State-of-the-art white hat SEO methods

Lukas Felbinger
Interactive Technologies
St. Pölten University of Applied Sciences
St. Pölten, Austria
It181505@fhstp.ac.at

***Abstract*—Presently, there are over 1.9 billion different websites on the Internet, therefore, the fight to appear on the first search engine result page (SERP) to increase the website´s traffic is becoming continuously harder. As a result, search engine optimization (SEO), the process of influencing the online visibility of a website in the search engine´s results, is gaining large recognition. This research paper focuses on state-of-the-art SEO techniques in order to improve a website´s page rank of a search engine (SE) long-term. Consequently, only methods that are approved by the search engine´s algorithm (white-hat) are stated and researched. Furthermore, the methods covered in this research are considered best practice for most modern search engines and not just for a specific one.**

The research concluded to the result that there is a vast variety of different factors which will influence the website´s page rank of search engines, but the first step for a successful SEO campaign should always be the keyword analysis. With a set of relevant keywords, it is possible to increase the visibility of a website with simple adjustments such as placing a title- and meta description tag and pairing them with the keywords. More on-page SEO methods being quality content, mobile accessibility and a custom 404 page can improve the rank even further by reducing the average bounce rate. In contrast, off-page methods are applied outside of the website´s scope but are as important. Incoming links from other websites are called backlinks or inbound links and are the most important off-page SEO method.

Keywords-search engine optimization; white hat SEO; search engines; page rank; search engine result page

I. INTRODUCTION

Finding relevant information on the Internet such as contact data, descriptions, facts or statistics is quite easy nowadays as a result of search engines (SE). A search engine uses complex algorithms to find proper results based on keywords along with phrases and lists them sorted by relevance for quick access. However, before the search engine is able to determine the most suited results, an agent called crawler looks at every URL on the web it has access to and collects keywords and phrases of each page, which are then stored in a database [1, p.5]. Given the fact that there are over 1.9 billion websites and about 4 billion Internet users, according to [2], online advertisement is essential for most companies. Since users hardly click on links beyond the first search engine result page (SERP) modern organizations try to improve their website´s ranking using search engine optimization (SEO). There are currently over 200 factors, which are considered by the complex algorithm that can determine whether a user visits the page [3]. Additionally, there are two major areas where improvements for a greater ranking can be applied: on-page optimization and off-page optimization [4].

Another type of online advertisement related to search engines is paid placements. This is operated by search engines in the form of sponsored results, which places an advertisement in a pre-specified region of a search engine result page. The placement fees charged by the search engines are directly related to the price of the relevant keywords, which is measured by cost per click (CPC) and the number of click-throughs the advertisement receives [5]. Whenever a better search result placement is achieved by paying a fee to the provider it is called inorganic; it is organic if the website warrants placement high in the SERP listing merely due to intrinsic merit [3].

SEO is getting more and more attention, because the price of keywords for the CPC in the search engine market has increased excessively over the years and it has been realized that organic search results are more engaging to users, because such results are considered more objective and unbiased than sponsored ones. As a result, SEO firms, which support companies in optimizing their web pages for better search results, are gaining momentum [6].

Even though SEO is a crucial part of online advertisement nowadays, there are still only few published academic research papers in this field. Therefore, this paper will focus on the various factors and methods of on-page as well as off-page optimizations and will then proceed to extrapolate the most important ones.

This research paper is structured in the following manner: Section II introduces the background of this research paper and covers the related work. Section III explains the anatomy of a modern search engine. Section IV comprises the different methods of SEO techniques and finally, section IV presents the conclusion and future scope of this paper´s research.

II. RELATED WORK AND BACKGROUND

A. Search Engines

Each search engine performs differently based on the implementation of its algorithm, but follows similar practices

such as: index words and their location and then give the user the possibility to search for results by using keywords or combinations of the indexed words [7].

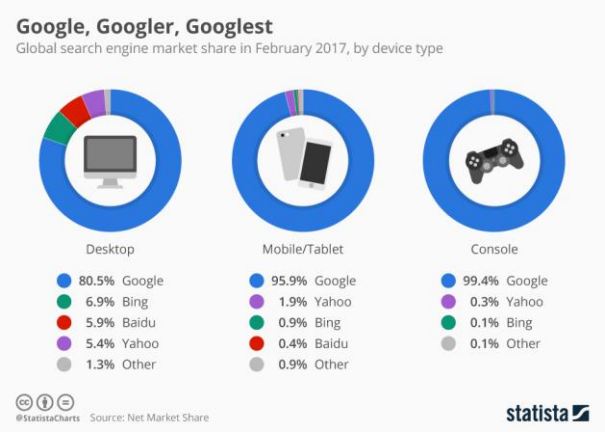


Figure 1. Market share of search engines
https://www.statista.com/chart/8746/global-search-engine-market-share (Jan, 2018)

On the desktop, over 90% of all search queries are handled by search engines from USA such as Google with 80.5%, Bing with 6.9% and Yahoo with 5.4%. The only search engine within the top 5 that is not from an English-speaking area is called Baidu, which is developed by a Chinese company. On mobile, Google has an even bigger market share with over 95% (Fig. 1). This is probably due to the reason, that Google is mostly natively implemented into their mobile operating system Android.

According to [8], even though the performance of the various search engines may differ, because of the different implementation of the applied algorithm, the vast majority of them rely on the same factors.

B. *SEO Methods*

SEO methods can be grouped into two areas: on-page and off-page optimization. On-page optimization covers the improvement of elements directly on the website, such as logical structure, correct navigation and required meta tags. In contrast, off-page optimization is mostly done after the website is already deployed and deals with techniques like indexing the website by using the search engine´s webmaster tools and managing link exchanges with other websites [9].

Furthermore, these optimizations can be divided into two major groups: black hat SEO and white hat SEO. While white and black hat SEO methods share the same goal to enhance a website's visibility, they use very different strategies. Black hat optimization is short-term success through quick increased traffic, due to the usage of unrelated keywords, doorway pages and cloaking. The search engine will punish the website by lowering its rank or completely de-indexing it. White hat SEO on the other hand is a long-term solution which relies on quality content, relevant backlinks and internal links [9], [10].

The most important factors elaborated by the case study [7] are the actual content of the page, the presence of keywords and qualitative links from a high number of external websites.

III. Search Engine Anatomy

A. *Query Interface*

The query interface is the visual part of a search engine which most people are familiar with. It is the first page users see when they visit a search engine to search for a specific term. While they were simple pages with a search box and a button to start the search in the past, nowadays most search engines capitalize on the real estate to them by adding personalized content. Modern search engines provide free e-mail accounts, news, sports, weather information and many other elements designed to make the users want to return to their personalized query interface [1, p.6].

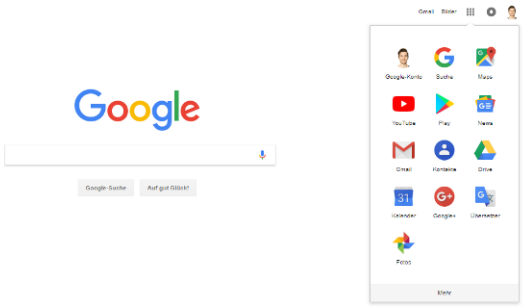


Figure 2. Query interface of Google

B. *Crawler*

Crawlers are lightweight and automated scripts that browse the Internet in order to download entire web pages or just portions of a single web page. The contents of these downloads are then indexed which allows the search engines to quickly match the searches to the index. Crawlers are also often referred to as robots or spiders [11, p.9-10].

In contrast to the query interface and search results pages, the crawler is behind the scenes and won´t be seen by the user. All the information it retrieves from web pages is cataloged according to the URL and are stored in the database afterwards [1, p.8-9].

Another crucial task of web crawlers is to delete information in the database when indexed websites no longer exist [12, p.141].

C. *Database*

The database is the place where search engines keep the information collected by the web crawlers. When users enter a search query into the search engine, it does not search the whole web for a proper match; it searches only in its own database of information about web pages that it has stored with the according URL. Since each search engine collects, stores and evaluates information of websites differently, the results are going to be different when typing the same search expression into multiple different search engines. There are different approaches of storing information in the database, for instance to store only parts of web pages or to store the complete content of every web page called *full text indexing*.

However, some search engines that claim to be full text indexed search engines omit common *stop words* such as "and", "it", "by" and "the" from their database [12, p.141-147]

D. Algorithm

Although all the previously mentioned parts are important, the search engine algorithm is the gear that makes everything work. The algorithm determines how the data is discovered by the user; it takes a phrase being searched for, evaluates several possible answers by sifting through the database that contains the keywords as well as the associated URL and then returns pages that contain the search query [1, p.11-12].

Due to the increase in popularity of SEO, especially methods categorized as black-hat SEO, it is important for the search engine to exclude "noises" such as keyword stuffing and link spamming in order to deliver quality results. This ability is called *algorithm robustness* [6].

IV. SEARCH ENGINE OPTIMIZATION

As a result of the findings in related work, this section will address the best practices for white hat SEO, which can be applied for the most common search engines. First, the on-page factors will be stated and afterwards the factors that are be applied off-page.

Beforehand the actual SEO, it is important to choose the right keywords, which are relevant to the page, but also where competition is low [7]. In addition, it is also possible to use online tools being Word tracker, Yahoo keyword selector, Google Adds keyword planer, Thesaurus, etc. to generate keywords. These tools return a large set of possible keywords by entering one keyword related to the website [9].

A. Title Tags

Title tags are perhaps the most important SEO tags on a website and their effectiveness can be improved even further by combining them with relevant keywords. Search engines use title tags to identify the topic of the web page. The best practice is to create a unique title for every site and make the title as descriptive as possible [1, p.98].

As stated by [13], the length of the title is another important factor and the recommended length is within 80 characters. Another suggestion for the length of the title tag presented by [14] is between 8 and 10 words.

B. Meta Description Tags

Meta descriptions serve as a short and accurate description of the web page in order to prompt the users to click on the snippets in the search results. The meta description should always describe the content honestly and be kept at 140 characters or less [13, p.262].

In combination with high ranked keywords associated to the content of the website, the description can increase the ranking even more. [16] applied the highest ranked keywords to 45 websites from different sectors in Jordan (Health, Business, Real Estate, Education, Construction, Beauty, Restaurants, Tourism, Agriculture) and investigated their increase in ranking after 2 months (Tab. 1).

Table 1. Sector ranking rate [16]

Sector	Ranking % after two months
Health	**75%**
Business	**60%**
Real Estate	**83%**
Education	**65%**
Construction	**62%**
Beauty	**73%**
Restaurants	**78%**
Tourism	**59%**
Agriculture	**55%**

C. Quality Content

Quality content, efficiency and the overall user satisfaction are essential factors, which influence the ranking of a website in search engines [13]. Bad implementation of one of those elements can directly result in an increase of the bounce rate. The bounce rate measures the percentage of users who visit only the first page of a website and then return to the SERP. If a website has a high bounce rate, the search engine assumes that the result was not a good response for that particular search query [13, p.547].

Besides in title and description, it is also important to use keywords in the content of the page. According to [8], a good practice is to use only one h1 tag and match it with the title tag of the page. Additionally, search engines favor rich and original content.

Usability has a big impact on SEO, especially in form of the initial loading time of the website. Users don´t like to wait an extended amount of time for the page to load [1, p.40]. A solution for this problem is lazy-loading, which allows to dynamically load content only as it is specifically requested.

D. Custom 404 Page

Whenever someone requests a page that is not available due to a typing error or a technical difficulty, the server answers with a 404 "page not found". If visitors encounter a 404 page, it is possible that they don´t want to make a second attempt to find the initially requested site. Accordingly, a custom error page can provide the user with feedback as well as with further actions to get closer to the destination [13].

In the results of [13] the customized 404 page has a significant effect on "Page Interest", which corresponds in the research paper to an increase of views on average per day, an increase of time spent on the website and a lower bounce rate.

E. Mobile Compatibility

Although the total visits of websites from mobile devices raised from 57% in 2016 to 63% in 2017 (Fig. 2), there are still fairly few publications about SEO with a focus on responsiveness of websites. Therefore, not only are there more web-users with mobile devices, the number additionally keeps growing. This lack of support for smaller

screens could explain the higher bounce rate from mobile devices compared to desktops (Fig. 3).

[7] observed that even though the website was optimized in terms of content, backlinks, keywords and titles, the bounce rate was still at 59.93%. Since 42% of the visitors are using handheld devices, they concluded that the high bounce rate was due to the lack of mobile-friendliness of the website. Furthermore, the bounce rate of their website is complementary to the average bounce rate of mobile devices in Fig. 3.



Figure 3. Total visits of desktop and mobile users in 2016 and 2017 https://www.stonetemple.com/mobile-vs-desktop-usage-study (Jan, 2018)

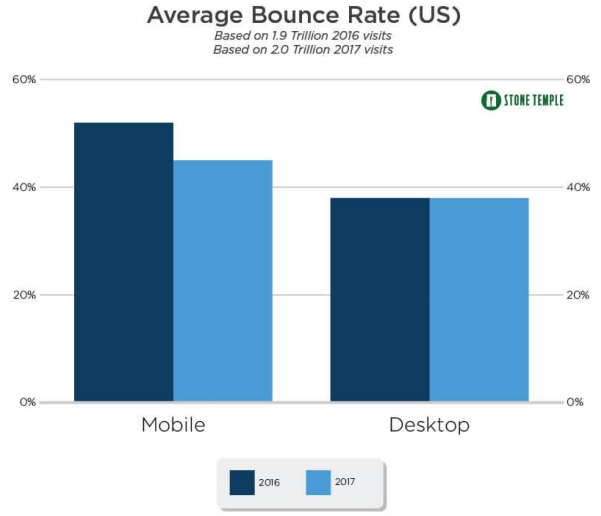


Figure 4. Bounce rates on desktop and mobile in 2016 and 2017 https://www.stonetemple.com/mobile-vs-desktop-usage-study (Jan, 2018)

F. *Keyword Density*

The keyword density describes the frequency of keywords compared to the total number of words on the page. The total number of words include the title, descriptions and overall content of the web page. This is important, because search engines use those keywords to categorize sites and determine the overall relevance for specific terms [13].

According to [13], the recommended keyword density is between 2% and 8%. In contrast, [9] states that an ideal keyword density for every page is a common myth and keywords should only be used in a way that makes sense.

G. *Off-page Optimization*

Off-page SEO revolves around different methods outside of the website´s scope to increase its ranking in search engines. [17] specifies backlinks (other sites linking to your website), also known as inbound links, as the most important factor of off-page optimization.

The relevance of backlinks is demonstrated in [7] where over 2500 visitors at the peak of the roughly 3500 users in January were coming from external sites. This results in a percentage of over 70%.

Su et al. [18] have stated that the quality of incoming links is much more important than the quantity. The Quality of the incoming link is measured by the page´s rank itself and the related keywords.

A web page without links is like an isolated island [14].

V. Conclusion and Future

This research paper has presented a wide variety of on-page and off-page methods to increase the rank of a website on the SERP. There are two major SEO groups, white hat- and black hat SEO, however, this paper has especially focused on techniques within white hat SEO, since it is long-term and approved by search engines. Furthermore, because the most essential factors for a successful search engine optimization can be applied on every modern search engine, this research doesn´t target a particular one.

The initial step for a substantial SEO is a comprehensive keyword analysis in order to find a set of keywords which is relevant to the web page as well as where the competition is fairly low. By entering the core keywords into online tools such as Word tracker, Yahoo keyword selector, Google Adds keyword planer and Thesaurus the keyword set can be quickly extended.

There are three main elements in a website where the previously established keywords should be embedded: title, meta description and content. The title tag helps the search engine to identify the content of the page and is the most important header tag for search engine optimization. The research has shown that the ideal length is within 80 characters or 8 to 10 words. Another important tag is the meta description tag, which describes the content of the site in a straightforward manner. Alike the title, the description should also be kept at a recommended length of about 140 characters. Although this research demonstrates that both title- and meta description tag have recommended lengths, it is advisable to not blindly trust those numbers and additionally keep a focus on the actual content of those tags.

While title- and description tags are primarily to differentiate between websites on the result pages through

snippets, the content of a website is the first thing users see when they click on such. The content influences the website´s page rank in two ways: bounce rate and keyword density. If the content has nothing in common with the search query, people will instantaneously head back to the SERP to find another website, which negatively affects the bounce rate and search engines will decrease the page´s rank for that specific keyword. Therefore, the keywords from the previous keyword analysis should be placed in a meaningful manner among the content. Further, each site should have a h1 that matches the title of the current page and a text

Presently, more people are visiting websites from mobile devices such as phones and tablets than from desktops. Consequently, search engines were further improved to evaluate the responsiveness of web pages, thus making mobile accessibility an essential part of SEO. Even though only few research papers are covering the mobile aspects of SEO at the current time, it is a known fact that mobile optimization has repercussions on the page´s rank. In addition, unresponsive pages increase the overall bounce rate of a website due to frustrated mobile users.

The results of this research have shown that a custom 404 page can help to keep the user on the website after an invalid request. A well-designed error page reduces frustration and gives the visitor multiple options to get back on track.

Lastly, there are SEO methods which are outside of the website´s scope called off-page SEO. The most important off-page SEO factor is called backlinks or inbound links. Quality backlinks can increase the page´s rank even more than some on-page optimizations. The idea behind this method is to find multiple other websites related to the content of the website that needs to be improved and have them contain incoming links. While it is still better to have backlinks from random sites than no backlinks at all, search engines favor quality over quantity.

Search engine optimization is an essential segment of web-development nowadays and will be even more important in the future as the number of websites will increase in the next years. The best time to start with SEO is as soon as the development of the web project starts, since elements such as title, description and content can be matched with the keywords from the start and don´t have to be readjusted later in the process. Once search engine providers start to make their algorithms more transparent to the average person, SEO will become more accessible and the quality of websites in general will increase.

REFERENCES

[1] J. L. Ledford and A. Edney, *SEO: search engine optimization bible, second edition*. Indianapolis, Ind.: Wiley, 2009.

[2] ‘Internet Live Stats, www.internetlivestats.com, (Accessed: Jan, 2019).

[3] V. N. Gudivada, D. Rao, and J. Paris, ‘Understanding Search-Engine Optimization’, *Computer*, vol. 48, no. 10, pp. 43–52, Oct. 2015.

[4] R. A. Malaga, ‘Worst practices in search engine optimization’, *Communications of the ACM*, vol. 51, no. 12, p. 147, Dec. 2008.

[5] B. Xing and Z. Lin, ‘The impact of search engine optimization on online advertising market’, in *Proceedings of the 8th international conference on Electronic commerce The new e-commerce: innovations for conquering current barriers, obstacles and limitations to conducting successful business on the internet - ICEC '06*, Fredericton, New Brunswick, Canada, 2006, pp. 519–529.

[6] Kai Li, Mei Lin, Zhangxi Lin, and Bo Xing, ‘Running and Chasing -- The Competition between Paid Search Marketing and Search Engine Optimization’, in *2014 47th Hawaii International Conference on System Sciences*, Waikoloa, HI, 2014, pp. 3110–3119.

[7] S. Krrabaj, F. Baxhaku, and D. Sadrijaj, ‘Investigating search engine optimization techniques for effective ranking: A case study of an educational site’, in *2017 6th Mediterranean Conference on Embedded Computing (MECO)*, Bar, Montenegro, 2017, pp. 1–4.

[8] W. Bin, G. Bingyun, L. Peishun, and L. Xiaoqing, ‘A study on tactics for college website at search engine optimization’, in *2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA)*, Shanghai, 2018, pp. 259–263.

[9] S. Duk, D. Bjelobrk, and M. Čarapina, ‘SEO in e-commerce: Balancing between white and black hat methods’, in *2013 36th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 2013, pp. 390–395.

[10] P. Patil Swati, B. Pawar, and S. Patil Ajay, ‘Search Engine Optimization: A Study’.

[11] S. Blankson, *SEO: how to optimize your web site for internet search engines*. U.K.: Blankson Enterpises Ltd., 2008.

[12] G. P. Schneider and J. Evans, *New perspectives on the internet: introductory*, Ninth edition. Boston, MA: Course Technology/ Cengage Learning, 2013.

[13] Fuxue Wang, Yi Li, and Yiwen Zhang, ‘An empirical study on the search engine optimization technique and its outcomes’, in *2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC)*, Deng Feng, China, 2011, pp. 2767–2770.

[14] C. Zhu and G. Wu, ‘Research and Analysis of Search Engine Optimization Factors Based on Reverse Engineeing’, in *2011 Third International Conference on Multimedia Information Networking and Security*, Shanghai, China, 2011, pp. 225–228.

[15] E. Enge, S. M. Spencer, and J. Stricchiola, *The art of SEO: mastering search engine optimization*, Third edition. Sebastopol, CA: O’REILLY, 2015.

[16] F. A. Zaghoul, O. Rababah, and H. Fakhouri, ‘Website Search Engine Optimization: Geographical and Cultural Point of View’, in *2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation*, Cambridge, United Kingdom, 2014, pp. 452–455.

[17] Jing Shi, Yan Cao, and Xue-jing Zhao, ‘Research on SEO strategies of university journal websites’, in *The 2nd International Conference on Information Science and Engineering*, Hangzhou, China, 2010, pp. 3060–3063.

[18] A.-J. Su, Y. C. Hu, A. Kuzmanovic, and C.-K. Koh, ‘How to Improve Your Search Engine Ranking: Myths and Reality’, *ACM Transactions on the Web*, vol. 8, no. 2, pp. 1–25, Mar. 2014.