

# Simplified Web-based SEO Score Counter for Increasing Web Page Rank

Raka Yusuf\*, Rusti Mentari, Harni Kusniyati

Faculty of Computer Science, Mercu Buana University, Jakarta, Indonesia

## ARTICLE INFO

### Article History :

Accepted: 10 Jan 2024

Published: 30 Jan 2024

### Publication Issue :

Volume 11, Issue 1

January-February-2024

### Page Number :

141-150

## ABSTRACT

SEO (Search Engine Optimization) is a technique intended for search engines to always index a website on the Internet with high value, so that each search process using certain keywords through search engines can display the website on the first page of search. This research will create a simple web-based application to calculate the ranking of a website from an SEO perspective. A website with a high SEO ranking is expected to appear on the first page of search results. This application uses the cURL function contained in the PHP programming language to read web pages. The result is SEO value as an initial picture of how a website can appear on the first page of search results. Search Engine Optimization, commonly known as SEO, is a highly effective technique that guarantees a website's continuous indexing by search engines at a high value. This results in the website being displayed on the top of search results when specific keywords are searched. Our research is dedicated to creating a web-based application that calculates a website's SEO ranking. This ranking is crucial to ensure that the website appears prominently in search results. The application employs PHP programming's cURL function to scrutinize web pages, providing an initial indication of a website's search result performance.

**Keywords:** *Web optimazion, SEO calculator, cURL technique*

## I. INTRODUCTION

According to the Merriam-Webster dictionary, the definition of a website is a group of World Wide Web pages usually containing hyperlinks to each other and made available online by an individual, company, educational institution, government, or organization [6]. Nowadays the term website is something familiar to the public. Not limited to only institutions or companies, every individual can also create a website for personal matters like hobby such as just writing or

blogging. Web technology has developed rapidly. Web development is mostly done by technology companies both small and large. This web development includes content development, web design, servers, network security, and other things.

Along with its development, a website is no longer used as a work or information tool, but now it can also be used to communicate more broadly, such as Twitter and Facebook. There is also an e-commerce website that is used as a meeting place for sellers and buyers.

There are so many competitions that occur in cyberspace world and require web developers to compete creating websites that has eligibility standards in various contents and business processes.

In developing websites that can show up on the first page of a search engine result page (SERP), developers can use a SEO (Search Engine Oprimization) technique. This is a technique that can make a website more widely known to the public. There are two types of SEO techniques, namely On-Page SEO and Off-Page SEO. On-Page SEO is a technique for optimizing website inside its pages by modifying certain factors as parts of a website page, such as titles, tags, contents that are relevant to the titles and other factors. Off-Page SEO is an optimization technique that is carried out from outside the website while still referring to certain search engine algorithms, for example increasing quality backlinks. The problems that often arise are the fact that website creators do not know how good the SEO rank of their websites in order to improve the quality of the website. They need to count SEO score of their website pages.

Today there are so many web-based tools that can completely count SEO scores, but people should pay for such services, especially when we use those services in daily manner. Few to mention are Woorank [7] and iWebChk [8]. This research will help web developers to count a basic SEO score of a website. Here we're going to use cURL method to read parts of a website page.

This paper will present how to design and build an SEO score counter website. Some websites will be used for experiments to compare with the score calculation results from other SEO scorers. Calculation of this score will be useful for developers, especially novice developers who need a quick calculation of the basic SEO score so that they can immediately improve the website they built without the need to register on paid SEO calculating websites.

In calculating the SEO score, here we will use the text reading on the website using the cURL technique, then calculate the HTML components using the SEO technique used by Woorank. As a start, this score counter will be tested on the Mercur Buana University website in Jakarta [11]. The Waterfall methodology is used as a guide for the steps in carrying out this research.

## II. METHODS AND MATERIAL

### A. Previous Researches

There are several factors that impact SEO score. Some of previous researchers say that the number of backlinks on a web is one of those factors [12,13,14]. This makes sense since those links tell us how valuable the web is. The other factors are the average user visit duration and bounce rate [13,15]. The higher visit duration causes the higher SEO score and the lower bounce rate causes the higher SEO score.

Some other researchers say that the number of clicks or page visits san also impact SEO score [13,16,17]. SSL certificate utilization and the low number of errors in W3C are also the factors that impact SEO score [18,19,20]. In this reasearch we will focus the discussion to the on-page SEO optimization. As discussed in later section, on-page SEO optimization focuses on the structure of HTML page.

### B. Internet

The Internet first appeared in the United States which was initiated by the Department of Defense in 1969, through the ARPA project or also called ARPANET (Advanced Research Project Agency Network). In the project they demonstrated that by using UNIX-based hardware and software, communication can be done over unlimited distances over telephone lines. The ARPANET project formed the forerunner of TCP/IP (Transmission Control Internet Protocol). They design the shape of the network for reliability, how much information can be transferred and all the standards.

At that time, the Internet was intended for military purposes, but over time the Internet developed for education and general use.

Internet is an acronym for Interconnection Networking. The internet comes from the Latin word "inter" which means between. The internet means connecting or a network between, so the definition of the internet is a connection between various types of computers and networks in the world with different operating systems and applications where the relationship takes advantage of communication advances (telephone and satellite) that use standard protocols in communicating, namely TCP/IP (Transmission Control/Internet Protocol) protocol.

The internet is a global network of world computers, large and very broad where every computer is connected to each other from country to country around the world and contains various kinds of information, ranging from text, images, audio, video and more.

### C. Website

The term website was first put forward by Sir Timothy John Berners-Lee and a website connected to a network first appeared in 1991. Tim's intention when creating a website was to make it easier to exchange and update information to fellow researchers at his place. work. On April 30, 1993, CERN (the place where Tim worked) informed that the WWW could be used free of charge by everyone.

In accessing information from the internet, users will go to an internet link address called a domain (Domain name/URL-Uniform Resource Locator) and will find information in the form of text, images, animation or sound in a medium called a website or site. web. This website is opened through a browsing program (browser) that resides on a computer. Browsing programs that can be used on a computer to be able to browse websites include Internet Explorer (IE), Firefox, Mozilla, Netscape, Safari and Opera.

The website consists of web pages, namely a document written in Hyper Text Markup Language (HTML) which can be accessed via the Hyper Text Transfer Protocol (HTTP) Protocol which is a protocol for conveying information from a central website to be displayed in front of users of information reading programs. is on the website.

### D. HTML

HTML is an acronym for HyperText Markup Language. HTML can be made in any editor, although there is special software intended to make work easier. The editors in question are NotePad, WordPad, MS Word, Excel and others. Making HTML text is almost the same as making other texts such as in MS Word. Formatting text in a document will immediately show the results.

HTML is used to build a web page. Even though many people call it a programming language, HTML is actually not a programming language at all, because as its name implies, HTML is a mark-up language. HTML is used to mark up a text document. These marks are used to determine the format or style of marked text.

### E. Search Engine Optimization

A search engine is a computer program designed to help find information on the internet. The Search Engine has an electronic database containing up to billions of scattered information site addresses. By typing in the keywords you want to search for, the search engines sort out the information in the database, after which several links will appear that will lead to the keywords previously entered. To speed up the search, search engines have certain methods, namely search algorithms. The speed of searching for information depends on the method used.

The World Wide Web is the first search engine application that uses robot technology to index web pages on a web server.

Danny Sullivan believes the term Search Engine Optimization was first used on July 26, 1997 by a spam

message posted on Usenet. At that time search engine algorithms were not necessarily complex so they were easy to manipulate.

Search Engine Optimization (SEO) is a series of processes that are carried out systematically with the aim of increasing the volume and quality of traffic through search engines to certain websites by utilizing the working mechanisms or algorithms of these search engines. The goal of SEO is to place a website in the top position or first page of search results based on certain targeted keywords. Logically, websites that occupy the top positions in search results have a greater chance of getting visitors. This opportunity is used by a number of parties to offer search engine optimization services for companies that have an internet business base.

A search engine is a computer program designed to help someone find files stored on a computer, for example on a public server on the Web (WWW) or on a computer itself. Search engines allow us to query media content with specific criteria (usually in words or phrases we specify) and obtain a list of files that meet those criteria. Search engines usually use the index to search for files after the user enters search criteria.

SEO is divided into two types:

#### On-Page SEO

On Page SEO can be said to be the foundation of a website, the better the On Page SEO settings on the website, the stronger the foundation will be. SEO On Page contains optimization techniques with certain factors on existing websites.

#### Off-Page SEO

Off Page SEO is optimization that is done from outside the website, or in other words not directly related to the website. On Page SEO can be factors that exist on your website, but are used in search engine algorithms when determining site rankings.

### F. SEO Strategy

SEO is a collection of strategies to improve the level at which a website is placed in a high position in the

search results list by using relevant keywords. Tags, links, navigation structure, and content are some of the elements that attract the attention of crawlers.

There are two techniques used in SEO assessment:

#### 1. White Hat SEO

White Hat SEO is believed by webmasters to be the safest technique in SEO. White Hat SEO is the best technique in marketing even though competition between webmasters is an obstacle. Web owners must think that content is king by providing precise and accurate information. Some of the techniques used in white hat SEO:

- a. Targeted information,
- b. Organize the web well,
- c. Providing meta tags,
- d. Create relevant keywords (title, meta description, meta keywords, heading h1 elements, text with “ “, alt image tags, title tags, links containing meta keywords).

#### Black Hat SEO

Black Hat SEO has proven to be effective in exploring the first page of search engines. However, this technique is considered a technique that is fraudulent or unnatural because it does not go through a process that is in accordance with usual optimization. Some of the techniques used in Black Hat SEO:

- a. Keywords Surfing.
- b. Cloaking.
- c. Invisible keyword.

There are causes and effects that arise when using deviant techniques. In general, this technique is a violation of ethics in optimizing a web. Even though the web is on the first page, after that the web will be blacklisted, then not indexed, then it will enter the search engine sandbox and may be banned.

### G. cURL

cURL is one of the most powerful PHP extensions. This stands for Client URL which functions to communicate or retrieve data on another server using various

protocols. cURL is a library created by Daniel Stenberg, which allows us to connect and communicate with various types of servers with various types of protocols. The most frequently used string functions and cURL functions are as follows:

1. *curl\_init* : function to start a cURL session
2. *curl\_setopt* : function selects multiple options
3. *curl\_exec* : this function to execute the transfer
4. *curl\_close* : function to close the session

### III. METHODOLOGY

This research will use software engineering methodology. Software Engineering is the creation and use of the principles of engineering expertise to get software that is economical, reliable and works efficiently on real machines [1].

In general, software engineers take a systematic and organized approach to their work because this is often the most effective way of producing high-quality software. However, this engineering actually involves selecting the most appropriate method for a given set of circumstances and a more creative, informal approach to development that may be effective in some circumstances. Informal development is well suited for the development of web-based e-commerce systems that require a combination of software and graphic design expertise.

The sequence of steps used in this study are as follows:

1. Analysis of software requirements

The process of gathering requirements is carried out intensively to specify software requirements so that it can be understood what kind of software the user needs. Software requirements specifications at this stage need to be documented.

Design

Software design is a multistep process that focuses on the design of software programs including data structures, software architectures, interface representations, and coding procedures. This stage translates software requirements from the needs analysis stage to a design representation so that it can

be implemented into a program at a later stage. The software design produced at this stage also needs to be documented.

Code Generation

Design must be translated into software programs. The result of this stage is a computer program according to the design that was made at the design stage.

Testing

Testing focuses on the software from a logical and functional perspective and ensures that all parts have been tested. This is done to minimize errors (errors) and ensure the output produced is in accordance with what is desired.

Support and maintenance

It is also possible for a software to change when it is sent to the user. Changes can occur due to errors that appear and are not detected during testing or the software must adapt to a new environment. The support or maintenance phase can repeat the development process starting from specification analysis for changes to existing software, but not for creating new software.

### V. RESULTS AND DISCUSSION

#### A. System Analysis

System analysis can be interpreted as a process to understand the existing system by evaluating and correcting several problems so that it becomes a complete and functioning system.

System analysis identifies the requirements that will be required in making this system. Therefore system analysis is the earliest stage so that in solving problems in a software according to the processes that will be applied from the beginning to finally becoming a unified system.

Website review is an assessment of a web-based application system that is useful for knowing the ranking of an application made by a user in order to optimize the performance of the website. That way, users get the pros and cons of the websites they want to test. There are many review websites that provide

these services for free (free) and are complete, but there is a specified time period for accessing them.

Based on the description above, the authors create a web-based system that is able to assess the performance of a website for free (free) and can be used without any specified time limit. The method used is cURL and the PHP programming language (Hypertext preprocessor).

### B. Problem Analysis

Problem analysis is a process of looking for what needs are needed by the system to be made. In designing this system the author faces several stages of problems that must be carried out, including:

1. Finding the location of the data that becomes the keyword for the Search Engine Optimization assessment on the website you want to test using the cURL method.
2. Retrieval of title data, h1, meta tag description, meta language, meta keywords.
3. Calculates the score obtained by each meta tag.

### C. Comparison

Web-based applications that function to assess the ranking of a website already exist in the internet world. Some of these websites include Woorank, Marketing Grader and IwebChk [8][9][10]. These applications provide detailed assessment results and information. However, these applications cannot be accessed for free. If you want to use it, the user is given a policy to pay.

In making this application the author chose Woorank as an application to be compared. Woorank can search for data, calculate and retrieve data contained on websites that are in the process of being assessed. System analysis will go through several processes including:

#### 1. Input

This application requires the address of the website to be entered and then assessed the level of performance that is owned by the web.

#### 2. Process

The process that will be carried out is searching and retrieving data on keywords which are the main objectives of SEO calculations such as title, h1, meta description, meta keywords, meta language. Output The output resulting from the process that has been carried out is a page that displays the scores and data contained in the content accompanied by a description of the assessment provisions set by www.woorank.com.

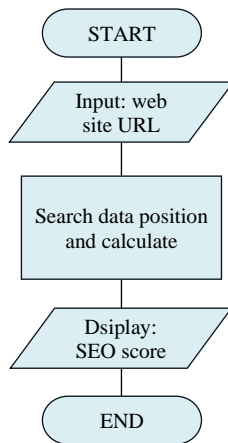
### D. Calculation Process

This process is preceded by a cURL that redirects to the address of the website to be tested. Then the next step is knowing the location of the data that is the goal and starting to do the calculations.

Based on the previous description and analysis, the outline of the process that takes place in general is depicted as shown in Fig 1. The process that will be described in this flowchart modeling is in the form of the initial step of the user entering the website address. After that, a search and calculation process will be carried out based on the variables contained in the keywords and proceed to the next stage, namely displaying a new page of assessment results from the system.

The flow of the data calculation process is a process in more detail than the continuation of the application flow process in general, namely:

1. Users are required to write the website address in the column provided on the first page.
2. The next process is the cURL function directed to the website address that was previously entered by the user.
3. When cURL has been redirected to the website, then start the process of calculating the title. In accordance with the provisions on Woorank, a good title has a character length between more than 10 and less than 70. If the title complies with these conditions, it will get a score of +3. Search for keywords for this title using the strpos function. The strpos function functions to find the position or location of the data we want to find.



**Fig 1.** General Application Flow Chart

After the above process, proceed to the next process, namely searching for H1 keywords. The function used to get the location or position of keywords is `strpos` as explained in point 3 before. If H1 is found, you will get a score of +3. However, if H1 is not found then the score obtained is 0. Then take the data contained in the H1 content and display it.

The next calculation is to detect the existence of the meta description and retrieve the content contained therein. A good meta description has a character length between more than 70 and less than 160. If the meta description is detected, it will get a score of +3, if it is not detected, it will get a score of 0. Then print the contents contained in the meta description.

The next process, still referring to the provisions of `www.woorank.com`, is to detect the presence of meta keywords and retrieve the content contained therein. The provisions set to determine the results of calculating the meta description range from more than 70 to less than 160. If meta keywords are detected, a score of +3 will be obtained. However, if it is not detected, it will get a score of 0. Then print the contents contained in the meta keywords.

The next calculation is to detect the presence of DOCTYPE. DOCTYPE is a useful declaration to identify the type of HTML document used so that the

browser can determine how to treat the code. In this calculation if there is a DOCTYPE it will get a score of +3. But if it is not detected then it will get a score of 0 and the results page will print "yes".

The next process is to detect the meta language. The meta language function is to determine the language syntax used in the web being tested. In this case if there is a meta language then it will get a score of +3. However, if it is not detected, it will get a score of 0. Print the content of the meta language.

Next, detect the presence of facebook and twitter. Same as the previous process. If Facebook and Twitter are detected, you will get a score of +3 then print the link. If the Facebook and Twitter keywords are not found, a score of 0 will be obtained.

The next process is to detect the presence of images used by the website being tested. If it is detected, it will get a score of +3. However, if the calculation does not detect an image, it will get a score of 0.

The next step is to find out whether the website uses `robot.txt` or not. `Robot.txt` is a file contained in the root directory of a website that functions to control certain pages from being indexed by search engines. In this calculation, if `robot.txt` is detected, the score obtained by the website is +3. However, if `robot.txt` is not detected, the score is declared 0.

The next step is to search for `sitemap.xml`. `sitemap.xml` is the architecture of a website to make it easier for search engines (search engines) to explore the contents of a website. Just like the calculations from the previous points, if there is a `sitemap.xml` on this website, it will get a score of +3. But if not then get a score of 0.

Those are 12 elements that count the most. There are many other elements that affect the SEO Score. However, in this simplified counter we only use the top 12 elements.

Code 1 shows an example of retrieving meta tag process. Other process are more or less similar to this.

**Code 1.** The meta description content fetch.

```

<?php
//checking meta description
$p1 = strpos($hasil, '<meta name="description"
content="')+strlen('<meta name="description"
content="');
$p2 = strpos($hasil, '"', $p1+1);
$des = substr($hasil, $p1, $p2-$p1);
$star2 = 0;
if (strlen($des)>=70 && strlen($des)<=160){
$star2 = "
<i class='glyphicon glyphicon-star'></i>";
};
?>
<a href="#" class="list-group-item"><?php echo
"Meta Description : ".$des; ?></a>
<a href="#" class="list-group-item"><?php echo
"Score : ".$star2; ?></a>

```

In code snippet 1 in line 1 there is a variable \$p1 containing strpos which is useful for searching for the location/position of the first letter in the content in the meta description. In line 2 the \$p2 parameter searches for the location of the last letter so that in line 3 the combination of the parameters \$p1 and \$p2 functions to print which characters \$p1 and \$p2 are found that are stored in \$desc. Line 5 is \$des more than 70 less than equal to 160 stored in \$star2.

Calculating the meta description is different from the title. The provisions set by Woorkank are if the contents of the meta description content are more than equal to 70 and less than equal to 160 then the content owned can be said to be good.

**E. Testing Result**

In order for this program to function properly, the authors tested the seo counter website application. In order to find out whether the application is running well or software, in conducting tests it can be defined

as not in accordance with the design that has been made.

The testing method using black box testing is testing that focuses on the functional specification of a set of input and output conditions as well as testing on the functional specifications of the program.

Errors found when using the black box method include:

1. Missing or incorrect function
2. Interface error.
3. Data structure errors or external database access.
4. Performance error
5. Initialization and termination errors.

Test scenarios are made aiming to find out the errors that occur during the testing process. In addition, this trial was carried out to ascertain whether the functions contained in these application modules were as expected.

The stages of this test include test scenarios, test results and perform an analysis of the results of previous tests. The testing method used in the testing process is the Black Box Testing method. Black Box Testing is to determine the steps in testing and check whether there are parts of the previous specifications that were not implemented.

Based on the test results, the authors obtain data from the testing process. The following is a verification of the test results:

1. The process of searching for content and retrieving data using cURL into the PHP programming language is as expected.
2. Certain websites still cannot index the content contained therein due to differences in the form of writing format used by each website.

The target website (Mercur Buana website) scores 80% SEO using this score counter.



## VI. CONCLUSION

As a result of the analysis of the SEO Score Counter Web application, the following conclusions can be drawn:

1. The cURL method can be used in searching and retrieving data/content within the source. As long as there are no changes to the writing of the content in it, there is little chance of failure.
2. The results of this simple calculation are enough to give an idea of how good a website's SEO score is.
3. It is very effective if the user does not know how far the website has a rating.

For future works, we are considering the following plans:

1. If in the future there are developers who want to develop these applications, they are expected to use different techniques.
2. Adding some ratings are still within the conditions provided by SEO.
3. Beautify the face to face (interface) to make it more attractive and interactive.

## III. REFERENCES

- [1]. Pressman, Roger S. 2002. *Rekayasa Perangkat Lunak Pendekatan Praktisi (BUKU SATU)*. Yogyakarta: Penerbit Andi.
- [2]. Digital Library – Universitas Komputer Indonesia [Internet], Retrieved June 7, 2020, from <http://elib.unikom.ac.id/>
- [3]. Eprints – Universitas Sebelas Maret [Internet], Retrieved June 8, 2020, from <http://eprints.uns.ac.id/>
- [4]. Unit Pelaksana Tes Potensi Akademik Bappenas [Internet], Retrieved June 8, 2020, from <http://www.upp-tpa.com/>
- [5]. Memahami Ecommerce, e-book [Internet], Retrieved June 9, 2020, from <http://kalamkata.com/>
- [6]. Jurnal Ilmu Komputer [Internet], Retrieved June 10, 2020, from <http://journal.ilmukomputer.org/>
- [7]. Supriyanto, 2008, *Teknik Informasi & Komunikasi SMP Kelas VII*, Hal: 26-50, Yudhistira, Yogyakarta.
- [8]. Woorank – Start Optimizing Your Website [Internet], Retrieved June 10, 2020, from <https://www.woorank.com/>
- [9]. Marketing Grader - Grade your marketing performance [Internet], Retrieved June 10, 2020, from <https://marketing-grader.com/>
- [10]. IwebChk - website SEO analysis & review tools [Internet], Retrieved June 10, 2020, from <https://www.iwebchk.com/>
- [11]. Universitas Mercu Buana – Jakarta [Internet], Retrieved June 10, 2020, from <https://www.mercubuana.ac.id/>
- [12]. Khan, M. N., & Mahmood, A. A distinctive approach to obtain higher page rank through search engine optimization. *Sādhanā*. 2018,43(3). <https://link.springer.com/article/10.1007/s12046-018-0812-3>
- [13]. Suksida, T. & Santiworarak, L. A study of website content in webometrics ranking of world university by using similarweb tool. *IEEE International Conference on signal and Image Processing (ICSIP)*. 2017. <https://ieeexplore.ieee.org/abstract/document/8124588>
- [14]. Zhang, S, & Cabage, N. Search engine optimization: Comparison of link building and social sharing. *Journal of Computer Information Systems*
- [15]. Narwal.A., & Chauhan,R.K. Web traffic perspective of private universities of Haryana *JIMS81-International Journal of Information Communication and Computing Technology*. 2018,6(1),352-356. <https://doi.org/10.5958/2347-7202.2018.00006.3>

- [16]. Chu, H.C., Yan,C.Y., Luo, Z.J., & Huang, X.C. The Improvement of Web Page Ranking on SERPs [paper presentation]. IEEE International Conference on Consumer Electronics Taichung, Taiwan. 2018. <https://ieeexplore.ieee.org/abstract/document/8448460>
- [17]. Husain, T., Sani, A., Ardhiansyah, M., & Wiliani, N. Online Shop as an interactive media information society based on search engine optimization (SEO). International Journal of Computer Trends and Technology (IJCTT). 2020,7(3), 53-57
- [18]. Google[Internet]. Retrieved January 20, 2021, from <https://developers.google.com/search/docs/advanced/security/https>
- [19]. Husain, T., Sani, A., Ardhiansyah, M., & Wiliani, N. Online Shop as an interactive media information society based on search engine optimization (SEO). International Journal of Computer Trends and Technology (IJCTT). 2020,7(3), 53-57
- [20]. Google. Browser compatibility [Internet]. Retrieved June 17, 2020, from [https://developers.google.com/search/docs/advanced/guidelines/browser-compatibility?hl=en&ref\\_topic=9460495&visit\\_id=637472627232142486-395217651&rd=1](https://developers.google.com/search/docs/advanced/guidelines/browser-compatibility?hl=en&ref_topic=9460495&visit_id=637472627232142486-395217651&rd=1)

**Cite this article as :**

Raka Yusuf, Rusti Mentari, Harni Kusniyati, " Simplified Web-based SEO Score Counter for Increasing Web Page Rank, International Journal of Scientific Research in Science, Engineering and Technology(IJSRSET), Print ISSN : 2395-1990, Online ISSN : 2394-4099, Volume 11, Issue 1, pp.141-150, January-February-2024. Available at doi : <https://doi.org/10.32628/IJSRSET23103215>