Techniques to improve web sites positioning in the World Wide Web

DOI: 10.46932/sfjdv5n1-019

Received on: November 10th, 2023
Accepted on: December 13th, 2023

David González Marrón
PhD in Computer Sciences
Institution: Tecnológico Nacional de México, Instituto Tecnológico de Pachuca
Address: Carretera México-Pachuca Km 87.5, CP: 42080
E-mail: david.gm@pachuca.tecnm.mx

Verónica Paola Corona Ramírez
Master in Information Technologies
Institution: Tecnológico Nacional de Mexico, Instituto Tecnológico de Pachuca
Address: Carretera México-Pachuca Km 87.5, CP: 42080
E-mail: veronica.cr@pachuca.tecnm.mx

Juan Alejandro Arrieta Zuñiga
Master in Software Engineering Management
Institution: Tecnológico Nacional de México, Instituto Tecnológico de Pachuca
Address: Carretera México-Pachuca Km 87.5, CP: 42080
E-mail: juan.az@pachuca.tecnm.mx

Roberto Hernández Pérez
Master in Software Engineering Management
Institution: Tecnológico Nacional de México, Instituto Tecnológico de Pachuca
Address: Carretera México-Pachuca Km 87.5, CP: 42080
E-mail: roberto.hp@pachuca.tecnm.mx

ABSTRACT
In this computer's era, characterized by the rapid proliferation of indexed documents across the World Wide Web, the imperative arises for the employment of diverse strategies to enhance the positioning of websites within search engines, including but not limited to Google, Yahoo, and Bing. The voluminous expanse of indexed documents on the web, exceeding 30 billion, underscores the need for a systematic approach. This paper introduces the Search Engine Optimization (SEO) technique a method designed to elevate the inherent attributes of websites, facilitating superior evaluation and ranking by search engines, Notably, search engines incorporate over 200 distinct features within their ranking algorithms, encompassing internal factors inherent to developed websites and external factors predominantly associated with usage frequency. This paper details the series of activities related to the SEO technique, as well as the measurement results obtained before and after its application.

Keywords: SEO, Sites Indexation, WWW.

1 INTRODUCTION
Web browsing is one of the main activities performed on the Internet. Nowadays, it is a very common activity to search for entertainment, to learn, to interact, to shop, among other activities. Because
of that, if services are offered through the Internet, it is essential to have an attractive, functional, friendly and, above all, a responsive website, that is to say, the website can be displayed correctly from any digital device with which the user accesses; from a desktop computer to any mobile device. Two of the main characteristics of a website are design and optimization. The design corresponds to the visual and functional part of the page; therefore, it is the first thing that a user observes when accessing the website, that is why it must be attractive. However, it is also very important that it facilitates internal navigation and offers the user what he/she is looking for. Optimization, on the other hand, applies internally to the website, to all the pages and elements that include text, images, videos and visual effects. To make the site fast it is necessary that both the code, structure and content of the website need to be optimized to be effective in maximizing the resources of the device to be used and obtain efficient performance in the web browsers used. This work focuses on the development of a website, from the layout, structure, development, optimization, and responsive design, up to the implementation of the fully functional site on a web server itself. The need to have a good positioning when searching on the Internet is essential nowadays, due to the fact that the size of the Web is constantly growing. Within the documents that reside on the Web, an existing classification is between the indexed Web and the non-indexed Web. The size of the non-indexed Web or Deep Web is estimated to be 9 times larger according to (Becerra, 2018) and (Zilman, 2019) compared to the size of the indexed Web (Kaspersky, 2021), but these are restricted accesses and generally require authentication. Within this we can mention mainly the databases of public and private organizations that require robust authorization credentials for access, there are also other sites that are not indexed in which they seek to preserve anonymity, to preserve freedom and protection against social, cultural or governmental repression, or to a lesser extent to commit illegal acts, these sites, where non-legal activities are performed, are known as the Dark Web (Gehl, n.d), although its size is only 0.01% of the total Deep Web, here are concentrated many illegal activities. In this work we analyze the indexed Web, since it is the one that most users use to perform activities of more generic use within the Web, access is done using different types of search engines such as Google, Bing, Edge, Yahoo and others. The size of the indexed Web, although much smaller than the non-indexed Web, is very large and is in continuous change. It is currently estimated that the size of the indexed Web is made up of almost 2 billion websites (Gray, 2018) and about 60 billion web pages (Kunder, 2023), so it is necessary for organizations seeking to consolidate their presence and relevance to be in the top positions. One widely used strategy is the use of advertising through paid ads on different sites, and another strategy is to improve the design of the sites so that they are more efficient, useful and secure.
2 METHODOLOGY DESCRIPTION

In the development of websites, agile methodologies are currently mainly used, due to their high efficiency (Gary, 2023), in this case the Scrum methodology was selected because with this methodology a lot of flexibility is obtained in the project, since it is relatively easy to adopt changes and new requirements during the development of the project. Transparency is benefited because everyone involved has an overview of the project, due to the continuous inspections required by the methodology, it is also possible to detect errors during the progress and through the development cycles or "Sprints" used by the development team, likewise it is possible to schedule a sprint when there is an important adjustment to be made. This is why this methodology allows the delivery of value in reduced periods of time.

There are many recommendations on the implementation of SEO techniques (Iggy, 2018), below are the main activities used in this work to achieve SEO, although these activities are not difficult to perform, they provide a better positioning as can be seen in the results section.

2.1 USING FRIENDLY URLS

The use of friendly URLs is one of the most important aspects taken into account for the optimization of a web page, since it is the gateway to the site.

This process consists of two parts:

Normalization: This is a set of Internet standards to facilitate the crawling of a web page, in order to avoid that the same URL (Uniform Resource Locator) belongs to another existing resource on the Internet. The most common processes are: conversion of uppercase to lowercase, removal of dots, removal of signs and unusable query parameters, it is also recommended that accents be removed for a more generic use in other languages.

Optimization: Optimization cannot be performed without first normalizing a URL. The most important aspects for optimization are: Generate short URLs, use the title as the URL path, remove duplicate URLs, remove file extensions, change dynamic URLs to static URLs. The latter is important for both search engines and users, since the simpler a URL is, the easier it will be for the user to remember it and for the Internet search engine to interpret and index it. A clear example of this can be seen in Figure 1.

Figure 1 Difference between Dynamic URL and Static URL

<table>
<thead>
<tr>
<th>Dynamic URL</th>
<th>Static URL</th>
</tr>
</thead>
</table>

Source: realized by the Authors to emphasize the specificity in static URL.
2.2 WEB DESIGN

Having a good Web Design consists of having a good location of the elements that make up a web page. The most relevant aspects to take into account for a good web design are:

**Markup** : This point is nothing more than the programming of the page in HTML language. When a search engine crawls a web page, it not only reads the text visible to the user, but also checks the coding.

**Tagging** : It refers to the process of defining styles of the elements of a web page. Although this part can be done from HTML, it is advisable to use independent style sheets, since this way a better control is maintained.

**Navigation structure**: Having a good navigation structure is fundamental to achieving a good SEO result. This helps users to find the sections quickly and helps Internet browsers to have a better view of the web page. An example of a good navigation structure is shown in Figure 2.

![Figure 2 Example of a correct navigation structure for a particular Web design](image)

Source: realized by the Authors to illustrate a good navigation structure

**Responsive design** : A website that has a good design both on computer screens and on smartphone or tablet screens is good to have a good SEO result (Responsive Design).

2.3 PAGE TITLES

When a browser analyzes a web page, the title of the web page is the first thing that is displayed, so having an "attractive" title for the user is essential for SEO. These are the main technical aspects to take into account for web page titles:

- Do not exceed 70 characters (based on Google's recommendations).
- To be included within the HTML `<title>` tags.
- Include the main keyword or phrase of the article.
2.4 CONTENT

Content optimization is paramount for SEO, and it starts with the page title and the content title.

- Web page title (inside the <title></title> tags)
  It must be brief and clear to the user
- Content title (inside the <h1></h1> tags)
  It can be the same as the previous one, but this one has the freedom to have more words to better explain the subject of the web page.

Many times, in a page there are several headings and subheadings, so it is recommended to make good use of these tags (<h2></h2>, <h3></h3>, <h4></h4>...).

As for the texts used, it is recommended:

- Have short paragraphs, but that explain in a good way the central idea of the page.
- Aligning texts to the left makes them easier to read, especially for people with vision problems.
- Use simple language and avoid the use of words that are difficult to understand.
- Use bulleted lists or numbers to define points, requirements, materials, ingredients, etc.
- Follow writing styles according to the nature of the subject matter of the content.

2.5 OPTIMIZED IMAGES

Images are elements that help to improve the idea and attractiveness within a web page and are a very important factor to provide good response times when displaying information. The image optimization process basically consists of the following two parts:

Image characteristics.

To have images that improve SEO we recommend the following:

- Reduce the weight of the image to its maximum expression.
- Name the file with a simple and clear description.
- It is recommended to sacrifice a little of the quality of the image to reduce its storage size (Vanderbilt, 2023), which allows to improve the loading time of the pages, reducing its costs. A heavy image increases costs to the user and the website owner.

Use of the image on the web page.

- This is focused on the HTML code, with the use of the attributes "title" and "alt". Which helps Internet browsers to better understand the purpose of an image in a web page with "title" it is possible to put a small title to the image, and with "alt" we can give a brief or broad description of it.
2.6 SITE MAP

The sitemap is a file with information about pages, videos and various files used on a website, as well as the relationships between them. It is of great importance to check the backlinks.

3 IMPLEMENTING METHODOLOGY

Once the web pages of the company's site were developed, optimization techniques began to be applied, starting with the search for "keywords", which are key words that are important for SEO positioning, since Internet search engines need to know what the content of a website is in order to be able to classify it (Gabbert, 2018). These words refer to both the content and the target audience. These words are located in the titles and paragraphs as well as in the name of the images and videos used in the website. In order to know which keywords are ideal for a website, a research must be done, since there are several aspects to take into account, such as search volume, SEO difficulty, payment difficulty and cost per click. To do this, there are several tools that help to analyze all this, as is the case of "ubersuggest" (Neipatel, n.d.), which, through a query provides the keywords that help to improve the SEO of a website, in addition to giving suggestions and synonyms of other keywords that help to "attract" new users. Once the definition of keywords is done, it is necessary to optimize images (Vanderbilt, n.d.). Several SEO experts recommend reducing the quality of the image up to 80% to achieve the ideal weight, these values vary according to the characteristics of each image and its importance in the web pages required. In addition, they indicate that converting images to the WEBP format developed by Google, helps to reduce the size of the images and thus reduce the loading time while retaining the highest possible quality, compared to formats such as JPEG or PNG. Images in JPEG format with an average weight between 300 and 400 kb. After converting them to WEBP format, sacrificing 20% of their original quality, images were obtained with an average weight between 70 and 100 kb. There were cases in which some images reached a final weight of 30 kb. Several tools can be found on the Internet to do this, one of them is the site "online-convert.com" (onlineconvert, 2023). On the other hand, it is important to name the images in a way that helps search engines to better understand your idea, allowing to have a better control of them. For example, we went from naming images with little meaning for example "img-001.jpg" to a content with more meaning "security-cameras.webp". The next recommended activity is to apply keywords in paragraphs and titles of the web site, this, with the purpose of increasing the interest of the users that visit the site. As well as images, it is important to optimize the files used in the site. By optimizing files such as style sheets, JavaScript (.js) and HTML (.html) documents, the loading time of web pages can be improved and accelerated. By reducing their weight and cleaning up extra code, the loading of these files is facilitated, which benefits the devices that consult these pages, mainly those with few resources. For this there are good tools, among these we can mention NorfiPc (Norfi, 2023). In order to apply the "friendly URLs",...
the ".htaccess" file is used, which, through regular expressions, converts a dynamic URL to a static one. An example of this file is shown in Figure 3.

![Figure 3 Example of the contents of a .htaccess](image)

Source: File realized by the Authors using commands to rewrite a dynamic URL into a Static URL generated with the Norfipc tool (Norfi, 2023)

One of the last steps in SEO optimization is the creation of a sitemap, which helps search engines easily crawl the site's pages, as well as the most recent and hard-to-find pages. An example of the main components of an XML sitemap is shown in Figure 4.

![Figure 4 Basic Example of XML Sitemap main components](image)

Source: realized by the Authors to crawl and index sites more effectively.

The last proposed activity is to review the modifications made using Google PageSpeed Insights, an open-source tool that indicates the performance in terms of optimization (Ivanovs, 2017), pointing out where corrections are needed. An example of its use can be seen in Figure 5.
4 RESULTS

The following bar graphs show a comparison between the results without SEO optimization and those obtained after the SEO optimization described in this work, the performance measurement is performed using Google PageSpeed Tools (Ivanovs, 2017) considering the main pages of a selected website, the desktop version and the mobile version were evaluated; in previous versions of the tool the aspects regarding responsive design, use of graphics and optimized elements were not considered, this can be seen in Graph 1 using the desktop version.

Graph 1: Ratings Comparison using Google PageSpeed Tools desktop version for the six pages that conform the system evaluated.

Source: By the Authors.

Graph 2 shows the comparative ratings of the pages in the mobile version.
Table 1 shows a comparison between the average scores of 6 pages analyzed, before and after optimization, evaluated in the desktop and mobile version of PageSpeed Tools.

As can be seen in the table above, both versions show a considerable improvement. In the desktop version there is an improvement of approximately 12% over the performance of the previous website and in the mobile version, due to the fact that the previous site did not have a fully responsive design, an improvement of over 70% is obtained on average.

**5 CONCLUSIONS**

As a conclusion to this project, the most relevant aspects are mentioned in chronological order. Before carrying out any project, it is very important to have a good plan of what you want to do and then carry out a study covering the most important details to achieve the proposed objective. A very important part of this project, and that allowed adequate development, was due to the use of the SCRUM methodology. Thanks to the fact that it allows us to work actively on the project in short delivery periods, it also provides a lot of flexibility to make changes, something very common in this type of development.

At the end of the project, we obtained a fully functional website, which complies with the following characteristics:
1. Optimized for cell phones.
2. Optimized for desktop computers.
3. Optimized for Tablets.
5. Direct links to contact the company.
7. Images and videos with pop-up formats.
8. Optimized resources; .css, .html, .js files.
9. SEO optimization.

Although characteristics one through six are the most related to aesthetic aspects of the site, this is of fundamental importance because if the site is liked by users, this will determine whether users stay or leave the site quickly. Characteristics seven and eight seek to achieve a good use of resources, and finally nine verifies the achievement of a good SEO optimization, which indicates a correct internal structure of the website. If a website is developed considering what the SEO technique proposes, it allows the website to obtain a privileged position within the first results in users' searches in web search engines. Once the SEO optimization is applied, the user can perceive an improvement in the performance of the site, however there are automated tools that make a more specific evaluation of various aspects that affect the performance of the sites, in this case we used one of the most complete and easy to use tools such as PageSpeed Tools which is a Google software to help optimize the performance of websites (Ivanovs, 2017).
REFERENCES


