

Review

# From Web Catalogs to Google: A Retrospective Study of Web Search Engines Sustainable Development

Mariusz Duka <sup>1</sup>, Marek Sikora <sup>1,2</sup> and Artur Strzelecki <sup>3,\*</sup>

<sup>1</sup> Faculty of Automatic Control, Electronics and Computer Science, Silesian University of Technology, Akademicka 16, 44-100 Gliwice, Poland

<sup>2</sup> Lukaszewicz Reasearch Network, Institute of Innovative Technologies—EMAG, Leopolda 31, 40-189 Katowice, Poland

<sup>3</sup> Department of Informatics, University of Economics in Katowice, 1 Maja 50, 40-287 Katowice, Poland

\* Correspondence: artur.strzelecki@ue.katowice.pl

**Abstract:** This study presents a review of search engines and search engine optimization and shows how the search engine landscape relates to sustainable development. We have used a narrative review research method and described three main topics: the past and present of web catalogs and search engines; current knowledge about the dominant types of search results presented in Google search; and methods of search engine optimization. Technical elements of important website areas related to technical website auditing are discussed. We summarize our research with several key findings on how web search engines are involved in sustainable development and offer a glimpse into the future use of web searching with the help of artificial intelligence chats and prompt engineering.

**Keywords:** web search engine; web catalog; search engine results page; search engine optimization; SEO; sustainable development



**Citation:** Duka, M.; Sikora, M.; Strzelecki, A. From Web Catalogs to Google: A Retrospective Study of Web Search Engines Sustainable Development. *Sustainability* **2023**, *15*, 6768. <https://doi.org/10.3390/su15086768>

Academic Editor: Leonidas Hatzithomas

Received: 22 February 2023

Revised: 27 March 2023

Accepted: 11 April 2023

Published: 17 April 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

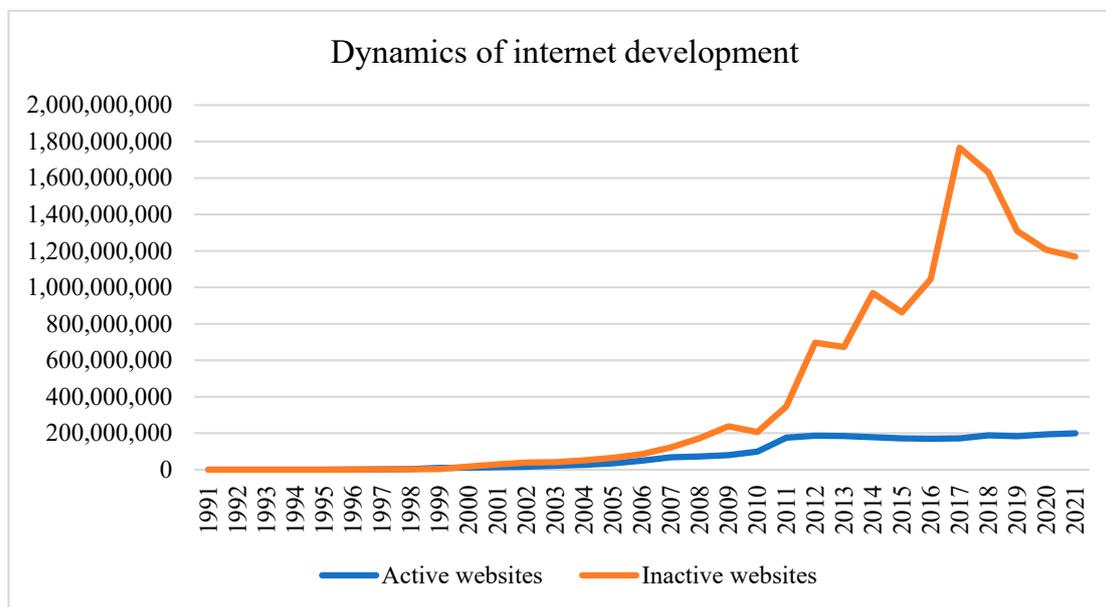
## 1. Introduction

Search engines play an integral role in the lives of many people, who often do not realize it. They have become powerful tools for obtaining useful information scattered across the internet [1]. A simple purchase on Alibaba or eBay usually begins with a search for a product. Likewise, booking a vacation often begins with a search for a hotel.

The internet, as we know it, existed for many years before the first webpage appeared. Internet Protocol (IP), as well as the suite of communication protocols on which the internet is built (TCP/IP), are technologies invented by Bob Kahn and Vint Cerf back in the early 1970s at the U.S. Department of Defense's Advanced Research Projects Agency (DARPA) [2]. The military nature of this technology made it unavailable to the civilian sector, except for local academic networks clustered in the common ARPANET network [3]. The internet service industries boomed when the TCP/IP protocol came into commercial use in the early 1990s. Fundamental at the time was the fact that Tim Berners-Lee had invented WWW technology in 1989. Berners-Lee used the existing HTTP and TCP/IP protocols to build a system consisting of the first web server and web directory, called the WorldWideWeb [4]. This system gave each webpage a unified address format (Uniform Resource Locator—URL), which then made it publicly accessible on the web. Berners-Lee also created HTML, based on SGML-CERN tags, for formatting textual content.

Since 1991, when the first webpage (info.cern.ch) was created [5], the HTML framework and web servers have appeared, and the internet space has been filled with more than 1.8 billion webpages on more than 233 million unique domains. UK-based Netcraft has been monitoring internet resources since 1995, taking into account the technologies used in websites and the software on hosting providers' servers. Netcraft calculates the number of active websites, i.e., those presenting specific content, based on analysis of websites' IP addresses and source code, excluding "under construction" websites, redirects, or domains

indicating identical content, e.g., with a “parking” service enabled at the registrar. Based on Netcraft’s periodic analysis, it is estimated that there are more than 200 million active sites on the internet. The dynamic of the development of internet resources is shown in Figure 1.



**Figure 1.** Dynamics of internet development from 1991 to 2021, compiled using data from netcraft.com.

The need for search engines or systems that assist users in finding specific information on the internet arose due to the increasing number of new websites. These systems collect data by analyzing webpage content and hyperlinked network topology. Many of the original search engines no longer exist, with new ones springing up in their place that are trying to catch up with the current leaders. Over the years, competition has led to the ability to search for multimedia content as well; this initially caused many challenges for algorithms, but today this functionality is built into almost every professional search engine [6].

There are well known problems of maintaining sustainability in a web environment. One of them is the need to ensure that web resources remain available and functional even after the organizations that created them cease to exist or have no resources to host them. There is a need to find ways to make web sites more sustainable, including transferring them to free hosting options, ensuring that they can be viewed and updated even without paid resources, and making them easy to manage and transfer ownership. The second problem is sustainability in web design, specifically how web designers and developers can apply sustainable practices in their work to minimize the environmental impact of websites [7]. It covers various aspects of sustainability such as energy efficiency [8], reducing materials consumption, and optimizing web performance. The third problem is the issue with practical solutions for creating sustainable websites, which include using green hosting [9], developing with modern tech stacks or updated plugins, designing for mobile, reducing images and videos, limiting fonts, and using fast and light website themes. These solutions can contribute to making websites more energy-efficient, accessible, and reducing inequality, which aligns with global sustainability goals.

In this study, we would like to review how sustainability was introduced by commercial web catalogs and search engines in past decades. There has been a lot of research on how search engines work and what their impact is on society. However, the topic of sustainable development for internet search tools has not been explored yet. The proposed method for achieving this goal is a narrative review.

We propose three general research questions to which we wanted to find answers: “what was the past and what is the present of web catalogs and search engines and how do they relate to sustainable development?”, “how has the presentation of search engine result pages changed and how it relates to sustainable development?”, and “what search engine optimization methods are used, and how do they contribute to sustainable development?”. These three topics are key concepts for which we will do the narrative review.

## 2. Materials and Methods

In this study we have used a narrative review research method. A narrative review is classified as a secondary research study since it uses existing primary research studies. A narrative review usually addresses one or more questions with a broad scope, and the selection criteria for the inclusion of the articles may not be specified explicitly. A narrative review does not have a strict protocol to be followed. A review’s design depends on its author and objectives. There is no consensus on the standard structure of narrative reviews, but their primary purpose is to deepen understanding of a certain research area by identifying and summarizing what has been published. Its general application includes exploring existing debates, appraising previous studies conducted on a topic, identifying knowledge gaps, and speculating on the latest interventions. Narrative reviews are also used to track and report on changes in an existing research field.

We searched the Web of Science and Scopus scientific databases for papers whose metadata included the keywords “web search engine” and “search engine optimization”. We set a timespan without any limit. These criteria resulted in around 530 published papers (after duplicates and papers not written in English were removed). We read all the abstracts; when the topic was “web search engine” or “search engine optimization”, we read the paper and took its results to present in this study. We narrowed the list of papers for this review to around 100. From the narrative review, we derived three general topics that are key concepts: “web catalogs and search engines”; “search engine result pages; and search engine optimization.

## 3. Results

### 3.1. Web Catalogs and Search Engines

Starting with the Archie search engine, which was launched in 1990 and is considered the first search engine [10], through W3Catalog, WebCrawler, and Lycos, the mid-1990s saw the emergence of many new search engines, such as Excite, AltaVista, and Yahoo! This was when their indexes, data [11], and the quality of their results [12] were first analyzed.

The most popular to this day is the Google search engine, which was established in 1998 and set new standards for indexing and ranking web resources with its PageRank algorithm [13]. After a year of operation, the number of indexed URLs in the search engine was more than 350 million [14], and this number rose rapidly to billions in the following years [15]. It is estimated that in 2000, the number of servers supporting the Google search engine was 25,000, and in 2010 the number had increased to 900,000 [16]. In the same year as Google, the DMOZ web directory (multilingual web directory, active from 1998 to 2017), which in its best period had a database of more than 5 million webpages, was established. Inclusion in the DMOZ directory was often equated with higher ranking in search engines, while prestige was added by the fact that selected specialists manually verified all added sites. The timing of the emergence of the first search engines and web directories also applies to social networking sites; for example, the “Six Degrees” service was established in 1997 [17].

Beginning with the first search engines, content from websites is downloaded [18] and indexed using bots (crawlers). Back then, due to the small number of sites and search engines, bots did not generate as much traffic as they do today [19]. Over the years, new algorithms have emerged to improve bot performance [20], and their behavior and the traffic they generate for websites [21] have been measured and compared [22]. Bots have also been designed for semantic content analysis [23]. Since bot activity can raise concerns

about the security and performance of webservers [24], many studies have isolated the traffic characteristics that distinguish bots from real users [25] and methods have been developed to automatically classify them. In addition, the increasing number of new websites has required the development of methods of classification [26], categorization [27], and bots' recommendations so that search results are as responsive as possible to specific queries [28].

Microsoft also took on the global search engine market, launching its MSN Search at the same time as Google, hoping that the high popularity of the Windows operating system would translate into an increase in the popularity of this new search engine. Unfortunately, despite increased marketing activity and name changes to Microsoft Live in 2006, Live Search in 2007, and finally to Bing in 2009, this search engine has failed to gain more than a 10% share in the search engine market, even despite generating more precise location-sensitive results than Google's search engine [29].

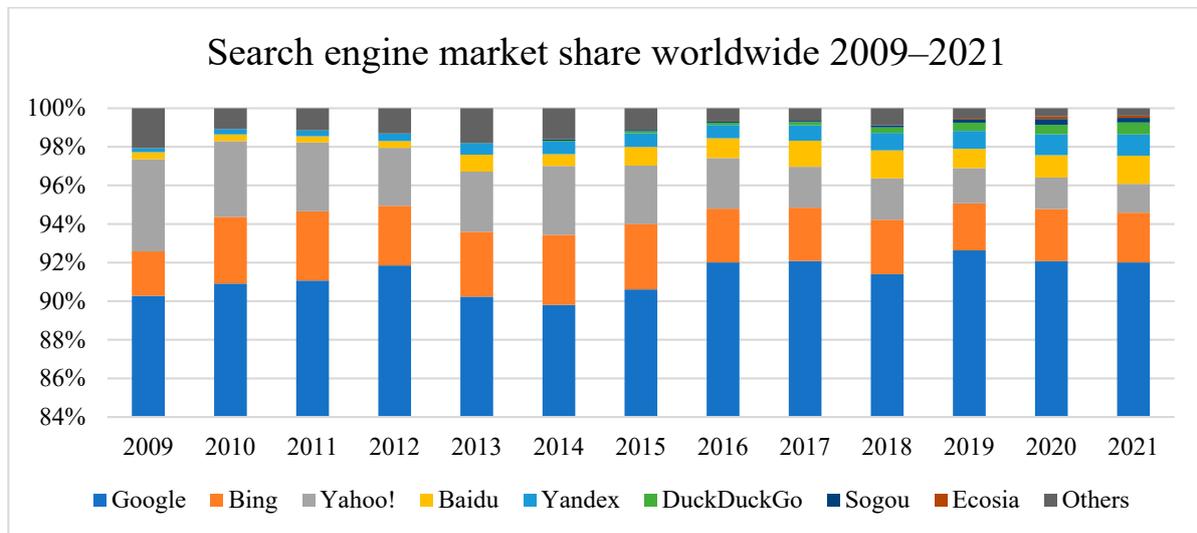
Although the Google, Bing, or Yahoo! search engines have been world leaders for years, there are local search engines that have historically had or still have a larger market share in some countries [30]. Such search engines certainly include Russia's Yandex, founded in 1997, and China's Baidu, founded in 2000. Yandex is sometimes more popular than Google in Russia due to its better understanding of Russian grammar, which sometimes results in more accurate search results for Russian websites. With the proliferation of Android and the Chrome browser, Google's search engine has gained an increasing share of the Russian market, particularly among younger and more advanced users. The only country where Google's search engine is not dominant is China. The most popular search engines in the Middle Kingdom include Baidu with a 65.6% share (based on StatCounter.com, August 2022), Bing with a 10.25% share, and Sogou with a 9.06% share, while Google's search engine share is only 3.15%. In the case of the Chinese market, Google's search engine has had a limited field of operation, mainly for political reasons [31], which has prevented it from developing significantly in this market.

The emergence of new search engines that try to compete in a market dominated by Google is extremely difficult and requires a great deal of ingenuity on the part of their developers [32]. It is currently difficult to imagine a viable global competitor to Google (or to Baidu in China). Against this background, the DuckDuckGo and Ecosia search engines are interesting projects; however, their market shares amount to just a few fractions of a percent, so they are still trying to find their way in the global search engine market. The launch of the DuckDuckGo search engine in 2009 was focused on privacy and not collecting any information that identifies the user. Unlike other search engines, when a particular search result is clicked, it redirects in a way that prevents the search keywords from being sent to the linked website. As a result, websites do not know how a given user found them on the internet. Similarly, the Ecosia search engine, which was founded in 2009, focuses on privacy and does not collect information about its users. It is worth noting that the idea behind its market presence was to donate a portion of its ad revenue to social organizations dedicated to planting trees in Brazil, Ethiopia, Burkina Faso, Madagascar, Senegal, and Indonesia, among others. Unlike other search engines that collect their own data, Ecosia shares results provided by Bing and Yahoo!

Global search engine market shares from 2009 to 2021 are shown in Figure 2. Yandex search engine data includes global and local searches on the yandex.ru domain. The AltaVista, Ask, and WebCrawler search engines have been omitted due to their residual share.

Not only have search engines and their ranking algorithms changed over the years, but how interfaces and search results are used has also changed. An analysis of eye-tracking studies on search engines conducted between 2004 and 2019 indicates a progressive change in the way search engine results are perceived in search engines [33,34]. Initially, the content was perceived in an F-like format; this was the foundation of the golden triangle concept, which included hyperlinks to the most visited webpages. Sponsored links, which had previously been located on the right side of search results, were eventually moved into this F-shaped space. An important conclusion of eye-tracking research analysis is that there is a

tendency to change search engines' interfaces by introducing new functionalities as a result of the progressive increase in the share of mobile devices used to search for information.



**Figure 2.** Search engine market shares worldwide 2009–2021, compiled from StatCounter.com data.

Searching on mobile devices is natural nowadays: according to a Google report, (the report covered 10 countries including the US and Japan) the prevalence of such searches exceeded the number performed on desktop computers in 2015. This trend continues today, with cell phones and tablets becoming the primary search devices. In the global ranking, mobile devices hold a 59.25% (based on StatCounter.com, August 2022) share of the number of searches, while desktop devices and tablets hold 38.53% and 2.22%, respectively. It is worth mentioning that mobile search capability became available in 1998 when the Wireless Application Protocol (WAP) was introduced. This was the first mobile web browsing feature to appear on cell phones. The development of wireless data transfer standards saw WAP technology displaced by GPRS, then EDGE and HSDPA, ending with LTE, 4G, and 5G. The proliferation of mobile devices, in effect, means that anyone can carry a small computer with them, allowing access to unlimited searching and information viewing. The strong position of the Android operating system is significant because, through its built-in Chrome browser, it reinforces Google's position as the leading search engine.

The development of search engines also includes the possibility of voice search instead of typing keywords into a web browser. While this possibility seems interesting, it is often criticized because, although search engine-generated voices sound realistic, few people want to listen to a computer read them a webpage of the top ten search results. Search engines try to improve the accuracy of answers with semantic search [35], which involves trying to understand natural language; however, given the number of languages, this is a very difficult task [36].

The future of search engines, in addition to being implemented in systems based on the Internet of Things (IoT) [37], will certainly involve the use of artificial intelligence [38], among other things, to compare and purchase services such as flights and hotel deals [39]. Products, shopping carts, and payments could be presented directly on the search results page. Search engines would receive a commission on the total transaction, not just a per-click fee.

Today, the global leader among search engines is undoubtedly Google, which, after 25 years of development, takes first place with a share of 92.0% (based on StatCounter.com, August 2022). The other search engines trail a long way behind: Bing (3.34%), Yahoo! (1.34%), Yandex (0.96%), Baidu (0.83%), and others (0.8%). At the same time, it should be noted that the presented figures were estimated based on traffic analysis of only a control group of websites. Given the peculiarities of the Chinese market in terms of the restrictions

imposed on search engines [40] and the massive population, it can be assumed with a high degree of probability that these figures for the Baidu search engine are underestimated.

### 3.2. Search Engine Result Pages

The search engine results page (SERP) that is displayed by search engines in response to specific user queries includes two types of results: free organic results or natural search results, the ranking of which is decided by the search engine algorithm; and paid results in the form of text and product ads. The ranking position in organic results is decided by the search engine's algorithm, while for paid ads the deciding factor is the platform's auction mechanism.

Over the years, SERPs have changed in appearance and form, mainly due to Google being a leader and setting trends in how search results are presented. In the early days of search engines, search interfaces were very simple: only a field for a search query and two buttons. In the *Internet 101* book, we can read "*Google ([www.google.com](http://www.google.com)) is a pure search engine—no weather, no news feed, no links to sponsors, no ads, no distractions, no portal litter. Nothing but a fast-loading search site. Reward them with a visit.*" [41]

Originally, SERPs contained a list of hyperlinks to pages containing the target keyword; initially, although search results were very similar across the most popular search engines [42], over the years user requirements forced the design of new and more functional forms of presenting results [43]. Search engine algorithms no longer simply respond to queries by analyzing the number of links recommending a particular page: nowadays, they primarily seek to understand the context and intent of the inquirer. In order to realize this, Google is implementing and developing various search types and features to make searching for information, places, or products more attractive and easier [44]. Google's search engine ranks above all its competitors in terms of the types of search results offered [45]. No other search engine has so many different types: the most common ones are the searchbox (1); rich snippets (2); direct answers (3); a knowledge panel (4); business profiles (5); local results on a map (6); special functions like translator, weather forecast, time, and calculator (7); and sponsored search results (8).

#### 3.2.1. Searchbox

This is an extension of the organic results, most often used for queries about a particular brand. It contains logos, hyperlinks to subpages, a searchbox to subpages within the resulting webpage, and information about the company or institution. According to the algorithm, the presentation form of hyperlinks is intended to help users get to the most valuable content quickly. The owner of a given company's website has no influence on the display of these elements, as the search engine algorithm decides.

#### 3.2.2. Rich Snippets

Rich snippets were introduced by Google in 2012 as a response to the changing way in which queries were created [46]. The goal of the changes was to generate valuable search result pages with interesting and credible content. The snippets display specially marked structural data in the code of a webpage, such as price and price range, rating, the number of reviews or votes, publication date, and author's name. The cooking industry has its own dedicated types of search results, augmented with thumbnail images of recipes so the user can also visually make a selection.

#### 3.2.3. Direct Answers

Direct answer snippets, also known as featured snippets, were introduced by Google in 2016 and consist of a concise answer in the form of a paragraph, list, or table. Featured snippets provide a quick answer to user queries that are formulated in the form of a question, thus allowing users to get the answer without having to visit the source website [47]. The most common types of direct responses are in paragraph form, as this type of response is the most readable and also the most convenient for voice search systems to read. The list

form appears most often in recipes, while tables appear most often in queries related to comparing flight prices or financial products data [47].

Obtaining a position of zero in search results, i.e., in a recommended snippet with a result, is desirable for web developers and emphasizes the expert nature of the site from which the snippet is taken. In order to increase the likelihood that a site's content will be used by Google in a recommended snippet, web developers should compose their content appropriately. It is also important for web developers to have a page on the first page of the SERP for a given keyword query [47].

#### 3.2.4. Knowledge Panel

A knowledge panel or knowledge graph is a collection of data that appears on the right side of search results. It provides a convenient form of information on famous people, music bands, historical figures, and scientific and cultural institutions. As with featured snippets, when deciding whether to include information in the knowledge panel, Google analyzes the structural data and ranking position of the website from which the knowledge panel is sourced [48].

#### 3.2.5. Business Profile

As knowledge panels are automatically generated by an algorithm, Google has enabled companies to manage their business information through company business cards that appear on the right-hand side of the free Business Profile platform. The ability to add information to a company business card requires registration by an authorized body representing the company [49]. A correctly prepared business card, enriched with photos and a detailed description of the business, has a significant impact on ranking positions in the Google search engine [50].

The main elements of a business card are buttons with hyperlinks to the website, Google Map directions, contact information, business hours, and complementary location information. A business card can additionally include photographs, a section with questions and answers, and user reviews added directly in the Business Profile panel or from social media profiles. There are also buttons to add a new review or photo.

#### 3.2.6. Local Results on a Map

The SERP on Google search can be supplemented with additional local results, most often for services in the immediate area (e.g., restaurants, hotels, local services), depending on the user's geographic location. The list of results is generated via Google Maps, which provides real-time traffic information and allows users to view maps and search for facilities. Google Maps uses a GPS system consisting of 27 satellites orbiting the Earth to determine the exact location of people and places. The list of local results also takes into account such information as the location on a map, the route to get there, a link to the company's website, user ratings, and phone number [51].

#### 3.2.7. Search Extensions

Google's search engine SERP is constantly evolving and providing users with new functionality. The goal of search engine developers is to provide features directly in the SERP in such a way that the user does not have to search for them on other websites. This can lead to a situation in which the purchase of products or services happens directly on the search engine site, which has an impact on the financial results of company owners.

Google allows the translation of text using its own translator, where it can choose language combinations from dozens of languages. In addition, the translator has a built-in voice function for learning pronunciation. Typing "city name weather" into Google will provide information about the current temperature and weather conditions in a specific city, as well as the forecast for the next few days, the probability of precipitation, and wind speed and direction. Google can easily provide information about the local time at any place in the world and perform mathematical calculations with the help of a calculator.

### 3.2.8. Sponsored Search Results

Sponsored search engine links are a popular and effective form of online advertising due to the quick results and high level of traffic they generate. Sponsored links are part of search engine marketing, which includes organic traffic and advertising. Advertising campaigns involve the purchase of advertising space on specific search engines and web portals. Search engine ads are divided into text and image ads. The main purpose of text ads is to increase traffic to a website, as evidenced by payment based on cost-per-click. In order to maximize their profits, search engines try to expose such ads by matching them to the broadest possible group of search keywords [52]. Graphic ads with video elements, on the other hand, are primarily aimed at building brand image and are billed on the basis of cost-per-thousand impressions [53].

The leading advertising platform, founded in 2000, is Google Ads (the name Google Ads (formerly AdWords) has been in effect since 24 July 2018), which allows sponsored links to be displayed in Google's search engine results and on Google AdSense collaborative sites, sold under the most popular cost-per-click and cost-per-thousand ad impression pricing models [54]. The position of an ad among other ads depends on the amount of the advertiser's declared price-per-click and the popularity of the ad, as determined by the click-through-ratio indicator, which denotes the percentage of people who clicked on the ad relative to the number of times it was displayed.

Search engine ads, in addition to a textual description and hyperlink, can additionally include an image, name, and the price of the advertised product. When the ad element is activated, shipping cost information is also displayed.

### 3.3. Search Engine Optimization

Search engine optimization (SEO) consists of activities that are designed to bring a website the highest possible ranking position in search engine results [55]. The scope of these activities is very broad and is related to the appropriate selection of keywords [56], optimal website content [57], the structure and optimization of HTML code [58], and graphical elements [59], as well as links to social media [60]. The most important goal of SEO activities is to improve the ranking position of a website for selected keywords in organic search engine results. In practice, SEO activities include the elements listed in Table 1.

**Table 1.** General SEO areas to optimize.

| Element                   | Description  |
|---------------------------|--|
| Content structure         | Editing of articles; category descriptions; offers; or products [61]   |
| Website structure         | Modification of navigation elements; internal linking [62]; structured data [63]   |
| Technical layer           | The speed at which a webpage is displayed in a browser [64]; the speed of downloading content by web robots; SSL encryption [65]; adaptation to mobile devices |
| User experience and trust | Adaptation of the site in accordance with the WCAG standard [66]; clarity and originality of content; number and placement of ads                              |
| Link building             | The number and quality of hyperlinks leading to the website from other sites [67]  |

We can divide the listed activities in Table 1 into those performed directly on the website (on-site) as well as those performed off-site [68]. On-site and off-site optimization should be a coherent and carefully planned strategy because it leads to a measurable effect in the form of a better ranking position in SERPs.

In practice, on-site SEO activities are fundamental to the process of optimization for search engines and should be performed first. The most advantageous option is to make appropriate modifications to the website code at the initial stage of its construction. It is crucial in the optimization process that all elements, such as structure, content, and the appropriate occurrence of keywords in it, affect its effectiveness from the moment it is

launched. In terms of optimization, existing websites can also be rebuilt. Optimization activities should be carried out periodically as search engine algorithm guidelines change over time [61]. The most important optimization elements within a website that significantly influence the position in search results are listed in Table 2.

**Table 2.** General on-site SEO elements to optimize.

| Element                    | Description   |
|----------------------------|---|
| TITLE element              | The title of a webpage is crucial because it helps users quickly evaluate the content and relevance of a result [69].   |
| Meta description tag       | The description or summary of a webpage is part of search engine results [70].  |
| Headers H1 to Hx           | Headers make it easier for the reader to get an idea of the subject matter and hierarchy of the content presented [71].   |
| Keywords                   | Keywords in the text, keyword selection, and density of occurrence [72].  |
| Attribute ALT              | The ALT attribute describes graphic elements; this information is used by indexing robots and screen-reader software to help blind users understand the content of images [73]. |
| Video                      | A detailed description of video files to make them easier to find [74].   |
| Internal linking structure | Internal linking makes it easier to navigate the website [75].  |
| HTML standards             | HTML code adapted for mobile devices [76], compliant with W3C standards [77].   |
| Friendly URLs              | Friendly URLs are usually addresses that include keywords and are short, simple, and easy to read [78].   |
| Page speed                 | Webpage display speed in web browsers [79].   |

Receiving hyperlinks (link building) from other websites, forums, blogs, Google business cards, or social media recommendations are all optimization activities that are performed off-site (off-site SEO) [80]. Acquired hyperlinks are treated by search engines as referrals and therefore have a significant impact on SERP ranking positions [81].

Company blogs, social media, friendly portals, sites that publish sponsored articles, and forums work well in the link acquisition process [67]. It is important that the content published in such places is valuable and contains hyperlinks related to the optimized website [82]. This process is time-consuming, especially when valuable links need to be acquired at minimal cost. Placing links too obtrusively in various places can result in the opposite effect.

By implementing and populating a Google business card with information, better visibility can be achieved, especially on mobile devices. Using geographic location, the search engine matches the best results to the user. An important factor affecting SERP ranking positions is also the name and age of the domain [83].

As off-page activities require experience, mistakes are sometimes made in obtaining hyperlinks from insecure sources [84] that are labeled as unsafe for users (Wang et al., 2017 [85]), or when using tools to artificially generate content [86], which is in violation of search engine policies [87]. The consequences of this can include removal of the website from search results [88].

The person in charge of optimizing a website, making changes to the site, and continuously monitoring the progress of these actions is called an SEO specialist. SEO specialists use many tools in the process of SEO diagnostics and search engine optimization. In practice, there is no perfect set of such tools to guarantee success. This is due to the fact that each SEO specialist creates such a set of programs that will meet their expectations in the long-term optimization process [89]. An additional factor affecting their selection is the financial resources that can be allocated to off-page optimization activities.

SEO tools allow for the analysis of key areas that determine effective optimization for search engines, thus making it possible to find errors in source code and missing elements in web content. This analysis, which is an SEO audit, focuses on issues related to the optimization, usability, and security of a website.

The most important areas of SEO auditing are analyzing a website's communication with web robots at the level of the robots.txt [90] and sitemap.xml [91] files; analysis of the website title and the meta tags responsible for the presentation of the website in search results [92]; analysis of WWW page structure and navigation, including internal and external hyperlinks and URL construction [75]; analysis of the content present on the WWW page, including its size, occurrence of keywords, duplicates [93]; analysis of webserver response statuses [94]; analysis of HTML code correctness [95]; analysis of performance, i.e., WWW page loading speed [96]; analysis of the occurrence of structural data [97]; analysis of WWW performance on mobile devices [76]; verification of WWW server logs [98]; analysis of hyperlink profiles [99]; and analysis of key competitors in organic results [100].

We can divide the tools that enable SEO auditing into two groups: the first group is tools offered by search engines, such as Google Search Console (the name Google Search Console (formerly Google Webmaster Tools) has been in effect since 20 May 2015) in the case of Google, and Bing Webmaster Tools in the case of Bing, both of which analyze the link profile, structure, and performance of a website; the second group is commercial tools, e.g., Ahrefs, Majestic, Semrush. These tools are used to analyze the ranking position in SERPs and the complexity of a webpage [101].

Google Search Console is an essential tool for the technical analysis of websites. It allows for faster indexing of subpages, helps to verify their technical status, collects information about problems encountered in accessing subpages, and verifies the status of the site map [102].

The information that Google Search Console provides includes [44] the number of impressions and clicks on hyperlinks from SERPs; the percentage of impressions that result in clicks on hyperlinks from SERPs; the average ranking position in SERPs; a list of the most common search queries that cause a given website to be displayed in SERPs; a list of the subpages most frequently displayed in SERPs; a list of the most popular countries from which queries originate; and a list of the devices that were used during searches.

Google Search Console provides information on errors encountered by indexing robots and the number of correctly indexed subpages in an entire domain. It makes it possible to manually define the URL of the sitemap, which is an XML-formatted text file containing a collection of hyperlinks to all subpages. This tool also provides knowledge of basic web metrics, i.e., information about the effectiveness of webpages based on data collected during use; this reveals good-quality URLs and poor-quality subpages that require improvement. When there are problems with resource inaccessibility, namely when a webserver returns a 404 error code, it is possible to temporarily or completely remove the URL from organic search results.

#### 4. Discussion

Our study of web search engines and catalogs has several effects on sustainable development. Access to knowledge, which is crucial for sustainable development, has been greatly facilitated by the development of search engines and web directories. Search engines have made it possible for people and organizations to make more informed decisions, participate in research and innovation, and connect with people who share their interests and objectives by making information more available and searchable. Another crucial component of sustainable development is the more effective use of digital resources, which has been made possible by the development of algorithms and technologies to enhance search engine performance. Search engines decrease the time and energy needed to find information by increasing the relevancy and accuracy of search results, which can help lessen the environmental impact of digital technologies.

The emergence of privacy- and sustainability-focused alternative search engines like DuckDuckGo, Ecosia, and Lilo demonstrates a growing understanding of the negative effects of digital technologies on the environment and society [103]. These search engines provide a more moral and environmentally responsible alternative to conventional search engines by encouraging sustainable behaviors like tree planting, water management, and reducing data collection [104]. Using search engines and the internet in general can help support sustainable development by giving users access to resources and information about sustainability and environmental protection. For instance, people can use search engines to find eco-friendly goods, learn about sustainable practices, and connect with environmental organizations. In a similar vein, businesses can access information on sustainability trends and best practices by using search engine optimization (SEO) techniques to promote their sustainable products or initiatives.

Google and other search engines have a big impact on the internet and how people access and share information. In order to promote websites and content that adhere to sustainable practices and values, search engines and their algorithms have a chance to support sustainable development [105]. Search engines could, for instance, give websites and content about sustainable living, renewable energy, and environmental conservation priority and promotion. Additionally, businesses and organizations that support sustainable practices might gain from managing their business information on Google's Business Profile platform, which may result in better search engine visibility and ranking positions. This might encourage sustainable behaviors and values and support sustainable development.

Since search engine optimization (SEO) focuses on raising websites' positions in search engine results, it has little direct bearing on sustainable development. The relationship between SEO and sustainable development, however, might be indirect. For instance, if a company or organization with a focus on sustainable development uses SEO techniques to optimize their website, they may improve their online visibility and draw more visitors to their site. Through engagement with a larger audience and promotion of their sustainable goods and services, they may be able to advance their sustainability objectives. Additionally, some of the on-site SEO components mentioned in the study, like mobile device optimization and adhering to W3C standards, can aid in building a more user-friendly and accessible website. Making sure that people of all backgrounds and abilities have access to information and resources about sustainable development may depend on this [106]. Additionally, off-site SEO activities like link building on social media and other platforms can aid in creating a network of connections and collaborations that may support sustainable development initiatives. A sustainable development organization, for instance, could collaborate with other businesses or organizations with a similar mission and use SEO to promote joint campaigns or events, drive traffic to each other's websites, and share resources and expertise.

The discussion focuses on the contribution of digital technologies to sustainable development and the need to weigh the advantages of greater information access against their environmental and social costs. We can ensure that digital technologies contribute to sustainable development and a more equitable and just world by encouraging innovation and sustainability in the design and use of search engines.

## 5. Conclusions

One of the most important conclusions is that the content and structure of the text on a website are crucial to achieving high search engine rankings. Websites with quality content will be displayed much higher in SERPs than those with content that is not substantive or is oversaturated with keywords. Appropriate and natural placement of keywords in the content helps attract the attention of users and allows search engines to categorize the subject matter accordingly. It is important that the text on a website contains information intended for real viewers, not for search engine bots. Linking a website's content through internal links speeds up the indexing process of an extensive site.

In the near future, we expect to use search engines by using voice commands instead of by entering search queries to receive a list of results. We can already see the first attempts to use voice assistants to interpret questions and return relevant results. This is based on language models. Several of the latest language models, such as GPT-3 and BERT, utilize the transformer neural network architecture, which was created and made available to the public by Google Research in 2017. This architecture enables the model to be trained to read text such as a sentence or paragraph, analyze the relationships between the words, and make predictions on what words may follow.

Recently, Google announced that it was creating an improvement in this field called LaMDA, which means “Language Model for Dialogue Applications” [107]. LaMDA builds on earlier Google research [108] that showed that transformer-based language models trained on dialogue could learn to talk about virtually anything. Unlike most other language models, LaMDA is trained on dialogue. Once trained, LaMDA can be fine-tuned to significantly improve the sensibleness and specificity of its responses.

With the release of more advanced language models to create text-to-text or text-to-image output, there is a concept in artificial intelligence called “prompt engineering.” Prompts are the short text descriptions that language models use to create responses. Soon, we expect to be able to talk to search engines in dialogue form and receive a response in natural language read by a voice assistant.

**Author Contributions:** Conceptualization, A.S.; methodology, M.D.; investigation, M.D.; resources, M.S.; data curation, M.D.; writing—original draft preparation, M.D.; writing—review and editing, A.S.; visualization, M.D.; supervision, A.S.; project administration, A.S.; funding acquisition, M.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research was financed by Department of Computer Networks and Systems of the Silesian University of Technology.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Jindal, V.; Bawa, S.; Batra, S. A Review of Ranking Approaches for Semantic Search on Web. *Inf. Process. Manag.* **2014**, *50*, 416–425. [CrossRef]
- Strawn, G.O. *Leadership in Science and Technology: A Reference Handbook*; SAGE Publications, Inc.: Thousand Oaks, CA, USA, 2012; ISBN 9781412976886.
- Perry, D.G.; Blumenthal, S.H.; Hinden, R.M. The ARPANET and the DARPA Internet. *Libr. Hi Tech* **1988**, *6*, 51–62. [CrossRef]
- O’Regan, G. *Giants of Computing*; Springer: London, UK, 2013; ISBN 978-1-4471-5339-9.
- Berners-Lee, T.; Groff, J.-F. WWW. *ACM SIGBIO Newsl.* **1992**, *12*, 37–40. [CrossRef]
- Tjondronegoro, D.; Spink, A. Web Search Engine Multimedia Functionality. *Inf. Process. Manag.* **2008**, *44*, 340–357. [CrossRef]
- Țichindelean, M.; Țichindelean, M.T.; Cetinã, I.; Orzan, G. A Comparative Eye Tracking Study of Usability—Towards Sustainable Web Design. *Sustainability* **2021**, *13*, 10415. [CrossRef]
- Manganelli, M.; Soldati, A.; Martirano, L.; Ramakrishna, S. Strategies for Improving the Sustainability of Data Centers via Energy Mix, Energy Conservation, and Circular Energy. *Sustainability* **2021**, *13*, 6114. [CrossRef]
- Karyotakis, M.-A.; Antonopoulos, N. Web Communication: A Content Analysis of Green Hosting Companies. *Sustainability* **2021**, *13*, 495. [CrossRef]
- Seymour, T.; Frantsvog, D.; Kumar, S. History of Search Engines. *Int. J. Manag. Inf. Syst.* **2011**, *15*, 47. [CrossRef]
- Bharat, K.; Broder, A.; Henzinger, M.; Kumar, P.; Venkatasubramanian, S. The Connectivity Server: Fast Access to Linkage Information on the Web. *Comput. Netw. ISDN Syst.* **1998**, *30*, 469–477. [CrossRef]
- Beg, M.M.S. A Subjective Measure of Web Search Quality. *Inf. Sci.* **2005**, *169*, 365–381. [CrossRef]
- Duka, M. Ranking of Websites Created with the Use of ISOWQ Rank Algorithm. *Inform. Autom. Pomiarowy Gospod. Ochr. Sr.* **2020**, *10*, 16–19. [CrossRef]
- Dimopoulos, C.; Makris, C.; Panagis, Y.; Theodoridis, E.; Tsakalidis, A. A Web Page Usage Prediction Scheme Using Sequence Indexing and Clustering Techniques. *Data Knowl. Eng.* **2010**, *69*, 371–382. [CrossRef]

15. Aiello, W.A. *Algorithms and Models for the Web-Graph*; Aiello, W., Broder, A., Janssen, J., Milios, E., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2008; Volume 4936, ISBN 978-3-540-78807-2.
16. Du, Y.; Li, C.; Hu, Q.; Li, X.; Chen, X. Ranking Webpages Using a Path Trust Knowledge Graph. *Neurocomputing* **2017**, *269*, 58–72. [[CrossRef](#)]
17. Makkar, A.; Kumar, N. User Behavior Analysis-Based Smart Energy Management for Webpage Ranking: Learning Automata-Based Solution. *Sustain. Comput. Inform. Syst.* **2018**, *20*, 174–191. [[CrossRef](#)]
18. Cho, J.; Garcia-Molina, H.; Page, L. Efficient Crawling through URL Ordering. *Comput. Netw. ISDN Syst.* **1998**, *30*, 161–172. [[CrossRef](#)]
19. Suchacka, G.; Cabri, A.; Rovetta, S.; Masulli, F. Efficient On-the-Fly Web Bot Detection. *Knowl.-Based Syst.* **2021**, *223*, 107074. [[CrossRef](#)]
20. Kumar, M.; Bindal, A.; Gautam, R.; Bhatia, R. Keyword Query Based Focused Web Crawler. *Procedia Comput. Sci.* **2018**, *125*, 584–590. [[CrossRef](#)]
21. Bai, Q.; Xiong, G.; Zhao, Y.; He, L. Analysis and Detection of Bogus Behavior in Web Crawler Measurement. *Procedia Comput. Sci.* **2014**, *31*, 1084–1091. [[CrossRef](#)]
22. Dikaiakos, M.D.; Stassopoulou, A.; Papageorgiou, L. An Investigation of Web Crawler Behavior: Characterization and Metrics. *Comput. Commun.* **2005**, *28*, 880–897. [[CrossRef](#)]
23. Batzios, A.; Dimou, C.; Symeonidis, A.L.; Mitkas, P.A. BioCrawler: An Intelligent Crawler for the Semantic Web. *Expert Syst. Appl.* **2008**, *35*, 524–530. [[CrossRef](#)]
24. Stevanovic, D.; An, A.; Vlajic, N. Feature Evaluation for Web Crawler Detection with Data Mining Techniques. *Expert Syst. Appl.* **2012**, *39*, 8707–8717. [[CrossRef](#)]
25. Tanaka, T.; Niibori, H.; Li, S.; Nomura, S.; Kawashima, H.; Tsuda, K. Bot Detection Model Using User Agent and User Behavior for Web Log Analysis. *Procedia Comput. Sci.* **2020**, *176*, 1621–1625. [[CrossRef](#)]
26. Chen, R.-C.; Hsieh, C.-H. Web Page Classification Based on a Support Vector Machine Using a Weighted Vote Schema. *Expert Syst. Appl.* **2006**, *31*, 427–435. [[CrossRef](#)]
27. Buber, E.; Diri, B. Web Page Classification Using RNN. *Procedia Comput. Sci.* **2019**, *154*, 62–72. [[CrossRef](#)]
28. Özel, S.A. A Web Page Classification System Based on a Genetic Algorithm Using Tagged-Terms as Features. *Expert Syst. Appl.* **2011**, *38*, 3407–3415. [[CrossRef](#)]
29. Wilkinson, D.; Thelwall, M. Search Markets and Search Results: The Case of Bing. *Libr. Inf. Sci. Res.* **2013**, *35*, 318–325. [[CrossRef](#)]
30. Van Couvering, E. The History of the Internet Search Engine: Navigational Media and the Traffic Commodity. In *Web Search*; Spink, A., Zimmer, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 177–206.
31. Jiang, M.; Okamoto, K. National Identity, Ideological Apparatus, or Panopticon? A Case Study of the Chinese National Search Engine Jike. *Policy Internet* **2014**, *6*, 89–107. [[CrossRef](#)]
32. Gao, R.; Shah, C. Toward Creating a Fairer Ranking in Search Engine Results. *Inf. Process. Manag.* **2020**, *57*, 102138. [[CrossRef](#)]
33. Strzelecki, A. Eye-Tracking Studies of Web Search Engines: A Systematic Literature Review. *Information* **2020**, *11*, 300. [[CrossRef](#)]
34. Lewandowski, D.; Kammerer, Y. Factors Influencing Viewing Behaviour on Search Engine Results Pages: A Review of Eye-Tracking Research. *Behav. Inf. Technol.* **2021**, *40*, 1485–1515. [[CrossRef](#)]
35. Lee, M.; Kim, W.; Park, S. Searching and Ranking Method of Relevant Resources by User Intention on the Semantic Web. *Expert Syst. Appl.* **2012**, *39*, 4111–4121. [[CrossRef](#)]
36. Sánchez, D.; Martínez-Sanahuja, L.; Batet, M. Survey and Evaluation of Web Search Engine Hit Counts as Research Tools in Computational Linguistics. *Inf. Syst.* **2018**, *73*, 50–60. [[CrossRef](#)]
37. Fagroud, F.Z.; Ajalloula, L.; Ben Lahmar, E.H.; Toumi, H.; Achtaich, K.; Filali, S. El IOT Search Engines: Exploratory Data Analysis. *Procedia Comput. Sci.* **2020**, *175*, 572–577. [[CrossRef](#)]
38. Derhami, V.; Khodadadian, E.; Ghasemzadeh, M.; Bidoki, A.M.Z. Applying Reinforcement Learning for Web Pages Ranking Algorithms. *Appl. Soft Comput.* **2013**, *13*, 1686–1692. [[CrossRef](#)]
39. Li, H. Internet Tourism Resource Retrieval Using PageRank Search Ranking Algorithm. *Complexity* **2021**, *2021*, 5114802. [[CrossRef](#)]
40. Jiang, M. Search Concentration, Bias, and Parochialism: A Comparative Study of Google, Baidu, and Jike’s Search Results from China. *J. Commun.* **2014**, *64*, 1088–1110. [[CrossRef](#)]
41. Lehnert, W.G. *Internet 101: A Beginners Guide to the Internet and the World Wide Web*; Addison Wesley Longman: Boston, MA, USA, 1999.
42. Dwivedi, N.; Joshi, L.; Gupta, N. Statistical Analysis of Search Engines (Google, Yahoo and Altavista) for Their Search Result. *Int. J. Comput. Theory Eng.* **2013**, *5*, 298–301. [[CrossRef](#)]
43. Behnert, C.; Lewandowski, D. Ranking Search Results in Library Information Systems—Considering Ranking Approaches Adapted from Web Search Engines. *J. Acad. Librariansh.* **2015**, *41*, 725–735. [[CrossRef](#)]
44. Strzelecki, A. Google Web and Image Search Visibility Data for Online Store. *Data* **2019**, *4*, 125. [[CrossRef](#)]
45. Höchstötter, N.; Lewandowski, D. What Users See—Structures in Search Engine Results Pages. *Inf. Sci.* **2009**, *179*, 1796–1812. [[CrossRef](#)]
46. Strzelecki, A.; Rutecka, P. The Snippets Taxonomy in Web Search Engines. In *Perspectives in Business Informatics Research*; Pańkowska, M., Sandkuhl, K., Eds.; Springer: Champagne, IL, USA, 2019; Volume 365, pp. 177–188. ISBN 978-3-030-31143-8.
47. Strzelecki, A.; Rutecka, P. Direct Answers in Google Search Results. *IEEE Access* **2020**, *8*, 103642–103654. [[CrossRef](#)]

48. Singhal, A. Official Google Blog: Introducing the Knowledge Graph: Things, Not Strings. Available online: <https://www.blog.google/products/search/introducing-knowledge-graph-things-not/> (accessed on 30 April 2020).
49. Lee, H. Viewing Local Organization Data from Google My Business. In *Hands On With Google® Data Studio*; Wiley: Hoboken, NJ, USA, 2020; pp. 173–219.
50. Durica, M.; Svabova, L. Improvement of Company Marketing Strategy Based on Google Search Results Analysis. *Procedia Econ. Financ.* **2015**, *26*, 454–460. [[CrossRef](#)]
51. Ciepluch, B.; Jacob, R.; Mooney, P.; Winstanley, A.C. Comparison of the accuracy of OpenStreetMap for Ireland with Google Maps and Bing Maps. In Proceedings of the Ninth International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Leicester, UK, 20–23 July 2010; p. 337.
52. Jansen, B.J.; Mullen, T. Sponsored Search: An Overview of the Concept, History, and Technology. *Int. J. Electron. Bus.* **2008**, *6*, 114–171. [[CrossRef](#)]
53. Etro, F. Advertising and Search Engines. A Model of Leadership in Search Advertising. *Res. Econ.* **2013**, *67*, 25–38. [[CrossRef](#)]
54. Asdemir, K.; Kumar, N.; Jacob, V.S. Pricing Models for Online Advertising: CPM vs. CPC. *Inf. Syst. Res.* **2012**, *23*, 804–822. [[CrossRef](#)]
55. Lewandowski, D.; Sünkler, S.; Yagci, N. The Influence of Search Engine Optimization on Google’s Results. In Proceedings of the 13th ACM Web Science Conference 2021, Virtual, 21–25 June 2021; ACM: New York, NY, USA, 2021; pp. 12–20.
56. Nagpal, M.; Petersen, J.A. Keyword Selection Strategies in Search Engine Optimization: How Relevant Is Relevance? *J. Retail.* **2021**, *97*, 746–763. [[CrossRef](#)]
57. Vállez, M.; Ventura, A. Analysis of the SEO Visibility of University Libraries and How They Impact the Web Visibility of Their Universities. *J. Acad. Librariansh.* **2020**, *46*, 102171. [[CrossRef](#)]
58. Hora, A. Characterizing Top Ranked Code Examples in Google. *J. Syst. Softw.* **2021**, *178*, 110971. [[CrossRef](#)]
59. Ferraz, R. Exploring Web Attributes Related to Image Accessibility and Their Impact on Search Engine Indexing. *Procedia Comput. Sci.* **2015**, *67*, 171–184. [[CrossRef](#)]
60. Xiang, Z.; Gretzel, U. Role of Social Media in Online Travel Information Search. *Tour. Manag.* **2010**, *31*, 179–188. [[CrossRef](#)]
61. Mavridis, T.; Symeonidis, A.L. Semantic Analysis of Web Documents for the Generation of Optimal Content. *Eng. Appl. Artif. Intell.* **2014**, *35*, 114–130. [[CrossRef](#)]
62. Shipley, T.G.; Bowker, A. Investigating Websites and Webpages. In *Investigating Internet Crimes*; Elsevier: Amsterdam, The Netherlands, 2014; pp. 293–314.
63. Abdou, M.; AbdelGaber, S.; Farhan, M. A Semi-Automated Framework for Semantically Annotating Web Content. *Futur. Gener. Comput. Syst.* **2018**, *81*, 94–102. [[CrossRef](#)]
64. Egri, G.; Bayrak, C. The Role of Search Engine Optimization on Keeping the User on the Site. *Procedia Comput. Sci.* **2014**, *36*, 335–342. [[CrossRef](#)]
65. Konidaris, A.; Koustoumpardi, E. The Importance of Search Engine Optimization for Tourism Websites. In *Innovative Approaches to Tourism and Leisure*; Katsoni, V., Velander, K., Eds.; Springer: Cham, Switzerland, 2018; pp. 205–218.
66. Centeno, V.L.; Kloos, C.D.; Fisteus, J.A.; Álvarez, L.Á. Web Accessibility Evaluation Tools: A Survey and Some Improvements. *Electron. Notes Theor. Comput. Sci.* **2006**, *157*, 87–100. [[CrossRef](#)]
67. Zhang, S.; Cabage, N. Search Engine Optimization: Comparison of Link Building and Social Sharing. *J. Comput. Inf. Syst.* **2017**, *57*, 148–159. [[CrossRef](#)]
68. Veglis, A.; Giomelakis, D. Search Engine Optimization. *Futur. Internet* **2019**, *12*, 6. [[CrossRef](#)]
69. Noruzi, A. A Study of HTML Title Tag Creation Behavior of Academic Web Sites. *J. Acad. Librariansh.* **2007**, *33*, 501–506. [[CrossRef](#)]
70. Craven, T.C. Variations in Use of Meta Tag Descriptions by Web Pages in Different Languages. *Inf. Process. Manag.* **2004**, *40*, 479–493. [[CrossRef](#)]
71. Pérez-Montoro, M.; Codina, L. The Essentials of Search Engine Optimization. In *Navigation Design and SEO for Content-Intensive Websites*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 109–124.
72. Pérez-Montoro, M.; Codina, L. SEO for Content-Intensive Sites. In *Navigation Design and SEO for Content-Intensive Websites*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 125–137.
73. Yalçın, N.; Köse, U. What Is Search Engine Optimization: SEO? *Procedia-Soc. Behav. Sci.* **2010**, *9*, 487–493. [[CrossRef](#)]
74. Choudhari, K.; Bhalla, V.K. Video Search Engine Optimization Using Keyword and Feature Analysis. *Procedia Comput. Sci.* **2015**, *58*, 691–697. [[CrossRef](#)]
75. Killoran, J.B. How to Use Search Engine Optimization Techniques to Increase Website Visibility. *IEEE Trans. Prof. Commun.* **2013**, *56*, 50–66. [[CrossRef](#)]
76. Pérez-Montoro, M.; Codina, L. Mobile Web and SEO. In *Navigation Design and SEO for Content-Intensive Websites*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 139–151.
77. Ismail, A.; Kuppusamy, K.S. Web Accessibility Investigation and Identification of Major Issues of Higher Education Websites with Statistical Measures: A Case Study of College Websites. *J. King Saud Univ.—Comput. Inf. Sci.* **2022**, *34*, 901–911. [[CrossRef](#)]
78. Ganapathy, A. Friendly URLs in the CMS and Power of Global Ranking with Crawlers with Added Security. *Eng. Int.* **2017**, *5*, 87–96. [[CrossRef](#)]
79. Ismail, A.; Abdallah, F. A Survey on Search Engine Optimization (SEO). *Int. J. Comput. Commun. Instrum. Eng.* **2017**, *4*, 61–67. [[CrossRef](#)]

80. Erdmann, A.; Ponzoa, J.M. Digital Inbound Marketing: Measuring the Economic Performance of Grocery e-Commerce in Europe and the USA. *Technol. Forecast. Soc. Chang.* **2021**, *162*, 120373. [[CrossRef](#)] [[PubMed](#)]
81. Fortunato, S.; Boguña, M.; Flammini, A.; Menczer, F. Approximating PageRank from In-Degree. In *Algorithms and Models for the Web-Graph*; Aiello, W., Broder, A., Janssen, J., Milos, E., Eds.; LNCS; Springer: Berlin/Heidelberg, Germany, 2008; Volume 4936, pp. 59–71.
82. Jeong, O.-R.; Oh, J.; Kim, D.-J.; Lyu, H.; Kim, W. Determining the Titles of Web Pages Using Anchor Text and Link Analysis. *Expert Syst. Appl.* **2014**, *41*, 4322–4329. [[CrossRef](#)]
83. Ziakis, C.; Vlachopoulou, M.; Kyrkoudis, T.; Karagkiozidou, M. Important Factors for Improving Google Search Rank. *Futur. Internet* **2019**, *11*, 32. [[CrossRef](#)]
84. Kim, D.W.; Yan, P.; Zhang, J. Detecting Fake Anti-Virus Software Distribution Webpages. *Comput. Secur.* **2015**, *49*, 95–106. [[CrossRef](#)]
85. Wang, R.; Zhu, Y.; Tan, J.; Zhou, B. Detection of Malicious Web Pages Based on Hybrid Analysis. *J. Inf. Secur. Appl.* **2017**, *35*, 68–74. [[CrossRef](#)]
86. Wright, C. Auditing Web-Based Applications. In *The IT Regulatory and Standards Compliance Handbook*; Elsevier: Amsterdam, The Netherlands, 2008; pp. 515–560.
87. Bello, R.-W.; Otopo, F.N. Conversion of Website Users to Customers-The Black Hat SEO Technique. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* **2018**, *8*, 29. [[CrossRef](#)]
88. Aswani, R.; Kar, A.K.; Ilavarasan, P.V.; Dwivedi, Y.K. Search Engine Marketing Is Not All Gold: Insights from Twitter and SEOclerks. *Int. J. Inf. Manag.* **2018**, *38*, 107–116. [[CrossRef](#)]
89. Schultheiß, S.; Lewandowski, D. “Outside the Industry, Nobody Knows What We Do” SEO as Seen by Search Engine Optimizers and Content Providers. *J. Doc.* **2020**, *77*, 542–557. [[CrossRef](#)]
90. Sun, Y.; Zhuang, Z.; Councill, I.G.; Giles, C.L. Determining Bias to Search Engines from Robots.Txt. In Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence (WI'07), Fremont, CA, USA, 2–5 November 2007; IEEE: New York, NY, USA; pp. 149–155.
91. Tomič, A.; Šupin, M. Increasing Website Traffic of Woodworking Company Using Digital Marketing Methods. *Acta Fac. Xylologiae Zvolen* **2019**, *61*, 153–161.
92. Wijaya, D.; Daniawan, B.; Gunawan, Y. Search Engine Optimization (SEO) As A Promotional Media On Google Search. *bit-Tech* **2021**, *4*, 31–39. [[CrossRef](#)]
93. Nen, M.; Popa, V.; Scurtu, A. The Computer Management—SEO Audit. *Rev. Int. Comp. Manag.* **2017**, *18*, 297–307.
94. Gek, D.; Kukartsev, V.; Tynchenko, V.; Bondarev, A.; Pokushko, M.; Dalisova, N. The Problem of SEO Promotion for the Organization’s Web Representation. *SHS Web Conf.* **2019**, *69*, 00122. [[CrossRef](#)]
95. Moreno, L.; Martinez, P. Overlapping Factors in Search Engine Optimization and Web Accessibility. *Online Inf. Rev.* **2013**, *37*, 564–580. [[CrossRef](#)]
96. Shenoy, A.; Prabhu, A. SEO Hub: Utilities and Toolsets. In *Introducing SEO*; Apress: Berkeley, CA, USA, 2016; pp. 103–117.
97. Katumba, S.; Coetsee, S. Employing Search Engine Optimization (SEO) Techniques for Improving the Discovery of Geospatial Resources on the Web. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 284. [[CrossRef](#)]
98. Gregurec, I.; Grd, P. Search Engine Optimization (SEO): Website Analysis of Selected Faculties in Croatia. In Proceedings of Central European Conference on Information and Intelligent Systems, Varaždin, Croatia, 19–21 September 2012; Hunjak, T., Lovrenčić, S., Tomičić, I., Eds.; Faculty of Organization and Informatics Varaždin, University of Zagreb: Zagreb, Croatia, 2012; pp. 211–218.
99. Vankov, N. Significance, Role and Principles of Website SEO for Digital Entrepreneurship—The Battle for the Top Google Rankings. In *The International Entrepreneurship: Trends, Challenges, Achievements, Proceedings of the Eighth International Conference, Varna, Bulgaria, 6–9 June 2017*; Todorov, K., Kolarov, K., Eds.; Bulgarian Association for Management Development and Entrepreneurship: Sofia, Bulgaria, 2017; pp. 192–209.
100. Evans, M.P. Analysing Google Rankings through Search Engine Optimization Data. *Internet Res.* **2007**, *17*, 21–37. [[CrossRef](#)]
101. Weiqing, C.; Yangyang, H.; Qiaofeng, Y.; Jiajia, C. Measuring Web Page Complexity by Analyzing TCP Flows and HTTP Headers. *J. China Univ. Posts Telecommun.* **2017**, *24*, 1–13. [[CrossRef](#)]
102. Lee, H. Using Google Search Console for Audience Insights. In *Hands on with Google®Data Studio*; Wiley: Hoboken, NJ, USA, 2020; pp. 135–171.
103. Palos-Sanchez, P.; Martin-Velicia, F.; Saura, J.R. Complexity in the Acceptance of Sustainable Search Engines on the Internet: An Analysis of Unobserved Heterogeneity with FIMIX-PLS. *Complexity* **2018**, *2018*, 6561417. [[CrossRef](#)]
104. Reyes-Menendez, A.; Saura, J.R.; Palos-Sanchez, P.R.; Alvarez-Garcia, J. Understanding User Behavioral Intention to Adopt a Search Engine That Promotes Sustainable Water Management. *Symmetry* **2018**, *10*, 584. [[CrossRef](#)]
105. Confetto, M.G.; Covucci, C. “Sustainability-Contents SEO”: A Semantic Algorithm to Improve the Quality Rating of Sustainability Web Contents. *TQM J.* **2020**, *33*, 295–317. [[CrossRef](#)]
106. Roumeliotis, K.I.; Tselikas, N.D.; Nasiopoulos, D.K. Airlines’ Sustainability Study Based on Search Engine Optimization Techniques and Technologies. *Sustainability* **2022**, *14*, 11225. [[CrossRef](#)]

107. Thoppilan, R.; De Freitas, D.; Hall, J.; Shazeer, N.; Kulshreshtha, A.; Cheng, H.-T.; Jin, A.; Bos, T.; Baker, L.; Du, Y.; et al. LaMDA: Language Models for Dialog Applications. *arXiv* **2022**, arXiv:2201.08239. [[CrossRef](#)]
108. Adiwardana, D.; Luong, M.-T.; So, D.R.; Hall, J.; Fiedel, N.; Thoppilan, R.; Yang, Z.; Kulshreshtha, A.; Nemade, G.; Lu, Y.; et al. Towards a Human-like Open-Domain Chatbot. *arXiv* **2020**, arXiv:2001.09977. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.