
An Effective SEO Techniques and Technologies Guide-map

Konstantinos I. Roumeliotis and Nikolaos D. Tselikas*

Communication Networks and Applications Laboratory, Department of Informatics and Telecommunications, University of Peloponnese, End of Karaiskaki Street, 22 100, Tripolis, Greece

E-mail: ntsel@uop.gr

**Corresponding Author*

Received 13 September 2021; Accepted 14 June 2022;
Publication 08 August 2022

Abstract

The paper analyzes from a technical point of view the search engine optimization (SEO) techniques and technologies, which lead to effective results up to date. More specifically, it examines the potential SEO alternatives, with ultimate target to increase the organic visitors of a website and climb it up on top of the ranking in respective search engine's queries. The main problem every website's owner has to solve is which of these SEO techniques have to be implemented in order to increase the organic traffic with the minimum budget. This paper aims to present the dominant SEO techniques and evaluate their effectiveness to organic traffic. This paper aims to present the way well-known websites apply SEO nowadays. Taking advantage of this information, Webmasters will know in advance which exactly SEO techniques should apply to their websites and in which specific order to gain higher search results, resulting to higher traffic.

Keywords: Search engine optimization, search engine optimization techniques, organic traffic, seo prototype tool.

Journal of Web Engineering, Vol. 21_5, 1603–1650.

doi: 10.13052/jwe1540-9589.21510

© 2022 River Publishers

1 Introduction

From the very first “document-based” Web 1.0 founded in 1990, proceeded to the “social and mobile” Web 2.0 founded in 1999 and shifted to the “semantic” Web 3.0, the WEB has rapidly grown in only three decades [1]. The same rapid growth occurred in web technologies [2].

In the past, the number of websites was restricted, so the web users knew in advanced the URL of the website they wanted to visit. In current Web, with almost up to 2 billion of live websites [3], search engines used as a way for users to find the information they are looking for. Search engines act as an intermediate between websites and web users, and could be considered as a bookmark that points to specific pages depending on the query.

Search engines are constantly trying to improve their algorithms in order to provide the best possible results to user searches [4]. However, they do not publicly reveal the criteria by which they categorize or classify the websites. Thus, SEO experts are trying to understand how search engine algorithms work through implementation and observation techniques. The Internet era is characterized by the abundance of information, so search engines have become an integral tool for Internet users [5]. According to several surveys, 61% of global Internet users are looking for online products, while 44% of them are using search engines to search for any kind of product [6]. However, people tend to be interested only in the first pages’ results, i.e., the results that are higher on searches [7]. While there are plenty of paid search marketing (PSM) ads, residing in the highest positions in search results, SEO techniques are trying to optimize the site’s ranking position to increase (without cost) their organic traffic through search results [8]. The aim of SEO is twofold. On one hand, SEO targets to display the website correctly on search engines, while, on the other, it aims at meeting the needs of its visitors [9]. More specifically, the term SEO refers to all these techniques used to optimize the website appearance, the webpage code and the webpage content as well [5]. By applying SEO techniques to a website, all usability, user experience (UX) and content quality are increased resulting to higher scores in the search engine rankings, respectively [10].

When implementing a SEO technique, individual changes and/or modifications are made at least in some parts of the website. These kind of modifications do not often lead to the expected results, but, if combined with more SEO techniques, they could result in significant improvement as regards the organic traffic of the website [11]. The SEO techniques and technologies that will be presented in this paper belong to the White Hat SEO

Table 1 SEO techniques and technologies summary table

SEO Techniques	SEO Technologies
2.1. Copywriting and Keyword Optimization	3.1. Serving Data over HTTPS
2.2. Crawlable URLs – SEO Friendly URLs	3.2. Schema.org and Structured Data
2.3. <meta> description – <meta> keywords	3.3. AMP – Accelerated Mobile Pages Project
2.4. Optimizing images for search engines	3.4. Apache Deflate and Gzip
2.5. Tags on Images and URLs	3.5. Page Caching
2.6. Rating and Review – Structured Data	3.6. Minify Js, Css
2.7. Sitemap, RSS, Robots.txt	3.7. Opengraph Protocol
2.8. Responsive Design	3.8. Google My Business
2.9. Breadcrumbs	3.9. Content Delivery Network (CDN)
2.10. Backlinks (Off-Page Optimization)	Links

techniques category, i.e., techniques that promote quality content and do not try to trap web crawlers, as Black Hat SEO techniques do [6, 12]. Black Hat SEO techniques usually attempt to increase the ranking of malicious pages in search engine results for popular search terms [13]. Just for the reader who wants to find extra information about Black Hat SEO techniques, Content Automation [14], Cloaking [15] and Guest Posting Networks [16], are listed among the most well-known of them, but they are out of the scope of this paper, since this kind of techniques are trying to violate the terms of use of search engines.

This research comes to fill the gap between theoretical SEO and technical – practical SEO. The paper starts with a detailed review of all SEO techniques from the existing literature, extending the study to SEO technologies that can help a website from a technical point of view to achieve a better presentation to both search engines and end-users who browse it.

Table 1 contains all the SEO techniques and technologies were reviewed in this paper.

The paper concludes with the invention of a SEO tool that, having the appropriate dataset, returns valuable outcomes to webmasters regarding the SEO strategy that should be implemented on their websites to gain higher results both in rankings and web traffic.

The remainder of the paper is organized as follows: Section 2 analyzes as a review on-page SEO techniques that optimize, in terms of presentation, a website both to its visitors as well as to search engines. These techniques and technologies give particular emphasis on the structure and the layout of the website, as well as to its content. Off-page SEO techniques are presented

in Section 2, too. Section 3 introduces new technologies that promote SEO, focusing on the special speed and security, as well as on the implementation of structured data in the website. Section 4 presents the most important of the presented SEO techniques and technologies used for the implementation of a web-crawler-like prototype SEO tool, that searches and checks for SEO techniques existence on the source code of live websites. A representative sample of the most well-known e-commerce and websites were collected and evaluated with this SEO tool. Section 4, three research questions are presented. Which specific SEO techniques have already been applied by the most well-known websites, which SEO techniques are the most important ones, and which SEO technique has to be implemented as a first priority? Having the results from the own-developed tool, a figure is created containing all the SEO techniques used by the most well-known websites and the adoption rate is presented for each SEO technique that has been applied. A SEO technique with a higher adoption ratio is used by more top-listed websites, which makes it more important according to the research hypotheses of Section 4. Finally, the paper concludes in Section 5 summarizing the results, and presenting readers with the importance of this research.

2 SEO Techniques & On-Page Optimization

On-page SEO is a set of techniques applied within the website's code [10]. On-page techniques are designed to make it easier for the visitor to browse the website, while offering useful content and of high quality. On the other hand, off-page SEO is a set of techniques applied outside the context of a website [17]. Off-page techniques are designed to influence rankings through search engine results pages (SERPs). Off-page SEO focuses on increasing authority on your domain through the act of getting links from other websites [17]. This section is primarily dedicated to the analysis of on-page SEO techniques, but at the end of the section the main off-page SEO techniques will be presented, too.

2.1 Content

Content is the most important ranking factor on a website [11]. The sustainability, the traffic and the rise of a website, depends mainly on its content [18]. The reason a searcher visits a website is in most cases to find an answer to one of his/her problems, needs or questions, for example, to find a book, to buy a pair of shoes, to find the address of a local business, to read the news, etc. All the potential solutions to his/her problem, need or question comprise

Table 2 Web manifest properties

Level	Keyword	Avg. Monthly Searches
Level 1	Apple	3.350.000
Level 2	Apple macbook (Medium-Tail Keyword)	49.500
Level 3	Apple macbook pro (Long-Tail Keyword)	40.000
Level 4	Apple macbook pro m1 (Long-Tail Keyword)	10.000

actually the content of websites. Without content, websites have no reason to exist, since none visitor would be able to find answers to any of his/her problems, needs or questions.

2.1.1 Keyword analysis – research

Keywords are the words imported by the user of a search engine to search for particular content on the Internet (search queries). Users type keywords based on their requirements and perceptions [19]. Keywords, through searches, are the most important factor that brings visitors to the website. SEO experts know in advance the keywords that should be used to rank their website high. However, a keyword analysis need to be done to find a set of keywords that will be then used to optimize the website [5]. Keyword analysis requires to determine which keywords are relevant to the website and at the same time are mostly searched in search engines. Google Adwords is an efficient tool to find out the most used keywords [20]. By typing a keyword, Google Adwords suggests similar keywords and classifies them according to their corresponding frequency appearances in searches.

Keywords could be organized in levels. Starting from Level 1 for more general keywords, up to Level 4 for very specific ones. As the levels increase from Level 1 to Level 4, the keyword-specificity also increases and presents fewer results in a search, as depicted in Table 2. On the contrary, as the level decreases from Level 4 to Level 1, the keyword becomes more general and presents more results in a search [9]. The above analysis concerns the query of a user in search engines, while, from the website's perspective, the more specific a keyword is, the easier to rank the website higher in searches is, as well. Typically, the so-called long-tail or medium-tail keyword makes it more likely to find the site higher in search results [11].

2.1.2 Articles & Content

The websites consist of text, images and multimedia (audio/video). Text can be articles, business information, terms of use, etc. Every text within

a website should be optimally structured, but not overoptimized, to attract not only visitors, but also search engines [21]. The Google search engine in Search Console Help has announced that the site should be primarily human first [22]. For this reason, the text of a website should have as its primary purpose to be useful to its visitors [23]. Once a text that meets the needs of the visitor is created, it can then be optimized to make it look better for search engines, too. When search engines parse a text or article they are not really aware of the actual content and what exactly it is dealing with. To fill this gap, a search engine notices the title of the article, meta tags and counts the number of keywords within a text and then understands what the article refers to. A digital marketing technique that deals with the process of creating a friendly text to both visitors and search engines is SEO copywriting, based on a set of rules that content's structure should follow, as shown in Table 3 [24].

Table 3 shows the essential guidelines under SEO Copywriting and describes how to display the keywords on a single page of the website. As far

Table 3 SEO copywriting structure

Rule Title	Copywriting Rules
Page Title	<ul style="list-style-type: none"> • The recommended Title Length is less than 78 characters and minimum 6 Characters [25]. • The target keyword should be included in Page Title and would be beneficial if used at the beginning of this [5],[26].
Page Description and Meta Description	<ul style="list-style-type: none"> • Meta Description must be at least 150 characters long and should not be longer than 312 characters [27]. • The target keyword should be included in Page Description and Meta Description [5, 26].
Main Content	<ul style="list-style-type: none"> • Main Content should consist of at least 500 words [28]. • The target keyword should be within the first 50 words of Content [28]. • The recommended Keyword Density is between 2% and 8% [25, 26]. • Content should be hierarchically structured correctly using h1, h2, and h3 headings [26]. • In the headings, it would be useful to add the target keyword – Keyword Distribution [5, 26]. • It would be useful for the target keyword to appear in bold and italic format where it is used in the text [26, 28]. • Internal links should be added to refer to the website itself [26]. • External links should be added which might explain the meaning of a word [26]. • Content must contain at least one image with SEO friendly URL and SEO friendly alt tags [28].

as the content is concerned, another concept that must be taken into account is keyword density, i.e., the frequency that a keyword displays between the text of a page. According to [25] and [5], the percentage of occurrence of the keyword within a page should be between 3% and 8% of the total number of words. Also, keyword distribution is the distribution of the keyword in the structure of an HTML document to deliver the best possible search results. A keyword can be placed in the page title, the meta keywords section, the meta description section, the page header tags (H1, . . . , H6), the internal links, the external links, the alt attribute of the images, and the URL [5]. It should be noted that the over-optimization of the content of the page should be avoided, because it may produce negative results, since the most important thing is to create unique content/articles that serve the researcher's goal and solve the problem that the searcher had when accessing the website [5].

2.2 Crawlable URLs – SEO Friendly URLs

Web crawlers work persistently in the back-end by following links to collect updated data from the World Wide Web and store them in large databases, which are used by search engines, like Google, to fulfil the search queries of searchers [26]. Indexing by a search engine is the process of creating an indexed list, consisting of websites [26]. Once a search request occurs, the search engine processes it and compares it to the data stored in its database. Finally, search engine algorithms undertake to display the most relevant results and present them to the user [12].

URL (Representational state transfer (RESTful) URLs or search-friendly URLs or user-friendly URLs) is a human-readable text that defines the structure of files within a web server. Each URL has three distinct parts and consists of the access protocol, the domain name, and a path [29]. The URLs of a website are an important ranking factor of the website and one of the most basic elements that search engines take into account so as to understand website's content and thus to relate it, or not, to a search query. A well-structured URL provides both visitors and search engines an easy way to understand the content of the page they will visit even before visit it. Search engines also assume that the target keyword is located in the URL keywords [30].

Furthermore, several times, users copy a URL to share it via email or through social media. If the URL is well written, legible and understandable, the more users will click on it [30].



Figure 1 Differences between SEO friendly and Not SEO friendly URL.

2.2.1 Differences between SEO URL & non SEO URL

Years earlier, most of the URLs were illegible, making it difficult for search engines to understand the content of the page. On the other side, SEO Friendly URLs contain the page path using words that describe the content separated by hyphens (–), as depicted in Figure 1 [31]. A common practice is to use the title tag of a page as a URL, since it usually contains the related to the search query keyword [11].

2.2.2 How SEO friendly URLs are created

In the not SEO friendly URL case, the corresponding URL syntax depends on the server-side scripting language. For example, by typing the above URL [Figure 1], the $id = 5$ clause implies an article having the value five as the unique primary key in a database's table, but there is nothing clear about its content. On the other hand, to create SEO-friendly URLs the developer has to put an effort using a scripting language in a combination with a web server software. There are multiple ways to achieve SEO-friendly URLs. Two of them are:

1. Many opensource CMS platforms, like Wordpress, use the front controller design pattern. The Front Controller pattern consists of a central controller that serves as a single point of entry into the application. Either using Apache, NGNIX or ISS web server software, the developer has to redirect all the traffic to a single point of entry using the appropriate URL Rewrite Module [32]. In this single point of entry, regardless of the programming language used, the central controller undertakes to break the given URL into sections that can help locating the content in the database.
2. An alternative way to achieve SEO-friendly URLs is to use the prior URL Rewrite Module to rewrite every single type of URL pointing to the appropriate scripting file. A rewrite rule using the .htaccess on Apache and PHP programming language is presented in Figure 2. In this example the rewrite rule matches the URL given by the browser,

```
Options +FollowSymLinks -MultiViews
RewriteEngine on
RewriteBase /
RewriteCond %{REQUEST_URI} !\.(?:css|js|jpeg|gif|png|ico|eot|ttf|otf|woff|html|php)$ [NC]
RewriteRule ^blog/([A-Za-z0-9-]+)/?$ blog.php?article=$1 [NC,L]
```

Figure 2 Source code from an .htaccess file.

```
<head>
<meta charset="UTF-8">
<title>Macbook Pro MJLQ2</title>
<meta name="description" content="Macbook Pro, 2.2GHz Intel Core i7 processor with
Turbo Boost up to 3.4GHz, the MacBook Pro delivers speedy performance.">
<meta name="keywords" content="macbook,apple macbook">
</head>
```

Figure 3 Appropriate <meta> description and <meta> keywords.

identifies the appropriate rule and redirects the user to the blog.php file with a GET value.

At this point we have to mention that URLs should not contain more than 2083 characters, in order to be visible by all browsers even in mobile searches [11, 33, 34].

2.3 <meta> Description – <meta> Keywords

By doing a search and observing one of the results, we initially see the title of the page, then the URL and finally the description. However, many times, although the page contains a meta description, the description of a result is not the same as the meta description. Search engines, and mostly Google, have started not to consider the <meta> description and <meta> keywords representative because they are often misleading on purpose [35]. The reason is that many times <meta> description is not created correctly to describe the exact content of the page. Advanced search engine algorithms understand whether <meta> description or <meta> keywords are relevant to the content and replace it with a text within the page they think is most relevant. It would be prudent to create a <meta> description for each page that describes the page content as precise as possible by using 50 to 160 characters [36]. For optimum viewing in searchers' queries the <meta> description could contain the target keywords, as depicted in Figure 3.

2.4 Optimizing Images for Search Engines

Each page, whether it contains an article or a text describing the activity of a business, includes many times images as well. Images contribute to the better ranking in search engines, because they can improve the user experience,

```
<!-- Not SEO Friendly Image Tag -->  
  
  
<!-- SEO Friendly Image Tag -->  

```

Figure 4 Image SEO filename & alt tag.

provided that they are in the proper form, i.e., in the appropriate file size first [37]. Through independent studies, an image that exceeds 100 kb file size is often difficult to view by users, although Internet speeds have improved in recent years [37, 38]. Many pages even today are slow to load, which prevents users from visiting them. Thus, an essential factor for the load time of a website is undoubtedly the file size of the images displayed on it [6]. Using Next-Gen formats such as JPEG 2000, JPEG XR, and WebP along with an Image CDN can increase a website's speed [37, 39].

2.5 Tags on Images and URLs

In addition to the file size of an image, a key role in SEO has the file name and the alt tags. Although algorithms have evolved by applying machine learning techniques to identify the content of images, alt tags and filename are considered more representative and indicative of the description of images on a website [40]. Apart from searching for websites, many search engine users choose to search for images in order to decide which website fits their needs. For both reasons, the file name should be relative to what the image is displaying, and the alt tag should describe what the image represents, too. Because the images are part of a page and are related to the content of the page, it would also be beneficial to contain the page title [Figure 4]. Alternative tags also contribute to better accessibility. Starting from the WCAG 2.0 the World Wide Web Consortium (W3C) has included in the accessibility guidelines the alternative tags. According to the success criterion 1.1.1 (Non-text Content), when using the `img` element, it is mandatory to include a short text alternative with the `alt` attribute [41]. Screen readers report alternative text in place of images, helping users with visual or certain cognitive disabilities perceive the content and function of the images [42].

Each page contains internal and external linking. A link could refer, for example, to an article or a product. When a user sees a link on a page, the user usually places the mouse over this link before it is clicked. Before clicking this link, the user would probably want to know where that link goes.

```
<!-- Not SEO Friendly URL Tag -->  
<a href="/product.php?id=2">  
  
<!-- SEO Friendly URL Tag -->  
<a href="/product/macbook-pro-mjlq2/" title="Macbook Pro MJLQ2">
```

Figure 5 Link SEO & title tag.

Similarly, search engines want also to know about the content of the landing page [33]. For all the above reasons, the title tag is necessary for both the description and the understanding of an internal or external link [Figure 5].

2.6 Rating and Review

Many users are looking for products and services on the Internet. Many of them finally buy the product or service based on a corresponding rating and/or review about it. User-generated content (UGC) is an effective tool for selling online products and services, because consumers would usually purchase an item after they read through all the personal information generated by other users on the platform [43]. The psychological reaction that makes users believe the reviews of other users of a website are correct is called social proof [44]. From the side of the website or e-commerce, social proof can significantly increase conversion rates by building a feeling of customer trust [45].

As a result of the above, building a rating and review system, each website can benefit from ratings and user reviews. Through the rating and review system users create valuable content for the website while encouraging other users to trust their choice. Search engines, in turn, reward the website which promotes user ratings and reviews., e.g., the Google search engine often shows user site reviews frequently in search results (Rich Results) [Figure 7].

In order for search engines to be in position to locate the rating system, the website should contain structured data describing the reviews [Figure 6].

Structured data can be written in three ways, i.e., with Microdata that are the widely used structured data, with Resource Description Framework in attributes (RDFa) or with JavaScript Object Notation for Linked Data (JSON-LD) [46]. Microdata is a set of Web Hypertext Application Technology Working Group (WHATWG) HyperText Markup Language (HTML) specification properties that are used to display metadata on existing page content. Search engines, web crawlers and browsers extract and process microdata from a website and use them to provide a richer browsing experience for

A

```

<!-- MicroData Rating and Review -->
<div itemscope itemtype="http://schema.org/Review">
  <div itemprop="itemReviewed" itemscope itemtype="http://schema.org/Product">
    
    <span itemprop="name">Macbook Pro MJLQ2</span>
  </div>
  <span itemprop="reviewRating" itemscope itemtype="http://schema.org/Rating">
    <span itemprop="ratingValue">4</span>
  </span> stars -
  <b><span itemprop="name">A good product.</span></b>
  <span itemprop="author" itemscope itemtype="http://schema.org/Person">
    <span itemprop="name">Bob Smith</span>
  </span>
  <span itemprop="reviewBody">The product is great.</span>
  <div itemprop="publisher" itemscope itemtype="http://schema.org/Organization">
    <meta itemprop="name" content="Washington Times">
  </div>
</div>
<!-- JSON-LD Rating and Review -->
<script type="application/ld+json">
{
  "@context": "http://schema.org/",
  "@type": "Review",
  "itemReviewed": {
    "@type": "Product",
    "image": "https://www.website.com/apple-macbook-mjlq2.jpg",
    "name": "Macbook Pro MJLQ2"
  },
  "reviewRating": {
    "@type": "Rating",
    "ratingValue": "4"
  },
  "name": "A good product.",
  "author": {
    "@type": "Person",
    "name": "Bob Smith"
  },
  "reviewBody": "The product is great.",
  "publisher": {
    "@type": "Organization",
    "name": "Washington Times"
  }
}
</script>

```

B

Figure 6 Structured data ratings and reviews – (A) microdata and (B) JSON-LD.

users [47]. RDFa is a W3C recommendation that adds a set of attribute-level extensions to existing HTML, XHTML to import metadata into web documents [48]. JSON-LD is a metadata encoding method using JSON [49]. Microdata and RDFa are applied directly to HTML by adding HTML properties. According to W3C use cases, JSON-LD have many advantages over RDFa since JSON-LD templates can be easily inserted as a single block of

Amazon.com: Apple 15 Inch MacBook Pro Laptop (Retina Display, 2.2 ...
<https://www.amazon.com/Apple-MacBook-Display...MJLQ2LL/dp/B00XZGMBVC> ▾
★★★★★ Rating: 4.3 - 296 reviews
PERFORMANCE: 2.2GHz Intel Core i7 processor with Turbo Boost up to 3.4GHz, the MacBook Pro delivers speedy performance. Includes 16 GB of memory so ...
[Apple Macbook Pro MJLT2LL](#) · [Apple MacBook Pro MJLQ2LL](#) ...

Figure 7 Ratings, reviews and feedback on search results.

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>http://www.website.com/article-url</loc>
    <lastmod>2018-12-01</lastmod>
    <changefreq>monthly</changefreq>
    <priority>0.8</priority>
  </url>
</urlset>
```

Figure 8 Sitemap example.

structured data within the <head> element of a web page [50]. More details on all of the above will be seen in Section 3.2.

2.7 Create Sitemap, Rich Site Summary or Really Simple Syndication (RSS) and Robots.txt

The sitemap is an XML file that provides information about the pages of a website [51]. By reading this file, search engines can better map the layout of a website. Each sitemap line describes to search engines, apart from the structure of the website and which pages are more important than others, how often other information is updated, too [Figure 8]. Without the sitemap, search engines should check the internal linking of the website to find out the layout, with the risk of creating a wrong understanding and indexing.

The RSS feed is also a decisive factor in improving the visibility of a website by search engines. The RSS format was originally developed by Netscape in the late 1990s for use on My Netscape Network, a customizable start page for Netcenter [52]. My Netscape Network is provided a simple RSS framework for websites to create channels that can then be added to the customizable start page [53]. Currently, the primary use of the RSS feed to websites is to inform visitors about the new data added to a website [54]. Typically blog sites use RSS to offer their users a more immediate update. To improve users' experience, RSS feed readers plugins have been created to help users get informed about developments in things that interest them.

```

<rss version="2.0">
<channel>
  <title>Website RSS feed</title>
  <link>https://www.websites.com</link>
  <description>The latest content of Website</description>
  <language>en-us</language>
  <copyright>Copyright (C) 2018 website.gr</copyright>
  <item>
    <title>Article Name</title>
    <description>Article Description</description>
    <link>
      https://www.website.com/article-url/
    </link>
    <pubDate>Sat, 01 Dec 2018 19:23:31 +0000</pubDate>
  </item>
</channel>
</rss>

```

Figure 9 RSS article example.

```

#Robots.txt on public_html                                #Robots Meta
User-agent: *                                             <html>
Allow: /                                                  <head>
Disallow: /administrator/                               <title>Administrator</title>
Sitemap: https://www.websites.com/sitemap.xml           <meta name="robots" content="noindex, nofollow">
                                                         </head>

```

Figure 10 Robots.txt and robot meta.

From the search engines' perspective, they are more likely to locate the most recent content on a website by following the RSS feed of the website, since it contains the information properly structured to be read programmatically by a search engine [6], [Figure 9].

The Robots.txt file is a text file created by webmasters to guide web robots on how to crawl the pages of a website [6]. Web robots, follow links, and visit millions of websites daily – spidering. When a robot arrives at a website its first action is to look for the robots.txt file. In case such a file is not available, the robot investigates the content of the site by crawling into it [55]. But, if the robots.txt file exists, it first read it and then enter as many pages it has access to. In Joomla CMS, by using the robots.txt file web administrators can allow access to web content and at the same time prohibit access to pages that do not want to be indexed, such as the administrator panel [56] [Figure 10].

2.8 Responsive Design

Currently, about half of Internet users' searches come from mobile smartphones and tablets [57]. When a user visits a website and finds it difficult to read the information it provides, he will give up soon. Using responsive design technologies, a website manages to facilitate access, reading, and, more generally, improve the user experience (UX) of its visitors [6, 58].



Figure 11 Google analytics – bounce rate – session duration.

By increasing all of the above, the visitor spends more time on the website, with the bounce rate decreasing and the average time spent on the website (Session Duration) rising [Figure 11].

Session duration and bounce rate are two of the most vital criteria the Google RankBrain algorithm uses to rank the websites in search engine results [18]. If the website is not mobile-friendly, it may make it difficult for visitors to read it and, thus leave it in a short while, reducing session duration and increasing the bounce rate.

In the early stages of responsive design, mobile pages were created as a separate page of the main website. The website using the user's HTTP headers identify the device the user uses to access the website. If the user accesses the website from a mobile device the mobile version of the website is delivered from the webserver. Instead, when the user entered the page from a computer, the website sent the main version of the website, using the dynamic serving technique [59]. The responsive web design [60] was developed, to solve the problem with the double websites. One of the leading responsive web design companies that provide ready-made templates for websites seen from each device is Twitter's Bootstrap [61] [Figure 12].

Google has created a tool called Mobile-Friendly Test. This tool checks whether the website looks appropriately on mobile devices and, if not, it informs the developer about the problems it encounters [62].

2.9 Breadcrumbs

Breadcrumbs are text that consist of links (usually internal) and are usually located at the top of the website as shown in Figure 13 [63]. It shows the visitor exactly where he/she is within a website. These links describe the path that a visitor should follow to reach for example a product. By clicking on one of the links, the visitor returns to the appropriate location.

As regards the visual design, an example of how breadcrumbs are created is shown in Figure 14.

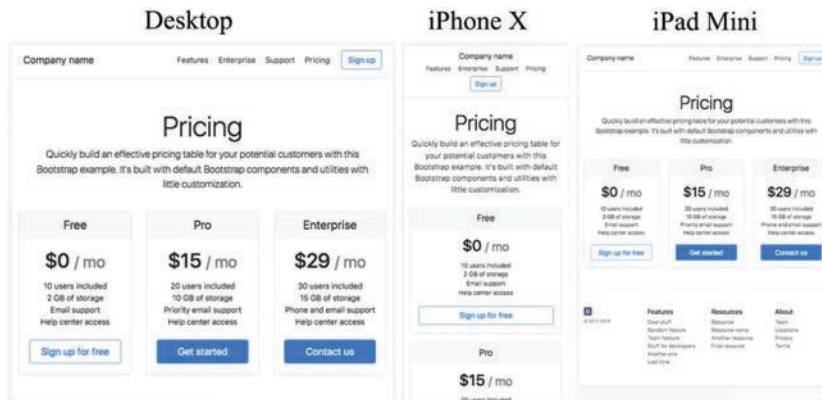


Figure 12 Responsive design for multiple devices – bootstrap.

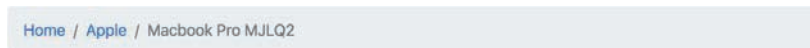


Figure 13 Breadcrumbs.

```
<!-- Breadcrumbs Visualization Bootstrap -->
<nav aria-label="breadcrumb">
  <ol class="breadcrumb">
    <li class="breadcrumb-item"><a href="/">Home</a></li>
    <li class="breadcrumb-item"><a href="/Apple/">Apple</a></li>
    <li class="breadcrumb-item active" aria-current="page">Macbook Pro MJLQ2</li>
  </ol>
</nav>
```

Figure 14 Using Bootstrap to Visual Design Breadcrumbs.

Structured data (Microdata, RDFa or JSON-LD) can be used to make optimal use of Breadcrumbs to help users and search engines understand and explore a site effectively [64], as illustrated in Listing 1 (for Microdata), Listing 2 (for RDFa) and Listing 3 (for JSON-LD), respectively.

```
<ol itemscope itemtype="https://schema.org/BreadcrumbList">

  <li itemprop="itemListElement" itemscope

    itemtype="https://schema.org/ListItem">

    <a itemprop="item" itemtype="https://schema.org/WebPage

      "

      href="https://website.com/">
```

```
        <span itemprop="name">Home</span></a>
        <meta itemprop="position" content="1" />
    </li>
    >
    <li itemprop="itemListElement" itemscope
        itemType="https://schema.org/ListItem">
        <a itemscope itemType="https://schema.org/WebPage"
            itemprop="item" href="https://website.com/Apple">
            <span itemprop="name">Apple</span></a>
        <meta itemprop="position" content="2" />
    </li>
    >
    <li itemprop="itemListElement" itemscope
        itemType="https://schema.org/ListItem">
        <span itemprop="name">Macbook Pro</span>
        <meta itemprop="position" content="3" />
    </li>
</ol>
```

Listing 1 Structured data – microdata to improve breadcrumbs

```
<ol vocab="https://schema.org/" typeof="BreadcrumbList">
    <li property="itemListElement" typeof="ListItem">
```

```

    <a property="item" typeof="WebPage"
      href="https://website.com/">
      <span property="name">Home</span></a>
    <meta property="position" content="1">
  </li>
  >
  <li property="itemListElement" typeof="ListItem">
    <a property="item" typeof="WebPage"
      href="https://website.com/Apple">
      <span property="name">Apple</span></a>
    <meta property="position" content="2">
  </li>
  >
  <li property="itemListElement" typeof="ListItem">
    <span property="name">Macbook Pro</span>
    <meta property="position" content="3">
  </li>
</ol>

```

Listing 2 Structured data – RDFa to improve breadcrumbs

```

<script type="application/ld+json">
  {
    "@context": "https://schema.org",

```

```
"@type": "BreadcrumbList",
"itemListElement": [{
  "@type": "ListItem",
  "position": 1,
  "name": "Home",
  "item": "https://website.com/"
},{
  "@type": "ListItem",
  "position": 2,
  "name": "Apple",
  "item": "https://website.com/Apple"
},{
  "@type": "ListItem",
  "position": 3,
  "name": "Macbook Pro"
}]
}
</script>
```

Listing 3 Structured data – JSON-LD to improve breadcrumbs

Visitors often get lost on a website. Breadcrumbs use links for the primary paths of the page by helping visitors find the right path for what they are looking for, improving by extension the overall user experience [63]. For instance, there are users who land on a product page through search engine

Apple MacBook Pro A1398 15.4" Laptop - MJLQ2LL/A | eBay
<https://www.ebay.com> › ... › Laptops & Netbooks › Apple Laptops ▼

Figure 15 Breadcrumb shows website hierarchy and categorization.

results. If the visitor wants to visit the category to which the product belongs, he should look in the menu among a vast array of categories to find the one that the product belongs to. Breadcrumbs are the solution to this problem, presenting the visitor the exact location of the product on a website, or even the category or the subcategory it belongs to.

Finally, Breadcrumbs, in addition to visitors, also give search engines a way to see how a website is structured [5]. The Google search engine rewards the use of Breadcrumb by displaying them in rich search results, as shown in Figure 15.

2.10 Backlinks (Off-Page Optimization)

A backlink is created when an external link of a referrer website points to a page of a receiver website. We assume that the receiver website contains a critical article while other websites prompt their visitors through a link to read the article. Then, all these links are backlinks. For search engines, when a website has many backlinks, it means it is reliable because many other websites trust it by suggesting their visitors follow the link on it. Nevertheless, not all backlinks have the same weight. A backlink from the New York Times certainly does not have the same influence as a backlink from a local business. SEO Experts rank websites according to their importance by using domain authority (DA) classification fluctuating between DA1 and DA100 (DA100 characterizes websites that are very important) [65].

Therefore, if a website contains many backlinks with a high domain authority (DA), search engines most likely will reward it by bringing it up in their searches. It is imperative to clarify that search engines don't use the DA metric to rank web pages to the search results. DA is a simulation metric made by SEO experts to model the search engines' algorithms. Another key ranking factor in addition to the domain authority (DA) of the referrer website is the relevance of the content between the referrer and the receiver website [11]. If the content of these two websites is relevant, the search engines evaluate it positively, something that does not happen in the case that the referrer and the receiver websites consist of irrelevant content, as shown in the example of Figure 16.

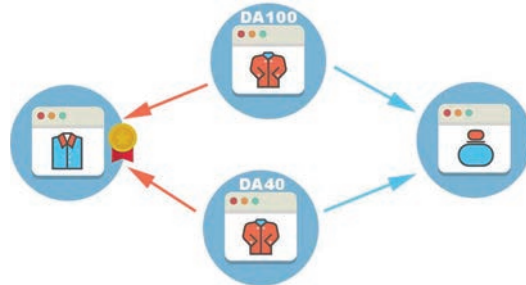


Figure 16 DA100 and DA40 clothing websites point backlink to related clothing website and not related perfume website.

There are different approaches to earn backlinks.

- Guest Posting or Guest Blogging backlinks are implemented by writing and submitting articles to relevant niche blog websites [17]. Those blogs accept valuable articles, which in turn allow the author to place a backlink pointing to his website.
- Article submission backlinks are created in the same way as Guest Posting backlinks. The main difference between them is that the article submitted in article directories instead of blog websites [66].
- Directory Submission or Link directory backlinks are created by submitting a website's name, description, and URL to popular link directory submission sites [17].
- Profile backlinks are implemented by creating profiles to social media platforms. The majority of the social media platforms accept user's URL while registering his/her profile [66].
- Social media sharing backlinks are created by social media users while posting a URL of their own website to their social media profiles [17, 66].
- Forum Posting backlinks are created while producing and commenting discussion topics on popular forums [66]. The majority of the forums allow users to redirect forums' traffic using URL to external resources as long as the source provides essential information solving a problem [66].
- Comment backlinks are created by commenting on an article or product page placing a backlink within the comment. Comments on a website are equally beneficial to search engines and visitors as mentioned in the Rating & Review (Section 2.6). On the other hand, earlier, SEO experts were trying to implement link building by making spam comments with

links in blog pages, which were known as blog comments. However, advanced search engine algorithms detect spam comments and punish not only the spammer website but also the spammed website [17].

- Question-and-answer (Q & A) comment backlinks. On many Q & A websites, like StackOverflow and Quora, there are usually three kind of visitors: (a) users who ask for help on a specific subject, (b) users who answer questions in the latter, and (c) visitors who find a solution to a problem by searching on search engines and looking at an already answered question. Backlinks from such websites are beneficial. However, to create a backlink to them, the webmaster will first have to answer some of the user questions. By answering a question, the webmaster has the option to add a link pointing to his website that proposes a solution to this specific problem. The link added is a backlink, which will not only be positively evaluated by the search engines but will also bring traffic to his website.

2.11 Paid Traffic (Off-Page Optimization)

Many websites cannot be waiting for a long time to bring organic visitors to their content. Organic visitors certainly cannot be reached in one day. It takes time, design, high-quality content, and SEO to attract organic visitors. The most common solution to this problem is to purchase paid traffic. Paid traffic practically drives visitors to the website for a fee. Many websites that have massive traffic, such as social networking pages, can send thousands of visitors to a website in just a few hours, depending on the money the website will pay for the paid traffic [18].

Search engines also offer paid traffic. The webmaster of a website can choose how many users want to visit his website by using a keyword and paying per click (PPC) advertising. The search engine, knowing that the website has paid to appear first in the searches for a particular keyword, places it on top of the search results. Following this process, the website increases its traffic through paid ads. However, to maximize results, targeted campaigns must be created consisting of several high-volume keywords.

Although paid traffic is not an SEO methodology, it may be necessary for websites that do not have the time-luxury to invest in SEO techniques. The combination of paid traffic and SEO techniques will certainly bring even better results.

2.12 Black Hat SEO Techniques

White hat SEO is a term used to refer to SEO strategies and techniques that operate within the rules and expectations of search engines and search users [67].

The effect of white hat SEO requires a lot of effort from website maintainers and takes effect slowly [68]. In order to achieve higher rankings in SERPs in less time, black hat SEO techniques have been created [13]. Black hat SEO techniques break the rules, promoting spam websites by trapping users and search engines [14]. Although this article is dedicated to reveal all the White Hat SEO techniques as a review, the most common black hat SEO techniques are also shortly presented below.

- Content Automation – Content Spinning. Websites have to create valuable content in order to survive. This content is evaluated by the search engines, and the web page is indexed to the SERPs to specific keyword queries. Content Automation is a procedure where multiple articles combined into one while an automation tool trying to spin the position of words to achieve a unique outcome [14].
- Cloaking is a black hat SEO technique used by the Webmasters to trap visitors and search engines by presenting different copy of content to users than the search engines [15]. In that way, they manage to achieve positions in the searches on specific high traffic keywords, while irrelevant content is presented to their users [15].
- Keyword Stuffing is also a common black hat SEO technique in which keywords are placed multiple times within the web page in an attempt to mislead search engines about the content [69].
- Private Blog Networks (PBN) is a group of high domain authority websites used exclusively for link building. Each PBN website links, by using backlinks, to the site they want to boost in the search results [70].

3 Technologies Promote SEO

3.1 Serving Data over HTTPS

Every day millions of users use the web to carry out important tasks, e.g., banking transactions, capital investment in the stock market, etc. Every user would definitely like the interaction with a website to be safe, just as his or her personal data.

HTTPS encodes and encrypts the connection between the user's browser and the web server that a website hosted, helping to ensure that data exchanged between them is not interrupted or intercepted. Any website that would like to encrypt traffic needs an SSL Certificate.

SSL Certificate can be installed on any web server to secure communication between visitors and the web server.

In 2014 Google search engine on its search central blog announced that security is a top priority for Google [71]. As a search engine that redirects searchers to websites wants to be sure that websites people access from Google are secure [71]. The same document provides detailed best practices to protect any website. Meanwhile, a warning from Google takes place cited that secure, encrypted connections is a signal for its search ranking algorithms, meaning that SSL certificates affect website's ranking on SERPs.

Google Chrome version 68 was the first to highlight websites that do not contain SSL certificates as not secure [72]. Nowadays, the majority of browser companies have developed user warnings on their user interface (UI) to protect their users from not SSL encrypted websites. The HTTPS or HTTP/2 Internet Protocol uses different technologies to make a website secure and faster by improving user experience (UX). For all the above reasons the use of an SSL certificate is considered of vital importance.

3.2 Schema.org and Structured Data

In 2011, some of the most popular search engines, like Google, Yahoo, and Bing created a collaborative community on Schema.org to promote standards for structured data on the Internet. "Schema.org" uses a vocabulary written with Microdata, RDFa, or JSON-LD to add information to web content. The content is understandable by search engines, providing them information on what the page is about. The most used item types are Organization, Person, Place, LocalBusiness, Product, and Offer. Each item type contains many properties, which in the case of an item type offer, is the name of the offer (`itemprop = name`), the image of the offer (`itemprop = image`), the color of the offer object (`itemprop = color`) as well as other properties.

Search engine bots browse every website and try to understand what each page presents. Thus, without the existence of structured data, from the phrase "Apple Macbook MJLQ2" a bot can assume that it is the title of the page, whereas from the "apple-macbook-mjlq2.jpg" an image that visualizes the content of the page, respectively. On the contrary, with the structured data, the search engine bots recognizes that the "Apple Macbook MJLQ2" is the

name of a product and “apple-macbook-mj1q2.jpg” a product photo [46]. Figure 17 demonstrates an example showing the use of the three alternatives for structured data, i.e. by using Microdata (A), RDFa (B), or JSON-LD (C), respectively.

In Figure 19, we also see that by applying structured data with Microdata or RDFa, each line of the page should be improved and the property – “itemprop” should be added. On the contrary, with JSON-LD, all structured data is within a script element located at the top of the page without having to correct all of the page code. To conclude, the most well-known structured data, i.e., Microdata, offer an optimal understanding of the content of the website by search engines and it is considered a valuable help for SEO [66].

3.3 AMP – Accelerated Mobile Pages Project

The AMP Project was created by Google in 2015 to improve the performance of websites on mobile devices. Essentially, each website can have a more “lightweight” version of its pages following the AMP standards. The pages that follow the AMP standards load instantly and that is the reason for selecting the lightning strike logo as the AMP logo on the AMP websites. When searching in Google from a mobile device, the most relevant AMP pages are proposed first, because they will be loaded faster [73].

Google locates and evaluates on a daily basis pages that follow AMP standards. As soon as Google evaluates a page of a website as AMP, it keeps the page code (cached copy) on its own servers so that it can appear even faster in searches. Although the AMP page will look high in searches, the visitor will see the content from the Google web server without adding its visit to the actual website traffic. Although the primary concern of SEO is to increase the traffic of a website, by using AMP technologies the added value comes to the brand of the business rather than to the organic traffic to the actual server, because more users view the content. Figure 19 shows the search results when a page has or has not been rated as AMP, respectively.

Each AMP page must have a specific format to pass the Google Bots AMP testing [73]. Initially, the page should contain Meta Data with structured data and more specifically JSON-LD. Navigation should be structured with the “amp-sidebar” component. Images should have a specific form, so does the price of the product [74].

When the AMP page is created, a canonical link and an amp link, respectively, as shown in Figure 19, should be added to the AMP header and the non-AMP page header. Canonical and amp links essentially inform

```

A
<!-- Product with MicroData -->
<div itemscope itemtype="http://schema.org/Product">
  <span itemprop="brand">Apple</span>
  <span itemprop="name">Macbook Pro MJLQ2</span>
  
  <span itemprop="description">The Apple MJLQ2 MacBook Pro with Retina
Display has a compact and suave design that accommodates immense power.
It has sleek dimensions and a narrow body that measures only 1.8cm in
height, which makes it easily portable.</span>
  <span itemprop="offers" itemscope itemtype="http://schema.org/Offer">
    <meta itemprop="priceCurrency" content="USD" />
    $<span itemprop="price">2299.99</span>
    Condition: <link itemprop="itemCondition"
href="http://schema.org/NewCondition" />New Condition
    <link itemprop="availability" href="http://schema.org/InStock" />In
stock! Order now!
  </span>
</div>

B
<!-- Product with RDFa -->
<div vocab="http://schema.org/" typeof="Product">
  <span property="brand">Apple</span>
  <span property="name">Macbook Pro MJLQ2</span>
  
  <span property="description">The Apple MJLQ2 MacBook Pro with Retina
Display has a compact and suave design that accommodates immense power.
It has sleek dimensions and a narrow body that measures only 1.8cm in
height, which makes it easily portable.</span>
  <span property="offers" typeof="Offer">
    <meta property="priceCurrency" content="USD" />
    $<span property="price">2299.99</span>
    Condition: <link property="itemCondition"
href="http://schema.org/NewCondition" />New Condition
    <link property="availability" href="http://schema.org/InStock" />In
stock! Order now!
  </span>
</div>

C
<!-- Product with JSON-LD -->
<script type="application/ld+json">
{
  "@context": "http://schema.org/",
  "@type": "Product",
  "name": "Macbook Pro MJLQ2",
  "image": ["https://www.website.com/img/macbook-pro-mjlq2.jpg"],
  "description": "The Apple MJLQ2 MacBook Pro with Retina Display has a
compact and suave design that accommodates immense power. It has sleek
dimensions and a narrow body that measures only 1.8cm in height, which
makes it easily portable.",
  "brand": {
    "@type": "Thing",
    "name": "Apple"
  },
  "offers": {
    "@type": "Offer",
    "priceCurrency": "USD",
    "price": "2299.99",
    "itemCondition": "http://schema.org/NewCondition",
    "availability": "http://schema.org/InStock"
  }
}
</script>

```

Figure 17 Structured data on product (A) Microdata (B) RDFa (C) JSON-LD.

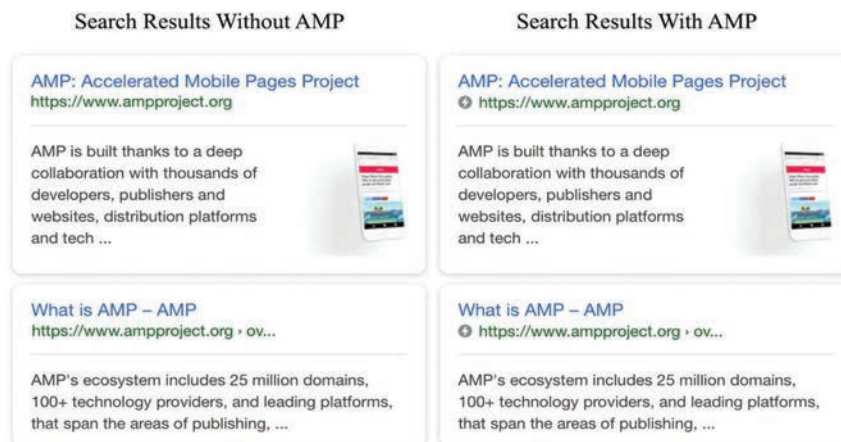


Figure 18 Search results without and with AMP.

```
//AMP Page Source
<!doctype html>
<html ⚡="" lang="en">
<head>
  <meta charset="utf-8">
  <link rel="canonical" href="https://www.website.com/product/macbook-pro-mj1q2/">
</head>

//Website Page Source
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="utf-8">
  <link rel="amphtml" href="https://www.website.com/product/macbook-pro-mj1q2/amp">
</head>
```

Figure 19 Canonical and Amp html link on header.

search engines, that besides the AMP page, there is also the non AMP page. Without the canonical and amphtml link, the AMP page can not be found by search engines [75].

To check if the generated page is an AMP Page, Google has created a tool [74]. The Washington Post writes for the AMP pages: “If our site takes a long time to load, it doesn’t matter how great our journalism is, some people will leave the page before they see what’s there”.

3.4 Apache Deflate and Gzip

There are cases of websites that are extremely slow to load, and as a result, users are discouraged and choose to leave. The loading time of a website – page load time is becoming more and more important for search engines when they visit the website to figure out if it is properly optimized [6].

The process of loading a website starts when the user using his browser requests an address – URL. The web server receives the request and sends back the requested page. If the page is too “heavy”, the web server will delay in sending it to the user’s browser, as a result, the user has to wait for a long time for the requested website to load. In this case, he/she may be waiting for the page to load, but he/she may look for another page that will give him/her the respective information more effectively and faster (Bounce Rate [25]). The average recommended Page Size is 150 kilobytes [25].

Two advanced and reliable solutions that can be applied to speed up websites are Deflate in Apache Server or Gzip on any other web server software solution [76]. Both cases compress the content in the web server and send it compressed back to the user’s browser. The browser, in its turn, undertakes the task of decompression and displays it to the user. With this technique, the user receives the page up to 70% faster. Since page load time will be shorter, users will no longer leave the website because of extremely slow loading and search engines will reward the site with a higher ranking [25, 77].

3.5 Page Caching

Client-side Page Caching also comes in its turn to increase the speed of loading a website, but instead of compressing the page on the web server and sending it to the browser, it takes the page from the web server and temporarily saves specific files in the user’s browser cache. The next time the user attempts to enter the page again, the browser will load the files that were stored the first time from the cache and requests to the web server only for the rest. By doing this, the second time the page will load more quickly.

There are also server-side caching approaches that create a cached copy of the produced HTML in a temporary place until the page is modified. The users get a cached copy of the page without waiting for the scripting language to retrieve data from the database. A great but not limited example of server-side caching is the TWIG template engine [78].

Client-side and Server-side caching approaches can work simultaneously without affecting each other. Both of them achieve a faster loading website which contributes to higher rankings [25, 79].

3.6 Minify Js, Css

In the same way, the images of a website should be optimized to load quickly, so should the other files on a website be as lightweight as possible to improve

Uncompressed jQuery 3.3.1	Compressed jQuery 3.3.1
<p style="text-align: right;"><i>Filesize: 272kb</i></p> <pre> /*! * jQuery JavaScript Library v3.3.1 * https://jquery.com/ * * Includes Sizzle.js * https://sizzlejs.com/ * * Copyright JS Foundation and other contributors * Released under the MIT license * https://jquery.org/license * * Date: 2018-01-20T17:24Z */ (function(global, factory) { "use strict"; if (typeof module === "object" && typeof module.exports === "object") { // For CommonJS and CommonJS-like environments where a proper `window` // is present, execute the factory and get jQuery. // For environments that do not have a `window` with a `document` // (such as Node.js), expose a </pre>	<p style="text-align: right;"><i>Filesize: 87kb</i></p> <pre> /*! jQuery v3.3.1 (c) JS Foundation and other contributors jquery.org/license */ function(e,t){"use strict";"object"==typeof module&&"object"==typeof module.exports? module.exports=e.document?t(e,!0):function(e) {if(!e.document)throw new Error("jQuery requires a window with a document");return t(e)}:t(e)} ("undefined"!=typeof window? window:this,function(e,t){"use strict";var n= [],r=e.document,l=Object.getPrototypeOf,o=n.slice ,a=n.concat,s=n.push,u=n.indexOf,l= { },c=l.toString,f=l.hasOwnProperty,p=f.toString,d =p.call(Object),h={},g=function e(t) {return"function"==typeof t&&"number"!=typeof t.nodeType},y=function e(t){return null!=t&&t===t.window},v= {type:!0,src:!0,noModule:!0};function m(e,t,n) {var i,o= (t=t r).createElement("script");if(o.text=e,n)fo r(i in v)n[i]&& (o[i]=n[i]);t.head.appendChild(o).parentNode.remov eChild(o);function x(e){return null==e? e+"":"object"==typeof e "function"==typeof e? l[c.call(e)] "object":typeof e}var b="3.3.1",w=function(e,t){return new w.fn.init(e,t),T/^[\s\uFEFF\xA0]+ [\s\uFEFF\xA0]+\$/g;w.fn=w.prototype= </pre>

Figure 20 Filesize uncompressed & compressed jQuery.

page load time and user experience (UX). Regarding JavaScript (.js) files and Cascading Style Sheets (.css) files, they often contain many empty lines, which increase significantly their size. Either by hand or with the use of online tools, you can minify the files, so they are “more lightweight”. Many times webmasters may have noticed “minimized” file types (e.g., jquery-3.3.1.min.js), which are minified JavaScript files. Minified files are most of the time up to 70% smaller in size than the not minified ones. An example of a file from the jQuery library before and after minification is shown in Figure 20.

3.7 Opengraph Protocol

Open Graph Protocol (OGP) helps each website become a rich object in a social graph [80, 81]. More specifically, OGP consists of a set of meta tags that allow each webmaster to give social media detailed information about the page. In turn, social media using this information can present in a better way a link on the page.

Open Graph (OG) are essentially meta tags, each one containing properties and contents, that are placed in the header of the page, preferably after its title. The most well-known properties are (og: title, og: description, og: type,

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="utf-8">
  <title>Macbook Pro MJLQ2</title>
  <!-- OG Tags -->
  <meta property="og:url" content="https://www.website.com/product/macbook-pro-mjlq2/" />
  <meta property="og:type" content="website" />
  <meta property="og:title" content="Macbook Pro MJLQ2" />
  <meta property="og:description" content="The Apple MJLQ2 MacBook Pro with Retina Display
  has a compact and suave design that accommodates immense power." />
  <meta property="og:image" content="https://www.website.com/img/macbook-pro-mjlq2.jpg" />
</head>
```

Figure 21 The Opengraph protocol example.

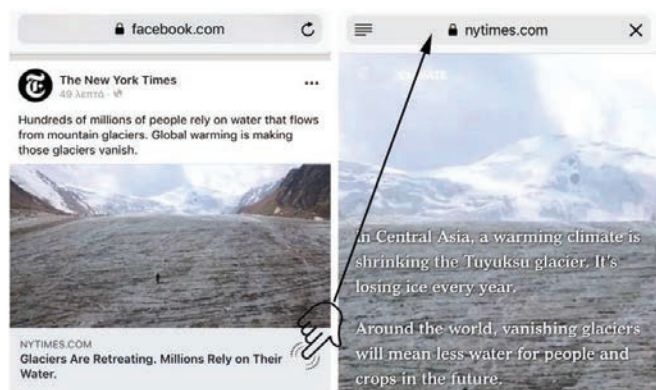


Figure 22 Facebook page post using OG, earn visitor to website.

og: url, og: image) [11]. OG is also considered structured data, like Microdata; however they can be applied either simultaneously or independently, too. The latter is depicted in Figure 21.

As regards SEO, OG meta tags are very important. Web crawlers are daily visiting millions of websites. In case they notice that the site frequently generates content, they visit it more often. Social media users are often creating new content, so web crawlers spend a lot of time looking for new content on social media. Thus, when the content of a website is republished in social media using OG, then it is more likely that this content will be indexed by search engines. Also, social media generates a lot of traffic, which results in republications gaining more clicks and websites attracting more visitors [Figure 22].

3.8 Google My Business

Google My Business is a free tool that companies can use to manage their online presence. By posting business information on Google My Business,

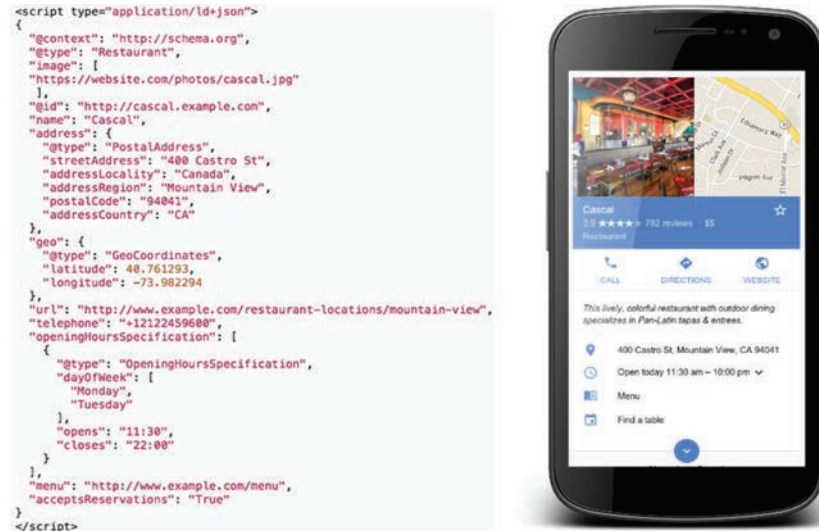


Figure 23 Google my business using structured data.

customers can find the company location more easily, interact with the company, and get informed about its activity [82].

When searching for services on the Google search engine, such as “create a website”, Google will first show Google My Business accounts from companies that create websites and are closer to the searcher. This way, the user will quickly receive the information and will be able to either visit the corresponding website or find the location of the company on the map effectively.

Google My Business for a company can either be created through the Google My Business website or by using structured data technologies within the company website, as shown in Figure 23.

Except for the promotion of the company, Google My Business allows its users to evaluate it as well; the more positive the ratings are, the more likely it is for a website to gain traffic. Finally, it should be mentioned that Google My Business is not a ranking factor; however, it directs traffic to the website.

3.9 Content Delivery Network (CDN) Links

A content delivery network (CDN) is a globally distributed network of proxy servers, known as Points of Presence (POPs), strategically placed around the world, so that end users can access the Internet content with low latency

```

<!-- Bootstrap CSS CDN -->
<link href="https://stackpath.bootstrapcdn.com/bootstrap/4.1.3/css/bootstrap.min.css">
<!-- jQuery JS CDN -->
<script src="https://code.jquery.com/jquery-3.3.1.min.js"></script>
<!-- Font Awesome CSS CDN -->
<link href="https://use.fontawesome.com/releases/v5.5.0/css/all.css">

```

Figure 24 Partial adoption of CDN (Bootstrap, jQuery and Font Awesome libraries).

and high Quality of Experience (QoE) [39]. Their purpose is to store and create cache copies of assets (HTML, JS, CSS, images, etc.) of websites. When a user visits the website, the assets are served by the POP that is closest to the user. As a result, the website is served faster to the end-user, because of the less propagation delay, due to the shorter physical distance, respectively. Many content providers provide CDN services, offering either monthly payment packages or pay-as-you-go packages such as Google’s Cloud CDN.

Many companies, although they desire to adopt technologies so as to improve the page load speed of their website, do not have or do not spend money to buy CDNs with enough bandwidth to cover the traffic of their website. A solution for them is offered by Cloudflare, which provides free CDNs with limited bandwidth [83].

On the other hand, websites can partially adopt the CDN using CDN URLs to load libraries such as jQuery CDN and Bootstrap CDN. A website could, for example, use the CDN provided by the “jquery.com” website to load the jQuery library without having to waste both space and traffic on its web server so visitors will receive the jQuery library from CDN Url, i.e., from the closest POP Server to them. Figure 24 presents a partial adoption of the CDN example of Bootstrap, jQuery, and Font Awesome CSS libraries.

PageSpeed is an equally important ranking factor for a website [84]. Google in its 2010 announcements informed webmasters through its Webmaster Central Blog that “Speeding up websites is of vital importance – not just to site owners, but to all Internet users. Faster sites create happy users and we’ve seen in our internal studies that when a site responds slowly, visitors spend less time there” [85].

4 SEO Techniques and Technologies Evaluation

The purpose of this article is to uncover which SEO techniques are more valuable for websites. Based on this, Webmasters will know in advance which

specific SEO technique should be applied first and which techniques are not commonly used among websites. The results should be considered as a SEO map for Webmasters to survive against the competition.

A representative sample of top-listed websites by their industry was collected from Alexa's website (2.000 entries). The 1689 websites on the list are online stores, and the rest are websites. The list is provided in a CSV format [86].

The sample's websites are top-listed in Alexa's ranking tool based on their daily traffic and unique visitors [87].

Research Questions

Three main questions should be answered by this research.

1. Which SEO techniques are the most important ones?
2. Which specific SEO techniques have already been applied by the most well-known websites?
3. Which SEO technique has to be implemented as a first priority?

Research Hypotheses

Six hypotheses were created which our research takes for granted:

1. The larger the sample the more representative the results.
2. Top-listed websites have more correctly applied SEO techniques.
3. Top-listed websites have entire digital marketing team to support them with specialized SEO experts who know exactly which SEO techniques are more effective and which of the SEO techniques must be implemented to acquire higher results in search engine results pages (SERPs).
4. SEO techniques used by top-listed websites are also the most effective.
5. The most common-used SEO techniques on our results should be applied first.
6. The most common-used SEO techniques on our results are most effective.

SEOmized Tool Implementation

Inspired by the SEO techniques and technologies thoroughly analyzed in Sections 2 and 3, the authors developed a prototype web crawler – web scraper tool, called SEOmized (Search Engine Optimized) [88]. SEOmized is

responsible to extract websites' HTML source code and identify which SEO technique(s), among the predefined set of 17 SEO techniques presented in Table 5, are also applied to the website. The source code of SEOmized is available in the GitHub repository [88].

Limitations and Application Programming Interface (APIs)

To capture results for website speed and mobile-friendliness two APIs have been integrated into the SEOmized functionality.

- Mobile-Friendly Test Tool API, created by Google, scans the given URL against responsive techniques and returns a list of mobile usability issues [89].
- PageSpeed Insights API, created by Google, measures the performance of a web page producing suggestions about the website's performance, such as page speed, accessibility, and SEO [90].

Both the API tools took as input the URLs of the dataset websites. Both the tools managed to return results for the majority of the sample websites. Mobile-Friendly Test Tool API, however, failed to scan 24 websites, so the responsive design inspections of these websites were held manually. In contrast, the PageSpeed Insights API tool did not return speed results for only two websites, which were tested with the corresponding Pingdom tool [91].

Crawling Top-listed Websites and Detecting SEO Techniques

The SEOmized tool was powered by the list of 2,000 top-listed websites retrieved by Alexa's website. Through a loop, each of the given websites crawled and applied SEO techniques detected and stored in SEOmized database. The database is provided in SQL format [92].

Processing the Results

In SEOmized database are now stored websites and the SEO techniques each one follows. Applied SEO techniques are defined as 1, and the techniques not found in websites' source code are defined as 0.

An illustration of the results is presented in Table 4.

Research Findings

A list of the most commonly used SEO techniques among top-listed websites and their usage percentage is presented in Figure 25.

Table 4 An illustration of the results in the SEOmized Database

seomized_id	26	72
website	https://case-mate.com/	https://www.batteryspace.com/
title	1	0
meta_description	1	1
heading1	0	0
heading2	1	0
url	1	1
image_alt	0	0
href_title	0	0
rss	0	0
robots	1	1
responsive	0	1
https	1	1
structured_data	1	0
amp	0	0
minify_css	0	0
minify_js	0	0
og_tags	0	1
page_load_time	1	1
SUM	8	7

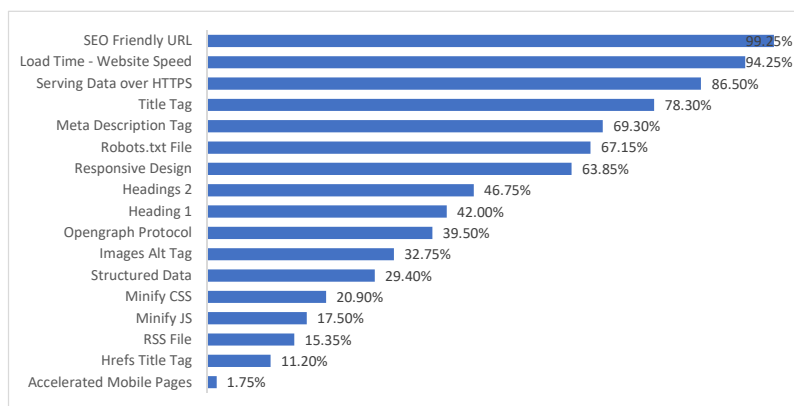


Figure 25 SEO techniques usage score among top-listed websites (%).

Results helped us to extract valuable information about the way top-listed websites are trying to apply SEO techniques to increase their organic traffic.

In order to quantify the results of SEOmized we introduced a corresponding metric, i.e., the SEOmized score, which is equal to the ratio of the

Table 5 Website Achievements against 17 SEO techniques

SEO Techniques	Ebay.com	Amazon.com	Walmart.com	Bookhampton.com
Title Tag	✓□	✓□	✓□	✗□
Heading 1	✓□	✓□	✗□	✓□
Headings 2	✓□	✓□	✓□	✓□
SEO Friendly URL	✓□	✓□	✓□	✓□
Meta Description Tag	✓□	✓□	✓□	✓□
Images Alt Tag	✗□	✗□	✗□	✓□
Hrefs Title Tag	✗□	✗□	✗□	✓□
RSS File	✗□	✗□	✗□	✓□
Robots.txt File	✓□	✓□	✓□	✓□
Responsive Design	✓□	✓□	✓□	✓□
Serving Data over HTTPS	✓□	✓□	✓□	✓□
Structured Data	✓□	✗□	✗□	✓□
Accelerated Mobile Pages	✓□	✗□	✗□	✓□
Load Time – Website Speed	1.3 seconds	2.1 seconds	1.2 seconds	3.9 seconds
Minify CSS	✗□	✗□	✗□	✓□
Minify JS	✗□	✗□	✗□	✓□
Opengraph Protocol	✓□	✓□	✓□	✓□
SEOmized Score	70,59%	58,82%	52,94%	94,12%

implemented SEO techniques and technologies detected in each website by SEOmized against the total ones supported by the tool, respectively.

Table 5 presents as a sample the results of three well-known e-commerce marketplace platforms (Ebay.com, Amazon.com, Walmart.com), as opposed to one less-known book store e-commerce (Bookhampton.com). At the end of the figure, the SEOmized score (percentage of total applied SEO techniques against the total ones supported by the tool) is calculated for each website.

Following the hypotheses made in the Research Hypotheses section, this research concludes as follows:

- 94.25% of websites have adopted technologies like apache deflate and gzip [section 3.4] to reduce their load time to under 5 seconds. More specifically, the overall average load time among top-listed websites is 1.88 seconds.
- As presented in section 3.1 the Chrome browser highlights as not secure all the websites that are not using SSL certificates. The response was immediate as 86.50% of websites have installed SSL certificates until today.

- Additionally, although search engines are evaluating positively websites that are using structured data, we noticed that 70.60% of the websites have not applied structured data in their code.
- In addition, although more and more visitors are using mobile devices to browse websites, only 63.85% of them have passed the responsive-design test.
- Table 4 shows the online store of a bookstore among the results (Bookhampton.com). This online store uses 94.12% of the SEO techniques that the SEOmized tool supports, unlike large online stores that have applied fewer SEO techniques. That is a significant example of smaller websites, following each and every SEO rule to climb to the search results.

All the above results is a first attempt that reveals the SEO techniques behind well-known websites. Having Figure 25 as a guide map, Webmasters can start applying to their websites these SEO techniques that have a higher adoption rate. SEO techniques like Accelerated Mobile Pages and HREFs title tags seem not preferred by top-listed websites thus they may need to be implemented in a later SEO stage.

Surely these numbers become outdated on a daily basis, but the overall trend points to a tendency for faster and safer WEB.

Discussion and Future Work

The data collected from this research contributes to a clearer understanding of how well-known websites construct their SEO strategy. Although the majority of the websites have applied SEO techniques for security and speed, they have failed to use certain techniques that the search engines consider important like the AMP, RSS, Minified files, and Structured data. An assumption could be that these techniques are difficult to apply to overcrowded websites. For instance, to create an AMP alternative page website owner has to create and maintain two versions of his/her website. So, it's convenient for top-listed websites not to use those SEO techniques that demand a lot of effort to maintain. Instead, they sacrifice on purpose the extra traffic that these techniques could possibly deliver to, and they are trying to gain it by using different methods, such as adding resources for customer support and satisfaction.

More and more techniques and technologies that promote SEO are created daily. SEO is now an infinite marketing industry that brings income to

developers and results to website owners. The SEOmized tool created for the needs of this paper can be extended with additional SEO techniques in the future, bringing more alternatives to web masters, which will lead to even greater results.

5 Conclusion

Webmasters have realized the importance of having their websites high in search results since the mid 1990s [12]. Thereafter they tried to optimize their websites for search engines with existing technologies or even develop new ones. SEO techniques through observation and experience have as a consequence evolved year after year. SEO has been and still is at the center of website promotion. This paper presents as a review many SEO techniques and technologies that have been invented by SEO professionals to date. The main focus of this article is to generate a guide-map for Webmasters regarding the SEO techniques that should be applied to their websites. Three main questions were answered throughout this research. Which SEO techniques are the most important ones? Which specific SEO techniques have already been applied by the most well-known websites? Which SEO technique has to be implemented as a first priority? To answer the above questions, we created a SEO tool which crawls a website and returns which specific SEO techniques it has applied. This tool was given as input a list of 2,000 top-listed websites and undertook to scan these websites and reveal which SEO techniques each one follows. Having the results stored in tool's database, we have created a figure which describes the SEO techniques usage score among top-listed websites [Figure 25]. That figure can be used as a guide-map by Webmasters since it responds all the questions examined by this article. Following the SEO plan implemented by the most well-known websites, every Webmaster without special SEO background will be able to apply the appropriate SEO techniques that will result in higher positions in SERPs and more traffic to their website.

List of Notations and Abbreviations

- Alt tags: alternative tags
- AMP: Accelerated Mobile Pages Project
- CDN: content delivery network
- CMS: Content Management Systems

- **Css:** Cascading Style Sheets
- **DA:** domain authority
- **Gzip:** Gzip is a file format and a software application used for file compression and decompression.
- **H1:** header 1
- **H6:** header 6
- **Htaccess:** files (or “distributed configuration files”) provide a way to make configuration changes on a per-directory basis.
- **HTML:** HyperText Markup Language
- **HTTP:** HyperText Transfer Protocol
- **HTTPS:** Hypertext Transfer Protocol Secure
- **Iframe:** The <iframe> tag specifies an inline frame. An inline frame is used to embed another document within the current HTML document.
- **Js:** JavaScript
- **JSON-LD:** JavaScript Object Notation for Linked Data
- **JSON:** JavaScript Object Notation
- **Kb:** kilobytes
- **Meta:** The <meta> tag defines metadata about an HTML document. Metadata is data (information) about data.
- **OGP:** Open Graph Protocol
- **POPs:** Points of Presence
- **PPC:** paying per click
- **PSM:** paid search marketing
- **Q & A:** question-and-answer
- **RDFa:** Resource Description Framework
- **RESTful:** Representational state transfer
- **RSS:** Rich Site Summary
- **SEO:** Search engine optimization
- **SERPs:** search engine results pages
- **SSL:** Secure Sockets Layer
- **UGC:** User-generated content
- **URL:** Uniform Resource Locator – web address
- **UX:** user experience
- **W3C:** World Wide Web Consortium
- **WHATWG:** Web Hypertext Application Technology Working Group
- **XHTML:** EXtensible HyperText Markup Language
- **XML:** eXtensible Markup Language

References

- [1] Fabien Gandon, Wendy Hall, 2022. A Never-Ending Project for Humanity Called “the Web”. WWW ’22: Proceedings of the ACM Web Conference (pp. 3480–3487).
- [2] Heuer, Jorg; Hund, Johannes; Pfaff, Oliver (2015). Toward the Web of Things: Applying Web Technologies to the Physical World. *Computer*, (v.48(5), pp.34–42).
- [3] Total number of Websites, <https://www.internetlivestats.com/total-number-of-websites/> (Last accessed: May 20th, 2022).
- [4] Vishwas Raval, Padam Kumar, 2012. SEReleC (Search Engine Result Refinement and Classification) – a Meta search engine based on combinatorial search and search keyword based link classification. In *IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM -2012)*, (pp. 627–631), Nagapattinam, Tamil Nadu, India
- [5] Zhou Hui, Qin Shigang, Liu Jinhua, Chen Jianli, 2012. Study on Website Search Engine Optimization. In *2012 International Conference on Computer Science and Service System*, (pp. 930–933), Nanjing, China.
- [6] Venkat N. Gudivada, Dhana Rao, Jordan Paris, 2015. Understanding Search-Engine Optimization. *Computing*, (pp. 43–52, vol. 48).
- [7] Mo Yunfeng, 2010. A Study on Tactics for Corporate Website Development Aiming at Search Engine Optimization. In *Second International Workshop on Education Technology and Computer Science*, (pp. 673–675), Wuhan, China.
- [8] Kai Li, Mei Lin, Zhangxi Lin, Bo Xing, 2014. Running and Chasing – The Competition between Paid Search Marketing and Search Engine Optimization. In *47th Hawaii International Conference on System Sciences*, (pp. 3110–3119), Waikoloa, HI, USA.
- [9] Joyce Yoseph Lemos, Abhijit R. Joshi, 2017. Search engine optimization to enhance user interaction. In *International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, (pp. 398–402), Palladam, India.
- [10] Nursel Yalçın, Utku Köse, 2010. What is search engine optimization: SEO?. *Procedia – Social and Behavioral Sciences*, (pp. 487–493, vol. 9).
- [11] Samedin Krrabaj, Fesal Baxhaku, Dukagjin Sadrijaj, 2017. Investigating search engine optimization techniques for effective ranking: A case study of an educational site. In *6th Mediterranean Conference on Embedded Computing (MECO)*, Bar, Montenegro.

- [12] Surbhi Chhabra, Ravi Mittal, Darothi Sarkar, 2016. Inducing factors for search engine optimization techniques: A comparative analysis. In 1st India International Conference on Information Processing (IICIP), Delhi, India.
- [13] L. Invernizzi, P. M. Comparetti, S. Benvenuti, C. Kruegel, M. Cova, G. Vigna, 2012. EvilSeed: A Guided Approach to Finding Malicious Web Pages. 2012 IEEE Symposium on Security and Privacy, (pp. 428–442), San Francisco, CA, USA.
- [14] Qing Zhang, David Y. Wang, Geoffrey M. Voelker, 2014. DSpin: Detecting Automatically Spun Content on the Web. In Network and Distributed System Security Symposium.
- [15] Luca Invernizzi, Kurt Thomas, Alexandros Kapravelos, Oxana Comanescu, Jean-Michel Picod, Elie Bursztein, 2016. Cloak of Visibility: Detecting When Machines Browse a Different Web. In IEEE Symposium on Security and Privacy (SP), (pp. 743–758), San Jose, CA, USA.
- [16] S. Savoska, Niko Naka, Violeta Manevska, Blagoj Ristevski, 2016. Search Engine Optimization on PHP based web pages in practice. Journal of Emerging Research and Solutions in ICT, (vol. 1, pp. 45–58).
- [17] Vikas M. Patil, Amruta V. Patil, 2018. SEO: On-Page + Off-Page Analysis. In International Conference on Information, Communication, Engineering and Technology (ICICET), Pune, India.
- [18] John B. Killoran, 2013. How to Use Search Engine Optimization Techniques to Increase Website Visibility. IEEE Transactions on Professional Communication (vol. 56).
- [19] Shruti Kohli, Sandeep Kaur, Gurrajan Singh, 2012. A Website Content Analysis Approach Based on Keyword Similarity Analysis. In IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, (pp. 254–257), Macau, China.
- [20] Grzegorz Szymanski, Piotr Lininski, 2018. Model of the Effectiveness of Google Adwords Advertising Activities. In IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), (pp. 98–101), Lviv, Ukraine.
- [21] Zuze H., Weideman M. 2013. Keyword stuffing and the big three search engines. Online Information Review. (vol. 37, no. 2, pp. 268–286).
- [22] Google Search Console Help, <https://support.google.com/webmasters/answer/7451184>, [Last accessed: May 20th, 2022].
- [23] Google Support, <https://support.google.com/webmasters/answer/40349>, [Last accessed: May 20th, 2022].

- [24] Imrul Kayes Sabbir, Rahma Akhter, 2018. Impact of copywriting in marketing communication, In BRAC Business School.
- [25] Fuxue Wang, Yi Li, Yiwen Zhang, 2011. An empirical study on the search engine optimization technique and its outcomes. In 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), (pp. 2767–2770), Dengleng, China.
- [26] Patil Swati P, Pawar B.V, Patil Ajay S, 2013. Search Engine Optimization: A Study. *Research Journal of Computer and Information Technology Sciences* (v.1(1) pp.10–13).
- [27] Radovan Dragić, Tanja Kaurin, 2012. Significance and Impact of Meta Tagson Search Engine Results Pages. *INFOTEH-JAHORINA* (vol. 11, pp. 751–755).
- [28] Vinit Kumar Gunjan, Pooja, Monika Kumari, Dr Amit Kumar,Dr (col.) Allam appa rao, 2012. Search engine optimization with Google. *IJCSI International Journal of Computer Science Issues* (vol. 9 issue 1, no. 3, pp. 206–214).
- [29] Tuan Ly Van, Duc Pham Minh, Thanh Le Dinh, 2017. Identification of paths and parameters in RESTful URLs for the detection of web Attacks. In 4th NAFOSTED Conference on Information and Computer Science, (pp. 110–115), Hanoi, Vietnam.
- [30] Cen Zhu, Guixing Wu, 2011. Research and Analysis of Search Engine Optimization Factors Based on Reverse Engineeing. In Third International Conference on Multimedia Information Networking and Security, (pp. 225–228), Shanghai, China.
- [31] Madlena Nen, Valentin Popa, Andreea Scurtu, Roxana Larisa Unc, 2017. The Computer Management – SEO Audit. Review of International Comparative Management. *Central and Eastern European Online Library*. (vol. 18, Issue 3 pp. 297–307).
- [32] Patzer, A., Moodie, M. (2004). The Front Controller Pattern. In: Moodie, M. (eds) *Foundations of JSP Design Patterns*. Apress, Berkeley, CA.
- [33] Sonya Zhang, Neal Cabage, 2013. Does SEO Matter? Increasing Classroom Blog Visibility through Search Engine Optimization. In 46th Hawaii International Conference on System Sciences, (pp. 1610–1619), Wailea, Maui, HI, USA.
- [34] Dr. Peter J. Meyers, <https://moz.com/blog/should-i-change-my-urls-for-seo>, [Last accessed: May 20th, 2022].
- [35] Best practices for creating quality meta descriptions, <https://developers.google.com/search/docs/advanced/appearance/snippet> (Last accessed: May 20th, 2022).

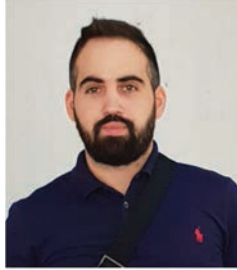
- [36] Meta Description – Moz, <https://moz.com/learn/seo/meta-description> (Last accessed: May 20th, 2022).
- [37] Choosing the right image formats, sizes and dimensions, <https://searchengineland.com/heres-what-you-need-to-know-about-image-optimization-for-seo-316046> (Last accessed: May 20th, 2022).
- [38] How to fix the issue when Image size is over 100 KB, <https://sitechecker.pro/site-audit-issues/image-size-100-kb/> (Last accessed: May 20th, 2022).
- [39] Huan Wang, Guoming Tang, Kui Wu, Jiamin Fan, 2018. Speeding Up Multi-CDN Content Delivery via Traffic Demand Reshaping. In IEEE 38th International Conference on Distributed Computing Systems (ICDCS), (pp. 422–433), Vienna, Austria.
- [40] Jeffrey P. Bigham, 2007. Increasing web accessibility by automatically judging alternative text quality. In Proceedings of the 2007 International Conference on Intelligent User Interfaces, (pp. 349–352), Honolulu, Hawaii, USA.
- [41] W3C H37: Using alt attributes on img elements, <https://www.w3.org/TR/WCAG20-TECHS/H37.html> (Last accessed: May 20th, 2022).
- [42] WebAIM Alternative Text, <https://webaim.org/techniques/alttext/> (Last accessed: May 20th, 2022).
- [43] Azlin Zanariah Bahtar, Mazzini Muda, 2016. The Impact of User – Generated Content (UGC) on Product Reviews towards Online Purchasing – A Conceptual Framework. *Procedia Economics and Finance*, (vol. 37, pp. 337–342).
- [44] Naveen Amblee, Tung Bui, 2011. Harnessing the Influence of Social Proof in Online Shopping: The Effect of Electronic Word of Mouth on Sales of Digital Microproducts. *International Journal of Electronic Commerce*, (vol. 16:2, pp. 91–114).
- [45] Hamim Hamid Gabir, Azza Z. Karrar, 2018. The Effect of Website’s Design Factors on Conversion Rate in E-commerce. In International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE), Khartoum, Sudan.
- [46] Structured Data, <https://schema.org>, [Last accessed: May 20th, 2022].
- [47] WhatWG, <https://html.spec.whatwg.org/multipage/microdata.html>, [Last accessed: May 20th, 2022].
- [48] RDFa 1.1 Primer – Third Edition, <https://www.w3.org/TR/2015/NOT-E-rdfa-primer-20150317/>, [Last accessed: May 20th, 2022].
- [49] Markus Lanthaler, Christian Gütl, 2012. On Using JSON-LD to Create Evolvable RESTful Services. In Proceedings of the Third International Workshop on RESTful Design, (pp. 25–32), Lyon, France.

- [50] RDF AND JSON-LD UseCases. https://www.w3.org/2013/dwbp/wiki/RDF_AND_JSON-LD_UseCases (Last accessed: May 20th, 2022).
- [51] Jatinder Manhas, 2015. Design and development of automated tool to study sitemap as design issue in Websites. In 2nd International Conference on Computing for Sustainable Global Development (INDIACom), (pp. 514–518), New Delhi, India.
- [52] Ben Hammersley, 2003. Content Syndication with RSS. O'Reilly Media, Inc. ISBN: 9780596003838.
- [53] James Lewin, 2000. An introduction to RSS news feeds Using open formats for content syndication. IBM developerWorks : Web architecture.
- [54] D. Ma, 2009. Offering RSS Feeds: Does It Help to Gain Competitive Advantage?. In 42nd Hawaii International Conference on System Sciences, (pp. 1–10), Big Island, HI, USA.
- [55] N Ragavan, 2017. Efficient key hash indexing scheme with page rank for category based search engine big data. In IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), Srivilliputhur, India.
- [56] Joomla Documentation, https://docs.joomla.org/Robots.txt_file, [Last accessed: May 20th, 2022].
- [57] Google Developers, <https://developers.google.com/search/mobile-sites/>, [Last accessed: May 20th, 2022].
- [58] Emmanouil Perakakis, Gheorghita Ghinea, 2017. Smart Enough for the Web? A Responsive Web Design Approach to Enhancing the User Web Browsing Experience on Smart TVs. IEEE Transactions on Human-Machine Systems (vol: 47, issue: 6).
- [59] Zhigang Hua, Xing Xie, Hao Liu, Hanqing Lu, Wei-Ying Ma, 2006. Design and Performance Studies of an Adaptive Scheme for Serving Dynamic Web Content in a Mobile Computing Environment. IEEE Transactions on Mobile Computing (vol: 5, issue: 12).
- [60] Fawwaz Yousef Alnawaj'ha, Mohammed Saeed Abutaha, 2018. Responsive web design commitment by the web developers in Palestine. In 4th International Conference on Computer and Technology Applications (ICCTA), (pp. 69–73), Istanbul, Turkey.
- [61] Wei Jiang, Meng Zhang, Bin Zhou, Yujian Jiang, Yingwei Zhang, 2014. Responsive web design mode and application. In IEEE Workshop on Advanced Research and Technology in Industry Applications (WARTIA), (pp. 1303–1306), Ottawa, ON, Canada.
- [62] Google Search, <https://search.google.com/test/mobile-friendly>, [Last accessed: May 20th, 2022].

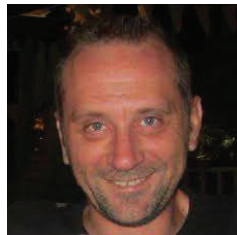
- [63] Norman E. Youngblood, Jo Mackiewicz, 2012. A usability analysis of municipal government website home pages in Alabama. *Government Information Quarterly*, (vol. 29, issue 4, pp. 582-588).
- [64] Google Search Central – Breadcrumb, <https://developers.google.com/search/docs/advanced/structured-data/breadcrumb> (Last accessed: May 20th, 2022).
- [65] Aanchal Kakkar, Rana Majumdar, Arvind Kumar, 2015. Search engine optimization: A game of page ranking. In 2nd International Conference on Computing for Sustainable Global Development (INDIACom), (pp. 206–210), New Delhi, India.
- [66] S. Sharma, S. Verma 2020. Optimizing Website effectiveness using various SEO Techniques. 7th International Conference on Signal Processing and Integrated Networks (pp. 918–922).
- [67] White Hat SEO: How To Rank by the Rules, <https://www.semrush.com/blog/white-hat-seo/> (Last accessed: May 20th, 2022).
- [68] X. Ma 2018. Research on Black Hat SEO Behaviour Measurement. IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), (pp. 1041–1045).
- [69] Bello, Rotimi-Williams & Noah, Otobo, 2018. Conversion of Website Users to Customers-The Black Hat SEO Technique. *International Journal of Advanced Research in Computer Science and Software Engineering*. (v. 8. pp. 29–35).
- [70] Tom Van Goethem, Najmeh Miramirkhani, Wouter Joosen, Nick Niki-forakis, 2019. Purchased Fame: Exploring the Ecosystem of Private Blog Networks. *Asia CCS '19: Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security* (pp. 366–378).
- [71] Google search central blog: HTTPS as a ranking signal, <https://developers.google.com/search/blog/2014/08/https-as-ranking-signal> (Last accessed: May 20th, 2022).
- [72] Stable Channel Update for Desktop, <https://chromereleases.googleblog.com/2018/07/stable-channel-update-for-desktop.html> (Last accessed: May 20th, 2022).
- [73] AMP Project, <https://ampproject.org>, [Last accessed: May 20th, 2022].
- [74] Is your AMP page or Web Story valid? <https://search.google.com/test/amp> (Last accessed: May 20th, 2022).
- [75] Make your pages discoverable, <https://amp.dev/documentation/guides-and-tutorials/optimize-and-measure/discovery/> (Last accessed: May 20th, 2022).
- [76] GNU, <https://gnu.org/software/gzip>, [Last accessed: May 20th, 2022].

- [77] Pratiksha H. Shroff, Seema R. Chaudhary, 2017. Critical rendering path optimizations to reduce the web page loading time. In 2nd International Conference for Convergence in Technology (I2CT), (pp. 937–940), Mumbai, India.
- [78] Twig Documentation, <https://twig.symfony.com/doc/3.x/tags/cache.html> (Last accessed: May 20th, 2022).
- [79] Apache, https://apache.org/docs/current/mod/mod_expires.html, [Last accessed: May 20th, 2022].
- [80] OGP, <https://ogp.me>, [Last accessed: May 20th, 2022].
- [81] Austin Haugen, 2010. The Open Graph Protocol Design Decisions. The Semantic Web – ISWC.
- [82] Google My Business, <https://support.google.com/business/answer/7091>, [Last accessed: May 20th, 2022].
- [83] The basics you need to protect and accelerate your site, <https://www.cloudflare.com/plans/free/> (Last accessed: May 20th, 2022).
- [84] Webmaster Central Blog, <https://webmasters.googleblog.com/2018/01/using-page-speed-in-mobile-search.html>, [Last accessed: May 20th, 2022].
- [85] Using site speed in web search ranking, <https://developers.google.com/search/blog/2010/04/using-site-speed-in-web-search-ranking> (Last accessed: May 20th, 2022).
- [86] SEOmized CSV file, <https://github.com/kroumeliotis/seomized/blob/main/seomized.csv> (Last accessed: May 20th, 2022).
- [87] Alexa by Amazon, <https://alexa.com/> (Last accessed: May 1st, 2022).
- [88] SEOmized tool, GitHub Repository, <https://github.com/kroumeliotis/seomized/tree/main/seomized>, [Last accessed: May 20th, 2022].
- [89] Mobile-Friendly Test Tool, <https://support.google.com/webmasters/answer/6352293?hl=en> (Last accessed: May 20th, 2022).
- [90] Pagespeedapi runpagespeed, <https://developers.google.com/speed/docs/insights/v4/reference/pagespeedapi/runpagespeed> (Last accessed: May 20th, 2022).
- [91] Website Speed Test | Pingdom Tools, <https://tools.pingdom.com> (Last accessed: May 20th, 2022).
- [92] SEOmized SQL file, <https://github.com/kroumeliotis/seomized/blob/main/seomized.sql> (Last accessed: May 20th, 2022).

Biographies



Konstantinos I. Roumeliotis received the master's degree in computer engineering from University of Peloponnese in 2019. He is currently PhD candidate at the Department of Informatics and Telecommunications, University of Peloponnese. His research areas include web technologies, technical search engine optimization and web accessibility.



Nikolaos D. Tselikas is an Associate Professor at the Department of Informatics and Telecommunications in the University of Peloponnese (UoP). He received both his Dipl.-Ing. and PhD from the School of Electrical and Computer Engineering of National Technical University of Athens in 1999 and 2004, respectively. His research interests include: services and networking, middleware and software engineering, software performance evaluation, location-based services, mobile applications and open source software. He has been involved in several research projects on the above areas and has published more than 70 papers in international journals, conferences and book chapters and two books on C programming language.

