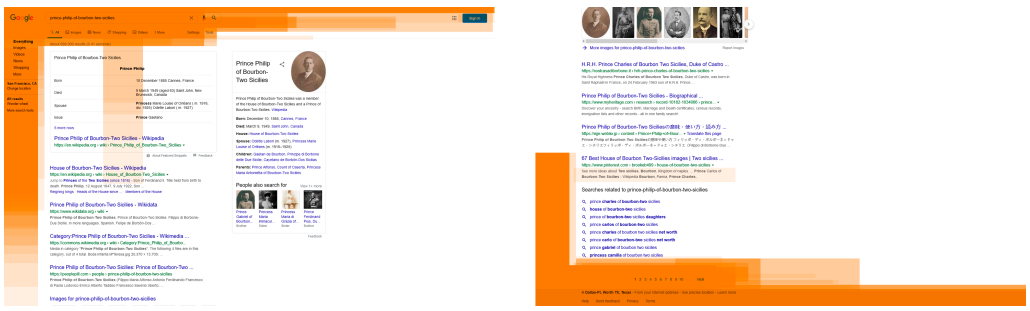


Figure 3.8: Navigation's footer demanding a bottom-aligned approach

To end the work with a more detailed contribution, the identifiers used in the detection of each element are also disclosed, which will be useful for future work on automating this detection and analysis. As such, a last cycle through all these identifiers was done to document the periods in which each identifier was associated with a certain element. This way, the identifiers table will associate each one to each element, having a third domain which is its time interval.

Finally, to entrench an overall analysis, we developed Python programs to scan all the dataset and generate graphs on interface areas and source codes' file sizes over time. The results from each analysis described above are detailed individually, referring to each search engine, and in comparison when applicable.

Chapter 4

SERP Elements Analysis

In this chapter, we address the evolution of all Search Engine Results Page (SERP) elements. Each element is addressed individually, and its analysis compares both Google and Bing, when applicable, and is structured as follows: **images** - showing some stages of the element's visual evolution; **definition** - stating the element's purpose; **content** - describing its composition over time; **design** - analysing graphics, text styling and colors; **positioning** - stating, when applicable, a common location in SERP layout; **design patterns** - indicating the application of certain user interface design patterns; **interface area** - analyzing the evolution of the element's area; **lifetime** - stating the time when the element was found present. Each element has a dedicated page at: https://bedgarone.github.io/serpevolution/elements.

Although our first appearance dates might be different from the specific dates on which search engines introduced these features, we expect these not to be considerably different. Besides, our goal is to provide an approximate date for these events.

4.1 Organic Results

Organic results are links and short descriptions related to the most relevant Web pages to the search. For each query, these results seek to satisfy users' search intentions. They are ranked based on relevance only, and the associated snippets are generated computationally through summarization algorithms [15].

Unlike all other types of results described below, organic results are mandatory elements in any Search Engine Results Page (SERP). Organic content is retrieved according to the crawling and indexing process. No one can pay to have a website listed as an organic result. Organic results are divided in regular and enriched results [49].

4.1.1 Regular Results

Regular results are the SERP basic entries, seen in Figures 4.1 and 4.2.

In Google, the **content** of regular results started as a basic block of a title, description, URL links for *similar* or/and *cached* pages for the result. The latter leads the user to a capture of the

result's source page, snapshotted by Google in the past. In 2013, these links were hidden in a dropdown, only visible by its arrow icon until now, while the source URL was placed below the title. From 2018, a link to translate the result was introduced. In 2019, the URL started to be displayed as a breadcrumb. Bing's element share almost every deatil, but it didn't present the *similar* option, and it kept the URL's place, while Google's one started to appear above the title from 2020.

As for Google's **design**, text styling was always applied in this element to enhance relevant terms in bold. The dominant colors are blue for title, green for URL, and black for description. The title was underlined until 2014, one year later for Bing. In 2020, the URL changed its color to gray, while Bing has kept its green color.

Its **positioning** is highly variable throughout the results container, for Google and Bing.

In Google, the *Breadcrumbs* **design pattern** is applied to this element since 2019. There is no design pattern applied in Bing.

Interface area was constant during Google's time range, whereas Bing's case suggest a slight increase starting from 2016. In the case of regular results, similar regardless of the engine, element area may be differentiated because of how long descriptions are before truncation.

Being a mandatory element in SERP, this element is **present** since the beginning of both search engines, lasting uninterruptedly until now.

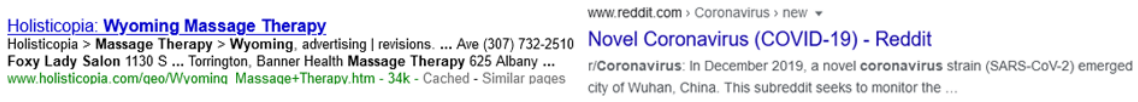


Figure 4.1: Google's regular result from 2003 (left) and 2020 (right)

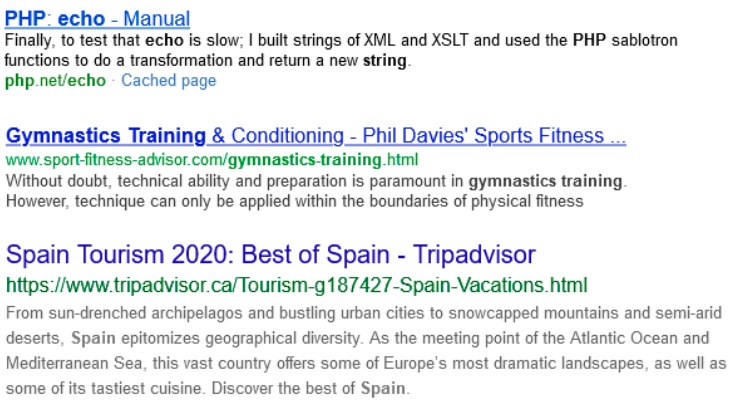


Figure 4.2: Bing's regular result from 2011 (top), 2013 (middle) and 2020 (bottom)

Complete information on regular results available at: https://bedgarone.github.io/serpevolution/elements/regular.

4.1.2 Enriched Results

Enriched results, seen in Figures 4.3 and 4.4, are organic entries similar to regular ones, with some elements to enhance the result. These elements improve user experience and give extra information to the user. Enriched results can have greater visibility and, in turn, a higher click-rate [49].

In Google, the extra **content** consisted initially of two sitelink columns, pointing to subpages of the result's domain. In 2010, each sitelink was enhanced with a short description, and in 2016, a search bar was introduced so that the user could search the result's website directly from SERP. Bing's sitelinks had no short description until 2014. However, from 2019, compared to Google, Bing's content has more extra components placed after the sitelinks. These can consist of a row of cards with image and title, a section with module tabs or a list of trending posts.

Design is coherent to *Regular Results* in both engines: in 2015, titles were no longer underlined. In 2020, Google's source URL also changed its color to gray and placed above the title.

In both cases, its **positioning** is consistently at the top of the results container.

In Google, the *Breadcrumbs* **design pattern** is applied to this element since 2019. In Bing, the *Cards* and *Thumbnail* design patterns are applied since 2019, whereas *Module Tabs* since 2020.

Google's case shows a constant **area** growth over time, with a considerable increase in 2014, when short descriptions where included. Bing presents a greater growth due to a higher completeness of this element nowadays.

The enriched results appeared for the first time in Google in 2008, continuing in 2009 and from 2012 to 2020. In Bing, it was **present** in 2010, 2014, 2015 and from 2017 to 2020.

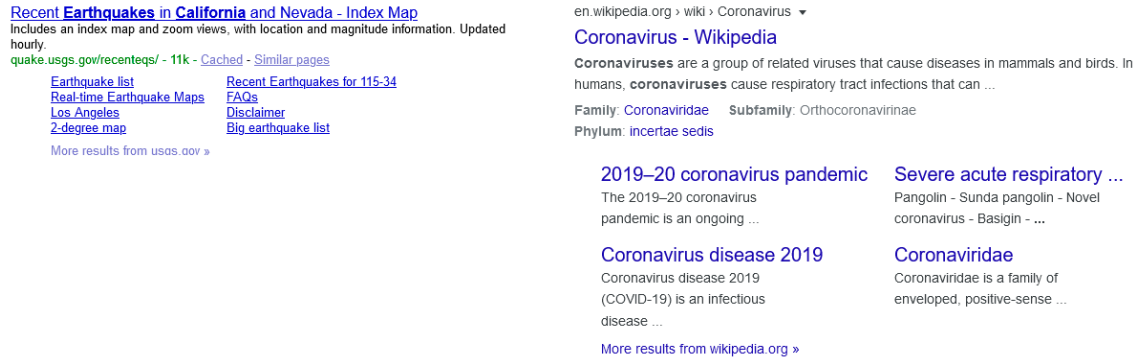


Figure 4.3: Google's enriched result from 2009 (left) and 2020 (right)

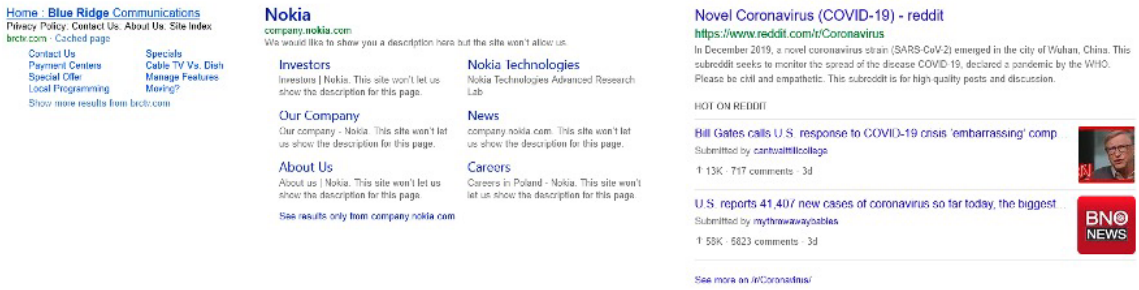


Figure 4.4: Bing's enriched result from 2010 (left), 2017 (center) and 2020 (right)

Complete information on enriched results available at: https://bedgarone.github.io/serpevolution/elements/enriched.

4.2 Sponsored Results

Sponsored results are short advertisements arranged alongside organic results, distinguished by a specific tag. These entries are more focused on commercial intent, combining relevance with revenue, and are usually manually crafted, a standard in the advertisement industry [15]. Sponsored results are a constant element in Search Engine Results Pages (SERP), being the most significant source of income for search engines [20]. These are presented in textual form or in a shopping carousel, with the addition of enriching elements, with images and reviews [38].

4.2.1 Textual Ads

Textual ads, seen in Figures 4.5 and 4.6, are counterparts of *Regular Results*, dedicated to sponsored content and are present in SERP since the very first day. Some ads placed in the right sidebar are shortened in width.

The **content** of this element and its evolution is identical to regular results, but marked initially with the tag 'Sponsored link' for Google and 'Sponsored sites' for Bing. From 2012, and one year later for Bing, some of the results would present, when applicable, simpler and less noticeable sitelinks than *Enriched results*. In 2020, some ads could have the shape of an actual *Enriched result*.

In early versions, regarding **design** until 2014, in both engines it would have a different background color to better distinguish the element from organic results. Scarce Google occurrences in 2017 also had a colored background. In 2014, the tag 'Ad' substituted the previous one. In Google, this tag began with a yellow background color, while from 2016, it was green. Contrarily, Bing's tag was always gray. Google's source URL also changed its color to gray and its positioning to above the title, as usual.

In both cases, textual ads' **positioning** is common at the top of the results container, at the bottom of the results container and at the top of the right sidebar.

No **design patterns** are applied to this element in both engines.

Textual ads have had a constant **area** size over time in Google's case, having increased slightly in 2012, before stabilizing. Contrarily, Bing's textual ads show considerable increase in area beginning in 2019.

In both engines, this element is **present** since the beginning, being constantly present until 2020, with scarce exceptions in Google in its early years.

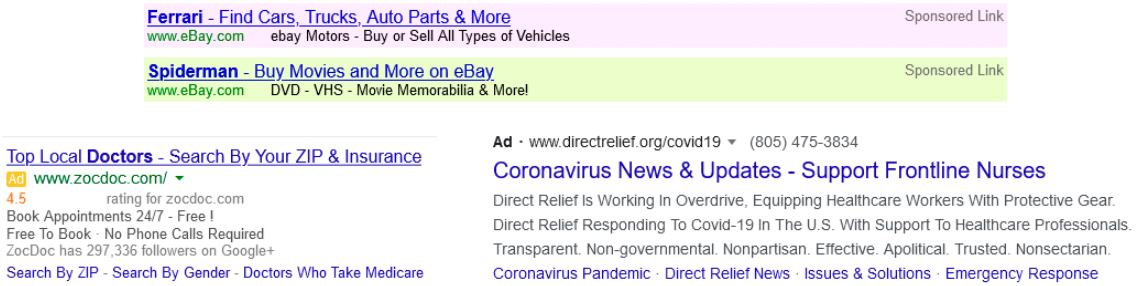


Figure 4.5: Google's textual ads from 2002 (top), 2014 (left) and 2020 (right)

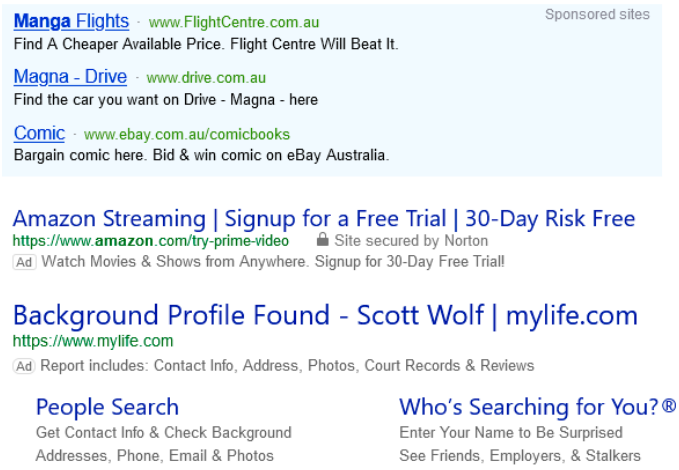


Figure 4.6: Bing's textual ads from 2010 (top), 2018 (middle) and 2020 (bottom)

Complete information on textual ads available at: https://bedgarone.github.io/serpevolution/elements/textualads.

4.2.2 Shopping Ads

The other type of ads, only found in Google and seen in Figure 4.7 - bottom, correspond to shopping content and are therefore called shopping ads. These can occur at the top or top-right of the SERP and are activated when the query is intended to be commercial [49]. These are very striking results, listing various information about each product (name, store, price, ratings, offers).

Shopping ads used to be exclusive of the right sidebar, where **content** consisted of displaying one to four results in a matrix. In 2018, each result was embedded in an individual card. In 2020, this element started to appear in the results container, with more width and less height, in a carousel of cards.

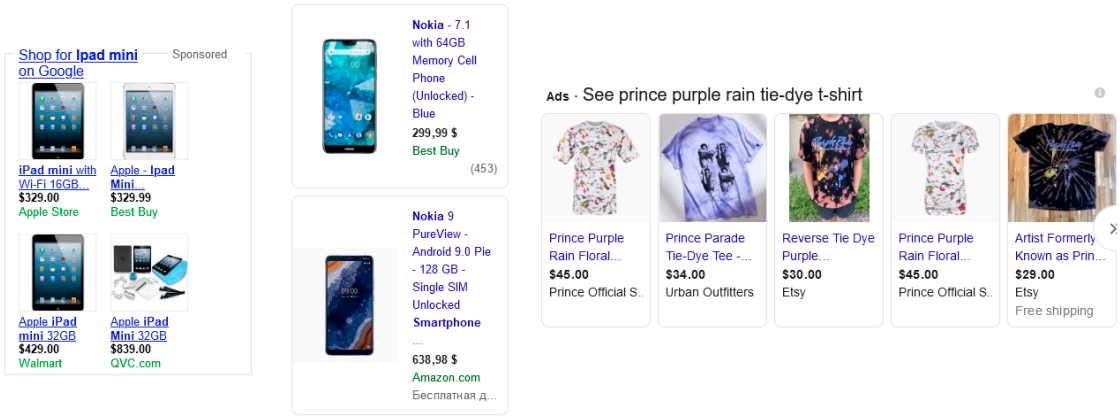


Figure 4.7: Google's shopping ads from 2013 (left), 2019 (center) and 2020 (right)

As for **design**, text styling was used until 2019 to enhance relevant words in bold. From 2020, the source's URL changed its color from green to gray, as usual.

Its **positioning** is usually at the top of the results container and, more frequently, at the right sidebar.

The *Cards* **design pattern** is applied to this element since 2019 and the *Carousel* since 2020. The *Thumbnail* design pattern is applied here since the element's beginning.

The element's **area** was constant until a recent increase since being recently displayed horizontally in a carousel.

The **shopping ads** appeared for the first time in 2013, lasting until now.

Complete information on shopping ads available at: https://bedgarone.github.io/serpevolution/elements/shoppingads.

4.3 Features

The interfaces have become more and more sophisticated, with the introduction of new elements, capable of making the user's search experience more direct and intuitive. In general, SERP were made up of organic results and advertising. In the second decade, articles such as Peter Meyers' [36] in 2013, begin to document a third category, features in Search Engine Results Pages. SERP features complement organic and sponsored results, and emerged because people process graphical information more easily [26]. New features include custom and synthesized content that attract the focus from the top rank organic results, attempting to provide answers to the query without just pointing to websites that might deliver that information.

Rosu [49] reinforces Google's effort to launch this type of features whenever it is possible to give the desired information the user without him leaving the SERP. Google, and other search engines, do it according to various features, which we now analyse.

4.3.1 Featured Snippets

Featured Snippets, seen in Figures 4.8 and 4.9, are answer boxes in which Google responds to a question-related query based on information taken from a page [75]. An algorithm determines the page that contains the most relevant information, returning a summary answer in the form of a paragraph, list, or table [38]. A video may also accompany it.

Content was initially made of a short paragraph with answering information for google. It evolved to a general layout, used until now, consisting of a larger paragraph, a thumbnail at the upper-right corner, and the title and link to the information's source, to where it is possible to navigate. The answer started also to be returned in the form of an ordered list or table. Bing launched a simpler version of this layout with paragraph and source, only. In Google, instead of an image, a video or a carousel of images may also accompany it. In 2018, when possible, the date of the source's publication was introduced after the information paragraph. In Bing, from 2020, the paragraph is just one of the options, as the content can now be displayed in a bullet list or accompanied by a carousel of images.

As for **design**, Google's element removed the underlined title in 2017 and changed the green URL to a gray breadcrumb URL in 2020, in accordance to the organic and sponsored results design. In Google and Bing, text styling is used in both paragraph and title, enhancing relevant words in bold.

In both engines, its **positioning** is consistently at the top of the results container.

In Google, the *Thumbnail* **design pattern** is applied to this element since 2016. In Bing, the *Carousel* and *Thumbnail* design patterns are applied to this element since 2020.

Google's Featured Snippets have a constant and smooth increase in area size over time, while Bing's increase occurs later in recent years.

Featured Snippets **appeared** for the first time in Google in 2016, and one year later in Bing, lasting until now.

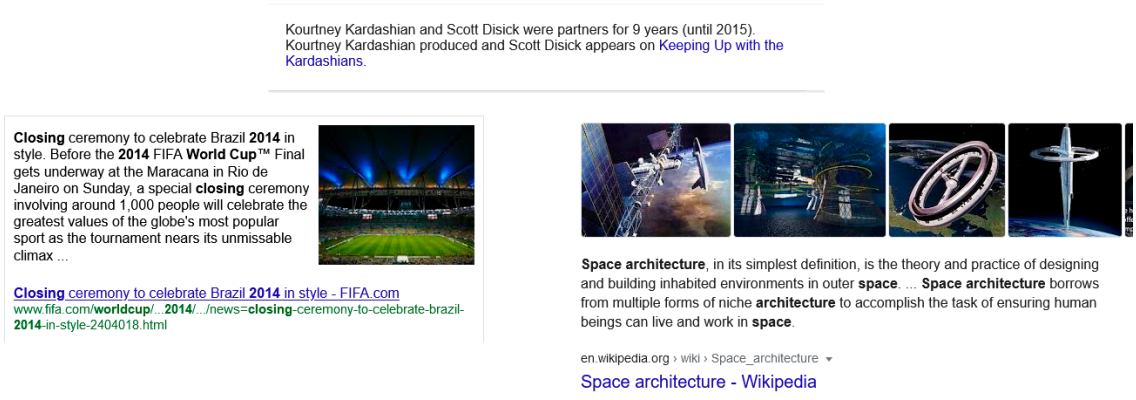


Figure 4.8: Google's Featured Snippets from 2016 (top), 2017 (left) and 2020 (right)

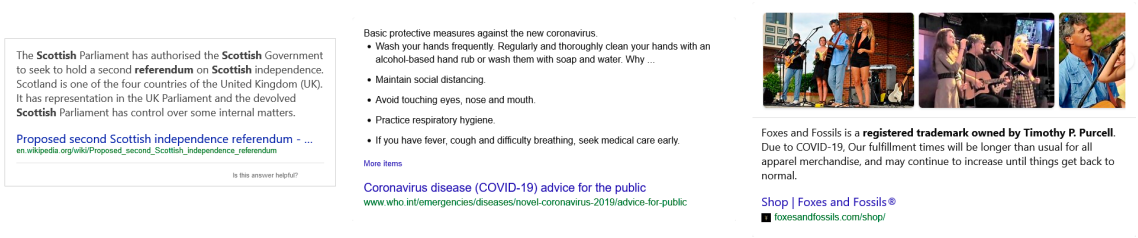


Figure 4.9: Bing's Featured Snippets from 2017 (left) and 2020 (center and right)

Complete information on featured snippets available at: https://bedgarone.github.io/serpevolution/elements/fsnippets.

4.3.2 Direct Answer Results

Direct Answer Results, exclusive to Google and seen in Figure 4.10, are instant answers with short and direct information to answer the user's question at the beginning of the SERP. It is considered public domain information, without credit to a particular source [49]. These answers can take different forms, such as date, weather, flights, sports scores, and results, scholarship, among many others. Some of these cases also allow interaction with the user, such as currency/unit conversion, translator, or calculator.

Being a recent element, with less than three years live, changes are slight. The basic **content** structure is a breadcrumb with categories and a row with a few words, the answer, but a thumbnail is normally present. In 2020, a timestamp of the information was introduced in this row. A second row at the bottom can be dedicated to a paragraph with complementary information, or to 'People also search' suggestions, each made of a smaller thumbnail, a hyperlinked title, and possibly a timestamp.

Design and **area** didn't change over time.

Its **positioning** is always at the top of the results container.

The *Breadcrumbs* and *Thumbnail* **design patterns** are applied to this element since it's beginning.

This element appeared for the first time in 2018, lasting until now.

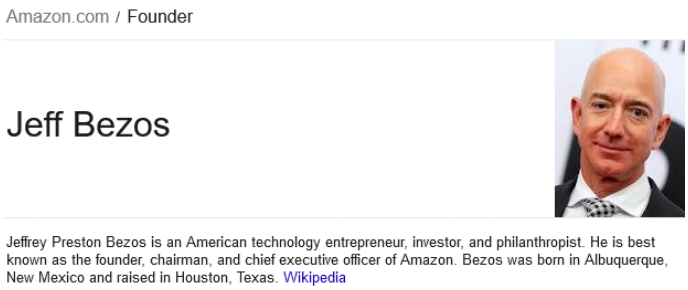


Figure 4.10: Google's Direct Answer result from 2018

Complete information on direct answer results available at: https://bedgarone.github.io/serpevolution/elements/directanswer.

4.3.3 Knowledge Panel

The Knowledge Panel, seen in Figures 4.11 and 4.12, is a dynamic feature that provides direct information in various formats within the same panel, pointing to related content [49]. For Sullivan [16], this is an easily recognized tool designed to help the user quickly understand a particular subject, facilitating a more in-depth search.

Content in Google's case consists of a panel with a top thumbnail of the subject, vertically followed by a title, website link if applicable, a resume paragraph normally by *Wikipedia*, a structured list of direct informations and a block of *People also search for*. During the following years, extra content was being introduced dynamically and highly dependant on the subject associated, such as Module Tabs, related people, CD albums, trophies, and huge variety of possibilities, proving the high variability from topic to topic. Bing's knowledge panel is highly similar to the Google's one, consisting of a panel with a top thumbnail of the subject, vertically followed by a title, some times a resume paragraph by *Wikipedia*, and a structured list of direct information, which can also be followed by dynamic and variable possibilities, in which Bing seems to include timelines whenever is possible.

The **design** of Google's element was stable over time, being released when Google had already abandoned the old and blueish lines of design. Text styling was always applied in this element, to enhance in bold categorizing terms. Contrarily, in Bing, text styling was not applied.

Its **positioning** is always at the right sidebar in both engines.

The *Grid of Equals* and *Thumbnail* **design patterns** are applied to this element since its beginning. The *Cards*, *Carousel* and *Module Tabs* design patterns were found applied to this element in 2018.

In both engines, the occupied area was constant from the beginning until 2018. The last two years enforced a sharp growth, effortlessly justified by longer panels nowadays, whose elements are increasingly variable and personalized.

The Knowledge Panel appeared for the first time in 2014 and one year later in Bing. It continued to be present in every single year since then.

Complete information on the knowledge panel available at: https://bedgarone.github.io/serpevolution/elements/kpanel.

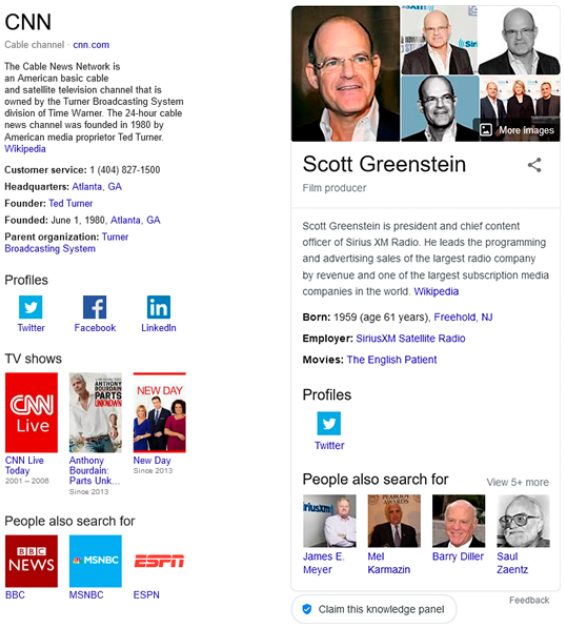


Figure 4.11: Google's Knowledge Panel from 2016 (left) and 2018 (right)

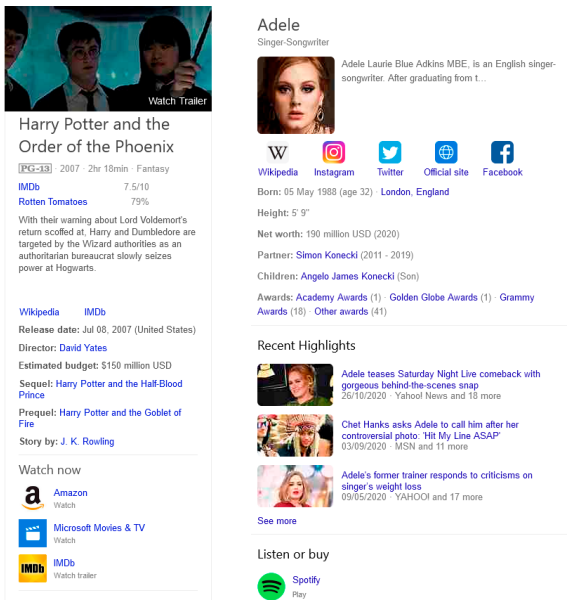


Figure 4.12: Bing's Knowledge Panel from 2016 (left) and 2020 (right)

4.3.4 Local Pack

The Local Pack, seen in Figures 4.13 and 4.14, is activated when the query is geographical, including the main related results from a map application [38].

In Google, the basic **content** structure consists of a title relating a map to the search query and a list of locations in a near location. Each result is embedded in a map preview, with its location, and relevant information. This information started to be the title, website, phone number and

a link to reviews. In 2013, a description, the complete location and the *similar & cached* links, like organic results, were added. In 2016, options to filter results by rating and schedule were introduced. Each result started to state if the place is open or closed, and had also buttons for the place's website or to get its directions, while in 2019, when applicable, a thumbnail image substituted those buttons. When focused on experience locations, like restaurants and hotels, it lists information about prices and testimonials. Although introducing an identical layout from 2020, where each row has two action buttons, one for visiting the respective place's website, and another for travelling to the location, Bing's initial structure was simple. It used, in partnership with TomTom, the image-map generator (which receives a position, area and language) to return an image of the respective map. Three links were presented, one to view a larger map, and two others to travel *to* and *from* the destination, all at *bing.com/maps*.

In terms of **design**, Google's map was initially a small image at the left of the results list. The image was later enlarged to full width and, in 2016, it changed from a static image to an embedded *Google Maps* instance, like later in Bing's case. Before 2016, text styling was always used in each suggestion to enhance relevant words in bold. From 2016, the Google's blueish usual design, full of borders and underlines, changed to softer and cleaner lines, as stated in the other elements. Text coloring is applied to both cases whenever showing information regarding its state of opening, using a gradient of colors to represent a closed, closing soon or a opened place.

Its **positioning** is highly variable throughout the results container, with tendency to be in the visible area or, in Bing's case, between the visible and scrolling areas.

The *Thumbnail* **design pattern** is applied to Bing's element since 2020, although it is not guaranteed that thumbnails accompany the results every time.

In both engines, the Local Pack constantly increased its **area** from the beginning until 2020.

The Local Pack was **present** for the first time in 2009, in Google, and scarcely before, when starting to appear frequently from 2018. There are presences of Bing's Local Pack in 2014 and 2020.

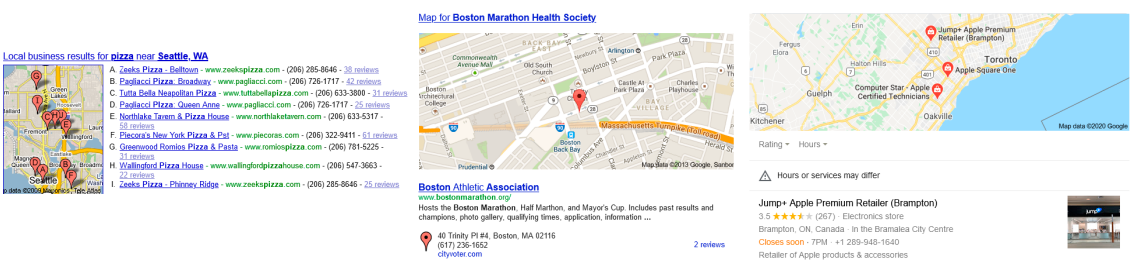


Figure 4.13: Google's Local Pack from 2009 (left), 2013 (center) and 2020 (right)

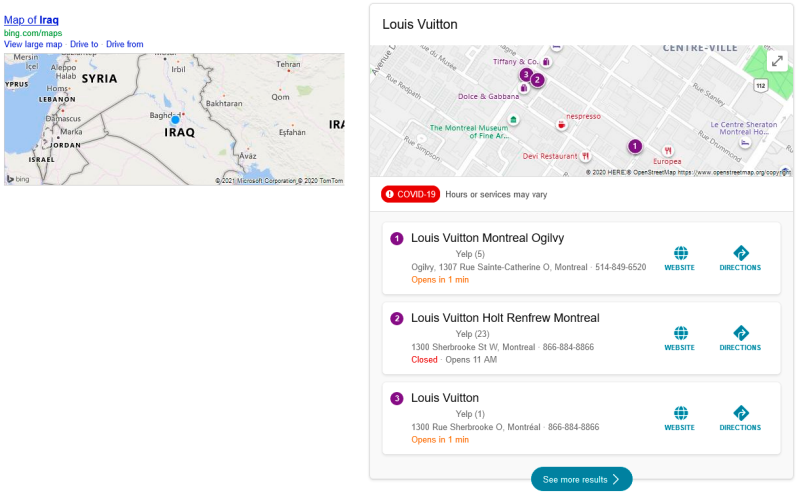


Figure 4.14: Bing's Local Pack from 2014 (left) and 2020 (right)

Complete information on the local pack available at: https://bedgarone.github.io/serpevolution/elements/localpack.

4.3.5 Image Pack

The Image Pack, seen in Figures 4.15 and 4.16, presents a set of images taken from various sources and launched when the search would have better results when accompanied by visual content [75]. The user is redirected to the search engine's images section when clicking on 'See all' similar buttons.

During most of the time, the **content** was a title associating images with the search query and a block of image thumbnails. Google included in 2014 a link for 'more images', which was always present in Bing's case, and a link to report images. In 2019, and one year later in Bing, a major change was introduced, consisting of a categories bar, a horizontal row of image and query related category buttons.

Google's **design** started with a considerable presence of blue colors, typical in Google's early interfaces, when images had a blue border. In 2014 this border was removed, the layout of a matrix appeared for the first time in 2018, but the main change was in 2019, where images started to be presented in a carousel. In 2020 the title, as usual in most Google elements, turned gray. The categories carousel changed its shape to a line that expands to a matrix according to the responsive disclosure pattern. As for Bing, in 2017, the underlined text was removed, and in 2019 bold text styling was removed while images also started to be displayed in a carousel.

In both engines, its **positioning** is highly variable throughout the results container.

In Google, the *Grid of Equals* **design pattern** is applied to this element since 2018, while *Thumbnail*, naturally, since its beginning. In Bing, the *Thumbnail* and *Grid of Equals* design patterns are applied since its beginning. In both cases, the *Carousel* is applied since 2019.

In both cases, the Image Pack **area** increased in 2019 and 2020 after the start of displaying images in carousel and, some times, in matrix.

This element **appeared** in Google SERP for the first time in 2006, but only after 2010 it started to appear frequently, lasting until now. In Bing SERP, it started to appear in 2015, continuing until 2020.

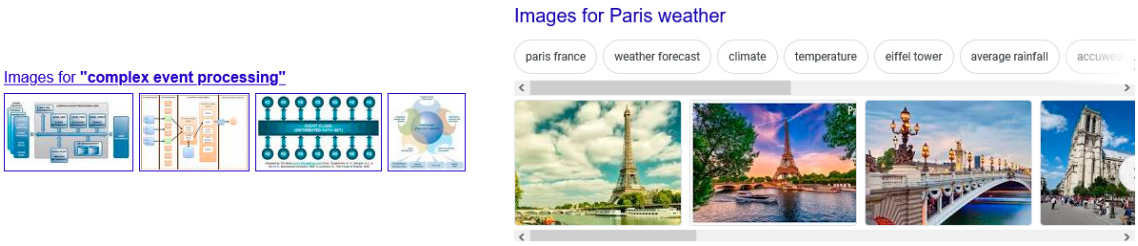


Figure 4.15: Google's Image Pack from 2010 (left) and 2020 (right)

Figure 4.16: Bing's Image Pack from 2015 (left) and 2020 (right)

Complete information on the image pack available at: https://bedgarone.github.io/serpevolution/elements/imagepack.

4.3.6 Video Pack

The Video Pack, seen in Figures 4.17 and 4.18, is a similar feature to the *Image Pack*, featuring content from video platforms like *YouTube*. This pack has also been documented in the form of a carousel [75, 49].

The **content**, in Google, was initially made of a title associating videos with the search query and a block of two videos, each accompanied by a thumbnail, title, timestamp, description and URL, while Bing's first layouts included more entries. In 2011, Google changed the structure to a vertical list of three videos, containing the same information except the description. In 2015, the principal layout, lasting until 2020, was introduced, while three or more videos were displayed individually in a carousel of cards. In the thumbnail, the video duration was included, as well as the account name in the card's body. The element's title is just 'Videos' since then. Bing's structure is very similar, substituting a previous video resolution badge with the number of views, above the source. In middle 2020, Google's layout returned back to showing videos in a vertical list, similarly to 2011, instead of a carousel. There is a button linking the user to a different page with other videos.

In Google's **design**, text styling was used in Google until 2011 to enhance relevant terms in bold. This latter version had, as seen in other elements, a color scheme fulfilled with blue, while the thumbnail had a solid blue border. The underlined links, typical in early interfaces, gave place to cleaner links without underline and only the video's title maintained the blue color. In Bing, the initial design was original, with creased square shapes, while the current one is clearly influenced by Google.

In both cases, its **positioning** is highly variable throughout the results container.

The *Carousel* **design pattern** is applied to this element since 2015, while *Thumbnail* since its beginning in both cases. The *Cards* pattern is present since 2015 in Google and 2019 in Bing.

The **area** of Video Packs were constant over time, having increased smoothly in the two last year, mainly in Bing.

This element was **present** in Google for the first time in 2010, lasting until now, except from 2012 to 2014. Bing introduced it in 2014, lasting until now.

Complete information on the video pack available at: https://bedgarone.github.io/serpevolution/elements/videopack.

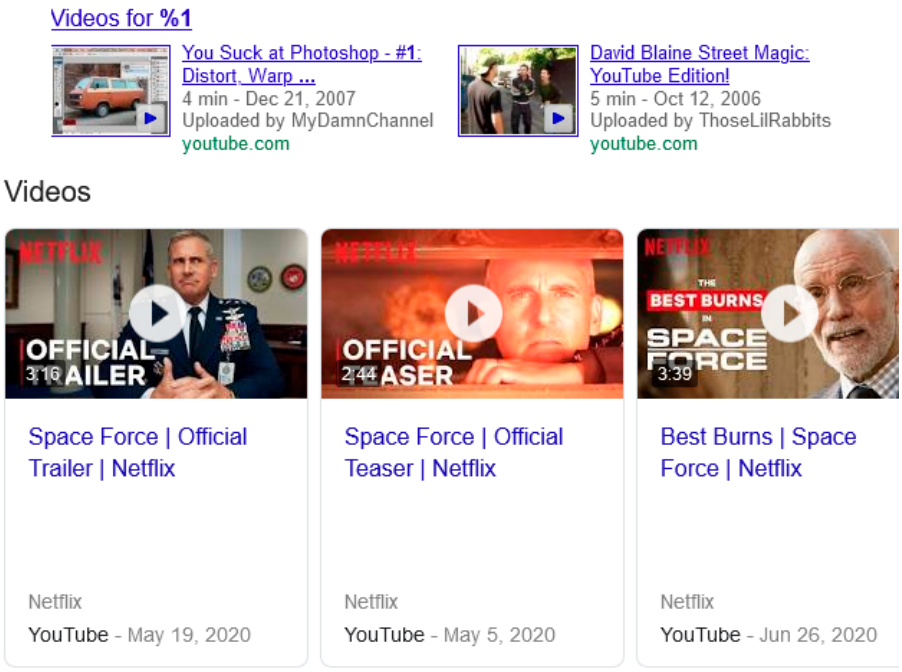


Figure 4.17: Google's Video Pack from 2010 (top) and 2020 (bottom)

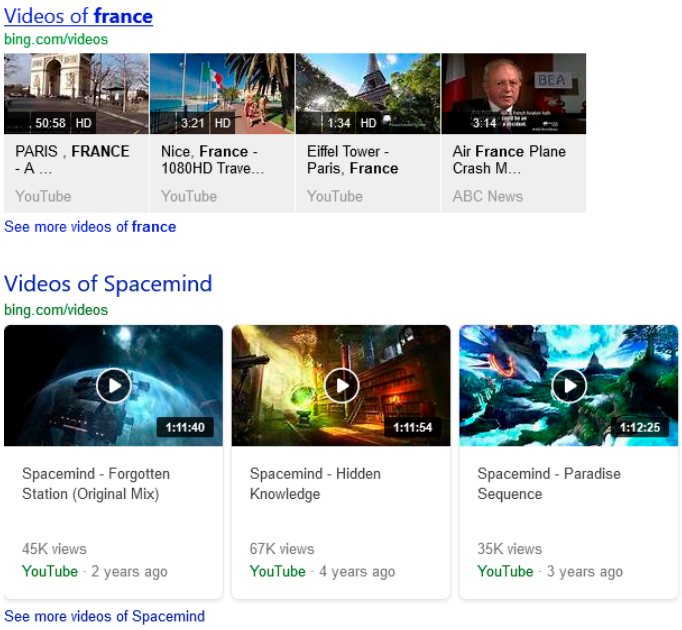


Figure 4.18: Google's Video Pack from 2015 (left) and 2019 (right)

4.3.7 Top Stories

The Top Stories, seen in Figures 4.19 and 4.20, are blocks of three or more news considered relevant to the query, placed in the form of a carousel [49]. By clicking 'View all', or similar buttons, the user is redirected to the news section of the search engine.

The element's **content** started with a vertical list of four or less news titles, each followed by the source's name and how long ago it was published. In 2006, a journal icon was placed at the left of the list and a link to 'today's top stories' was introduced. In 2013, the icon was substituted by a thumbnail for the first news result, being that the most important news in the element. The latter was complemented with an extract of the news, while the rest stayed the same. In 2020, the graphics was majorly altered to display the results in a carousel of cards. However, the content was simplified to only present, for each result, a thumbnail, title, source and how long ago it was published. In Bing's case, the element is currently more sophisticated. In 2010 it emphasized the first news item, including a brief description of the content. In 2014, this top position started to include a thumbnail when possible, which carried over to all items in 2017. From 2020, three layouts were introduced. A shorter format, with a horizontal list of three cards, each accompanied by image, title, source and time of update. Another layout presents a highlight to a main news, in a horizontal card that occupies the entire width, followed by a block with three other news items as in the previous layout and, below them, a block of cards with Quotes related to the theme. Each quote indicates the person, their role, what they said, and the source of the news. The third layout is identical to this one, changing the quote block with an opinion block, with each opinion being a regular news card.

As for Google's **design**, as usual, color scheme was mainly blueish and fulfilled of blue borders and underlines. These were removed after 2020 with the softer gray colors for additional information but still blue titles. Bing's design, however, was always similar to Google's Top Stories.

Its **positioning** is mainly in the Google's visible area, while in Bing it is highly variable throughout the results container.

In both cases, the *News Stream* **design pattern** is applied to this element since its beginning. However, in this element, relation to the query appears to be more relevant than publication time since it is no longer possible to observe any chronological order. The *Thumbnail* design pattern is applied to this element since 2012 in Google and 2017 in Bing, while *Cards* and *Carousel* since 2020 in Google.

The **area** of this element increased considerably after the transition to the cards, in 2020. In Bing's case, this increase is bigger than Google's due to the size of the extra content elements already described.

This element **appeared** for the first time in Google in 2004, but only after 2011 it started to appear frequently, lasting until now. In Bing, it was found in 2010, 2014 and 2017 before a considerable presence in 2020.

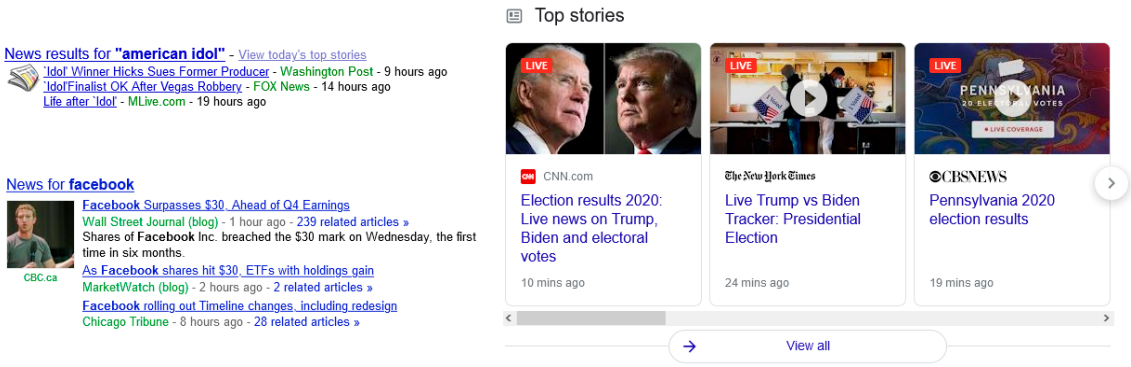


Figure 4.19: Google's Top Stories from 2006 (left-top), 2013 (left-bottom) and 2020 (right)

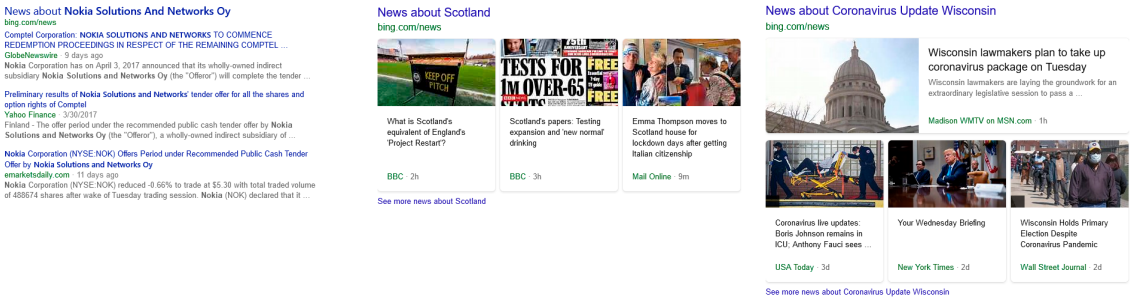


Figure 4.20: Bing's Top Stories from 2017 (left) and 2020 (center and right)

Complete information on top stories available at: https://bedgarone.github.io/serpevolution/elements/topstories.

4.3.8 Carousel

A Carousel, seen in Figures 4.21 and 4.22, is a line of several cards, accompanied by a name and highlighting an image, which presents related results at the top of the SERP that are part of the same category (such as discs or personalities). It is possible to scroll through more cards without leaving the SERP. In addition to presenting brief information directly, and sometimes in an orderly manner, Moran and Goray [38] point that this also serves as a navigation tool, since each card takes the user to the result's respective SERP.

Initially, in Google, it had a dark gray background and each result was based of a squared image, a title, and a subtitle, when applicable. Categories were figuring as a breadcrumb at the top. The **content** maintained over time, while the **design** changed in 2017, with a brighter background and an individual card for each result. The same background color contrast happened in Bing. The first example, despite being sparsely populated, exemplifies a horizontal line of items with image, title and brief pertinent information. In 2020, each card's content was improved. Figure 4.22's second example shows a list of schools, representing one of many dynamic possibilities for structuring content, with the school name, type, address and even a rating if applicable.

In both engines, its **positioning** is always below the search query bar and above the results container, filling the entire width.

The *Carousel*, and *Thumbnail* **design patterns** are applied to this element since its beginning, while *Cards* since 2016 in Google and 2020 in Bing. The *Breadcrumbs* pattern was also applied in Google since the element's beginning.

Considering the scarce appearances of this element in the dataset, it is possible to conclude that the **area** was constant over time in Google, while in Bing it suggests that it can vary according to the height of the items' cards.

Google's Carousel **appeared** for the first time in 2015, lasting until 2018 in the dataset, although it is still present today [49]. In Bing, it was found twice in the dataser, in 2017 and 2020.

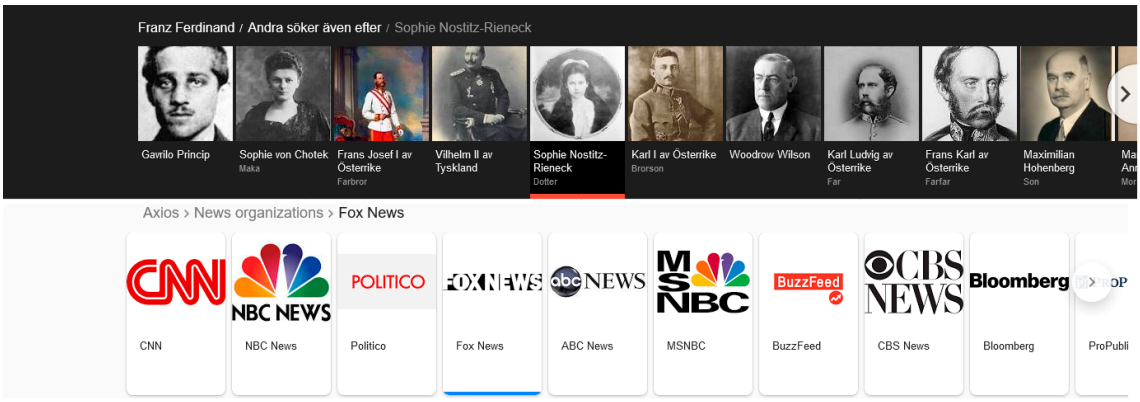


Figure 4.21: Google's Carousel from 2016 (top) and 2018 (bottom)

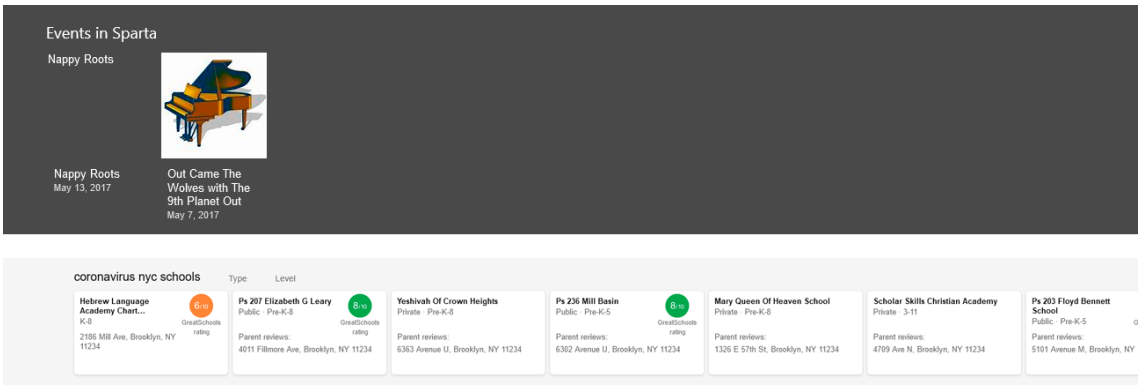


Figure 4.22: Bing's Carousel from 2017 (top) and 2020 (bottom)

Complete information on the carousel available at: https://bedgarone.github.io/serpevolution/elements/carousel.

4.3.9 Carousel Grid

The Carousel Grid, exclusive to Google and seen in Figure 4.23, is a *Carousel* displaying **content** in a matrix of cards instead of a line, positioned at the same place. In this case, the cards are simpler and smaller to avoid increasing the element's height in excess. Results could be accompanied with a thumbnail, but the same decision was applied to every result.

The **design** is related to the *Carousel* from 2017 and no evolution was noticed over time.

Its **positioning** is always below the search query bar and above the results container, filling the entire width.

The *Breadcrumbs*, *Carousel* and *Grid of Equals* **design patterns** are applied to this element since its beginning.

The **area** was irregular over time, because its height was dependant on the number of grid lines.

This element appeared for the first time in 2017, except for 2019.

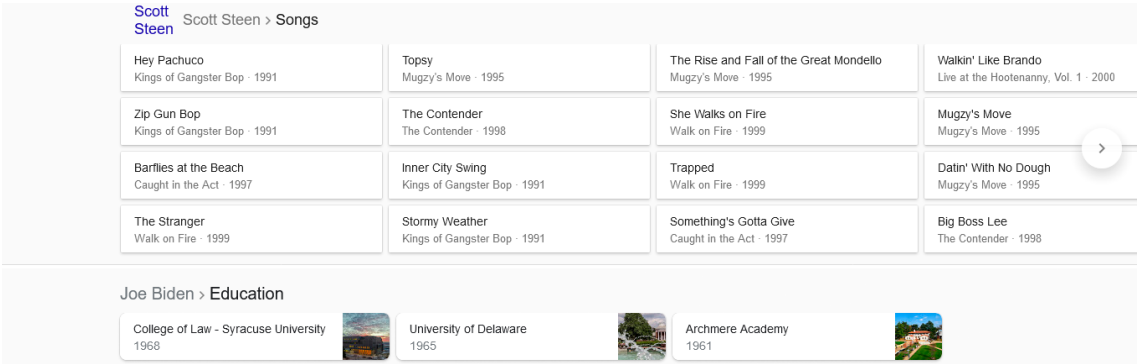


Figure 4.23: Google's Carousel Grid from 2018 (top) and 2020 (bottom)

Complete information on the carousel grid available at: https://bedgarone.github.io/serpevolution/elements/carouselgrid.

4.3.10 People Also Ask

People Also Asked, seen in Figures 4.24 and 4.25, is an accordion of some questions suggested by the search engine and which are related to the search query [75]. Each expanded element features a featured snippet answering the element's question, complementing the information with the link from where the excerpt is taken [38].

In terms of **content**, no changes can be noticed, as the element has kept its shape, four accordion questions, untouchable over time.

In 2019, a soft **design** change was made to improve readability, by making text padding larger and, therefore, increasing the distance between text and borders. Bing's element is completely influenced by the Google's one, as the only noticeable difference is the text styling of the elements title, in uppercase and in a different color.

In both engines, its **positioning** is highly variable throughout the results container.

The *Accordion* **design pattern** is applied in Google and Bing since the element's beginning, freeing up space and allowing the user to access the content only if interested in the answer to the suggested question.

Bing's element show a clear consistency in **area** size due to the small element lifetime. Google's element had a slight increase in area after the mentioned changes in spacing in 2019.

In Google, it was **present** since 2014, lasting until now, and was introduced by Bing a half decade later, in 2020.

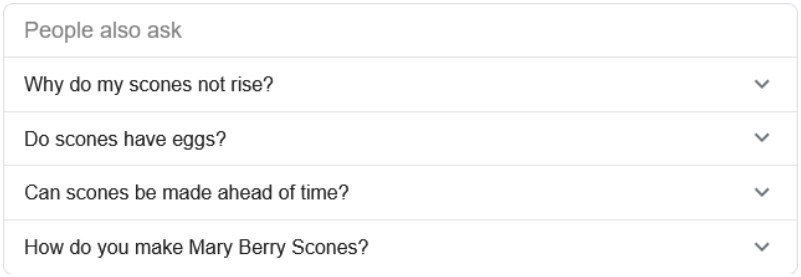


Figure 4.24: Google's People Also Ask from 2019

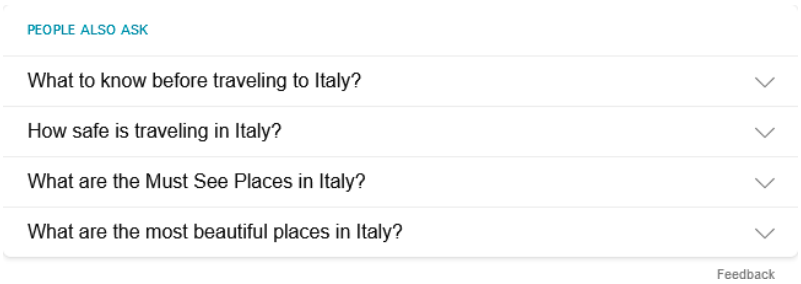


Figure 4.25: Bing's People Also Ask from 2020

Complete information on people also ask available at: https://bedgarone.github.io/serpevolution/elements/pplalsoask.

4.3.11 Related Searches

Related Searches, seen in Figures 4.26 and 4.27, are a common element on SERP from a very early stage and offer suggestions for related searches, usually through the suggestion of similar queries. Each link takes the user to the respective SERP.

Content was diversified just in terms of how many suggestions would appear and what would be its layout. Each suggestion of search is a hyperlinked title, pointing to its respective SERP. Initially, in Google, it was organized in a matrix of columns. In 2011 this schema was reduced to two columns, but for suggestions with longer text, it could be displayed in just one column, a list. In Bing, when the element was in the sidebar, the options were listed in a single column, but after moving to the results container, in 2014, options appeared in two columns. In 2020, the schema was reduced again to a single column.

As for **design**, text styling was always used in Google in each suggestion to enhance relevant words in bold, while in Bing this started in 2013. In Google, until mid 2020, suggestions were blue and, until 2014, underlined. A search icon was applied to each entry in 2020. Later, a new version changed the graphics, making each entry a button, with solid gray background, search icon and title in black. In Bing, after the appearance of options in two columns and underlined in 2014, the shape of the element would be similar until 2019, where items became more spaced for better reading. Although some captures present it before, it is in 2020 that the list is resumed in a single column, keeping the previous spacing.

In Google, the element's **positioning** was always at the bottom of the results container, while in Bing it was in the left sidebar until 2014, and then at the bottom of the results container.

There is no **design pattern** applied in both engines.

Related Searches **area** increased constantly in both engines over time. This growth is due to the options spacing that tended to enlarge, while on Google the appearance of icons and colored backgrounds enforced even more this increase.

This element **appeared** for the first time in Google in 2008, lasting until now, except for 2010. In Bing, it was present since the beginning, except for 2011.

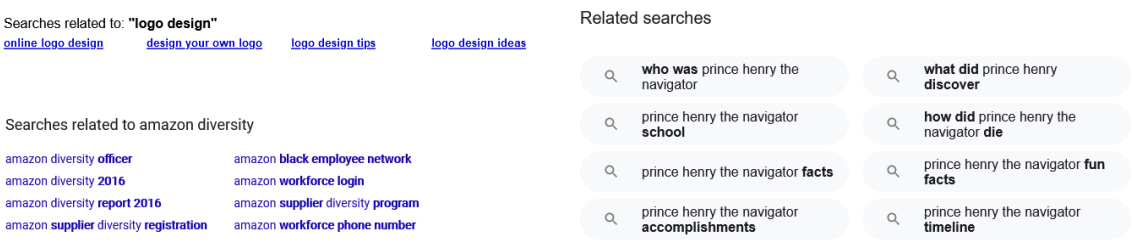


Figure 4.26: Google's Related Searches from 2008 (left - top), 2017 (left - bottom) and 2020 (right)

Figure 4.27: Bing's Related Searches from 2010 (left), 2017 (center) and 2020 (right)

Complete information on related searches available at: https://bedgarone.github.io/serpevolution/elements/relatedsearches.

4.3.12 Twitter Pack

The Twitter Pack, seen in Figures 4.28 and 4.29, presents recent tweets related to the search query.

The initial structure for **content** in Google was a header with the twitter account username and the account URL, followed by two recent results. Each result could be accompanied by a thumbnail and included the tweet, how long ago it was published and a link to the tweet's page.

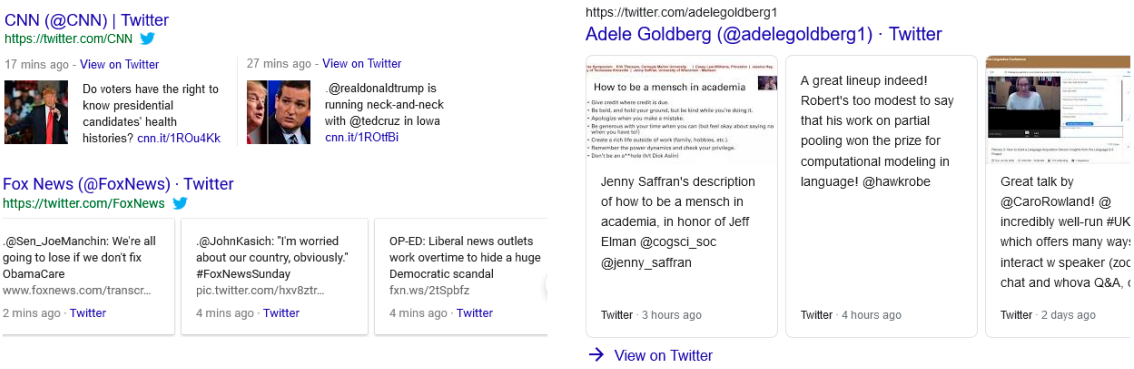


Figure 4.28: Google's Twitter pack from from 2015 (left - top), 2018 (left - bottom) and 2020 (right)

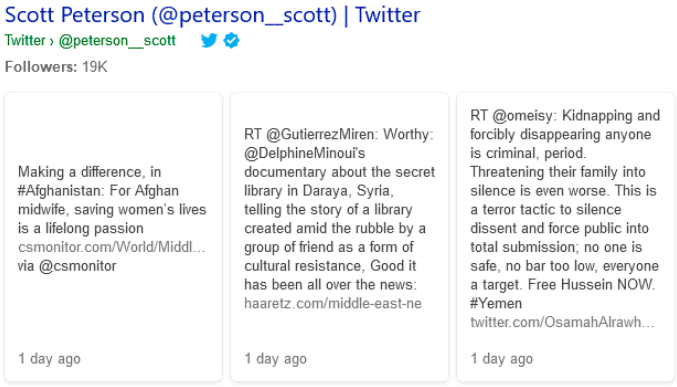


Figure 4.29: Bing's Twitter Pack from 2019

In 2017, tweets started to be three or more, in a carousel, and its **design** was improved, displaying items in individual cards, without showing an image. In 2020, the structure still maintained but the thumbnail for each tweet returned. However, Bing's element appears once in the dataset, thus, positioning and area evolution are not referred. It displays the person's name, Twitter username, link to profile, rounded number of followers, and a row of three cards with recent tweets.

Its **positioning** is highly variable throughout the results container.

In Google, the *Cards* and *Carousel* **design patterns** are applied to this element since 2017, while *Thumbnail* since its beginning. In Bing, the *Cards* pattern is applied.

Google's element increased its **area** constantly since the beginning.

In Google, it **appeared** for the first time in 2015, lasting until now, while in Bing, only once in 2019.

Complete information on the twitter pack available at: https://bedgarone.github.io/serpevolution/elements/twitterpack.

4.3.13 Recipe Cards

The Recipe Cards, seen in Figure 4.30, appeared for the first time in 2020, and no evolution was noticed over the year. Its positioning is frequently at the top of the results container. More cards appear after clicking on 'Show more'. The *Cards*, *Responsive Disclosure*, *Thumbnail* and *Grid of Equals* design patterns are applied to this element since its beginning, the latter after the element is being disclosured. Responsive Disclosure is used when the user is intended to be in the same page while exploring information or proceeding with different actions, being able to see part of the information at first [70]. The interface is being revealed (disclosured) as the user proceeds.

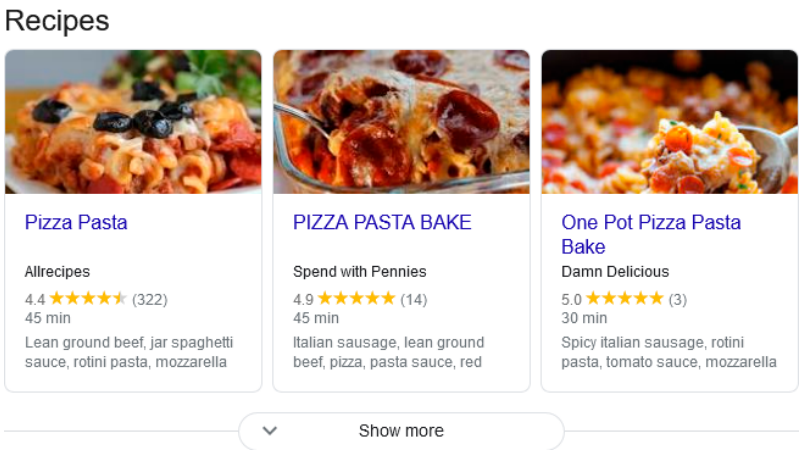


Figure 4.30: Recipe Cards from 2020

Complete information on recipe cards available at: https://bedgarone.github.io/serpevolution/elements/recipes.

4.3.14 Category Hierarchy

The Category Hierarchy, exclusive to Google and seen in Figure 4.31, is a no longer existing element that presented one or more categories related to the query, taking advantage of breadcrumbs to locate these categories in more general areas.

Its **positioning** was always at the top of the results container.

The *Breadcrumbs* and *Categorization* **design patterns** are applied to this element since its beginning.

The element's **area** and **design** were constant during its lifetime.

Category Hierarchies were **present** in 2003 and 2004.

Figure 4.31: Google's Category Hierarchy from 2004

Complete information on the category hierarchy available at: https://bedgarone.github.io/serpevolution/elements/categoryhierarchy.

4.3.15 Covid-19 Left Panel

In 2020, queries related to the Coronavirus pandemic led to highly sophisticated SERP versions. In order to categorize the content, the left panel came back to SERP in the shape of Covid-19 Left Panel, exclusive to Google, as shown in Figure 4.32. It allows the user to consult various tabs of updated information regarding Covid-19 panorama and prevention.

The **content** consists of a column of categories that change the SERP results container to specific information when each button is selected.

In mid-2020, the **design** was changed, removing the red gradient in the background and changing the selected tab colors to blue.

Its **positioning** is always in the left sidebar.

The *Categorization* **design pattern** is applied in this element since the its beginning.

The element's **area** was constant during its lifetime.

This element was **present** in 2020.

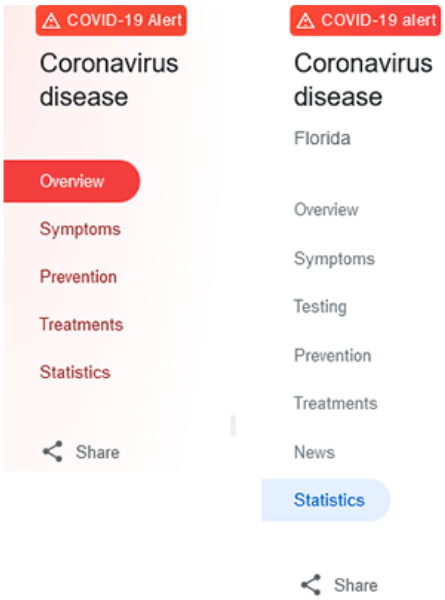


Figure 4.32: Google's Covid-19 Left Panel from 2020

Complete information on the covid-19 left panel available at: https://bedgarone.github.io/serpevolution/elements/covidpanel.

4.3.16 Search Refinement

The Search Refinement, exclusive to Bing and seen in Figure 4.33, is a recent element that provides query-refinement options to the user.

Its **content** consists of a row of query-related search options, similar to related searches, but appearing earlier, at the top of the results container.

Its **design** is simple and consists of a button shape for each suggestion. Text styling is not applied for enhancing new or changed words.

It's **positioning** is always between the search query bar and the results container.

The element's **area** was constant during its lifetime.

This element was **present** since 2018.

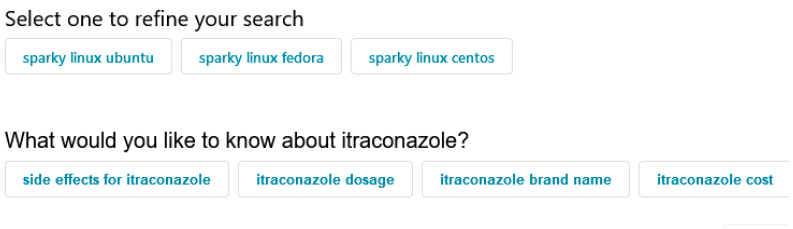


Figure 4.33: Bing's Search Refinement from 2018 (top) and 2020 (bottom)

Complete information on search refinement available at: https://bedgarone.github.io/serpevolution/elements/searchrefinement.

4.3.17 Categories Bar

The Categories Bar, exclusive to Bing and seen in Figure 4.34, presents a row of buttons referring to query-related categories, which can lead the search to a more restricted set of information.

Its **content** consists of a varied number of categories that are usually an integral part of the initial query's subject, such as the history, points of interest and geography of a country; cast, episodes and awards of a series; or the age, career and speeches of a politician.

Its **positioning** is always below the tabs bar.

The *Categorization* and *Thumbnail* **design patterns** are applied to this element since its beginning.

The element's **area** and **design** were constant during its lifetime.

This element **appeared** for the first time in 2020.

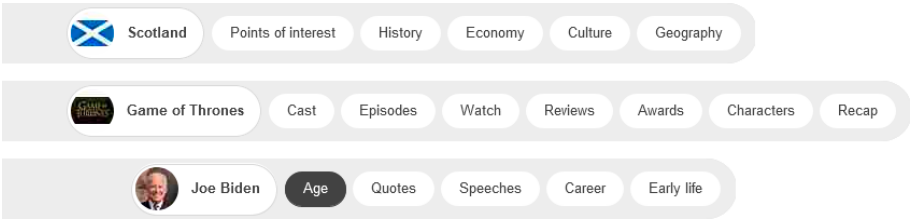


Figure 4.34: Bing's Categories Bar from 2020

Complete information on the categories bar available at: https://bedgarone.github.io/serpevolution/elements/categoriesbar.

4.3.18 Travel Destinations

The Travel Destinations, exclusive to Bing and seen in Figure 4.35, presents a set of three cards with main travel destinations when the query is dedicated to a country.

As for the **content**, each card features an image of the place, its city and a brief indication of what is highlighted in the photograph, usually a monument.

Its **positioning** is usually in the scrolling area.

The *Cards* and *Thumbnail* **design patterns** are applied to this element since its beginning.

The element's **area** and **design** were constant during its lifetime.

This element **appeared** for the first time in 2020.

Complete information on travel destinations available at: https://bedgarone.github.io/serpevolution/elements/traveldestinations.

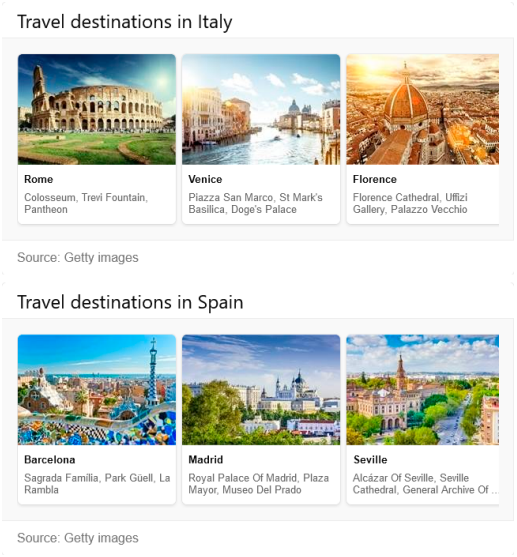


Figure 4.35: Travel Destinations from 2020

4.4 Summary

In sum, every element listed in Section 2.2.2 was found in Google, but *Shopping ads*, *Direct Answer Results* and *Recipe Cards* are inexistent in Bing. Elements in this section that are new to that list, and are exclusive to each engine, are summed up Table 4.1.

Table 4.1: Exclusive SERP elements found in the dataset

Element	**Description**	**Engine**
Carousel Grid	A *Carousel* displaying content in a matrix of cards instead of a line, positioned at the same place	Google
Category Hierarchy	No longer present element that stated in breadcrumbs one or more categories related to the query	Google
Covid-19 Left Panel	Allows the user to consult various tabs of updated information regarding Covid-19 panorama and prevention	Google
Categories Bar	Presents a row of buttons referring to query-related categories	Bing
Travel Destinations	Presents a set of three cards with main travel destinations when the query is dedicated to a country	Bing
Search Refinement	Provides query-refinement options to the user	Bing

Chapter 5

Overall Analysis

This chapter analyzes results from an overall perspective. We synthesize how Search Engine Results Pages (SERP) layout and overall design evolved, highlight the most important changes, and discuss what led to the increase of SERP area and how files number and size varied over time.

5.1 Layout

In this section, we address the results from the colored interface breakdown, by analysing the layouts designed for each category. This process overlays images from different captures, only colored at spaces occupied by specific elements, with a high level of transparency, as described in Section 3.3. Whenever common areas overlap, transparency level reduces and coloring becomes more intense. The complete results from this layout overlappings are available at the *Layout* section in the website: https://bedgarone.github.io/serpevolution/layout.

Category **welcome & identity** encompasses elements with the search engine's visual identity, such as the logo, and other graphic elements that welcome the user, such as slogans, or that are associated with the brand's identity. It was noted in Figure 5.1 that both search engines place their identity in SERP without welcoming elements. Nevertheless, their pattern is almost identical; each logo has kept its position and size with rare variation. Exception applies to a small part of Google interfaces with a larger distance between content and the left of the screen, appearing considerably shifted to the right.

As for **search statistics**, which relates to elements that display statistics about the retrieved results, they represent another group of similar elements with identical placement in both engines, appearing consistently below the search query or navigation bar, either left-aligned, right-aligned, or both (justified).

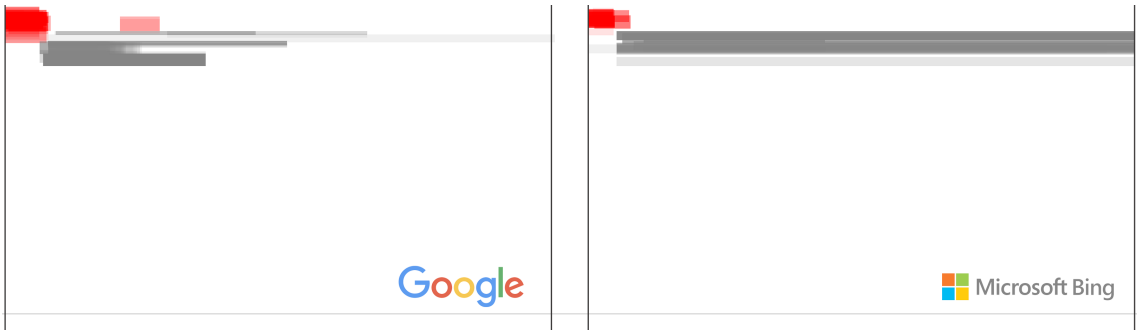


Figure 5.1: Layout overlapping for welcome, identity (red) and search statistics (gray)

Category **navigation & user inputs** embraces all the links that the user can use to navigate within the website structure, as well as the elements where the user can introduce inputs, such as the search query bar example. It is possible to notice in Figure 5.2 a significant presence of the left navigation bar, although not present during several years in both cases. The query bar has also marked its place at the top of both pages, left-aligned, but appeared in early phases at the bottom as a duplicate. On the opposite side, right-aligned, there is a usual space for sign-in and user account information. At the bottom of the page, two areas are noticeable: pagination, aligned to the left; and the footer, at the very end, covering all the available width. Thus, no differences can be seen in this category, with the usual exception for Google interfaces from 2010 to 2012 that where adjusting elements' position according to the screen's width and its central axis, rather than a left alignment. This lead the left navigation bars to being slightly shifted to the right. The end result is trimmed at the middle and should be seen as two separate results, inferring a variable height of the gray bar.

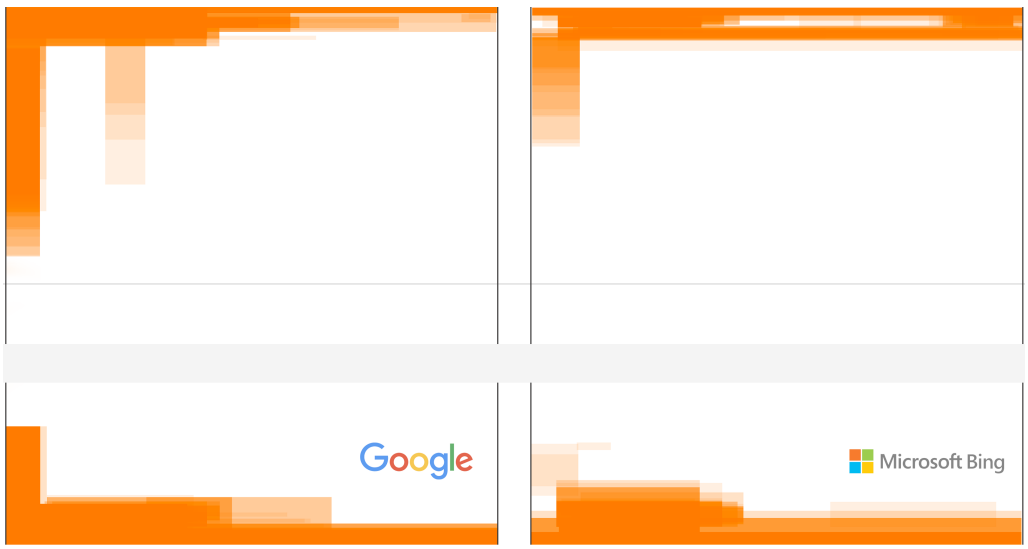


Figure 5.2: Layout overlapping for navigation & user inputs

The positioning results of **organic results**, seen in Figure 5.3, reveal an expected strong presence of frames in the main results container, with a greater focus on the visible area and

initial part of the scrolling area. Over time, it is noticeable that SERP pages have increased their height due to the vertical decrease of color intensity, revealing the appearance of results in lower areas of the page. In both cases, part of the results is slightly shifted to the right, referring to interfaces with a left side navigation bar. Content was left-aligned during all the history of Bing. In Google's case, contrarily, it is possible to observe two other very tenuous sets apart from the strongest container. One with more centered results due to the interfaces referred in the previous category. Another small set of frames cover the entire width of the interface, not because the content was that large but because any div of the first Google SERP had a width of 100%.

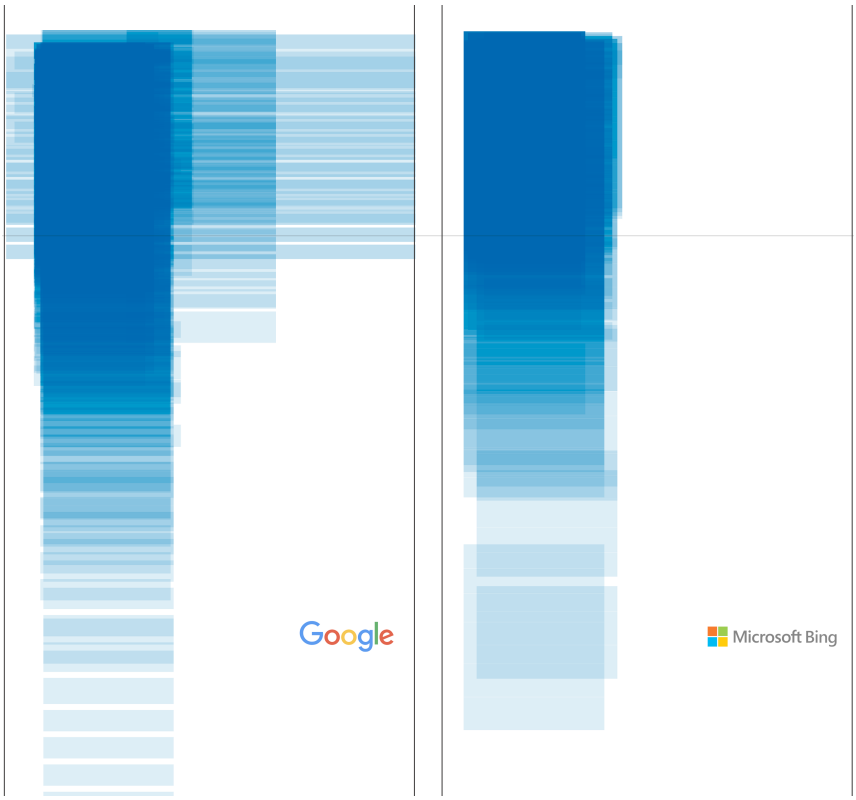


Figure 5.3: Layout overlapping for organic results

Regarding **sponsored results** (Figure 5.4), it is possible to identify three major advertisement areas in both SERP: top ads, right ads and bottom ads. Bing is more stable and consistent when limiting those areas, with emphasis for a higher incidence on bottom ads, comparing to Google's layout. These seem to appear more often due to a higher vertical variation in that position. Google's bottom adds are more centered on the scrolling area's initial part. Again, contrarily to Bing, Google advertisements in the right sidebar move horizontally over time because some interfaces force this sidebar to be responsive to the screen's width and closer to the right edge. Top adds are the most common positions for ads in both cases, although Google is the one having a higher vertical variation. Older Google interfaces place advertisements that occupy almost the entire width, characterized by a fully colored div, as referred in previous categories.

SERP Features's layout in SERP, seen in Figure 5.5, is similar in both SERP, particularly in

their central results container, as many features are similar to regular results, with larger height. It is not possible to determine a most common localization in this container, since they appear in almost every possible position throughout visible and scrolling areas. SERP features share place in the sidebar with advertisement, being recently way more present than the latter. In both cases, most frames with high height values correspond to the well-known knowledge panel. A third horizontal area is noticeable, normally assigned to the carousel. The result suggests that Bing's placement for this bar is variable whereas Google's one is more consistent. The same behaviour applies to the knowledge panel.

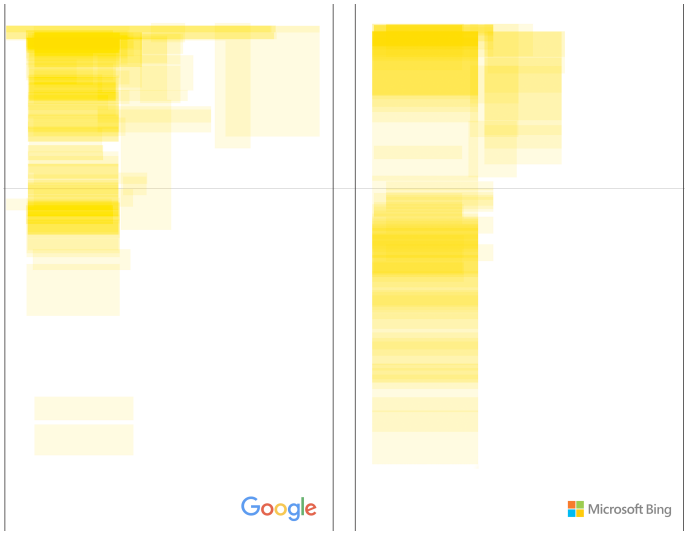


Figure 5.4: Layout overlapping for sponsored results

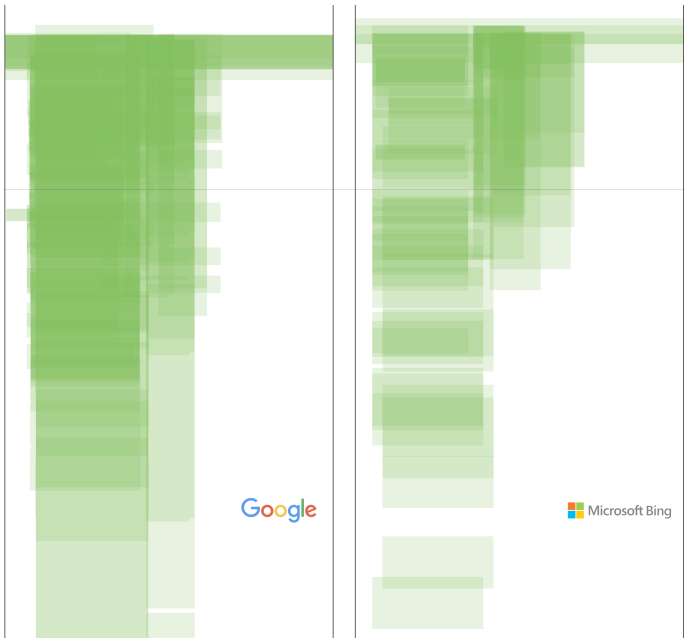


Figure 5.5: Layout overlapping for SERP features

5.2 Visual Evolution

As stated before, the first stage of the dataset analysis was made upon a visual selection from a monthly sample from the dataset. Figure 5.6 and Figure 5.7 are the first outcome from this dataset sample, displaying how Google's and Bing's **overall interfaces** evolved in terms of design. In these figures, we include the main versions of Search Engine Results Pages (SERP) interfaces based on major noticeable changes. Timelines with larger resolution are available at the website's *Design* section: https://bedgarone.github.io/serpevolution/design.

Google's initial interfaces were made with simple HTML Document Object Model (DOM) trees, with few depth levels. The first interface design, traced in 2000, differentiated the sponsored results with a colored background. A second search query bar existed at the bottom of the page, and the user could change the number of results presented. This latter feature was removed in the following interfaces. The one traced from 2000 to 2004 revealed a right block of results, exclusive to sponsored ones. It marked the appearance of the first bar with tabs directing the user to other types of content (e.g. images and news). The fourth interface design, traced from 2010 to 2012, the content was not left-oriented, but varied in a spaced manner depending on the screen's width. It introduced a sidebar on the left, containing tabs to manage the results, but some of these tabs were duplicated due to another tabs bar in the navbar.

Significant aesthetic changes occurred in 2012. The fifth interface design relates to an important period for Google, the launch of the Knowledge Graph, forcing the right column to complement information with the knowledge panel and compete for its space with sponsored results. Some modifications were found earlier in the dataset during those years, suggesting some interface testing by Google, the sixth interface design. However, the main and solid design, close to what we have online today, began at the end of 2018, the seventh interface design. As noted before in some elements' graphics, this interface focused on modernizing its lines.

Despite being live during half of the time, Bing have still promoted 5 major interface changes.

Bing's first interface design, traced in 2010, would differ from the others for having two search query bars serving the same effect. It was also characterized by its left sidebar, called Explore Pane, that enables users to easily navigate various categories of results within a consistent location [67]. There one could find a triggered Reference tab, where Bing used to present semantically indexed Wikipedia content about entity related searches [30].

Bing's second interface design was traced between 2012 and 2014. The interface brought cleaner results, and was marked by the removal of the left pane enabling an easier scan in both desktop and mobile environments [50]. While the second search query bar, placed at the bottom, also disappeared, related searches changed their position to the right pane and/or to the bottom of the page, as the main tabs were placed at the top navbar.

Bing's third interface design, traced between 2014 and 2018, maintained the top tabs bar and presented an even cleaner design. The overall design was upgraded to mark the recent introduction of the main knowledge feature *Snapshot*, equivalent to Google's *Knowledge Graph*,

to help users with all the supporting context needed, and the Pole Position, a surface area at the top of the page [2].

Bing's fourth interface design was released earlier for testing, as few captures are dated from 2016, 2017, and 2018, intersecting the period of the previous one, before appearing in huge quantity in 2019. Thus, it was traced between 2016 and 2019, having as a major change, the repositioning of the main tabs, that started (until today) being localized under the search query bar. Following improvements, traced until the end of 2020, are just of a graphic design nature, updating the interface to more modern lines.

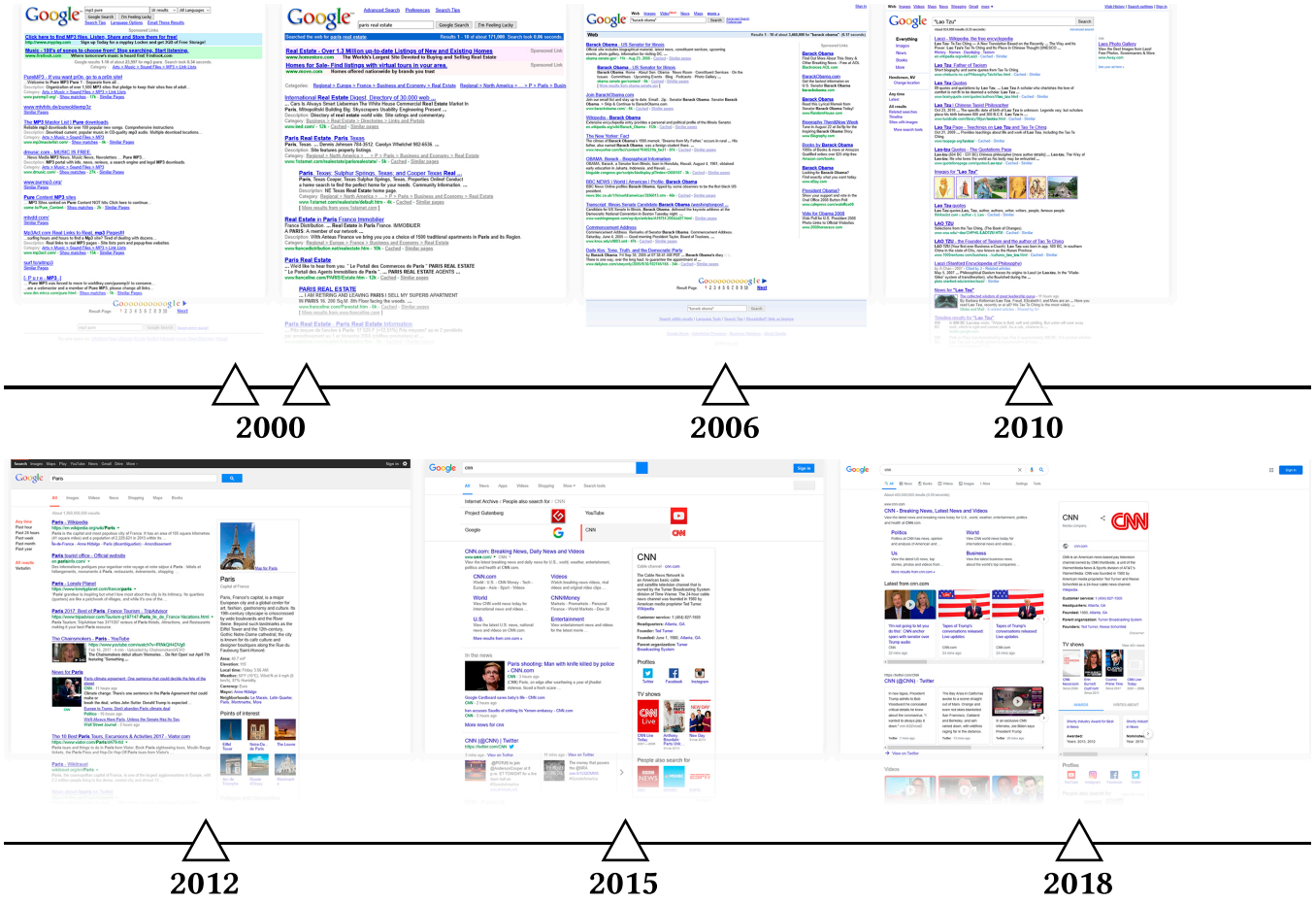


Figure 5.6: Interfaces' visual evolution over time (Google)

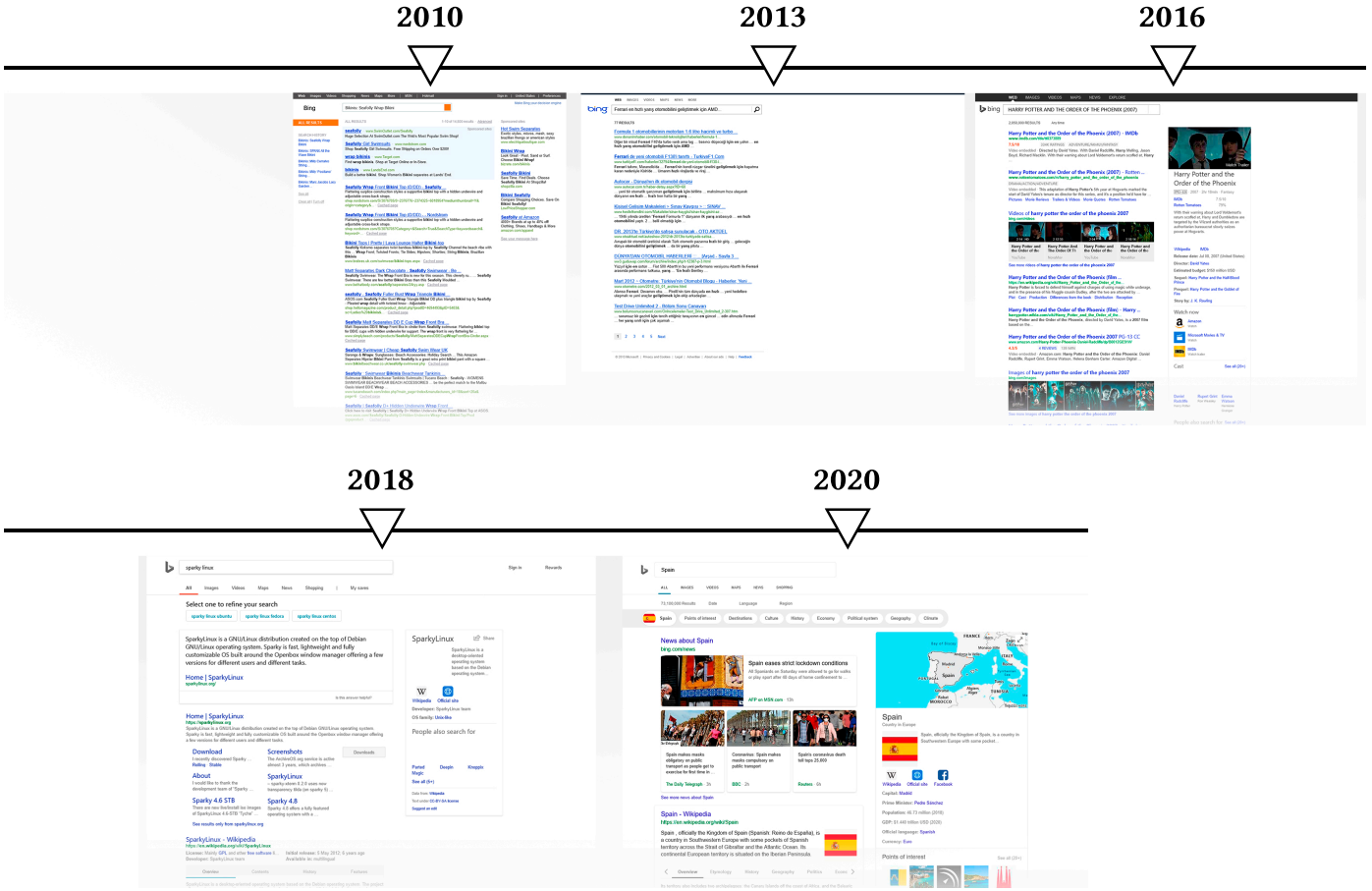


Figure 5.7: Interfaces' visual evolution over time (Bing)

In terms of **Identity**, we can see how Google and Bing logos evolved in Figures 5.8 and 5.9, respectively. Google's is clearly the most consistent, as it has kept its main components since the beginning: the letters "Google" sequentially colored with blue, red, yellow and green. Only soft design changes were applied to enforce modern lines. In Bing's case, its logo underwent design renovations and varied between the presentation of the 'Bing' letters, just the icon of the initial 'b' and even the inclusion of the *Microsoft*'s identity, using the latter's visual to harmonize all the logo.

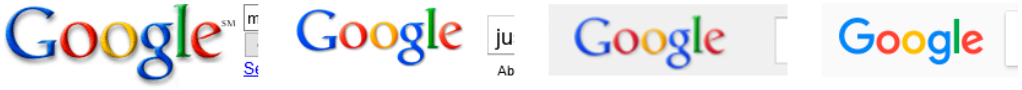


Figure 5.8: Google's logo in 2000, 2011, 2013, and 2020 - from left to right

Figure 5.9: Bing's logo in 2013, 2016, 2018, and 2020 - from left to right

As for **Search Statistics**, they varied in both engines accordingly to the main interface changes seen earlier in this section. This evolution can be seen in Figure 5.10 for Google and Figure 5.11 for Bing. Two components can be identified: the number of results, present in both cases, and the time that lapsed when retrieving the results, exclusive to Google. In fact, all the Bing versions present a simple sentence in dark color and white background. Contrarily, Google has went further when embedding the element each interface design, making use of background colors and text styling. However, if we restrict the analysis to the common period, from 2010 to 2020, both elements are similar.

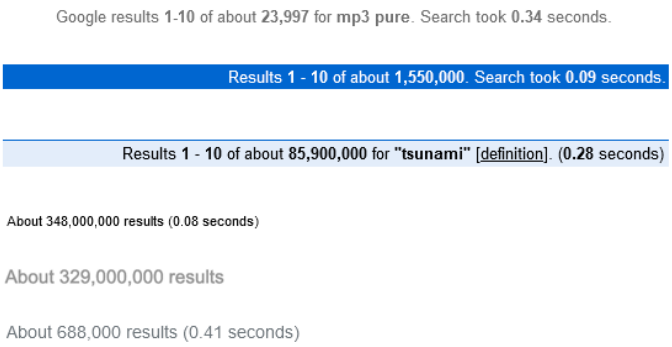


Figure 5.10: Google's results statistics in 2000, 2004, 2006, 2011, 2013, and 2020 - from top to bottom

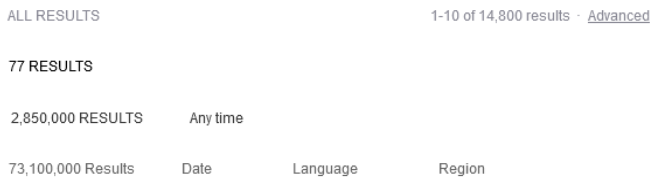


Figure 5.11: Bing's results statistics in 2010, 2013, 2018, and 2020 - from top to bottom

Navigation & User Inputs elements play a major role on the evolution of Search Engine Results Pages (SERP) interfaces. Figure 5.12 displays the main stages of how Google's search query bar and its surroundings evolved. In 2000, the main buttons for this category were 'Google Search' and 'I'm feeling lucky'. It was possible to select, directly in a dropdown, how many results should be shown, and to search within the results. There was also a tab with different links to other Web search engines for the user to search that same typed query, and an option to email the results retrieved. Other lesser relevant links would point to SERP experience (e.g., language, search tips). In 2001, the tabs bar was introduced in the shape of *Module Tabs*. The modules were Web, Images, Groups, and Directory. In 2002, the News tab was added. In 2006, these tabs started appearing above the search query, and the module tabs pattern was no longer applied. The Video, Maps, Froogle, and More tabs were introduced in this year. In 2008, the tabs bar went to the very top of the page, in a navbar, when Shopping and Gmail tabs were included. In 2010, the left sidebar was introduced, complementing the interface with other tabs and other information such as location and results filtering. At this stage, Bing would have been launched. In 2012, the bottom query bar was removed. In 2013, the tabs bar was displaced to underneath the main query bar, although the navbar still maintained some of the links, changing others. At the same time, in the left sidebar, filtering results by theme and by time were the only options available before this sidebar was removed in 2014. Finally, in 2015, the tabs bar below the query bar became the only existing tabs bar.

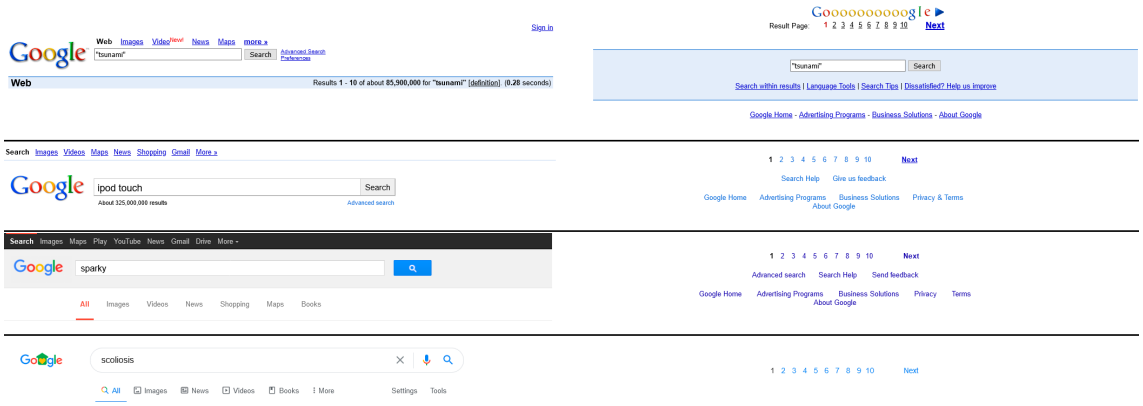


Figure 5.12: Google's navigation & user inputs elements from 2006 (top), 2012 (mid-top), 2017 (mid-bottom) and 2020 (bottom)

As seen before, Bing's case started with similar structures to Google. Figure 5.13 displays the

main stages of how the search query bar and its surroundings evolved. The tabs bar was placed initially in the top navbar, like Google in the same period. The MSN and Hotmail tabs were exclusive of this first period until 2013. The left sidebar was also present from the beginning, in coherence with Google's interface of the same period. There, the search history was listed, with previous entered queries, and action links to see all the history entries, to clean them or to disable this functionality. Next to search statistics, there was a link to advanced settings. In 2013, the bottom query bar was removed, as well as the shopping, MSN and Hotmail tabs. The left sidebar also disappeared. In 2016 the More tab changed its name to Explore. Finally, in 2017 the tabs bar was placed below the query bar, until now, although the previous design was maintained throughout the year. This year marked also the return of the Shopping tab, replacing the Explore one.

Regarding the bottom part of the interfaces, one can notice that Bing's footers and pagination were always simpler and less interactive than Google's, which used to have a variety of navigation links until 2020, contrarily to Bing's focus on pagination, scarcely surrounded by other elements. Those links are mainly of two types: informative or search settings.

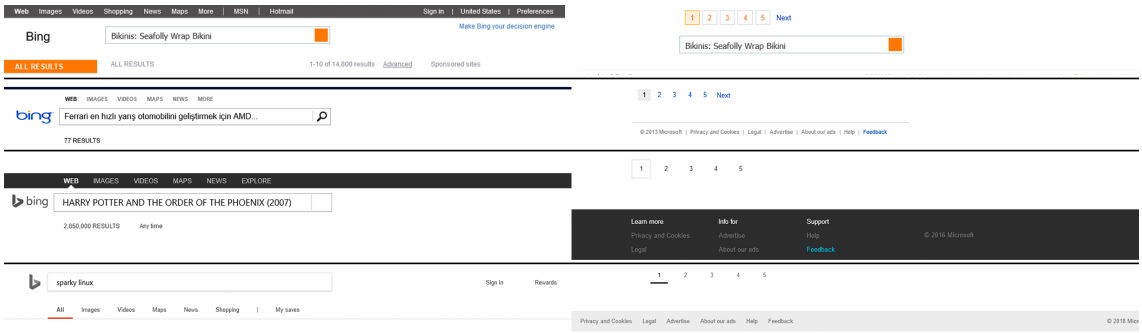


Figure 5.13: Bing's navigation & user inputs elements from 2010 (top), 2013 (mid-top), 2016 (mid-bottom) and 2018 (bottom)

5.3 Temporal Evolution

Search Engine Results Pages (SERP) have always had a large variety of elements, spread along with their layout, each with its evolution and **active time periods**, as described in Chapter 4. In this section, we sum up the detected presence of each SERP element and synthesize relevant changes to SERP identified during this study, both over time.

Figure 5.14, where each black cell means the element was found present in the respective year in column, is referent to both Google and Bing. Each colored cell marks the first appearance of the element. It is noticeable how SERP features have emerged in the last decade, contributing to a matrix full of element possibilities in the last few years. In this work, any comparison between the two search engines must be aware of the respective time windows. Unlike Google, whose data stretches back to 2000, Bing's data only sees 2010 onwards. Thus, it is also possible to conclude that over the course of these 11 years, Bing has increased the number of elements in

its SERP. In fact, many of them were inspired by previous Google releases, which justifies that the latter remains at the forefront of SERP innovation, coupled with worldwide market share leadership by a vast difference. Besides, considering the same interval for both, from 2010 to 2020, Google's filled cells relate to Bing's in a 1.65 ratio, approximately. If we restrict the analysis to the very last year, the difference is highly reduced, since Google offers 17 SERP elements, whereas Bing offers 15. A timeline regarding the presence of elements in SERP is available at: https://bedgarone.github.io/serpevolution/timeline/2010.

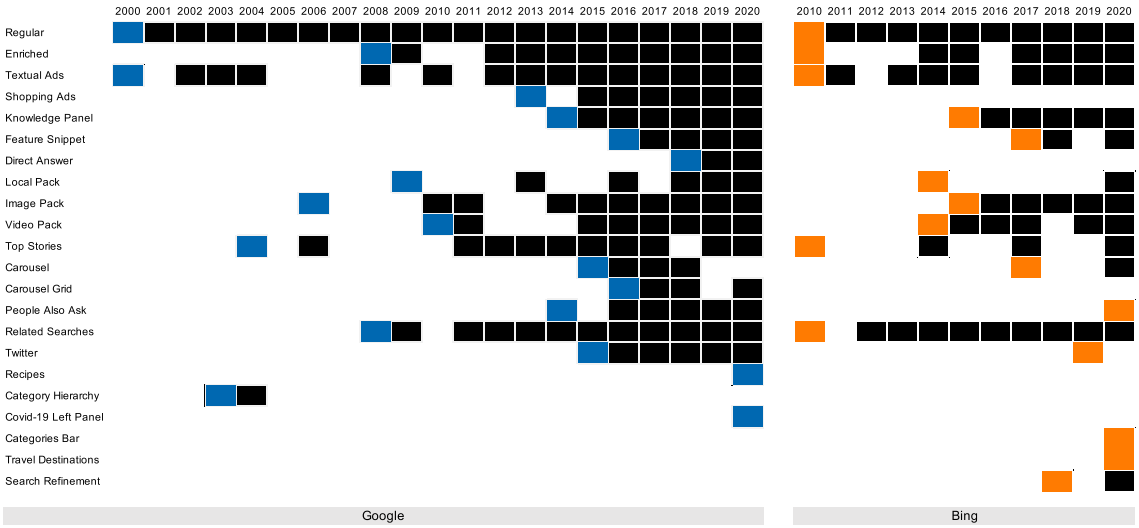


Figure 5.14: Detected presence of Google (left) and Bing (right) SERP elements over time

The most relevant changes to SERP, whether due to the entry of new elements or to significant changes in navigation and user input options, are identified in a two-decade timeline of SERP in Figure 5.15 and Figure 5.16, for Google and Bing respectively. These timelines are also available at: https://bedgarone.github.io/serpevolution/timeline.

In fact, Google created the SERP format, very early in 2000, and started taking its first steps, testing and considerably modifying the structure during the initial decade. This involved studying the placement of sidebars, shaping sponsored content, introducing the tabs bar, modifying and moving it to different locations, until including, as we saw earlier in this work, content from other tabs in the general SERP. With Google reaching solid ground, Bing enters the scene in 2020, producing a competing product, structurally similar to the live version of the Google SERPs. From that moment on, the path has been done almost in parallel, whose changes are usually made in advance by Google and followed by Bing, taking as example the tabs bar placement below the search query, the removal of the colored ads background, removal of the left sidebar or the introduction of the carousel, for example.

Hence, we could substantiate what Höchstötter and Lewandowski [26] said regarding Google being always at the forefront of presenting sophisticated pages, and being a pioneer in implementations. Interestingly, this author wrote at a time when the SERP only included simple features. This reign continued for another decade, as seen we've seen in this study.

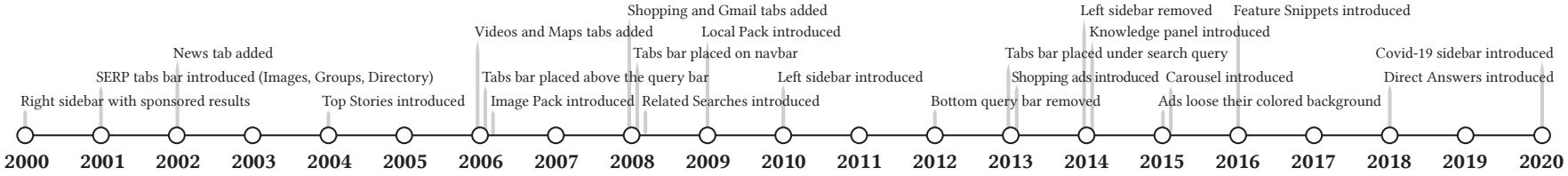


Figure 5.15: Highlights of Google's SERP overall evolution

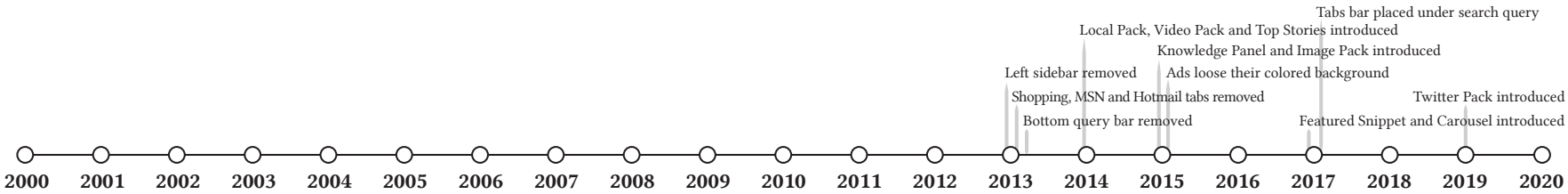


Figure 5.16: Highlights of Bing's SERP overall evolution

5.4 User Interface Design Patterns

Almost every element in Search Engine Results Pages includes well-known user interface **design patterns** to improve the user experience. These are listed in the website's *Patterns* section available at: https://bedgarone.github.io/serpevolution/patterns. In this section, we synthesize this appliance and compare the job by both engines.

Figure 5.17 maps each Google and Bing elements with each user interface design pattern already addressed. Each cell is complemented with the start date of that appliance. Older SERP elements by Google make later use of design patterns for individual improvement, whereas some recent elements may have arisen from the need to apply a design pattern solution. In those cases, we can trace the design pattern since the element's beginning. Regarding Bing, it is possible to notice a less intensive application of these patterns. In particular, it is highlighted the non-use of *Breadcrumbs* and *Progressive Disclosure* but also a lesser focus on *Carousel*.

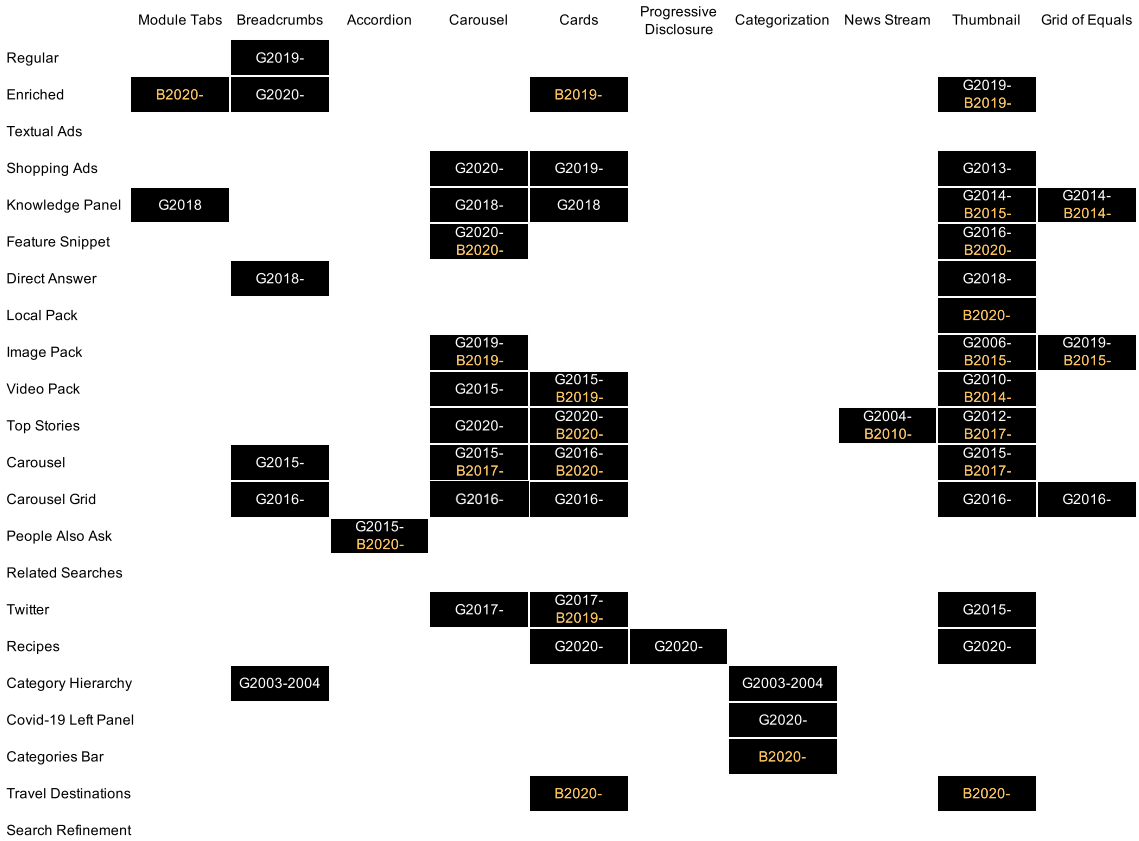


Figure 5.17: User Interface Design Patterns and time of appliance to Google (white) and Bing (orange) SERP elements

5.5 Interface area size

We calculated the area of all screenshots in the dataset to analyze its evolution over time, seen in Figure 5.18. It shows the evolution of **interface area** per month (dots) and per year (line),

measured in rem squared units. A rem unit, in most browsers, equals to 16px, the default font size of an HTML element. Each entry in the chart corresponds to the average area per month for all captures in the dataset, and it was considered the full height of the capture. The results show an increase, close to exponential. This is due to the appearance of SERP features, that have been adding extra content to SERP, thus, making them bigger over time, and also to the increase of the screen's width, already addressed in Section 3.2. Both graphs show a similar growth, but Bing's area was almost always slightly lesser than Google's. Months without values, as indicated in Table 3.1, not have captures in the dataset.

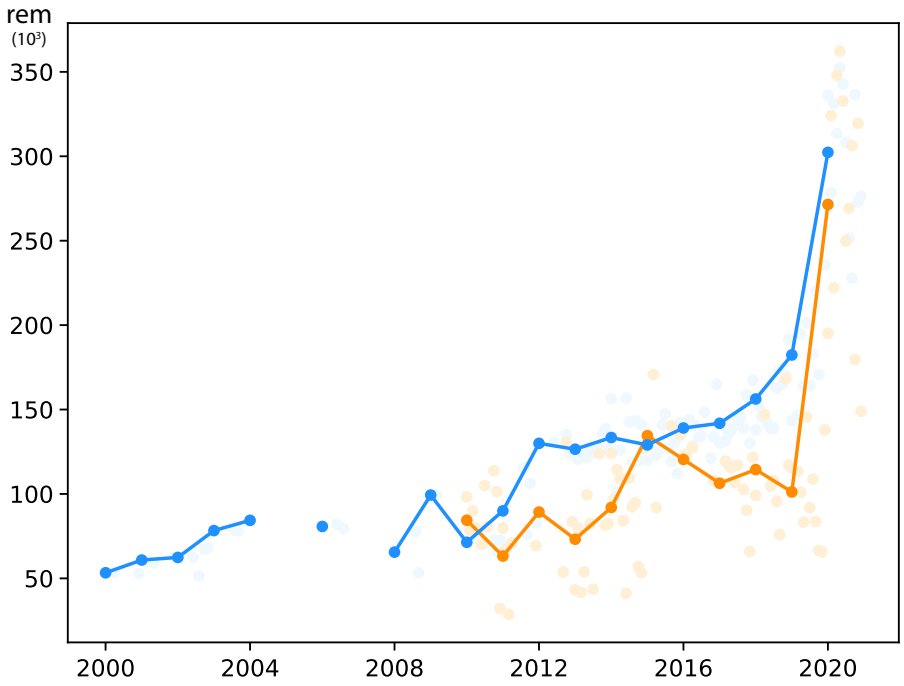


Figure 5.18: Interface area evolution for Google (blue) and Bing (orange)

5.6 Files size

We made a similar approach to study the variation of file's size regarding the entire dataset. Figure 5.19 represents the SERP captures' **file size** evolution and the added size of both the HTML file and the associated files' folder. In the Google's case, size results accompany the evolution of the interface area as seen in Figure 5.18, expressing a steep rise in the last few years. In Bing's case, this growth exists but its intensity if weak. From the comparison of these results, it is relevant to notice that Google's SERP source code is now heavier than Bing's. Besides, the difference when adding the files folder size is even bigger. One may argue that Microsoft is more successful in SERP efficiency while others may justify this difference with richer SERP by Google.

Nevertheless, as seen in this work, by 2020 both SERP are becoming simillar in terms of element varieties.

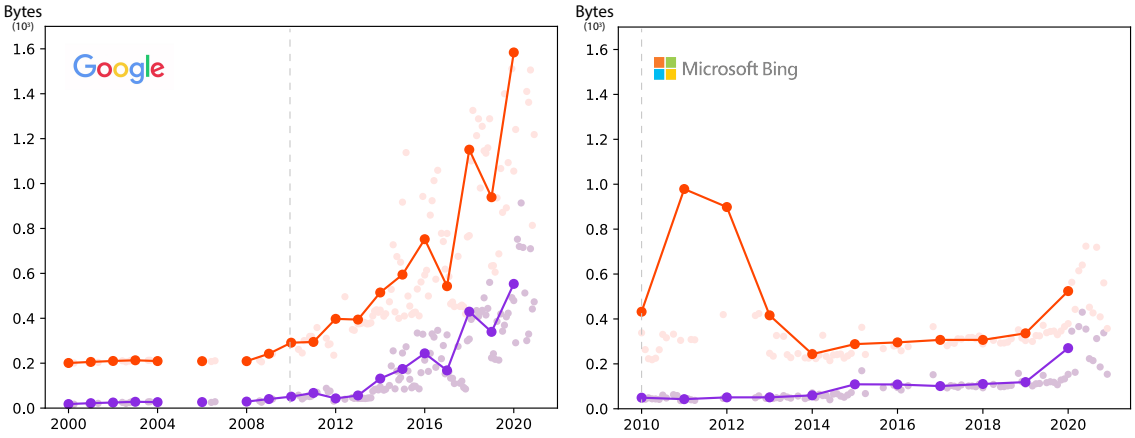


Figure 5.19: Source code size (violet), Source code + Files folder size (orange) evolution for Google (left) and Bing (right)

This increase cannot be clearly related to an increase in the number of associated files, as seen in Figure 5.20, since later values share similar values with first years of SERP existence. Nevertheless, results suggest that SERP sought to reduce the number of associated files, having both achieved this aim in an initial stage. This number started to rise again because interfaces evolved and demanded more images and graphics, that can somehow increase the files needed to load a SERP. However, similarly to the previous Figure, Google's growth is more intense and consistent, whereas Bing present a lower tendency for 2020.

Figure 5.20: Average number of files associated for Google (green) and Bing (yellow)

Chapter 6

Conclusions

Using Google and Bing as case studies, we examined how Search Engine Results Pages' (SERP) user interfaces evolved over two decades of existence. While existing research has relied on the actual states of these interfaces, we have updated and improved the analysis with an evolution perspective, addressing SERP elements, their positioning, layout, size and patterns.

We extracted and provide a dataset with 7,000+ SERP captures that include HTML versions and their screenshots. Interfaces have kept track of web development's evolution, which allowed a better-automated extraction of information for this study in more recent years, compared to rudimentary HTML early versions.

We have shown that SERP constitution has become more diversified over time, especially in recent years, to provide information effectively while keeping the user in the results page. SERP element's range has been growing continuously, although steeper in the last few years. Google was responsible for launching itself firstly on the market, in the beginning of the century. We have seen that Bing comes out a decade after with an interface similar to its big competitor, and we've shown how many of its elements were influenced by Google and appearing briefly *a posteriori* after Google's initial launch.

Despite some innovation, patent on the unique elements of each engine, we saw that the layout of both interfaces is similar for the various categories of elements analyzed. In a process that has been intensifying over time, the composition of Bing interfaces has come to resemble Google's and, as of 2020, the number of SERP elements is almost equal.

We also saw how and when user interface design patterns were applied to SERP elements, in which Google presents a greater number of elements reinforced with these solutions. In Bing, it was possible to conclude a lesser use of design patterns, noting how recent these strategies occur in its elements, only from 2017 onwards.

Finally, we have shown how interface area increased almost exponentially, noting that Google's SERP area was always slightly higher in the common comparison period. In terms of source code and files' weight, Bing presents less heavy pages, especially regarding associated files, where there is a greater efficiency in the storage of graphic information that has recently become as demanding as the competitor's.

6.1 Future Work

Regarding the extraction of interfaces, unfortunately, despite that Internet Archive has billions of varied pages captures, these highlighted by their diversity, SERP were not captured consistently for many years as, for example, search engine homepages were. Achieving the right interfaces for an evolutionary analysis requires an additional effort in order to study how to efficiently collect and reach more of these captures, while testing other methods to find the most representative queries for triggering SERP features over time.

Web search engine variety is beneficial for studying SERP elements. Despite a lower rank in market share, there will certainly be exclusive bets for new interface elements from other search engines not incorporated in this work. Extending the analysis to these search engines could lead to more element contributions.

Accompanying the evolution of web development is the coding of SERP elements. Its detection is a complex process that could be studied further, especially for the automation of this detection, while analysing SERP structure and hierarchy.

This study's target was the desktop versions of SERP. Nowadays, many advocate a web development paradigm focused primarily on the mobile version. It is of great interest to replicate this analysis in portable versions such as tablets and smartphones, which are other common means for consulting the information provided by Web search engines.

Finally, noting that interfaces are static, it is not possible to analyze and assess user interaction. However, this is a very important field for the user experience. Thus, it is valuable to complement this analysis with an overview of the current state of interfaces concerning their interaction and usability.

Appendix A

Element Identifiers

Organic results

Regular results	
.g	2003-2020
p	2000-2002

Enriched results	
.nrgt	2015-2020
td div.sld	2012-2014
div.std table.slk	2008, 2009

Sponsored results

Textual ads	
#tads	2008, 2010, 2012-2020
#tpa1	2002-2004, 2006
#tpa2	2002, 2003, 2006
#taw1	2002-2004
td[bgcolor="#FFF5F6"]	2000, 2002-2004
td[bgcolor="#D4F0FC"]	2000
td[bgcolor="#D7FFD7"]	2000

Shopping ads	
.DALGre	2020
#rhs_block .ts	2011, 2013, 2015-2019

Features

Knowledge Panel	
.knowledge-panel	2016-2020
.kp-wholepage	2020
.kp-blk	2014-2020
.kno-kp	2014-2020
#rhs_block .hp-xpdbox	2014-2017, 2019

Featured Snippets	
#res .NFQFxe.oHglmf	2017-2020
#knocube	2016, 2018
#res .hp-xpdbox	2016-2018
.slk	2008, 2009, 2012-2017, 2019

Direct Answer Results	
.kp-blk.EyBRub.fm06If	2019, 2020
#res .N6Sb2c.i29hTd	2018-2020

Local Pack	
.xERobd	2016, 2018-2020
.g > div.mlmcm	2013
table.ts	2008-2020

Image Pack	
.LnbJhc	2019, 2020
#NzwoZe	2019
#imagebox_bigimages	2014-2019
#imagebox	2010, 2011

Video Pack	
.COEoid	2015-2020
#videobox	2011
.g.videobox	2010, 2011

Top Stories	
g-scrolling-carousel.F8yfEe	2020
.g > div > table.ts	2011-2019
#newsbox	2011
p.e	2003, 2004, 2006

Carousel	
.klbar	2015-2018

Carousel Grid	
.vsQKGc	2016-2018, 2020

People Also Ask	
.kp-blk.cUnQKe	2014, 2016-2020

Related Searches	
#bres > #bzMwOe	2020
#brs	2008, 2009, 2011, 2014-2020

Twitter	
g-scrolling-carousel.rQgnxe	2017-2020
g-scrolling-carousel._mip	2017
g-flippy-carousel._nGh	2015, 2016

Recipe Cards	
.MmzWWe	2020

Category Hierarchy	
p.e	2003, 2004

Covid-19 Left Panel	
#Yf1RJc	2020

Figure A.1: Google elements' identifiers

Microsoft Bing

Organic results

Regular results	
[class="b_algo"]	2014-2020
#results [class="sa_wr"]	2010-2014
#results [class="sa_cc"]	2010-2014

Enriched results	
#rpd	2020
.b_vlist2col.b_deep	2014-2015, 2017-2020
.b_overhangR	2018
.sb_vdl	2010-2014

Sponsored results

Textual ads	
.b_ad	2014, 2015, 2017-2020
.b_adBottom	2015, 2017-2020
.sb_adsWv2	2010, 2011, 2013, 2014
.sb_adsW	2010
.sb_adsW2	2010

Features

Knowledge Panel	
#b_context > li.b_ans > .b_entityTP	2015-2020

Featured Snippets	
#d_ans	2017, 2018, 2020

Local Pack	
#lMapContainer	2020
#maps_container	2014

Image Pack	
.b_rich > .iaplanner	2018-2020
.b_rich > #iaplanner	2015-2017

Video Pack	
#serpvidans	2020
#vidans2	2014-2017, 2019, 2020

Top Stories	
.ans_nws	2020
#ans_news	2017
.ans_bd	2010, 2014

Carousel	
#ent-car-exp	2017,2020

People Also Ask	
.rqnaContainerwithfeedback	2020

Related Searches	
.b_rs	2014-2020
.sb_title.rrrs_ttl	2012-2014
#sw_rel	2010

Twitter	
.twcwr	2019

Categories Bar	
.ent-dtab-content	2020

Travel Destinations	
#destCar	2020

Search Refinement	
.b_cnvsug	2018, 2020

Figure A.2: Bing elements' identifiers

References

[1] Desktop Screen Resolution Stats Worldwide | StatCounter Global Stats. https://gs.statcounter.com/screen-resolution-stats/desktop/worldwide.

[2] The next phase. https://blogs.bing.com/search/2013/09/16/the-next-phase/, journal=Bing, publisher=Microsoft, author=Ripsher, Lawrence, Sep 2013.

[3] The size of the World Wide Web (The Internet). https://www.worldwidewebsize.com/, 2021.

[4] Marco Allegretti, Yashar Moshfeghi, Maria Hadjigeorgieva, Frank E. Pollick, Joemon M. Jose, and Gabriella Pasi. When Relevance Judgement is Happening? pages 719–722, 2015.

[5] Mostafa Alli. SERP-level disambiguation from search results. *IC3K 2015 - Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, 1(Ic3k):627–636, 2015.

[6] Khalil Amjad and Fadi Alrub. A comparison of search engine's features and mechanism. (November):1–8, 2014.

[7] Peter G Anick. Adapting a Full-text Information Retrieval System to the Computer Troubleshooting Domain. Technical report.

[8] Jaime Arguello, Wan Ching Wu, Diane Kelly, and Ashlee Edwards. Task complexity, vertical display and user interaction in aggregated search. *SIGIR'12 - Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 435–444, 2012.

[9] Ricardo Baeza-yates and Berthier Ribeiro-neto. Modern information retrieval: the concepts and technology behind search. *Choice Reviews Online*, 48(12):48–6950–48–6950, 2011.

[10] Loren Baker. Google and Yahoo Users Prefer Organic Search Engine Results. https://www.searchenginejournal.com/google-and-yahoo-users-prefer-organic-search-engine-results/525/, 2004.

[11] Georg Buscher, Susan Dumais, and Edward Cutrell. The good, the bad, and the random: An eye-tracking study of ad quality in web search. *SIGIR 2010 Proceedings - 33rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 42–49, 2010.

[12] Alex Chris. Top 10 Search Engines In The World (2021 Update). https://www.reliablesoft.net/top-10-search-engines-in-the-world/, 2020.

[13] Cisco. Cisco Annual Internet Report (2018–2023) White Paper. https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html#Executivesummary, 2020.

[14] Douglass R. Cutting, David R. Karger, Jan O. Pedersen, and John W. Tukey. Scatter/gather: a cluster-based approach to browsing large document collections. *SIGIR Forum (ACM Special Interest Group on Information Retrieval)*, (May 1996):318–329, 1992.

[15] Cristian Danescu-Niculescu-Mizil, Andrei Z. Broder, Evgeniy Gabrilovich, Vanja Josifovski, and Bo Pang. Competing for users' attention: On the interplay between organic and sponsored search results. *Proceedings of the 19th International Conference on World Wide Web, WWW '10*, pages 291–300, 2010.

[16] Danny Sullivan. Google's Knowledge Graph and Knowledge Panels. https://blog.google/products/search/about-knowledge-graph-and-knowledge-panels/, may 2020.

[17] Alan Dix, Janet Finlay, Gregory D. Abowd, and Russell Beale. *Human Computer Interaction*. Pearson Prentice Hall, 3 edition, 2003.

[18] Victoria Edwards. Google in 1998: New Easter Egg Takes You Back in Time. https://www.searchenginewatch.com/2013/09/26/google-in-1998-new-easter-egg-takes-you-back-in-time/, 2013.

[19] Israt Jahan Farzana. Towards a user-centric entity search Israt Jahan Farzana. (March 2013), 2014.

[20] Anindya Ghose and Sha Yang. An empirical analysis of sponsored search performance in search engine advertising. In *WSDM'08 - Proceedings of the 2008 International Conference on Web Search and Data Mining*, pages 241–250, New York, New York, USA, 2008. ACM Press.

[21] Google. Google Launches Self-Service Advertising Program. http://googlepress.blogspot.com/2000/10/google-launches-self-service.html, 2000.

[22] Google. Information about Sitelinks. https://developers.google.com/search/blog/2006/09/information-about-sitelinks, 2006.

[23] Zhiwei Guan and Edward Cutrell. An eye tracking study of the effect of target rank on Web search. *Conference on Human Factors in Computing Systems - Proceedings*, pages 417–420, 2007.

[24] Marti A. Hearst. *User Interfaces and Visualization. Modern Information Retrieval.* 1999.

[25] Marti A Hearst and Jan O Pedersen. Reexamining the Cluster Hypothesis: Scatter/Gather on Retrieval Results *. In *Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '96*, New York, New York, USA. ACM Press.

[26] Nadine Höchstötter and Dirk Lewandowski. What users see - Structures in search engine results pages. *Information Sciences*, 179(12):1796–1812, 2009.

[27] InternetLiveStats. Google Search Statistics - Internet Live Stats. https://www.internetlivestats.com/google-search-statistics/, 2021.

[28] Bernard J. Jansen, Danielle L. Booth, and Amanda Spink. Determining the user intent of web search engine queries. *16th International World Wide Web Conference, WWW2007*, (May):1149–1150, 2007.

[29] Thorsten Joachims, Laura Granka, Bing Pan, Helene Hembrooke, and Geri Gay. Accurately Interpreting Clickthrough Data as Implicit Feedback. *ACM SIGIR Forum*, 51(1):4–11, 2017.

[30] Mark Johnson. Researching with bing reference. http://blogs.bing.com/search/2009/07/27/researching-with-bing-reference, Jul 2009.

[31] Jon Penland. The History of Web Search Engines: What Came Before Google? - WhoIsHostingThis.com. https://www.whoishostingthis.com/resources/history-search-engines/, apr 2020.

[32] Yvonne Kammerer and Peter Gerjets. How the interface design influences users' spontaneous trustworthiness evaluations of web search results: Comparing a list and a grid interface. *Eye Tracking Research and Applications Symposium (ETRA)*, 1(212):299–306, 2010.

[33] M.D. Kickmeier and Dietrich Albert. The Effects of Scannability on Information Search: an Online Experiment. *Proceedings of the 17th British HCI Group Annual Conference*, 2:33–36, 2003.

[34] Michele Kimpton and Jeff Ubois. Year-by-year: From an archive of the internet to an archive on the internet. In *Web Archiving*, pages 201–212. Springer Berlin Heidelberg, 2006.

[35] Peter Lyman, Hal R. Varian, Kirsten Swearingen, Peter Charles, Nathan Good, Laheem Lamar Jordan, and Joyojeet Pal. How much information? 2003. Technical report, 2003.

[36] Peter J. Meyers. Mega-SERP: A Visual Guide to Google - Moz. https://moz.com/blog/mega-serp-a-visual-guide-to-google, 2013.

[37] Silvana Drumond Monteiro, Cecílio Merlotti Rodas, and Silvana Aparecida Borsetti Gregório Vidotti. A busca e o eye tracking: um olhar semiótico sobre o knowledge graph. *Em Questão*, 26(1):304–326, 2020.

[38] Kate Moran and Cami Goray. The Anatomy of a Search-Results Page. https://www.nngroup.com/articles/anatomy-search-results-page/, 2019.

[39] Peter Morville and Jeffery Callender. *Search Patterns: Design for Discovery*. O'Reilly Media, Inc., 1st edition, 2010.

[40] Kevin Mullet and Darrell Sano. *Designing Visual Interfaces: Communication Oriented Techniques*. Prentice-Hall, Inc., USA, 1995.

[41] Scott Nicholson, Tito Sierra, U. Yeliz Eseryel, Ji Hong Park, Philip Barkow, Erika J. Pozo, and Jane Ward. How much of it is real? Analysis of paid placement in web search engine results. *Journal of the American Society for Information Science and Technology*, 57(4):448–461, 2006.

[42] Jakob Nielsen and Marie Tahir. *Homepage Usability: 50 Websites Deconstructed*. New Riders Publishing, USA, 2001.

[43] Peter Norvig. English letter frequency counts: Mayzner revisited. http://norvig.com/mayzner.html.

[44] Sérgio Nunes, André Almeida, Bruno Giesteira, Fernanda Pinto, and Paula Oliveira. Interface evaluation of Portuguese and international news websites. *MCCSIS 2007 - IADIS Multi Conference on Computer Science and Information Systems - Proceedings of Intelligent Systems and Agents 2007, Interfaces and Human Computer Interaction 2007, Computer Graphics and Visualization 2007*, (November 2005):109–116, 2007.

[45] Vicki L. O'Day and Robin Jeffries. Orienteering in an information landscape. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '93*, pages 438–445, New York, New York, USA, 1993. Association for Computing Machinery (ACM).

[46] Laura Ortiz-Chaves, Eduardo Martínez-De-Pisón, Germán Cancela-López-Carrión, Juliano Gonçalves-De-Vasconcellos, Cristòfol Rovira, and Mari Carmen Marcos. AdWords, imágenes y ceguera a los banners. Un estudio con eye tracking. *Profesional de la Informacion*, 23(3):279–287, 2014.

[47] Lori Phelps. Active Documentation: Wizards as a Medium for Meeting User Needs. In *Proceedings of the 15th annual international conference on Computer documentation - SIGDOC '97*, New York, New York, USA. ACM Press.

[48] Zuzana Pinkosova, William J. McGeown, and Yashar Moshfeghi. The Cortical Activity of Graded Relevance. *SIGIR 2020 - Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, (i):299–308, 2020.

[49] Larisa Rosu. The Anatomy of a Google Search Results Page in 2020 - Advanced Web Ranking. https://www.advancedwebranking.com/serp/, 2020.

[50] Sally Salas. Bing gets a new look. https://blogs.bing.com/search/2012/05/02/bing-gets-a-new-look, May 2012.

[51] Gerard Salton. *Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*. Addison-Wesley Longman Publishing Co., Inc., USA, 1989.

[52] Tefko Saracevic. Relevance reconsidered. *Proceedings of the second conference on conceptions of library and information science (CoLIS 2)*, 201:18, 1996.

[53] Barry Schwartz. Google Knowledge Graph: Worldwide, Carousel & Instant . https://www.seroundtable.com/google-knowledge-graph-expanded-15543.html, 2012.

[54] Barry Schwartz. Google Testing Dropping Local Carousel. https://www.seroundtable.com/google-carousel-gone-19021.html, 2014.

[55] Barry Schwartz. Google AdWords Automated Extension For Shopping Ads (PLAs) Live. https://www.seroundtable.com/google-adwords-automated-extension-for-shopping-ads-plas-20779.html, 2015.

[56] Barry Schwartz. Bing Testing Twitter Carousel In Search Results. https://www.seroundtable.com/bing-twitter-carousel-22051.html, 2016.

[57] Barry Schwartz. Google Search Console Add Rich Cards Section. https://www.seroundtable.com/google-search-console-add-rich-cards-22099.html, 2016.

[58] Barry Schwartz. Google Launches Dynamic Loading People Also Ask. https://www.seroundtable.com/google-dynamic-loading-people-also-ask-23399.html, 2017.

[59] Yonik Seeley. What is Faceted Search? http://web.archive.org/web/20140316215652/http://searchhub.org/2009/09/02/faceted-search-with-solr/, 2009.

[60] Ahmed Seffah. The evolution of design patterns in hci: from pattern languages to pattern-oriented design. 01 2010.

[61] Ahmed Seffah and Mohamed Taleb. Tracing the evolution of hci patterns as an interaction design tool. *Innovations in Systems and Software Engineering*, 8, 06 2012.

[62] Tom Seymour, Dean Frantsvog, and Satheesh Kumar. History Of Search Engines. *International Journal of Management & Information Systems (IJMIS)*, 15(4):47, 2011.

[63] Ben Shneiderman. Designing the user interface strategies for effective human-computer interaction. *ACM SIGBIO Newsletter*, 9(1):6, mar 1987.

[64] Ben Shneiderman, Donald Byrd, and W. Bruce Croft. Sorting out searching a user-interface framework for text searches. *Communications of the ACM*, 41(4):95–98, apr 1998.

[65] Amit Singhal. Introducing the Knowledge Graph: things, not strings. https://blog.google/products/search/introducing-knowledge-graph-things-not/, 2012.

[66] Gregory Smyth. The Importance of Search Engines. https://www.inetasia.com/resources/articles-the-importance-of-search-engines.html.

[67] Martin Stoddart. Explore pane – a ribbon for search. https://blogs.bing.com/search/2009/06/17/explore-pane-a-ribbon-for-search, Jun 2009.

[68] Danny Sullivan. Searching With Invisible Tabs. https://www.searchenginewatch.com/2003/12/02/searching-with-invisible-tabs/, 2003.

[69] Teoalida. Screen resolution statistics. https://www.teoalida.com/webdesign/screen-resolution/, 2019.

[70] Jenifer Tidwell. *Designing Interfaces*. O'Reilly Media, Inc., 2010.

[71] Stella D. Tomasi. Using cues to forage for information on the Web. *Journal of Systems and Information Technology*, 16(4):296–312, 2014.

[72] Anders Toxboe. Design patterns. http://ui-patterns.com/patterns.

[73] Edward Tufte. The Visual Display of Quantitative Information / E.R. Tufte. *American Journal of Physics*, 31, 2001.

[74] Steven Vaughan-Nichols. History of search engines: The internet before Google | HPE | HPE. https://www.hpe.com/us/en/insights/articles/how-search-worked-before-google-1703.html.

[75] Kristen Vaughn. SERP Features in 2019: The Complete Guide | KoMarketing. https://komarketing.com/blog/serp-features/, 2019.

[76] J Wardini. Google Search Statistics – How Many Google Searches Per Day? https://serpwatch.io/blog/how-many-google-searches-per-day/, 2019.

[77] Bo Xing and Zhangxi Lin. The Impact of Search Engine Optimization on Online Advertising Market. pages 519–529, 2005.

[78] Yunwen Ye and G. Fischer. Supporting reuse by delivering task-relevant and personalized information. (May 2014):513–523, 2003.