Search Engine Results Ranking System and Techniques to Increase Website Visibility *

Marek Kuruc
Slovenská technická univerzita v Bratislave
Fakulta informatiky a informačných technológií
xkuruc@stuba.sk

30. september 2023

Abstract

Search engines are powerful tools that help users find information on the web. They use various algorithms and techniques to crawl, index, optimalize and rank web pages based on their relevance and quality. However, not all web pages are equally visible to search engines, and some may be buried under millions of other results.

Therefore, web content creators and webmasters need to optimize their websites and content to increase their visibility and reach their target audience.

This document aims to explain the methods for increasing the website visibility, by first providing an overview of how search engines work, what factors affect their ranking, how are search engine sorting and optimalizing found results, in order to give an comprehensive explonation on how to align the website goals with the user needs and expectations. Then, it will describe some of the best practices and strategies that can be applied to improve the web design, content, keywords, links, and social media presence of a website. Finally, we will discuss some of the challenges and limitations of search engine optimization (SEO) and how to measure its effectiveness.

By following these guidelines, web content creators and webmasters can enhance their online presence and attract more visitors to their websites.

1 Introduction

In the contemporary digital landscape, search engines serve as the cornerstone of information discovery on the World Wide Web. These sophisticated tools use intricate algorithms and methodologies to traverse, catalog, optimize, and rank the vast expanse of web pages in accordance with their relevance and quality. Yet, not all web pages enjoy equal visibility. Many remain hidden beneath an avalanche of millions of other results, challenging the prospects of being discovered by users seeking information.

*Semestrálny projekt v predmete Metódy inžinierskej práce, ak. rok 2023/24, vedenie: Ing. Mohammad Yusuf Momand, MSc.

Consequently, web content creators and webmasters find themselves confronted with the request to enhance the visibility of their websites. This endeavor is essential for achieving alignment between the goals of a website and the expectations and needs of its target audience.

This document embarks on a comprehensible explanation of the mechanisms and techniques that underlie the augmentation of website visibility. Initially, it provides a foundational understanding of the inner workings of search engines, elucidating the factors influencing their ranking processes and the techniques employed for sorting and optimizing the results they yield. With this groundwork in place, we delve into an examination of the methods available to harmonize a website's objectives with the demands of its users.

Subsequently, this document delineates a compendium of best practices and strategies, encompassing aspects such as web design, content optimization, keyword selection, link building, and the cultivation of a robust social media presence. By elucidating these methodologies, we endeavor to empower web content creators and webmasters with the knowledge and tools necessary to fortify their digital presence.

Nonetheless, it is crucial to recognize that the realm of search engine optimization (SEO) is not without its challenges and constraints. We shall, therefore, broach the subject of the limitations inherent to SEO and explore methodologies for gauging its effectiveness, and at the end we will discuss the potential future of SEO.

2 How does seach engine search for information

In the following section, we will discuss the mechanisms of search engines. We will examine how they interpret user queries and navigate the vast digital landscape to retrieve relevant information. This process involves complex algorithms and data structures, which we will explain in detail

2.1 Crawling

Crawling is the process used by search engine web crawlers (also known as bots or spiders) to visit and download a page and extract its links in order to discover additional pages. Pages known to the search engine are crawled periodically to determine whether any changes have been made to the page's content since the last time it was crawled. If a search engine detects changes to a page after crawling a page, it will update its index in response to these detected changes. [1, p. 245] (paragraph B), [14] (Introduction) (Search Engines)

Search engines use their own web crawlers to discover and access web pages. All commercial search engine crawlers begin crawling a website by downloading its robots.txt file, which contains rules about what pages search engines should or should not crawl on the website. The robots.txt file may also contain information about sitemaps- this contains lists of URLs that the site wants a search engine crawler to crawl. [14, p. 150] (Robots Exclusion Protocol)

Search engine crawlers use a number of algorithms and rules to determine how frequently a page should be re-crawled and how many pages on a site should be

indexed. For example, a page that changes on a regular basis may be crawled more frequently than one that is rarely modified. [1, p. 248] (discussion)

When a search engine bot crawls a web page, it reviews all content and code that it can find. This includes plain text, images and alt text, links, etc. Crawlers note any links found on a site and crawl those pages too. In this way, site owners can create a link path for crawlers. [14, p. 154] (history of bias)

2.2 Indexing

After a page is crawled, the next step is to index its content. The indexed page is stored in a giant database, from where it can later be retrieved. Essentially, the process of indexing is identifying the words and expressions that best describe the page and assigning particular keywords to the page. The frequency of words is evaluated as well. For a human it will not be possible to process such amounts of information but generally search engines deal just fine with this task. [3, p. 2](Classical search) [3, p. 2](A. Search engines)

Key words are then adjusted to their basic form (root word), their prefixes, suffixes, replacements, adverbs and other unnecessary parts are removed. So indexing can work more efficiently. [3, p. 2](Classical search)

2.3 Query Processing

This is how search engines interpret the search terms entered by users. The query is generally taken from a UI search box where the user enters the search query. The query may be pre-processed or transformed using techniques like spell checking, query suggestions, or query expansion (adding more terms to the query, to be more precise).

The resulting string / modified Query, is further compared with the indexes in the databases and with the information gathered by crawlers. With this process, web engines try to match users intentions [16, p. 2](Implementation of the distributed query processing system)

2.4 Paid results

Besides organic results, search engines also display paid results, which are essentially advertisements.Search engine will first start matching similarities with paid ad indexes and then will continue with regular list of indexes, that's why these results are usually distinguished from organic results but are also relevant to the user's query.

These paid results (ads) typically appear at the top of relevant search engine results pages (SERPs). [2, p. 2](Paid search advertising) [2, p. 2](Case study)

The whole process is nicely illustrated in the following picture. [9, p. 2]

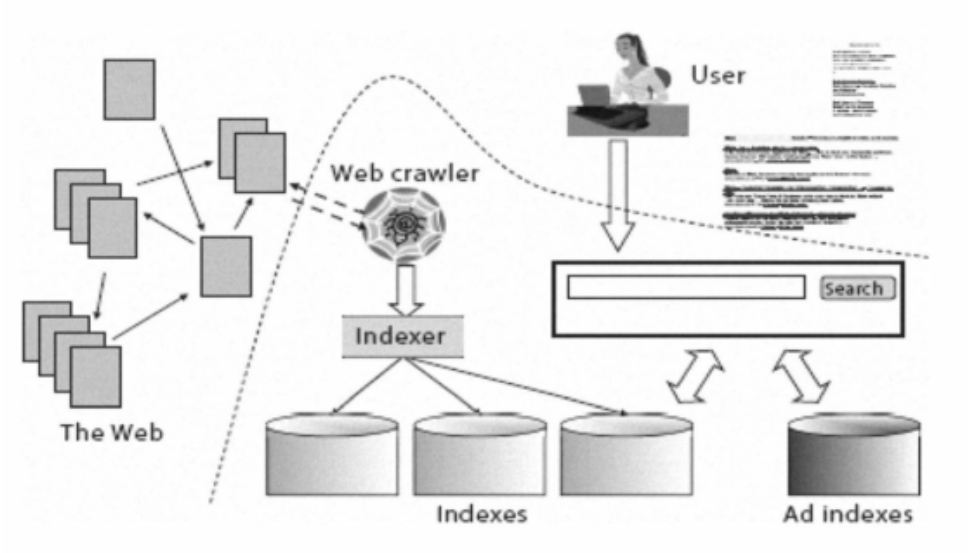

Figure 1: The various components of a web search engine.

3 How does seach engine search Rank results

Search engines analyze information across the web network and then rank them based on their own set qualitative measures. Every search engine use different methods.
In the past, it was far easier to rank higher in Search Engine Result Pages SERPs because it was easier to surpass the quality check measures. Presently search engines cognitive abilities have grown exceptionally and give more value to content published by established brands and less to new digital marketers. [12, p. 1](Introduction)

3.1 Search Engine Algorithms

There are many different ranking algorithms with hundrets of different parameters unique to every search engine. Some of the most popular search engines ranking algorithms include [13, p. 3](Introduction):

For example PageRank algorithm: The ranking of page is gotten through whether referenced by other pages, how many pages are referencing this page and whether referenced by some important pages.

HillTop algorithm determine the importance and relevance of the page through users input the query keywords in the query box. Searching result will be returned out the list page called “experts’ document” [15, p. 1699](Overview Other Sorting Technologies)

Panda:This algorithm update was designed to reduce the prevalence of low-quality, thin content in Google’s top search results and to reward unique, compelling content.

Mobilegeddon: This update was to ensure that mobile-friendly pages rank higher on mobile searches. [13, p. 3](google SEO Algorithms)

The more these measures match with your target, the higher the chances of your site getting ranked high on search engines

3.2 Popularity

Popularity in the context of search engines often refers to the number of backlinks a webpage has. The more backlinks a webpage has from reputable sources, the more popular it is considered to be.
In addition, search engines like Google use hundreds of factors to determine both the relevance and popularity of webpages. These factors include things like page speed, mobile-friendliness, and content quality.
A popular and relevant website is likely to rank higher in search engine results, leading to increased visibility and potentially more traffic. [6, p. 676] (Introduction) [6, p. 677](The proposed stochastic model)

3.3 Search engine companies

Searchers using more than one search engine will likely have noticed that for a given query, the competing SERPs tend to show different rankings, and indeed often show different sites entirely. Every search engine has its own ranking system. For example Google Scholar will prefer educational websites and articles, on the other hand, Google Chrome will prefer more commercial sites. [10, p. 53] (Search Engine Companies and Programmers)
[11] [8] [5] [4] [7]

References

[1] Ioannis Avraam and Ioannis Anagnostopoulos. A comparison over focused web crawling strategies. 2011.

[2] Tobias Blask. Do specific text features influence click probabilities in paid search advertising? 2014.

[3] Kamal EL GUEMMAT and Sara OUAHABI. Towards a new educational search engine based on hybrid searching and indexing techniques. 2019.

[4] Ashwini R. Gaurkhede and D. R. Dhotre. 2014 international conference on power, automation and communication (inpac). 2014. url: https://ieeexplore.ieee.org/document/6981154.

[5] S. Geetharani and M. Soranamageswari. Location-based ranking method (lbrm) for ranking search results in search engines. 2016. url:https://ieeexplore.ieee.org/document/7727140.

[6] Hiromitsu Hama, Pyke Tin, Thi Thi Zin, and Takashi Toriu. A stochastic model for popularity measures in web dynamics. 2010.

[7] Ayad Hameed Mousa, Mowafak Mohsen, Ali Alnasrawi, and Intedhar Nasir. Imuw-app: An instrument for measuring the usability of web applications. *Indonesian Journal of Electrical Engineering and Computer Science*, 24:1183, 11 2021.

[8] K. B. Jones. 2010. url:https://www.usetech.org/wp-content/uploads/ebooks/Search-Engine-Optimization-Your-Visual-Blueprint-for-Effective-Internet-Marketing.pdf.

[9] Ayman E. Khedr and Mohamed M. Reda Aly. An integration framework for search engine architecture to improve information retrieval quality. 2010.

[10] John B. Killoran. How to use search engine optimization techniques to increase website visibility. *IEEE Transactions on Professional Communication*, 56(1), 2013.

[11] John B. Killoran. How to use search engine optimization techniques to increase website visibility. *IEEE Transactions on Professional Communication*, 56(1), 2013. url:https://ieeexplore.ieee.org/document/6463486.

[12] Asad Nadeem, Musa Hussain, and Ahsan Iftikhar. New technique to rank without off page search engine optimization. 2020.

[13] Akshita Patil, Jayesh Pamnani, and Dipti Pawade. Comparative study of google search engine optimization algorithms: Panda, penguin and hummingbird. 2021.

[14] Yang Sun, Ziming Zhuang, Isaac G. Councill, and C. Lee Giles. Determining bias to search engines from robots.txt. 2007.

[15] Renyuan Wang, Yan Chen, Taoying Li, and Yingying Yu. The optimization of search engines ranking technology based on grey system. 2013.

[16] Gongxin Wu and Yujia Ge. A framework for distributed query processing based on web services. 7, 2010.