

ANKARA YILDIRIM BEYAZIT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES



EXPLORING SEARCH ENGINE OPTIMIZATION FACTORS

M.Sc. Thesis by
Melike SARIKAYA

Department of Computer Engineering

December, 2021

ANKARA

EXPLORING SEARCH ENGINE OPTIMIZATION FACTORS

A Thesis Submitted to

The Graduate School of Natural and Applied Sciences of

Ankara Yıldırım Beyazıt University

**In Partial Fulfillment of the Requirements for the Degree of Master of Science
in Computer Engineering, Department of Computer Engineering**

by

Melike SARIKAYA

December, 2021

ANKARA

M.Sc. THESIS EXAMINATION RESULT FORM

We have read the thesis entitled “**EXPLORING SEARCH ENGINE OPTIMIZATION FACTORS**” completed by **MELİKE SARIKAYA** under the supervision of **ASST. PROF. DR. MUSTAFA YENİAD** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Asst. Prof. Dr. Mustafa YENİAD

Supervisor

Prof. Dr. Fatih Vehbi ÇELEBİ

Jury Member

Asst. Prof. Dr. Yılmaz AR

Jury Member

Prof. Dr. Sadettin ORHAN

Director

Graduate School of Natural and Applied Sciences

ETHICAL DECLARATION

I hereby declare that, in this thesis which has been prepared in accordance with the Thesis Writing Manual of Graduate School of Natural and Applied Sciences,

- All data, information and documents are obtained in the framework of academic and ethical rules,
- All information, documents and assessments are presented in accordance with scientific ethics and morals,
- All the materials that have been utilized are fully cited and referenced,
- No change has been made on the utilized materials,
- All the works presented are original,

and in any contrary case of above statements, I accept to renounce all my legal rights.

Date: 10.12.2021

Signature:

Name & Surname: Melike SARIKAYA

ACKNOWLEDGMENTS

Firstly, I would like to express my sincere gratitude to my supervisor, Asst. Prof. Dr. Mustafa YENİAD for his tremendous support and motivation during my study. His immense knowledge and precious recommendations constituted the milestones of this study. His guidance assisted me all the time of my research and while writing this thesis.

I would like to express my gratitude to my dear mother Dilek SARIKAYA, who brought me to this day, motivated me and supported me in all circumstances, and my dear father Mehmet SARIKAYA, who was my guide with her hard work and perspective on life.

Finally, I would like to thank my friends for their support and motivation.

2021, 10 December

Melike SARIKAYA

EXPLORING SEARCH ENGINE OPTIMIZATION FACTORS

ABSTRACT

A search engine is a system that allows users to enter queries to get information or find a result and lists the results according to these queries. It enables people who want to bridge between people and data to obtain information over the internet. Search Engine Optimization (SEO) is the activity of optimizing web pages or entire sites to make them search engine friendly and thus rank higher in search results. There are different methods for search engines to access effective information.

In this study, first of all, the development process of search engines, search engine types and working principles are mentioned. The stages of search engine optimization are explained one by one. How web crawling, indexing and ranking is done are examined. Then, algorithms used in search engine optimization are mentioned. PageRank, TrustRank and HillTop algorithms, which are revolutionary algorithms in search engines, were compared. Important algorithm updates published by Google were reviewed and their strengths and weaknesses were compared. Ahrefs, Moz and Semrush, which are search engine optimization tools, were examined and their prominent features were compared. Finally, the performance of selected university websites in search engine optimization tools is shown. Domain views, organic search results, keyword analyzes, backlink analyzes were performed and the results were compared with graphics.

Research and analysis results have shown that websites will have a high position in search engine results when they apply appropriate optimization techniques in line with their own needs. It has been determined that the appropriate keyword selection and the number of reliable backlinks are important search engine optimization factors. Since search engines have a dynamic structure, it is important for website owners to keep themselves updated.

Keywords: SEO, search engine optimization, search engine, Google, search engine algorithms

ARAMA MOTORU OPTİMİZASYON FAKTÖRLERİNİN KEŞFİ ÖZ

Arama motoru, kullanıcıların bilgi almak veya bir sonuç bulmak için sorgu girmelerini sağlayan ve bu sorgulara göre sonuçları listeleyen bir sistemdir. İnsanlar ve veriler arasında köprü oluşturmak isteyen kişilerin internet üzerinden bilgi edinmelerini sağlar. Arama Motoru Optimizasyonu (SEO), web sayfalarını veya tüm siteleri arama motoru dostu hale getirmek ve böylece arama sonuçlarında daha üst sıralarda yer almak için optimize etme etkinliğidir. Arama motorlarının etkili bilgilere erişmesi için farklı yöntemler mevcuttur.

Bu çalışmada öncelikle arama motorlarının gelişim sürecinden, arama motoru türlerinden ve çalışma prensiplerinden bahsedilmiştir. Arama motoru optimizasyonunun aşamaları tek tek anlatılmaktadır. Web taraması, indeksleme ve sıralamanın nasıl yapıldığı incelenmiştir. Ardından arama motoru optimizasyonunda kullanılan algoritmalarından bahsedilmiştir. Arama motorlarında devrim niteliğinde algoritmalar olan PageRank, TrustRank ve HillTop algoritmaları karşılaştırılmıştır. Google tarafından yayınlanan önemli algoritma güncellemeleri incelenerek güçlü ve zayıf yönleri karşılaştırılmıştır. Arama motoru optimizasyon araçlarından Ahrefs, Moz ve Semrush incelenmiş, öne çıkan özellikleri karşılaştırılmıştır. Son olarak, seçilen üniversite web sitelerinin arama motoru optimizasyon araçlarındaki performansı gösterilmektedir. Domain görünümleri, organik arama sonuçları, anahtar kelime analizleri, geri bağlantı analizleri yapılmış ve sonuçlar grafiklerle karşılaştırılmıştır.

Araştırma ve analiz sonuçları, web sitelerinin kendi ihtiyaçları doğrultusunda uygun optimizasyon tekniklerini uyguladıklarında, arama motoru sonuçlarında üst sıralarda yer alacaklarını göstermiştir. Uygun anahtar kelime seçimi ve güvenilir geri bağlantı sayısının önemli arama motoru optimizasyon faktörleri olduğu tespit edilmiştir. Arama motorları dinamik bir yapıya sahip olduğu için web sitesi sahiplerinin kendilerini güncel tutmaları önemlidir.

Anahtar Kelimeler: SEO, arama motoru optimizasyonu, arama motoru, Google, arama motoru algoritmaları

CONTENTS

M.Sc. THESIS EXAMINATION RESULT FORM.....	ii
ETHICAL DECLARATION	iii
ACKNOWLEDGMENTS.....	iv
ABSTRACT	v
ÖZ	vi
NOMENCLATURE.....	ix
LIST OF TABLES	x
LIST OF FIGURES.....	xi
CHAPTER 1 - INTRODUCTION.....	1
1.1 Aim of the Thesis	1
1.2 Research Questions and Outcomes	2
1.3 Thesis Organization	2
CHAPTER 2 - BACKGROUND.....	4
2.1 Search Engine.....	4
2.1.1 Crawler-based Search Engines.....	5
2.1.2 Directory-based Search Engines	6
2.1.3 Hybrid Search Engines.....	6
2.1.4 Vertical Search Engines	6
2.2 Search Engine Optimization Techniques	7
2.2.1 On-Page SEO	8
2.2.2 Off-Page SEO.....	15
2.2.3 SEO Technical Stages	17
2.3 Search Engine Optimization Algorithms	18
2.3.1 PageRank Algorithm	18
2.3.2 TrustRank Algorithm	20
2.3.3 HillTop Algorithm.....	21
2.4 Google Algorithm Updates	22
2.4.1 Google Panda Update.....	23
2.4.2 Google Penguin Update	24
2.4.3 Google Hummingbird Update	25
2.4.4 Google Pigeon Update	27

2.4.5 Google Mobilegeddon Update	28
2.4.6 Google RankBrain Update	30
2.4.7 Google Medic Update	32
2.5 Search Engine Optimization Tools	34
2.5.1 Ahrefs	34
2.5.2 Moz.....	35
2.5.3 Semrush.....	37
2.5.4 Nibbler.....	38
CHAPTER 3 - LITERATURE REVIEW.....	39
CHAPTER 4 - METHODS AND APPLICATION.....	44
4.1 Analysis of University Websites with Semrush.....	44
4.2 Analysis of University Websites with Moz	49
4.3 Analysis of University Websites with Nibbler.....	53
CHAPTER 5 - EXPERIMENTAL RESULTS.....	56
5.1 Findings.....	56
5.1.1 Advantages of SEO	57
5.1.2 Disadvantages of SEO.....	57
5.2 SEO Recommendations	58
5.3 Comparison of SEO Tools	59
5.4 Comparison of Analysis Results on University Websites.....	61
CHAPTER 6 - CONCLUSION & FUTURE WORK.....	65
6.1 Conclusion.....	65
6.2 Future Work	66
REFERENCES.....	67
CIRRICULUM VITAE	73

NOMENCLATURE

Acronyms

ALT	Alternative Tag
CTR	Click Through Rate
DOC	Document
E-A-T	Expertise, Authority and Trust
HTML	Hyper Text Markup Language
HTTP	Hyper Text Transfer Protocol
HTTPS	Hyper Text Transfer Protocol Secure
IP	Internet Protocol Address
PDF	Portable Document Format
SEF	Search Engine Friendly
SEO	Search Engine Optimization
SERP	Search Engine Results Page
SSL	Secure Socket Layers
TLS	Transport Layer Security
TXT	Text
URL	Uniform Resource Locator
XML	Extensible Markup Language
YMYL	Your Money Your Life

LIST OF TABLES

Table 2.1 Comparison of PageRank, TrustRank and HillTop Algorithms	21
Table 2.2 Comparison of Google developed algorithm updates	22
Table 5.1 SEO tools comparison	60
Table 5.2 Analysis results with Semrush Tool.....	62
Table 5.3 Analysis results with Moz Tool.....	63
Table 5.4 Analysis results with the Nibbler Tool.....	64



LIST OF FIGURES

Figure 2.1 Search engine history.....	4
Figure 2.2 Usage rate of search engines.....	5
Figure 2.3 SEO techniques.....	7
Figure 2.4 One-way and two-way link building	16
Figure 2.5 SEO technical stages.....	17
Figure 2.6 PageRank scores of web sites and their connections.....	19
Figure 2.7 A web page traffic report after the Panda update	24
Figure 2.8 Web page examples of penalized by Penguin algorithm.....	25
Figure 2.9 Impact of Hummingbird update on search rankings.....	26
Figure 2.10 Google local search after Pigeon update.....	28
Figure 2.11 Impact of Mobilegeddon update on top 10 ranking webpages.....	30
Figure 2.12 How RankBrain uses user experience signals	31
Figure 2.13 Sites affected by Medic update.....	32
Figure 2.14 Web page examples of affected by Medic update	33
Figure 2.15 Ahrefs SEO tool overview.....	35
Figure 2.16 Moz SEO tool overview	36
Figure 2.17 Semrush SEO tool overview.....	37
Figure 4.1 Domain overview of Ankara Yıldırım Beyazıt University website with Semrush.....	45
Figure 4.2 Organic research of Ankara Yıldırım Beyazıt University website.....	45
Figure 4.3 Domain overview of Ankara University website with Semrush	46
Figure 4.4 Domain overview of Oxford University website with Semrush.....	47
Figure 4.5 Traffic analysis comparison of university websites.....	48
Figure 4.6 Backlink analysis comparison of university websites	49
Figure 4.7 Domain overview of Ankara Yıldırım Beyazıt University website with Moz	50
Figure 4.8 Domain overview of Ankara University website with Moz.....	50
Figure 4.9 Domain overview of Oxford University website with Moz	51
Figure 4.10 Discovered and lost linking domains of Ankara Yıldırım Beyazıt University website.....	51
Figure 4.11 Discovered and lost linking domains of Ankara University website.....	52

Figure 4.12 Discovered and lost linking domains of Oxford University website.....	52
Figure 4.13 Technical SEO report of Ankara Yıldırım Beyazıt University website with Nibbler.....	53
Figure 4.14 Technical SEO report of Ankara University website with Nibbler	54
Figure 4.15 Technical SEO report of Oxford University website with Nibbler	55



CHAPTER 1

INTRODUCTION

1.1 Aim of the Thesis

With the widespread use of the Internet, the information on the Internet has increased exponentially. The increase in the accumulation of knowledge day by day has also been the reason for the emergence of search engines. With the advancement of technology, it has become important to reach information in the fastest and easiest way. The development and continuous updating of search engines has made it the first medium used by people to access information.

The increase in the size of information on the Internet, the development of search engines, has provided great convenience in reaching information. However, it has been discussed how the results obtained through search engines are related to the information the user is really looking for. As a solution to this problem, search engine optimization techniques have been developed to achieve the most relevant results for the queries users are looking for.

Within the scope of this study, the development process of search engines was examined, the types of search engines were mentioned, the optimization techniques developed for search engines were summarized and the stages they went through were explained one by one. The features of the algorithms developed to rank the results are specified and their advantages and disadvantages are compared with each other. The main purpose of the study is to contribute to the people who will do research on this subject when they want to develop a new search engine by providing information that can be basic. Another feature is that it is informative in that individuals or organizations can move their web pages to the top of the search results.

1.2 Research Questions and Outcomes

There are many studies about search engine optimization. Quick access to information, access to the most relevant and efficient resources, e-commerce, website design is always a critical issue when researching or applying. The subject of this study is one of the topics that can be updated every day, constantly changing and being developed. A detailed literature review on search engines, search engine optimization and optimization tools are thought to help me conduct my research more effectively and efficiently. The main questions that arise from the research problem and need to be answered in order to achieve my goals and objectives are:

1. How do search engines rank websites?
2. What are the techniques and processes used in search engine optimization?
3. What algorithms do search engines use?
4. What are the tools used in the SEO process? How to use these SEO tools to optimize websites?
5. How is the SEO performance of existing websites measured using SEO tools?
Can it be analyzed with competitor websites?

The study is focused on search engine indexing techniques and how search engines rank websites. It is aimed to provide an overview of which criteria are more important. It is planned to examine the search engine optimization techniques in depth, to investigate the contributions of these techniques to the websites and to analyze the tools that facilitate these processes. In the study, it is briefly explained how to apply SEO methods and how to achieve high web traffic.

1.3 Thesis Organization

This thesis consists of six chapters. The first part contains the aim and scope of the study. The remainder of the thesis is organized as follows:

Chapter 2 contains the theoretical and technical information in the study. Search engine types, search engine optimization techniques, search engine optimization

algorithms, algorithm updates carried out by Google and search engine optimization tools are detailed in this section.

Chapter 3, recent studies about research are examined in this chapter.

Chapter 4, optimization analyzes, and comparisons applied on different websites using the search engine optimization tools Semrush, Moz and Nibbler are included in this section.

Chapter 5, the findings related to the study are summarized and the results drawn from the analysis studies are explained in this chapter.

The final chapter, Chapter 6, contains the conclusions of this thesis and suggestions for future work.

CHAPTER 2

BACKGROUND

2.1 Search Engine

A search engine is a system that allows users to enter queries to get information or find a result and lists the results according to these queries. It enables people who want to bridge between people and data to obtain information over the internet.

The development process of search engines is shown in Figure 2.1. Many search engines have been developed since 1990. In 1998, the Google search engine entered our lives and became the most used search engine.

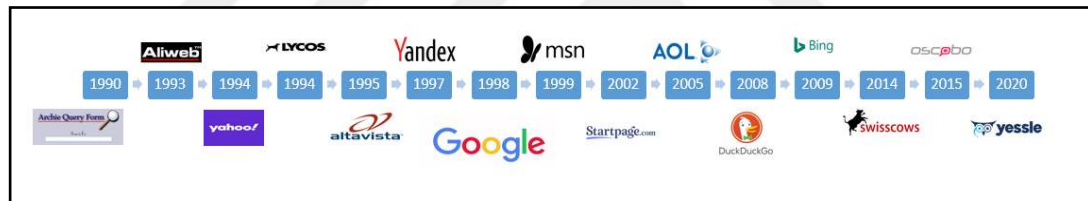


Figure 2.1 Search engine history

The popularity of the Google search engine continues today. The graph, which was created with the data obtained by Statcounter regarding the search engine usage rates in the period of August 2021 and comparing the usage rates of search engines in the world and in Turkey, is as in Figure 2.2 [1]. While the Google search engine usage rate is 74.57% in Turkey, it is by far the leader in the world with a rate of 92.03%.

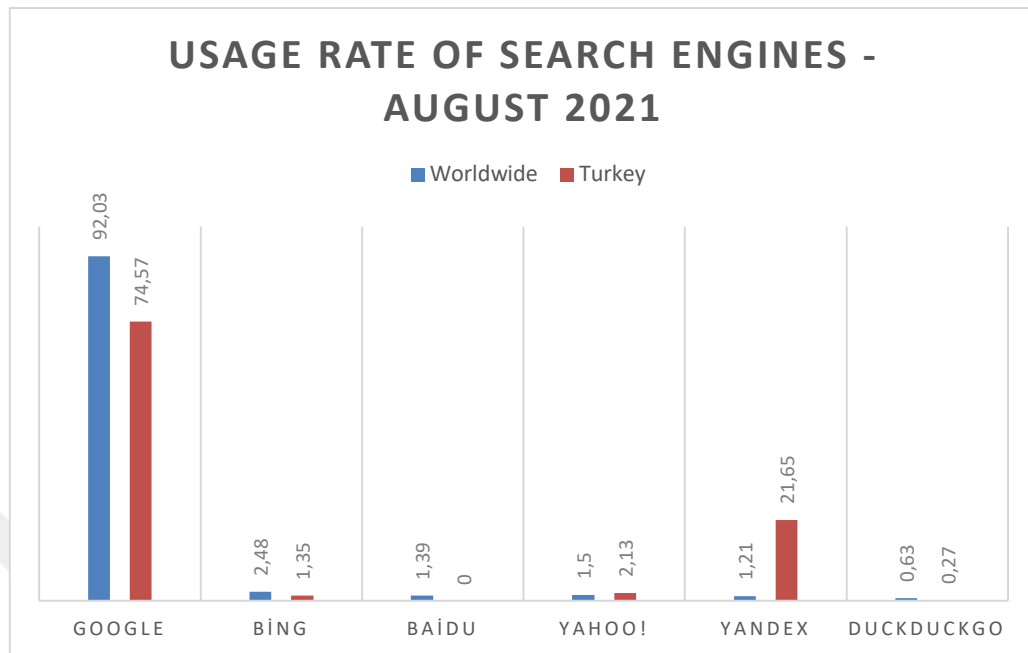


Figure 2.2 Usage rate of search engines [1]

Search engines, in a sense, are software that index web pages on the internet. For a search engine to access and index data on web pages on the internet, it must first reach all web pages on the internet. For this purpose, search engines using different approaches have been developed. In general, four basic architectures are used in search engines.

2.1.1 Crawler-based Search Engines

The most common type of search engine is crawler-based search engines. In crawler-based search engines, there are programs called spiders or crawlers in the first place. These programs periodically surf the internet, read the content of the website and list all the links in a central database. The web spider crawls web pages on the internet and users search the data generated by the web spider. These lists are catalogs made by the search engine's indexer. These indices are like a set that holds web page contents found by crawling by the web spider. The web spider works with a working principle that constantly visits the pages on the internet and the links within these pages [2].

Unchanging methods are used within the framework of the principle of continuous operation. Therefore, if any changes are made to the web page, the priority of that web page in the search results will be affected. In addition to these, it should be noted that for the visited web page to be indexed, the page must be visited by the search spider, and this process can take place within a certain period. It is not possible for the search engine to remember that web page and show it among the results until it is indexed. Google, Yahoo and Bing are examples of crawler-based search engines.

2.1.2 Directory-based Search Engines

Another name is “Human-powered Search Engines”. A human decides whether the web page to be indexed will be useful or not and adds it to the system indexes. It is preferred for non-text content such as images and videos. It is advantageous in terms of search result quality, as each page is reviewed for relevance and content before being indexed. However, it is disadvantageous in terms of time spent. It is used by many search engines such as Open Directory and LookSmart.

2.1.3 Hybrid Search Engines

A hybrid search engine is a type of search engine that uses more than one search engine structure together. It allows the search results to be viewed from records obtained using both crawler-based and directory-based search engine features. The search structure is not different from other search engines. Search is made according to the keyword received from the user. MSN Search is an example of a hybrid search engine. It provides both crawler-based results and human-powered listings.

2.1.4 Vertical Search Engines

Vertical search engines, when a search is desired on a certain topic, it extracts the results that are not related to the topic and ensures that completely relevant search results are displayed. By indexing data specific to topics within the scope of the search engine, it displays web pages that contain the information that users are looking for. Users only want to view the pages that contain the information they are looking for,

but horizontal search engines are not very successful at this point. Vertical search engines have been developed as a solution to this problem. Kosmix, Scirus, Libra, MALDO, Trovit, and Medstroy are examples of these search engines.

2.2 Search Engine Optimization Techniques

Search engine optimization (SEO) is the activity of optimizing web pages or entire sites to make them search engine friendly and thus rank higher in search results. There are different methods for search engines to access effective information.

SEO techniques are some basic strategies and ethical steps to consider when developing a website. However, this process is a long-term process that continues if the website exists. SEO techniques allow search engines to tell what a website is about. It also allows the targeted user to visit the website. It is possible to work from many different angles for optimizing search engines. As a basic process, on-page and off-page optimizations are made as seen in Figure 2.3. Then, the transition to the technical optimization stages is made.

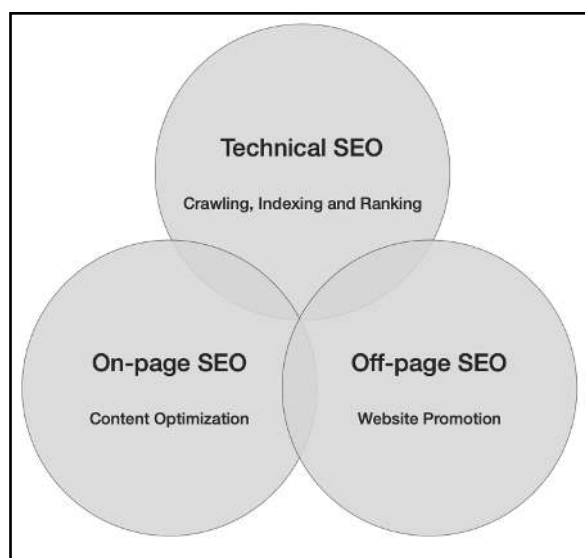


Figure 2.3 SEO techniques

2.2.1 On-Page SEO

On-page SEO is the process of making the website discoverable by the search engine. It is also called on-site optimization. It is about what is done during the design and development process of the website with the aim of improving SERPs. It has many factors such as keywords, meta tags, headers, URL structure [3].

2.2.1.1 Page Title

This section is an important place where "meta" information about a web page is found and affects ranking. The title tag, located in the <head> section of the HTML document, represents the main keywords and subject of the web page. The following syntax is an example of the <title> tag in HTML.

```
<title> Your web site Title </title>
```

The page title appears in the title bar in the web browser as well as in the title of the search results. The ideal number of characters is in the range of 50-80. It is very important to increase the click through rate (CTR) of any ad or website.

2.2.1.2 Meta Description Tag

The meta description tag contains concise information about the content of a web page. It is the section displayed after the title on the search engine results page (SERP). It is important because it can be used as a method in terms of being preferred by search engine visitors and conveying the marketing message. The meta description tag is in the <head> section and comes after the <title> tag. The syntax is as follows:

```
<meta name= "description" content= "your website description">
```

Important keywords can be used in the meta description tag to increase the click through rate (CTR) of a website. The ideal number of characters is in the range of 150-200.

2.2.1.3 *Meta Keywords Tag*

The meta keywords tag is the tag that indicates which keyword the website or content is focusing on. It allows crawler-based search engines to index it. The most important point when writing meta keywords tag is to write the most relevant keywords for search engines first.

`<meta name= "keywords" content= "list of keywords separated by commas (,)">`

2.2.1.4 *Meta Robots Tag*

Robots is a tag that tells search engines whether the page or the links in its content are requested to be indexed by search engine spiders and crawlers. Normally there are two parameters in the robots tag and the syntax is as follows.

`<meta name = "robots" content= "none | noindex | nofollow">`

The following values are optionally added to the page:

Index: It means the page is allowed to be indexed.

Noindex: It means the page is not allowed to be indexed

Follow: It means that it is desired to follow the linked links within the page.

Nofollow: It means that the linked links within the page are desired to be followed.

Archive: It means that a screenshot of the page will be created for preview.

Noarchive: Means that the page will not be screenshotted for preview.

None: Indicates that the page is requested to be ignored.

2.2.1.5 Robots.txt File

The robots.txt file is the file that the spiders send to the search engine which addresses should be indexed and not indexed. Thus, irrelevant links are not scanned, and page not found errors that may occur are avoided. The meta tag option tells bots not to index individual files, while robots.txt can be used to restrict access to all directories.

Before the web browser crawls a website, it finds and reads a robots.txt file for that site. The robots.txt file is in the root directory of the website and contains information about how the website it belongs to should be crawled by the search engine. It consists of one or more rules. Each rule blocks or allows the browser to access a particular file path on that website.

The syntax of a simple robots.txt file is as follows.

```
User-agent: Googlebot
Disallow: /seo/

User-agent: *
Allow: /

Sitemap: http://www.example.com/files/sitemap.xml
```

The following keywords are added to the file depending on the content of the rule:

User-agent: The public name of the client and it is mandatory. It must be added to the beginning of each rule. The asterisk (*) makes it a general rule for all crawlers.

Disallow: Specifies the page or directory information that is not allowed to be scanned.

Allow: Specifies the page or directory information that is allowed to be scanned.

Sitemap: It is the location of the sitemap of the website.

2.2.1.6 Heading Tag

Heading tags are very important because they are informative about the subject of the web page. The words to be searched in search engines must be included in the heading

and subheading. Targeted keywords should be selected and placed in heading tags. There are heading tags according to importance from H1 to H6. The H1 tag has the most importance, while the H6 tag has the least importance. The syntax for the heading tag is as follows:

```
<h1>Heading Title</h1>  
<h2>First subheading</h2>
```

2.2.1.7 ALT Tag

Images or graphics cannot be indexed by search engines. An ALT tag (alternative tag) is used to mark visual elements (images or graphics) on a web page. It is very important to use ALT tags in image search processes for the web page to be found or to appear in the top ranks. It makes it significantly easier to find the web page by the search engine. The text inside the ALT tag is actually a short description of the unloaded image. The syntax is as follows:

```
<img src= "pic.jpg" width= "100" height= "78" alt= "keyword text"/>
```

2.2.1.8 Sitemap

A sitemap helps the search engine index all web page URLs on a website and other files such as PDF, TXT or DOC files. It is a list of URLs used on the website. It is the shortest and most effective way to tell the search engine what pages the website contains. Sitemaps make it easier for search engines to crawl websites and produce more appropriate results. In general, the sitemap is created in two different ways, either as an HTML sitemap or as XML sitemap. If websites are large and dynamic, sitemaps become more necessary.

HTML sitemaps are known as traditional sitemaps. They used to be designed to make it easier for users to access important web pages.

XML sitemap developed by Google is the most widely used sitemap method today. The XML sitemap was designed in response to the indexing problems of large and complex structured websites. Yahoo, Bing, Ask and AOL search engines have also agreed to support the XML protocol over time. Thus, a single sitemap was created for all search engines, providing a great advantage.

An XML Sitemap has 4 core elements for each link but can be expanded:

- URL location: The absolute URL.
- Last modified: When the page was last updated.
- Change frequency: How often the page is updated (never, yearly, monthly, weekly, daily, hourly, always).
- Page priority: The importance of this page on the website. A number from 0.0, 0.1, 0.2 to 1.0.

A sample code for sitemap.xml is given below:

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>http://www.example.com/foo.html</loc>
    <lastmod>2021-06-12</lastmod>
    <changefreq>weekly</changefreq>
    <priority>0.7</priority>
  </url>
</urlset>
```

2.2.1.9 URL Structure

Search engine crawlers use URLs to find, access, and rank site content for search engine results. This means that URLs play an important role in SEO. Developing URLs simple and static is considered the most ideal approach. Because readable URLs

reflect a website's understanding. If the URL contains keywords based on the page content and is easily crawlable, it will rank higher in search engine rankings.

Search Engine Friendly (SEF) URLs are URLs designed to meet the needs of users and searchers. Specifically, SEO-optimized URLs are short and keyword-rich. The URL structure is as simple as possible, and the content is understandable. Static and readable URLs can be crawled and tracked by search engine spiders, making them easier to navigate.

URLs that are difficult to understand are often referred to as dirty URLs or URLs that are not Search Engine Friendly. A complex and illogical URL structure forces the browser to search for hard-to-discover information. Dynamic URLs containing special characters cannot be crawled by search engine spiders.

The rules for a URL to be SEF are listed below.

- The URL should be relevant to the page content. The user should have an idea about the page content from the URL information.
- The URL should be as short and descriptive as possible.
- It should contain keywords, but excessive use of keywords should be avoided.
- A hyphen (-) can be used as needed to separate words. Do not use underscores (_), spaces or other special characters.

SEF URL: *"https://domain.com/url-optimization-tips"*

non-SEF URL: *"https://domain.com/url_optimization_tips"*

- Lower case should be preferred. Capital letters should not be used. In some cases, capital letters cause duplicate URL problem and loss of ranking.

SEF URL: *"https://domain.com/url-optimization-tips"*

non-SEF URL: *"https://domain.com/URL-Optimization-Tips"*

- Whenever possible, it is better to use static, unchangeable URLs. The use of special characters (such as &, %, ?), which are common features of dynamic

URLs, should be avoided. Static URLs are easier to read for both users and search engines.

SEF URL: <code>"https://domain.com/url-optimization-tips"</code> non-SEF URL: <code>"https://www.domain.com/bags.php?productid=1&colorid=2"</code>

- Https protocol should be used.

SEF URL: <code>"https://domain.com/url-optimization-tips"</code> non-SEF URL: <code>"http://domain.com/url-optimization-tips"</code>

- Attention should be paid to the use of subdomains. It should be used when necessary and its scope should be well determined when choosing. Subdomains appear before the domain name. They are often used to separate different areas of a website based on their functions. An example URL for correct subdomain usage:

<code>"https://blog.domain.com/"</code>

- Unless you have a specific reason to use a subdomain, it's best to use subfolders instead. An example URL for subfolder usage:

<code>"https://domain.com/category/subcategory"</code>
--

2.2.1.10 SSL/TLS Certificate

SSL (Secure Socket Layers) is an encryption protocol that provides an encrypted connection between a client and a server. This encryption ensures that data between the web server and internet browsers is unreadable by a third party.

TLS (Transport Layer Security) is a newer and improved encryption protocol that uses stronger encryption algorithms than SSL and offers unmatched privacy performance.

HTTPS is the secure extension of HTTP. Websites that have installed TLS/SSL certificates use the HTTPS protocol to establish a secure connection with the server. The purpose of SSL/TLS technology is to securely transfer sensitive information such as personal data, payment or login details.

Data security is an important concern for users and search engines. Google uses SSL as a ranking factor in its PageRank algorithm, one of its most important algorithms. In this case, all other things being equal, a website with SSL will rank higher in Google search engine results than a website without SSL [4].

2.2.2 Off-Page SEO

Off-page SEO is all kinds of promotional and sharing activities that are not related to the content of the website and are made outside the site. It is applied after on-page optimization techniques. Off-page optimization techniques are not completed in one go, they are applied gradually and aim to provide better and long-term results. Some off-page optimization techniques are available for thorough crawling and indexing of dynamic websites. It has factors such as external links and social media use [5].

2.2.2.1 Directory Submission

Submitting website links to directories is a good way to get efficient traffic. These directories break down the submitted link into different categories and sub-categories according to its relevance. It's a good method to get traffic from directories when the website is newly created or doesn't have many links.

2.2.2.2 Link Building

Link building is one of the most important off-page SEO techniques. A website often has external links that provide access by other websites. The higher the quality of these linking websites, the more traffic they bring and positively affect search engine rankings. Links from high-ranking websites are more important than how many links come to a website. Websites that receive well-known links are positively affected in terms of ranking, while bad and poor-quality links can reduce the website's ranking. Quality inbound and outbound links are much more useful and successful. Outgoing and incoming connection types are configured gradually, and this process can take a long time.

Link building can be done in two different ways, one-way and two-way link building such as Figure 2.4. A one-way link is a link that leaves the target web page but does not link back to that web page; the reverse is called a two-way link. Both types of links are important in off-page SEO strategies.

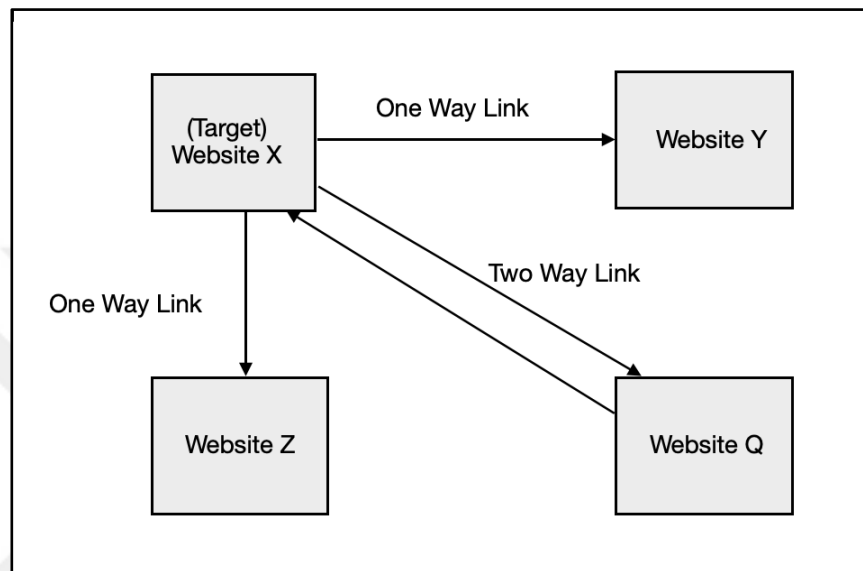


Figure 2.4 One-way and two-way link building

Dead Link

Dead links are external links that no longer exist on websites. Another name is broken links. Clicking on these hyperlinks displays an error message. Dead links reduce the performance, ranking and visitors of any website.

A few situations that can cause dead links to appear are listed as follows:

- Reorganization of the website.
- Changing the Internet Service Provider.
- Typos in the link.
- Grammar errors in the link.
- Deletion of the web page.

There are many tools available for detecting and repairing dead links. Dead links can be fixed on the website without any loss of information. Thus, the number of visitors to the website increases and it rises in the search rankings. Sitechecker [6], Semrush [7], W3C Link Checker [8] and Dead Link Checker [9] are important online tools that detect dead links. By entering the URL information of the website or a single page, all broken links in its content are analyzed. It generates a report that includes URL information with broken link, the reason for broken link, and suggestions to solve the problem.

2.2.3 SEO Technical Stages

In the more technical part of search engine optimization, there are three different stages. These; web crawling, indexing and ranking operations (Figure 2.5).

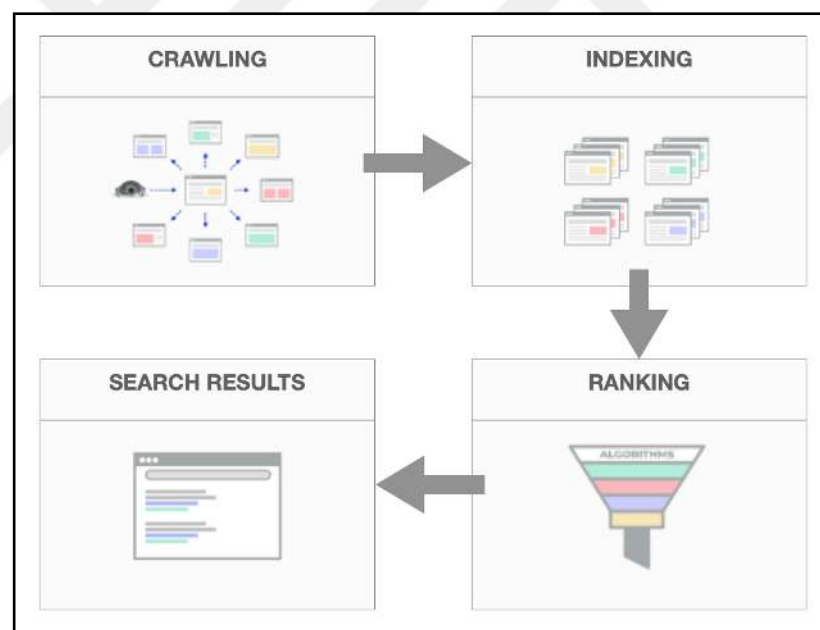


Figure 2.5 SEO technical stages

2.2.3.1 Web Crawling

It is the discovery process of search engines through browsers or software called spiders in order to find new and updated content. Images, videos, articles, files, music,

etc. It scans and collects information in many different formats and structures under different directories [10].

The browser fetches pages from a website. Then they reach new websites by following the links on this page. In this way, it creates a pathway for accessing the trillions of interconnected website pages that exist. They start with the trusted website and expand their browsing on the web by following the links they see on the pages they visit.

2.2.3.2 Indexing

It is the process of adding information about a web page to a search engine's index. A directory is a database containing information about pages crawled by search engine spiders.

Before search engines decide which web pages from the index to show in what order, they apply algorithms to help them rank those web pages.

2.2.3.3 Ranking

It is the process of ranking search results according to their relevance. It tries to find the most relevant answer, considering many factors such as the user's location, language, and device. Many algorithms have been developed for the ranking process.

2.3 Search Engine Optimization Algorithms

2.3.1 PageRank Algorithm

It is the algorithm that is the reason for the emergence of the Google search engine. It is based on a method of measuring the importance of a page. It is the numerical scoring algorithm given to the pages by considering certain factors [11]. These scores given to the pages are taken as a basis by the search engines in measuring the relevance of the pages with the search query and ranking the sites according to their quality. Using PageRank instead of all other factors such as title tag and keywords gives more efficient results.

Web pages with a high page rank value rank better in search results and gain reputation. If this value is low, it means that the page is not prominent enough. There are multiple factors that determine the PageRank value. These:

- If the number of inbound links to a web page is high, the PageRank score of that web page will also be higher.
- If inbound links to a web page have a high PageRank score, that web page will also have a high PageRank score.
- However, if an inbound link to a web page has a high PageRank score and many outbound links at the same time, the PageRank score for that web page decreases.

Figure 2.6 shows how the PageRank score of web pages from A to G is calculated. Each arrow in this figure represents a hyperlink from one web page to another.

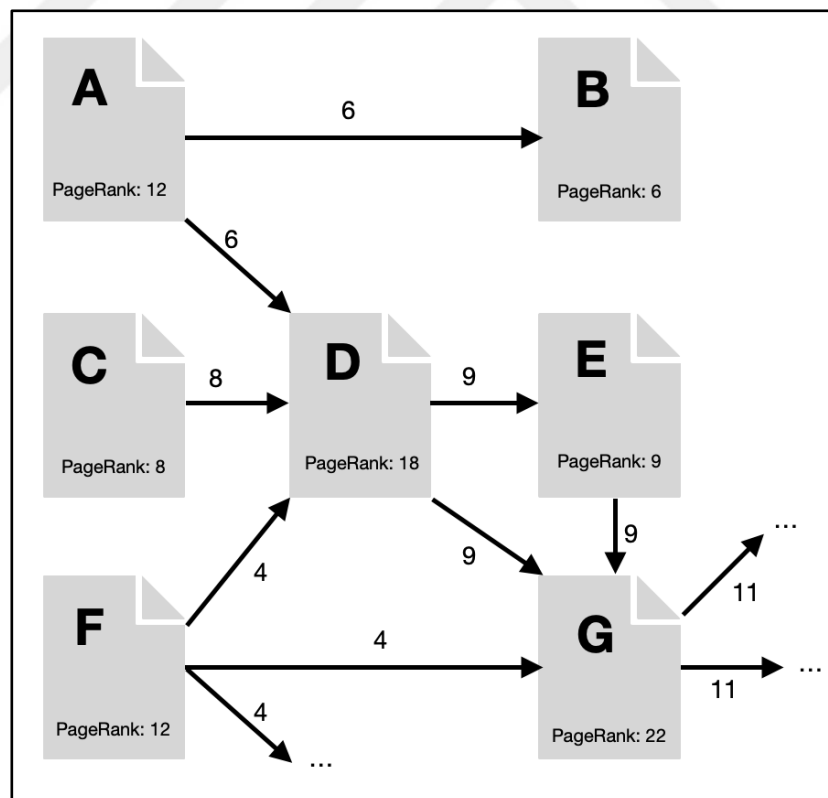


Figure 2.6 PageRank scores of web sites and their connections

The original PageRank algorithm was developed by Sergey Brin and Larry Page [12]. The formula of PageRank Algorithm is given in Equation 2.1.

$$PR(A) = (1 - d) + d \left(\frac{PR(T1)}{C(T1)} + \dots + \frac{PR(Tn)}{C(Tn)} \right) \quad (2.1)$$

d: A damping factor that can be adjusted around 0 and 1,

PR(A): PageRank of page A,

PR(Tn): PageRank of page Tn which connects to page A,

C(Tn): Number of outgoing links on page Tn

2.3.2 TrustRank Algorithm

TrustRank serves to distinguish between trusted sites and spam sites. It is a semi-automatic process. The sites that are considered the most reliable by experts are manually marked and these sites are called root sites. It continues by calculating the TrustRank values of the websites to which the links inside the root sites are pointing. The further away from these trusted root sites, the lower the TrustRank's reliability. The opposite of this situation is also possible and is called Anti-TrustRank. The closer a site is to sources of spam, the more likely it is to be spam [13].

The TrustRank algorithm and methodology were developed by Zoltan Gyongyi, Hector Garcia-Molina, and Jan Pedersen. The basic equation of the TrustRank Algorithm is as shown in Equation 2.2 [14].

$$t = \alpha \times T \times t + (1 - \alpha) \times d \quad (2.2)$$

t: The TrustRank score vector,

α : The decay factor,

T: The transition matrix (in which T (i, j) is the probability of following the link from page j to page i),

d: The normalized trust score vector for the seed set

Many factors are effective when calculating the TrustRank value. A few important factors can be listed as follows.

- Age of the page
- PageRank of backlinks to the page
- Age and quality of backlinks to the page
- Keeping the page up to date and consistent
- Not too much blocked by users
- Security certificate

2.3.3 HillTop Algorithm

It works in conjunction with the PageRank algorithm. Its distinguishing point from other algorithms is that it identifies website pages by deciding which of the linked sites is the expert. By creating quality backlinks, it ensures that the most relevant pages rank higher [15].

The comparison of PageRank, TrustRank and HillTop algorithms is shown in Table 2.1.

Table 2.1 Comparison of PageRank, TrustRank and HillTop Algorithms

	PageRank	TrustRank	HillTop
Connection Source	Good pages have many important links.	Good pages only indicate good links.	Only expert pages point to good links.
Initial Scoring	The average is taken.	1 or 0	The algorithm is applied.

2.4 Google Algorithm Updates

Google algorithms are a complex system used to deliver the best possible results for any query in the shortest amount of time. It uses many algorithms and ranking factors to present web pages ranked by relevance in search engine results pages.

While Google initially offered only a few updates to its algorithms, it now makes countless changes every year. However, most of these updates do not attract attention because they are light. In addition, Google is also making major algorithmic updates. Panda, Penguin, Hummingbird, Pigeon, Mobilegeddon, RankBrain and Medic updates are compared in Table 2.2.

Table 2.2 Comparison of Google developed algorithm updates

	 Panda (2011)	 Penguin (2012)	 Hummingbird (2013)	 Pigeon (2014)	 Mobilegeddon (2015)	 RankBrain (2015)	 Medic (2018)
Scope	Content of websites	Backlinks	Question phrases instead of keywords	Local search	Mobile versions	Based on machine learning	E-A-T of pages
Successful Aspects	Quality content	Natural website links	Use of natural language processing Voice call	Optimized website for local search	The website is mobile compatible	Quality content Quality user experience	High E-A-T
Weaknesses	Copy, Plagiarism Thin content	Spammy links Irrelevant links	Low quality content Keyword stuffing	Poorly on and off-page SEO	Non-mobile-friendly web pages	Weak Content Poor user experience	Low E-A-T
Impact On Search Queries	11.8%	3%	90%	50%	14%	3-5%	Up to 50% impact on health sites

2.4.1 Google Panda Update

The goal behind the Google Panda update was to highlight high-quality content and filter out sites that may be of lower quality. Low-quality pages are those created by people who do not have enough of the content the user wants to reach or who are not experts in the content. If the site has copied pages or has plagiarized content, it is filtered by the Panda algorithm. Panda also aims to filter out duplicate content.

A lot of websites were affected after the Google Panda update. All websites, including large-scale sites, suffered from SERP and thus SEO visibility. While the Panda update did not work site-wide, certain URLs were affected, meaning that some sites' landing pages lost visibility.

The Panda update addressed several issues occurring in the SERPs [16]. Some of these are:

Poor content: Web pages with irrelevant and fixed text, each page consisting of only a few sentences.

Duplicate contents: Content that is copied and created on many pages when browsing the Internet. They are web pages that are usually created by copying the same text or with little modification. This problem is also encountered between the pages of a website itself.

Low quality content: Site content that cannot or has little contribution to its readers due to the lack of information it contains.

Lack of authority/trustworthiness: Web pages created by unverified sources, whose owner is unknown.

High ad-to-content ratio: Pages that contain more paid ads than information.

Sites affected by the Google Panda update face penalties. If a site is penalized for this update, it first experiences a steady decline in search traffic and then a stabilization [17]. If any URL of a website is penalized by Panda, the entire site is also affected.

Google provides the traffic report in Figure 2.7 as an example of a web page penalized by the Panda update.

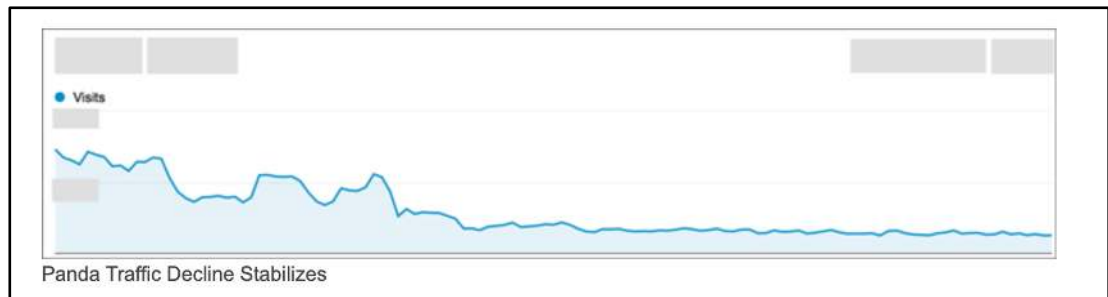


Figure 2.7 A web page traffic report after the Panda update [17]

2.4.2 Google Penguin Update

In 2012, Google released the Penguin update to its ranking algorithm to optimize search results [18]. With its launch, it made a huge impact on many websites. The purpose of the Penguin update is that it only deals with inbound links to a website, i.e., backlinks. It is not concerned with which sites the link goes to, but which sites the links come from. It filters out unnatural, spam websites or irrelevant links. Penguin has been revamped several times since its introduction and became part of Google's core algorithm.

Google's stance is that every external link to a website should be natural and should be given because its content is indeed of good quality. When Google's Penguin algorithm detects an attempt to rank through unnatural link building, it penalizes the site for the targeted keyword and/or keyword phrase [19]. Methods such as paid links, link rental, link exchange negotiations, automatic link creation with SEO tools and link request are considered as web spam by Google. Rapid link growth is also perceived by Google as web spam. Google provides the traffic report in Figure 2.8 as an example of a web page penalized by the Penguin update.

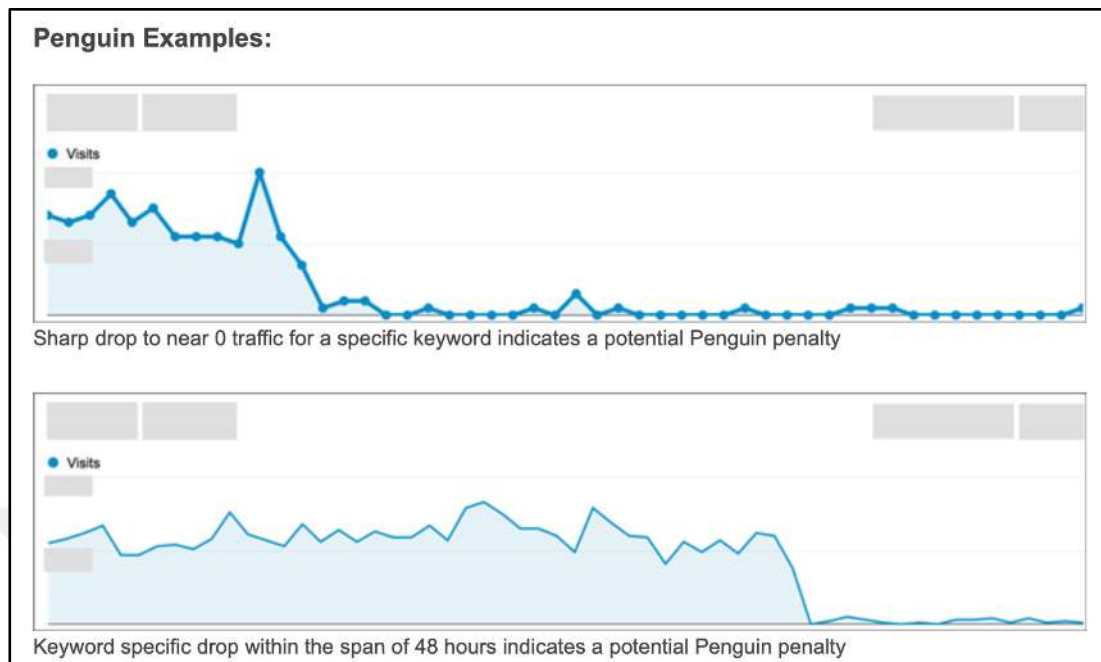


Figure 2.8 Web page examples of penalized by Penguin algorithm [20]

One of the important differences for the Penguin update is that it is web page specific. While a page from the same website may be penalized and drop in search rankings, other pages may have good ranking performance.

While the Penguin algorithm update creates a negative situation for web pages with unnatural links, it provides an advantage by providing a rapid rise in search rankings in terms of web pages with quality and natural links.

2.4.3 Google Hummingbird Update

Developed by Google in 2013, the Hummingbird update is revolutionary as it helps people make sense of the words typed into the query field. This developed algorithm not only matches the words in the query with the words in the search results, but also understands what is being asked and matches it with the most related results. The Hummingbird update went into effect about a month before it was announced and globally affected 90 percent of search queries. Figure 2.9 highlights the fluctuations in search rankings prior to the update's announcement date of August 21.

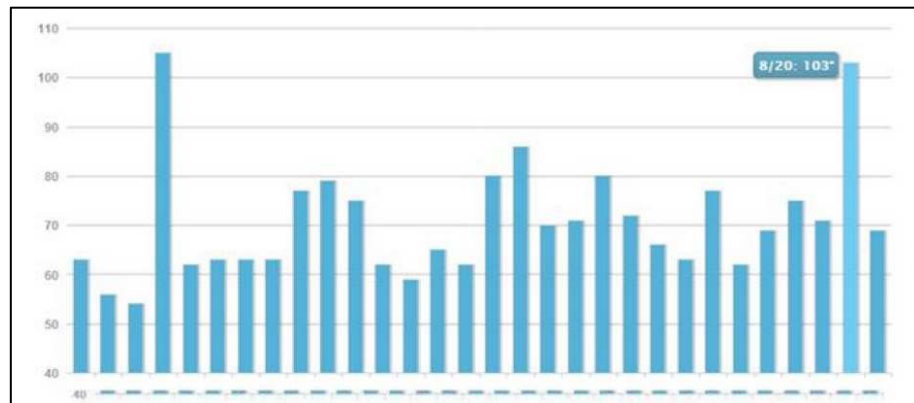


Figure 2.9 Impact of Hummingbird update on search rankings [21]

Hummingbird is focused on a few key features.

Semantic Search: The most difficult point since the use of search engines is that it is very difficult for the user to reach the desired result if they do not have any idea about the searched subject. At the core of the Hummingbird application lies the concept of meaning. Semantic search refers to the concept of improving search results by focusing on user intent and how a search topic relates to other information in a broader sense or relevance to the content [22]. Instead of a set of keywords, semantic search focuses on determining what a user's search word really means, and then returning relevant results. Hummingbird tried to solve this situation by focusing on synonyms and thematic issues. Hummingbird lets people know what they don't know and select results that will help users find what they're looking for. Using natural language processing to make sense of people's queries in sentences and producing results shows that Google is one step closer to being human.

Conversational Search: With the Hummingbird update, Google shows that it has taken a step towards dominating the inevitable rise of voice search [23]. An update like semantic search and Hummingbird is designed to work effectively in the natural language world. A few years after the Hummingbird update was released, the voice search feature emerged. Considering that the voice search feature is on the rise, we can say that the semantic search feature that comes with Hummingbird is effective at this point. Because searches are now made in the form of voice search and with a meaning.

With Hummingbird, it is aimed to give the most appropriate results for these searches. In fact, Google stressed that the Hummingbird algorithm is important precisely because users expect more natural and conversational interactions with a search engine.

2.4.4 Google Pigeon Update

The Google Pigeon update has been revolutionary for local search algorithms. The way local companies rank in organic search has changed with this update. For the first time in 2014, Google updated and even redesigned its algorithm with Pigeon to focus on local results [24]. It made major changes to its core algorithms. In this way, Google tried to combine local search results with web results and was successful.

Pigeon's goal was to provide the most relevant and useful results when a user searches, based on both the relevance of the searched word and the radius of that user's location. Pigeon algorithm works by eliminating results that are duplicates between organic and combined web page results [25]. It affects local searches and location-based searches. It enables to find exactly the information sought according to the location. It aims to provide more relevant, useful and accurate search results for local search queries. For example, when searching for a query like "Yenimahalle pizza", a user in the Batıkent area of Yenimahalle will get local results downgraded to that neighborhood instead of citywide results. It provides many advantages. Some of those:

- With this update, Google has increased its capacity to see typos that occur during the search process. In this way, in the search process made with typos, there was no problem in being found in the SERPs [26].
- Synonym recognition capacity is also increasing with this update.
- Local websites rank better in search results, especially for restaurants.
- Local businesses get an increase in web traffic and, as a result, potential new customers.
- With this update, Google has also updated its distance measurement tools and maps, resulting in more accurate results for local searches than ever before.

- It provides a superior search result for users searching with Google places or Google maps platform.

With the Pigeon algorithm, local search is allowed to resemble traditional organic search as much as possible, while also acting on a location-based basis. In other words, when a user searches for a word without specifying a location, results are generated based on that user's location. It offers a user experience as shown in Figure 2.10.

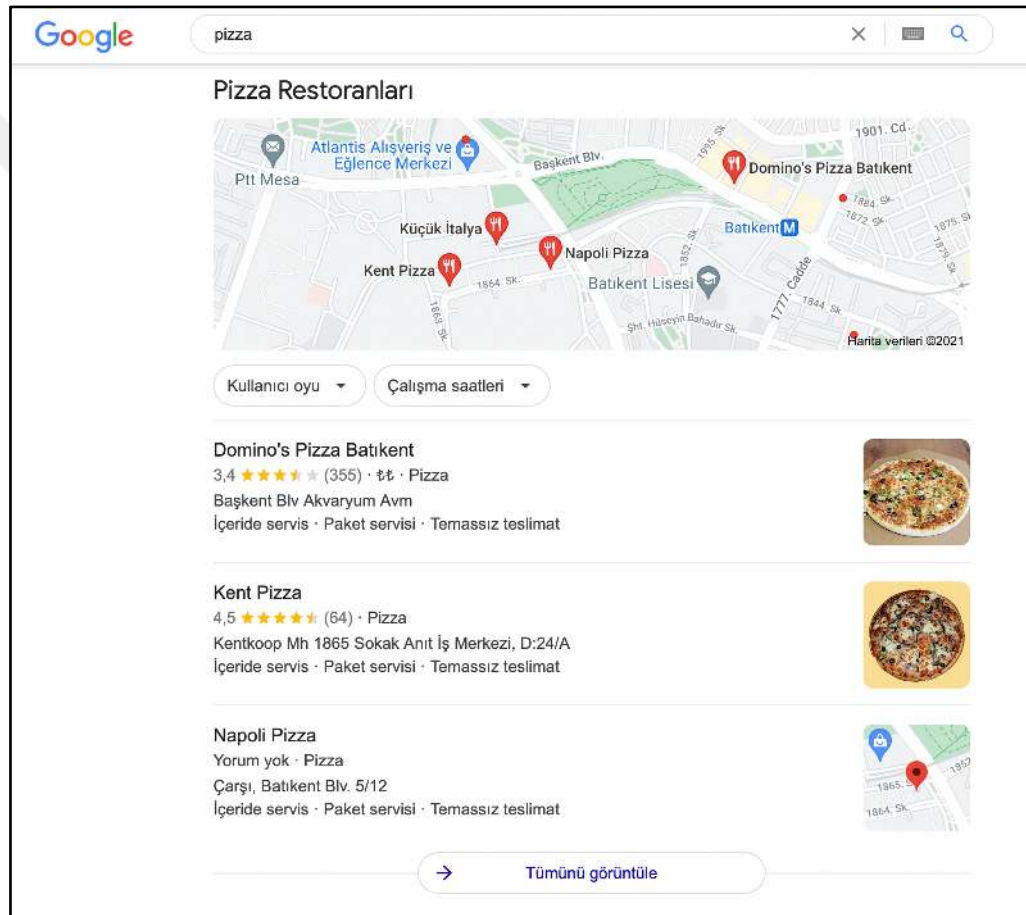


Figure 2.10 Google local search after Pigeon update

2.4.5 Google Mobilegeddon Update

With the Mobilegeddon algorithm update announced in 2015, Google made mobile compatibility a ranking factor for search results for the first time. With this update, Google shifted websites from desktop to mobile version and forced all sites to be

mobile friendly. Today, Google ranks websites based on how user-friendly they are and how fast their mobile versions are. When Googlebot crawls a site to evaluate its mobile friendliness, it checks [27]:

- Does the website avoid using software that is not common on mobile devices?
- Is the text readable without zooming?
- Is the content sized appropriately? (So, users don't have to pan or zoom horizontally.)
- Are the connections placed wide and far enough? (So, they can be easily touched.)

The main impact of this update was to prioritize websites that display well on smartphones and other mobile devices. The change did not affect calls made from a desktop or laptop computer [28].

Since only mobile search results will be affected by the content of the algorithm update, it was assumed that the impact would be low. However, it was assumed that what Google really wanted was to draw attention to "mobile optimization". General features of Google mobile update [29]:

- Mobile compatibility is becoming a ranking factor.
- The algorithm is URL-based and not an entire domain.
- It affects the results of a search made with a smartphone or mobile compatible device.
- Google automatically recognizes if a page is mobile-friendly.

With the update, many non-mobile-friendly sites lost their rank in searches on mobile devices. According to a study conducted by Stone Temple Consulting, the effect of this update on the top 10 websites is shown in Figure 2.11 [30].

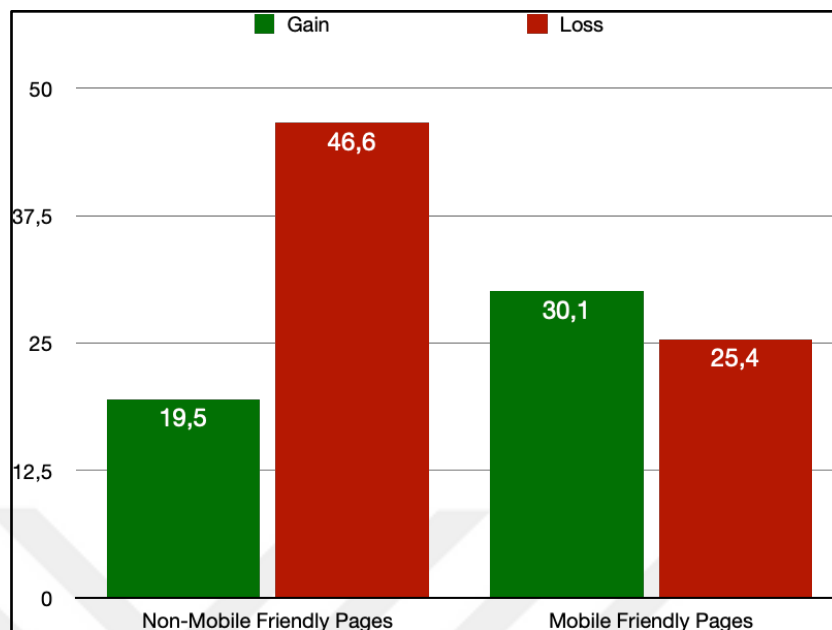


Figure 2.11 Impact of Mobilegeddon update on top 10 ranking webpages

2.4.6 Google RankBrain Update

The RankBrain algorithm is a major update that uses artificial intelligence and machine learning to rank search results. Thanks to this algorithm, Google processes and makes sense of search queries. Prior to this update, the entire Google algorithm was hand coded. Thanks to RankBrain, the algorithm is self-adjusting. The algorithm works as follows. First, it looks at how people interact with the search results, learning the meaning behind the query and the information people are looking for. Afterwards, if the search results are liked by the users, the algorithm will continue to be used, if not, it will be withdrawn [31]. How RankBrain uses user experience signals is illustrated in Figure 2.12.

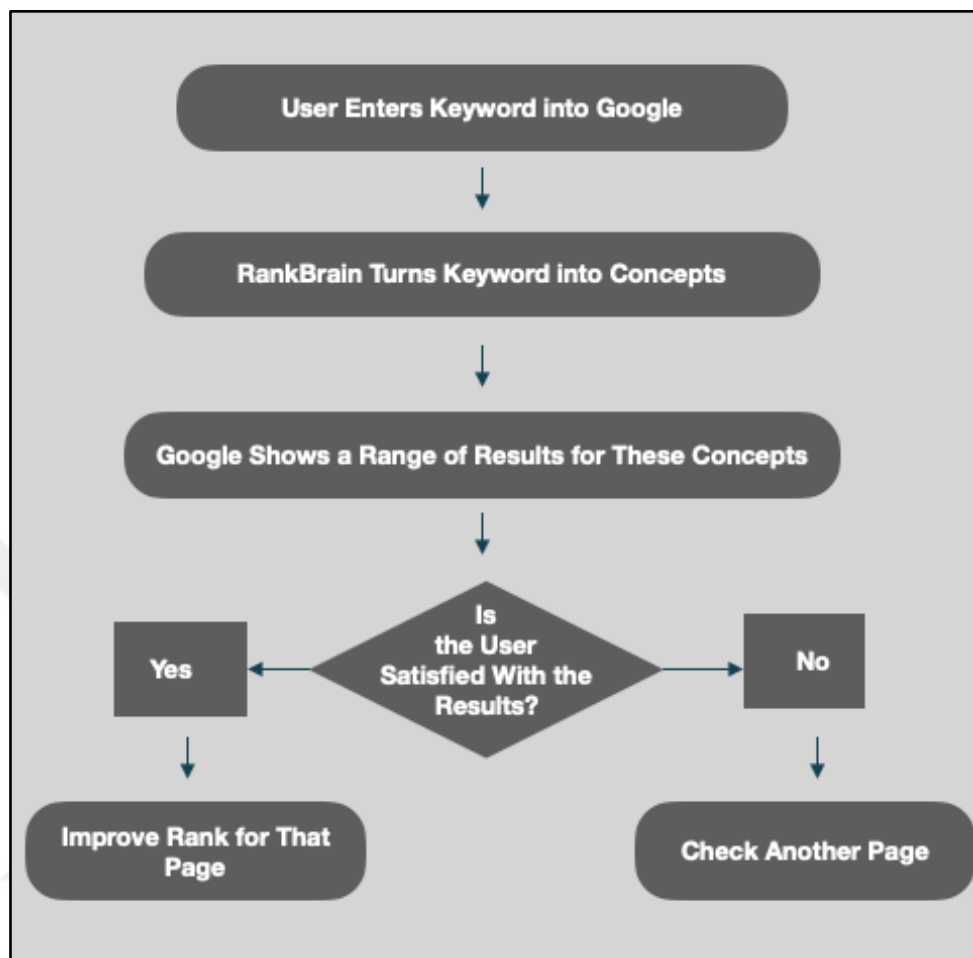


Figure 2.12 How RankBrain uses user experience signals

This technology was developed to help Google process about 15% of queries entered every day that the search engine has never seen before [32]. RankBrain has become the third most important ranking signal since its inception and is now included in most queries to Google.

Although RankBrain was developed to focus on unique queries that were not previously queried, Google has stated that this part of the algorithm is now an indispensable part of organic search.

2.4.7 Google Medic Update

The Google Medic update is a Google algorithm update released in late 2018. While the update affected many industries, the Your Money, Your Life (YMYL) web pages were the most affected [33].

"Your Money, Your Life" pages are classified as pages that could potentially affect users' health, happiness, or wealth. These pages are subject to higher scrutiny by Google search crawlers. Some of the pages that are considered YMYL pages are: news and current affairs sites, citizenship, government and law sites, financial sites, e-commerce sites, health and safety related sites, public/official sites, etc. [34].

This update was named "Google Medic" because more than 42 percent of affected sites are in the fields of medicine, health, fitness or healthy lifestyle (Figure 2.13) [35].

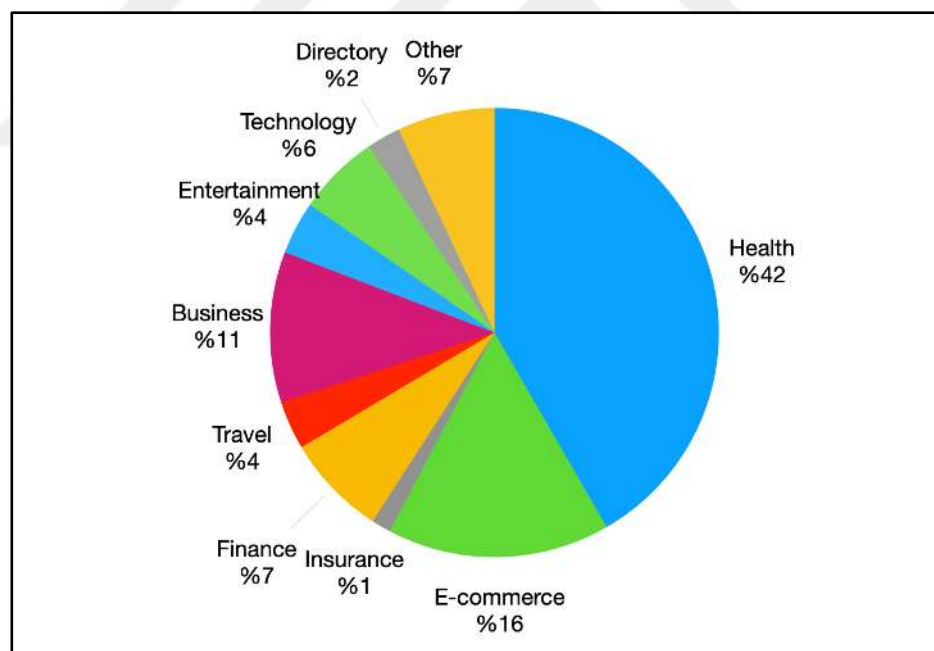


Figure 2.13 Sites affected by Medic update

The Medic update focused on the expertise, authority and trust (E-A-T) of pages that have the power to influence users' happiness. It aimed for websites to highlight and enhance these features. It penalized websites with no or low expertise, authority, or

trust signals (E-A-T), and rewarded websites with high E-A-T. The change in traffic of four different sites after the Medic update is shown in Figure 2.14. It is clearly observed that the first two sites have high E-A-T, and the other two sites have low E-A-T. It is also analyzed that the visit rates of sites with low E-A-T values after the update decreased by 75%.

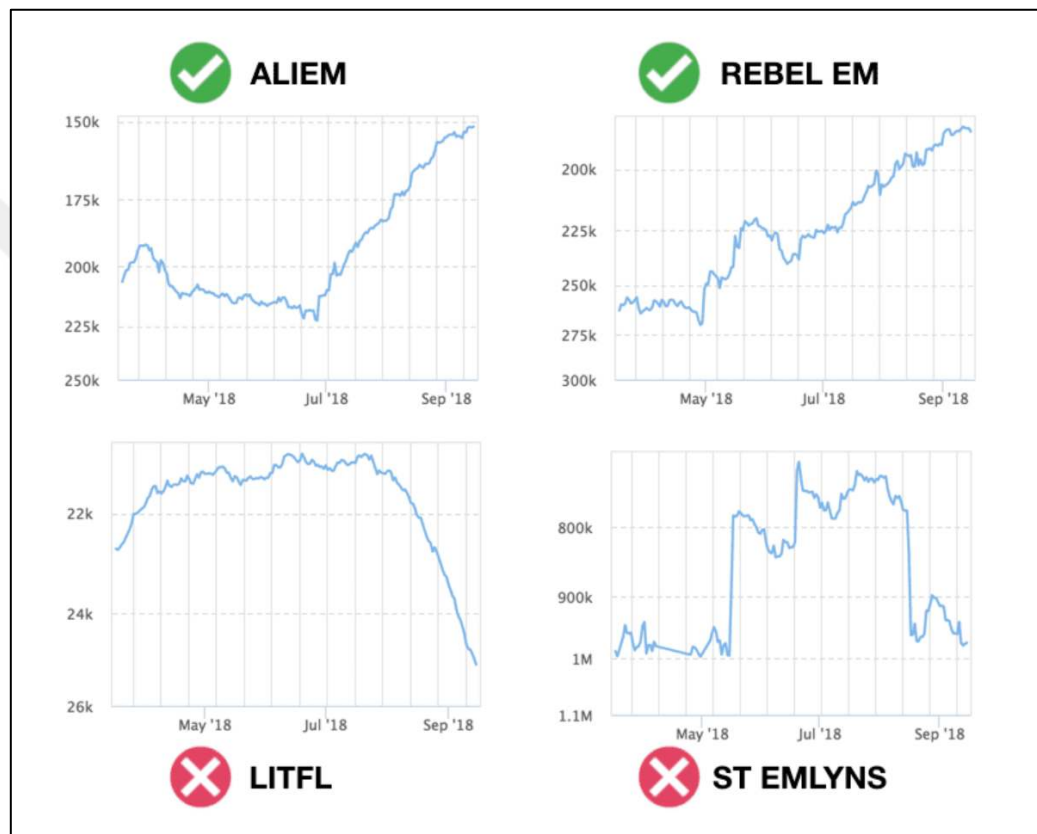


Figure 2.14 Web page examples of affected by Medic update [36]

Google has certain criteria for a website to be classified as high E-A-T content.

- If it is a medical website, it should be written by individuals or organizations with appropriate medical expertise.
- If it is a news website, it should be produced with journalistic professionalism that is descriptive and understandable for users.

- If it is a web page containing scientific topics, the topics covered should be based on well-established scientific ideas and should be prepared by scientific experts.
- If it is a web page with financial topics, its sources and advice should be reliable and regularly updated.

The most realistic solution to avoid being adversely affected by the Medic update is to clean up the content of the website and improve E-A-T.

2.5 Search Engine Optimization Tools

SEO tools are software that allow users to analyze the technical features of their websites, identify and organize keywords and plan strategies around them, and improve their organic search rankings. It is a set of tools that evaluates many methods such as backlink analysis, content analysis, keyword analysis and competitor analysis. Since it comes in different packages in terms of its features, its benefits vary.

Site optimization for search engines is very important. However, if the site is optimized and its position in the search engine does not change after a certain period of time, the website needs to be reanalyzed. This is where SEO tools come into play.

2.5.1 Ahrefs

Ahrefs is one of the most used online SEO tools. It is paid. It analyzes the site and highlights the parts that need improvement. Performs keyword analysis, ranking tracking, competitor analysis [37]. The general view of the menus and functions of Ahrefs is as in Figure 2.15.

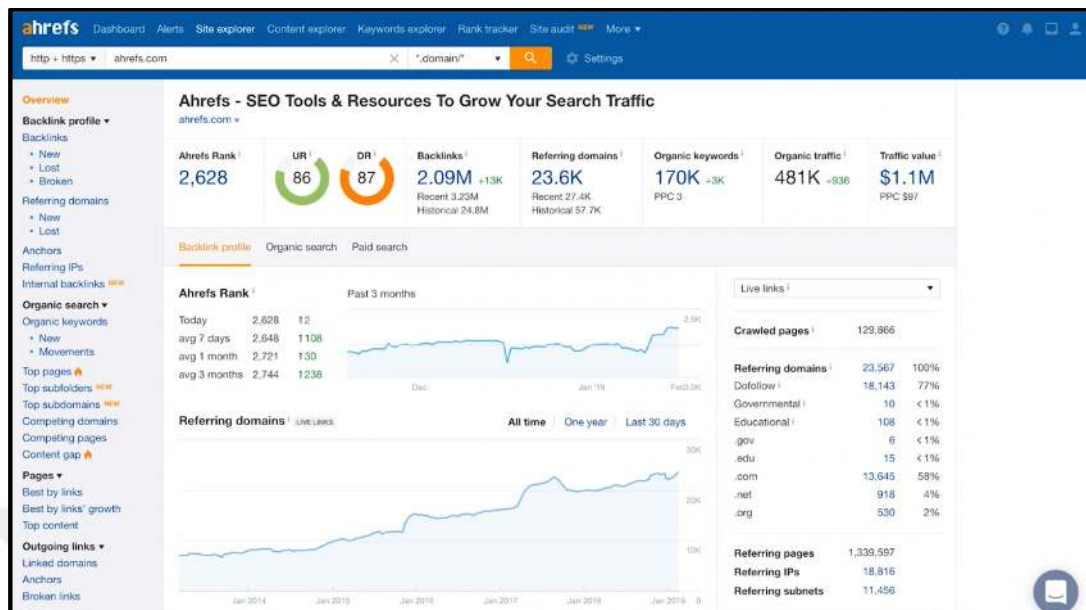


Figure 2.15 Ahrefs SEO tool overview [38]

Some of the prominent features of Ahrefs SEO tool are as follows.

- **Site Audit:** This tool allows to improve the technical SEO performance of the website.
- **Site Explorer:** This tool allows analysis of competitor sites by entering URL information.
- **Keyword Explorer:** Generates keyword idea for domain name. It allows to observe the performance in the search results by using different keywords.
- **Content Explorer:** It allows to find the most popular content by keyword or subject.
- **Keyword Tracker:** It shows the ranking in the search results according to the keyword information. It can be filtered by country/region, desktop/mobile.
- **Click data:** It helps to understand the estimated clickthrough rate.

2.5.2 Moz

Moz is a full featured SEO tool that can help users optimize their websites and get more traffic. It has the ability to closely monitor the algorithm changes that Google

regularly releases and to reflect this quickly on the data. It is one of the most important advantages of website owners to keep them updated in this way. It is a professional toolkit for SEO professionals, small businesses and agencies to use. The general view of the menus and functions of Moz is as in Figure 2.16.

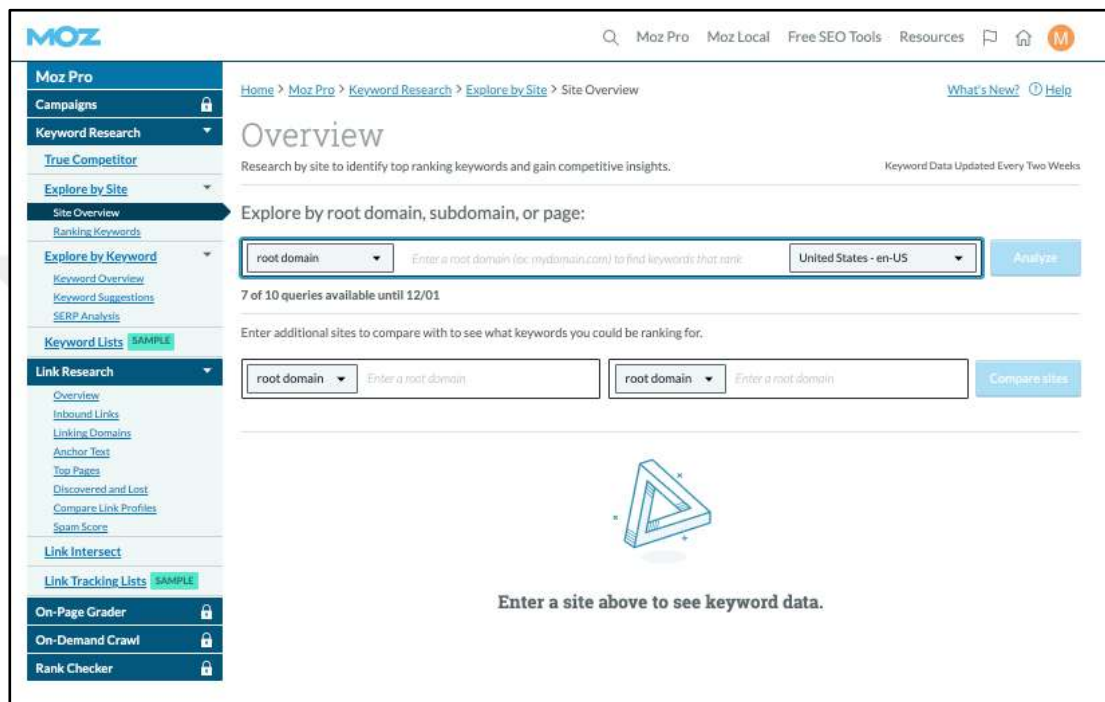


Figure 2.16 Moz SEO tool overview [39]

Some of the prominent features of Moz SEO tool are as follows.

- **Keyword Research Tool:** It provides the top-ranking keywords to be found by entering the domain name.
- **Site Audit Tool:** It is used to measure the performance of the website in search results. It shows errors, scans for problems and gives suggestions.
- **On-Page SEO:** Provides optimization suggestions when keyword or URL information is entered.
- **Backlink Research:** Shows backlinks to the website.
- **Generate Reports:** Allows creating website-specific analytics reports.

2.5.3 Semrush

Semrush is one of the best SEO tools on the market. It is a comprehensive analysis program with many functions that helps in the development of websites. It offers many different features to stand out among its competitors and provides a comprehensive SEO review. It shows the level of search traffic and the number of backlinks a site has. Through this tool, trending keywords are checked for competitor backlinks on social media. It also allows checking low-quality links. That's why Semrush is a complete SEO tool that helps to rank websites higher in search results. The general view of the menus and functions of Semrush is as in Figure 2.17.

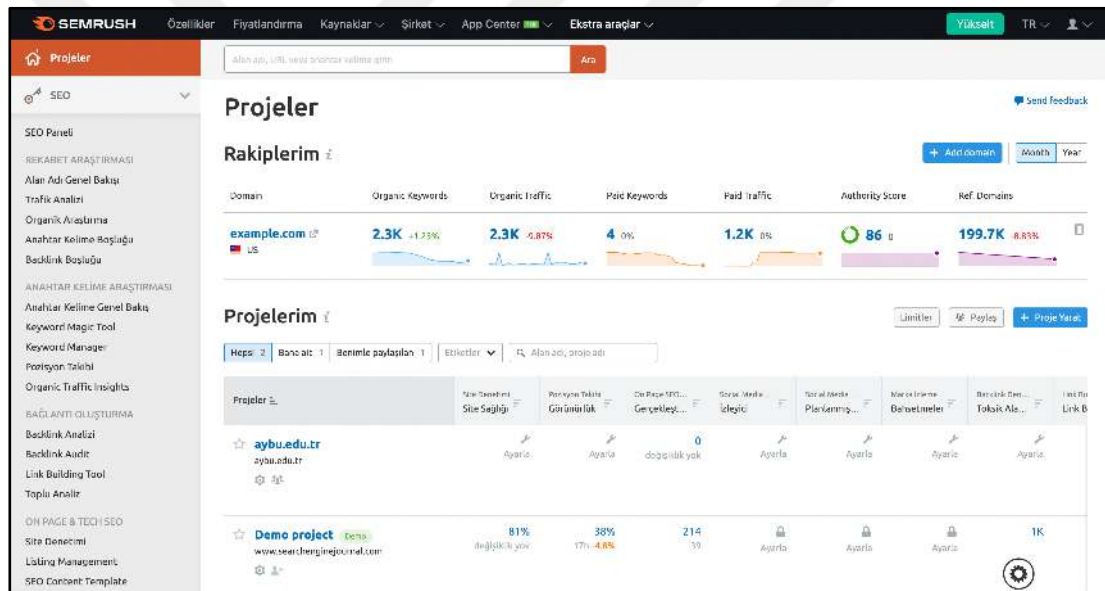


Figure 2.17 Semrush SEO tool overview [7]

Some of the prominent features of Semrush SEO tool are as follows.

- **Keyword Research Tool:** Performs performance analysis according to different keywords.
- **Competitive Research:** Provides to find competitor sites by domain name. Shows organic or paid keywords and content used by competing sites.
- **Local SEO Monitoring:** Shows website status based on location.

- **Visitor Analysis:** It provides to see the visitor source, language, number of monthly visitors and the pages that attract the most visitors.
- **Ad Tracking:** Click-through rates on website pages.
- **Social Media:** Allows advertising directly on social media platforms.
- **Content Analysis:** Provides practical tips for improving content.
- **Report Generation:** Allows to generate custom reports for individual projects, different domains, URLs or keywords.
- **Mobile:** Shows performance on mobile devices.

2.5.4 Nibbler

Nibbler is a free online optimization tool designed to help users run performance tests on websites. It is a website checker that tests the technical SEO of websites and provides actionable advice. It's fast and simple. The tool evaluates the website URL information out of 10 in the following areas:

- **Overall:** The overall rating, response time and content quality for the website
- **Accessibility:** Ease of use for mobile and disabled users as well as regular users
- **Experience:** User experience and satisfaction
- **Marketing:** How popular the website is, its ranking in search results
- **Technology:** The technical quality of the website, the technology and tools used

CHAPTER 3

LITERATURE REVIEW

In the literature on search engine optimization, there are many books, articles, magazines, and authoritative websites. Studies in which different approaches are developed for the techniques, algorithms and technologies used in search engine optimization are examined in this section. The main goal in these studies is based on how to achieve better ranking in search engine results. While researching within the scope of this thesis, many articles were examined, and progress was made in line with this information.

Kılıç developed a web-based vertical search engine that can be customized by the user, considering the characteristics of Turkish [40]. Horizontal search engines do not consider the specifics of the language in which documents are written, except for a few languages, such as English, where most of the content is written. It is aimed to index Turkish documents more effectively by ensuring that Turkish characters are processed correctly on the pages, the document recognizes the language in which it is written, and the roots of the symbolized text are found. Kılıç tested the sub-components of the vertical search engine and the general search engine separately and found that the search engine increased the efficiency of accessing information.

Patil et al. examined page ranking algorithms, search engine optimization techniques, and black hat SEO techniques [41]. Website ranking in search results is highly dependent on how SEO is implemented. White hat SEO techniques return quality content. These techniques produce slow but long-lasting results. These are beneficial for both users and search engines. Black hat SEO techniques give fast results in a short time and the site can be penalized if the search engine detects the unethical activities of the site. The aim of the study is to raise awareness and encourage more research in this area.

Eğri examined the techniques offered by search engines as clues for search engine optimization in his thesis [11]. PageRank and TrustRank are emphasized, and the basic concepts involved in search engine optimization are explained. Google's platforms that can be used for search optimization are mentioned. The updates developed by Google on the scanning methods of the sites are discussed and the changing factors in search optimization techniques are explained with these updates. Within the scope of the thesis, search engine optimization techniques were applied on a site and these applications were successful, and the site managed to rank first in the search results. As a result, it is emphasized that the primary factor to be considered during optimization is not the search engine, but rather the users.

Horasan used genetic algorithm in the optimization process in order to reduce the cost of search engine optimization for institutions [42]. The main purpose is to compare the software created with the genetic algorithm with the traditional method and to reveal its differences. In its content, titles of web technologies, working principles of search engines, considerations in search engine optimization, what genetic algorithm is and how it is used in search engine optimization are explained in detail. For a sample application, a website was developed, and website optimization was provided by genetic algorithm. In the conclusion part, it has been observed that the system applied on 100 different content increases the indexed pages and the queries ranked in the index. For this reason, this system is presented as an alternative to traditional search engine optimization.

Abdulwahid examined in detail how open source-based search engines work in his study [43]. It has been mentioned what search engine and web browsing are and the relationship between them. Algorithms used in web browsing were examined. By giving information about Apache Nutch and Apache Lucene, its architecture was explained, then the application was developed and tested.

Kayakökü carried out his work aiming to develop a mobile-based patent search engine [44]. The features that the patent search engine should have been determined, a new patent search engine design was made within the framework of these features, and a

mobile-based prototype was developed in accordance with this design. The patent system, the technical information of this system is explained, several existing patent search engines are examined and compared in many ways. Adequacy of patent search engines was examined, and a new mobile-based patent search engine was developed in line with the deficiencies.

Yuniarthe has done research on the use of Search Engine Optimization (SEO) and artificial intelligence technology [45]. Artificial intelligence is divided into three methods: evolutionary computing, fuzzy logic and classifiers, and statistical models. Depending on these methods, Yuniarthe has obtained a series of prototypes by researching scientific articles in the literature describing the application forms of artificial intelligence in SEO. These; Fuzzy Inference System, Polidoxa, Commercial Packages, Support Vector Machine application and K-Nearest Neighbor Algorithm application. He explained how these artificial intelligence techniques will be applied on SEO and the related results. He offered suggestions to analysts who will implement SEO optimization methods with artificial intelligence-based devices.

Ankalkoti examined search engine optimization tools and techniques in his study [46]. The stages that search engines go through while working are explained and the types of search engines are examined. On-page and off-page SEO, which are search engine optimization techniques, are mentioned, and white hat SEO, gray hat SEO and black hat SEO techniques are examined. Google Webmaster Tool, Google Analytics Tool, Google Ad Words, Google Trends and Google Recommendations, which are SEO tools developed by Google, are detailed. Ankalkoti explained the importance of the search engine optimization process and touched on what needs to be done for the pages to rank higher in the search results.

Türeç in her master's thesis, discussed the application of search engine optimization on websites that work with the logic of "collecting-combining" and the analysis of these applications [47]. In this context, a site that works with the "collection-combining" logic has been created and search optimization techniques and data optimization techniques, which are also explained in the thesis, have been applied on

the site. As a result of the study, while the number of visitors, the number of page views and the first place in Google have increased during the search engine optimization period; as a result of the termination of off-site optimization, a decrease was observed in these metrics. Türeç suggested that off-page SEO should be continued as a recommendation.

Ulah et al. conducted a review and comparative study of different page ranking algorithms available for search engine optimization [48]. They mentioned that a new optimization technique is needed to improve the current search engine optimization in increasing page rankings. This article also explores some of the improvements needed to overcome the current problem in this area by testing on some real case data. This article is a comparative analysis examining various studies on the page ranking algorithm. It is thought that the analysis results will help the relevant organization to decide on the appropriate algorithm that can be used in search engine optimization.

Dhruva and Patel created a research paper focusing on the impact of search engine algorithms on ranking [49]. This article has examined the factors that affect ranking such as keywords, on-page optimization, off-page optimization, and website content from SEO techniques. They gave information about page numbering algorithms and made an application for a website.

Karlık explained the working architecture of crawler programs, which are bot software used by search engines, and developed a web crawler application that works in parallel [50]. Then, he explained the indexing architecture of search engines and developed an indexing application that works in parallel. How to search in the indexing infrastructure of search engines has been mentioned and a search application with a web interface has also been developed.

Karabıyık performed text classification with machine learning methods, in his thesis on vertical search engine [51]. It has designed a restrictive web spider so that the search engine only finds academic publications. In the application, the text classification process was carried out in order to find the search results in the most accurate way. The results obtained by applying Plain Bayesian and Support Vector Machine

classifiers, which are the most used classifiers in this field, are compared. As a result of the study, it has been determined that the developed application provides vertical search engine principles by eliminating irrelevant results.

In order to measure the changes caused by on-site search optimization in a website traffic, Koç analyzed a website that he found weak in terms of search engine optimization, identified its weaknesses and applied step-by-step optimization techniques [52]. After the improvements, measurements were made again. As a result of these measurements, the mass obtained from the search traffic of the website increased proportionally, the site loading speed decreased, the bounce rate decreased and there was an increase in the number of page visitors between the two periods.

Vuran examined how effective SEO methods are for search engine results page (SERP) rankings [53]. Google search engine was chosen as the experimental environment. He examined Google search engine algorithms and analyzed SEO criteria and methods. The performances of SEO methods are shown comparatively.

In general, most of the studies in the literature are optimization techniques that aim to rank higher on Google's search results page. It has also been emphasized that the algorithms of Google and other search engines are closed boxes. However, the point of particular emphasis in the traditional search optimization method is that the optimization is not directly for search engines; must be done for the user. In addition, it is necessary to see optimization as a process and take it as a part of the work.

CHAPTER 4

METHODS AND APPLICATION

In this section, some features are tested in order to see the results of the analyzes made with SEO tools. This study was conducted in order to compare the success rates in the search results with the results of the analysis. Semrush, Moz and Nibbler online search engine optimization tools were preferred for analysis. Although the Ahrefs tool is popular, it has not been used for analysis because it does not have a trial version.

First, a university website was selected. In order to see the performance of the site in the search results, various reports were produced by using the features of SEO tools such as keyword analysis, backlink analysis and organic traffic. These features were compared with competing university web pages.

4.1 Analysis of University Websites with Semrush

Domain overview data of Ankara Yıldırım Beyazıt University's official website was created with Semrush. This analysis was made by entering the URL information of the relevant website into the online Semrush tool. The analysis report, which includes analysis results such as authority score, organic search traffic, paid search traffic, number of backlinks, and percentage of searches by country, is shown in Figure 4.1. The authority score of Ankara Yıldırım Beyazıt University website was 59, organic search traffic was 306.3K, paid search traffic was 132. This analysis includes data from around the world. It is seen that the most calls are made from Turkey with a rate of 98%.

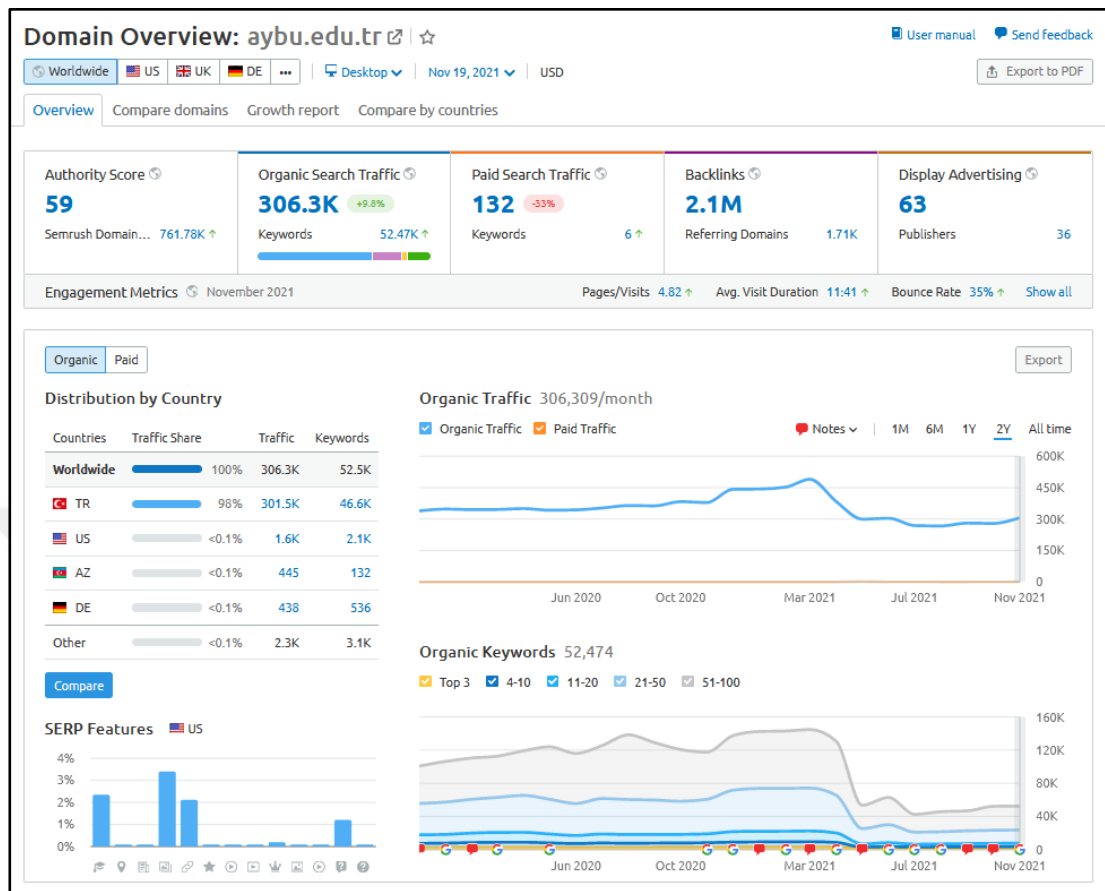


Figure 4.1 Domain overview of Ankara Yıldırım Beyazıt University website with Semrush

Traffic in organic searches, keywords, distribution graph by previous months and search data by country are shown in organic research analysis in Figure 4.2. It is observed that database growth increased towards the end of May 2019.

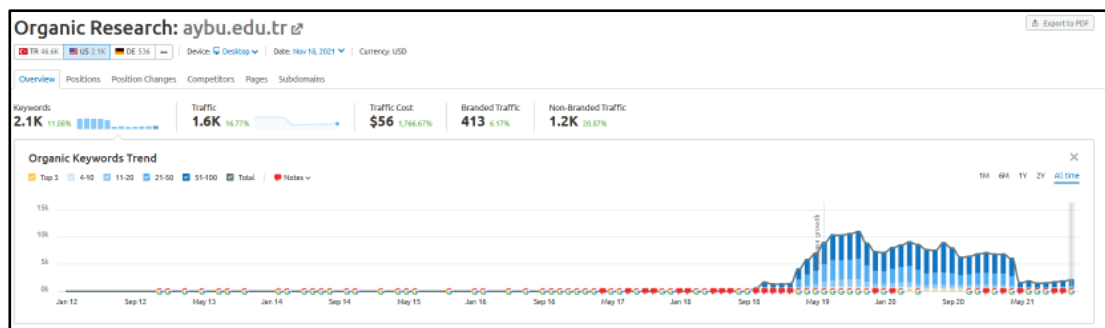


Figure 4.2 Organic research of Ankara Yıldırım Beyazıt University website

The domain overview data displayed by entering the URL information of the Ankara University official website was created with Semrush. Authority score is 71, organic search traffic is 4.1 million, and the number of backlinks is 5.2 million. The distribution of search traffic over the last few months is also charted. It is shown in Figure 4.3.

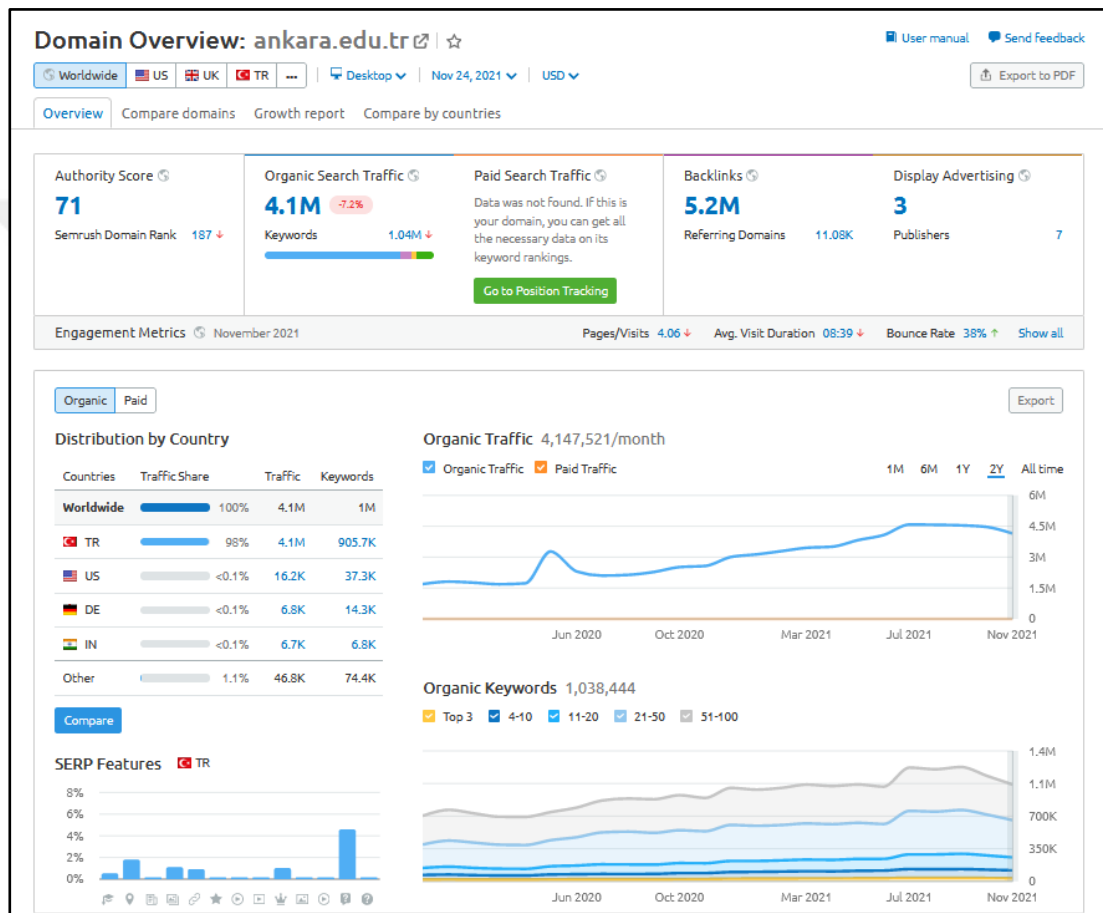


Figure 4.3 Domain overview of Ankara University website with Semrush

The analysis was made in the Semrush tool by entering the URL information of the official website of Oxford University. Domain overview data is shown in Figure 4.4. Authority score is 83, organic search traffic is 5.2 million, paid search traffic is 7200. It has also been analyzed that the most searched country is United Kingdom.

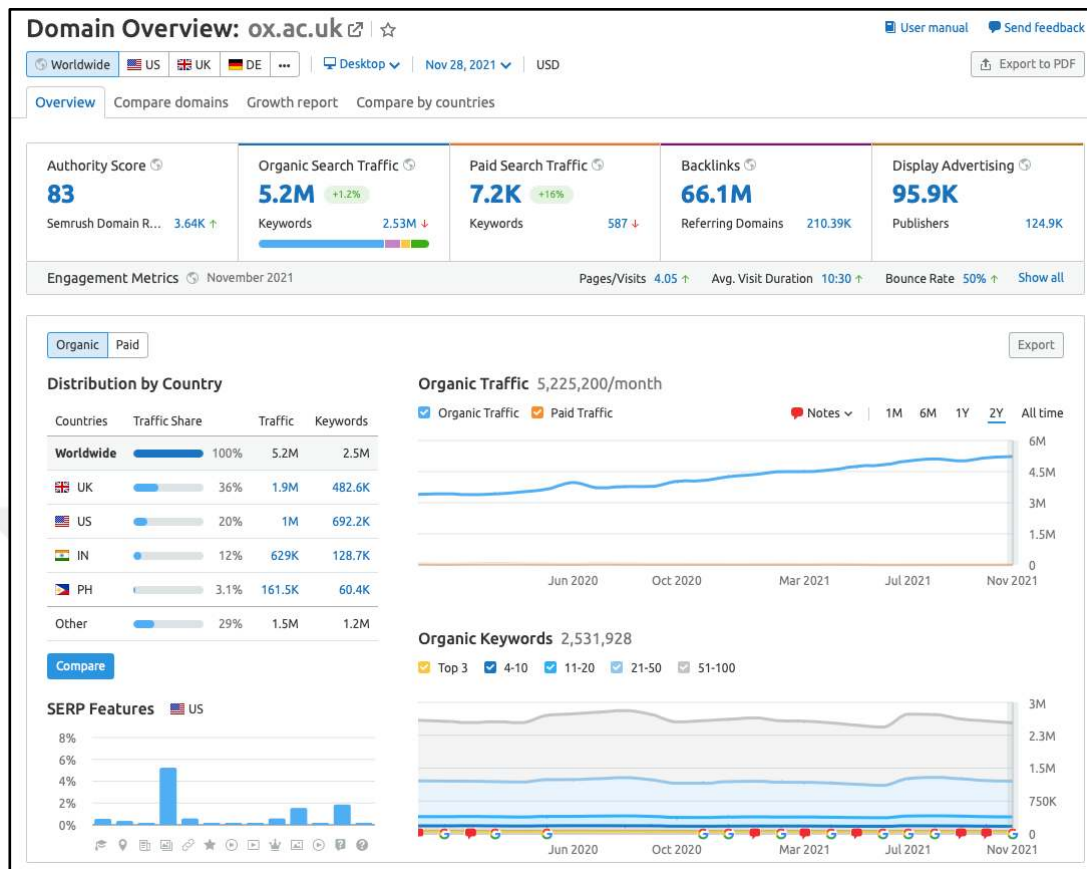


Figure 4.4 Domain overview of Oxford University website with Semrush

Comparative traffic analysis report with university websites is as in Figure 4.5. The comparison was made by entering the URL information of Ankara Yıldırım Beyazıt University, Ankara University and Oxford University websites and clicking the “Compare” button. Up to 5 competing sites can be compared. It summarizes the number of visits by domain, unique visitors, average visit duration, and bounce rate. Oxford University's official website is well ahead with 9.6 million visits. The low bounce rate on the official websites of Ankara Yıldırım Beyazıt University and Ankara University is positive. The number of Ankara Yıldırım Beyazıt University in terms of average visit duration is high and it is increasing gradually.

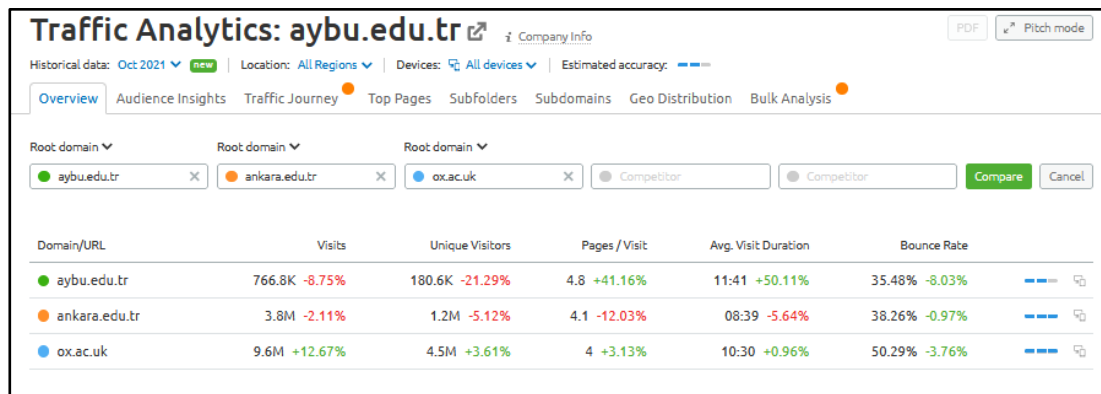


Figure 4.5 Traffic analysis comparison of university websites

Backlink analysis of Ankara Yıldırım Beyazıt University is shown in Figure 4.6. Backlink comparison was also made by entering URL information of competing university websites. Data such as authority score, referring domains, backlinks were compared with the websites of different universities.

In terms of authority score, aybu.edu.tr is at 59 points and increasing, while Oxford University has the highest score with 83 points. While aybu.edu.tr has 2.1 million backlinks, ankara.edu.tr has 5.2 million and oxford.ac.uk 64.8 million. The number of referring IPs is 1.7K for aybu.edu.tr, 10.7K for ankara.edu.tr and 149K for oxford.ac.uk. Monthly visit rates and keyword counts were also compared. The number of keywords is 54.1K for aybu.edu.tr, 1.1M for ankara.edu.tr and 2.6M for oxford.ac.uk. Backlink types and link attributes comparisons are also available in the charts.

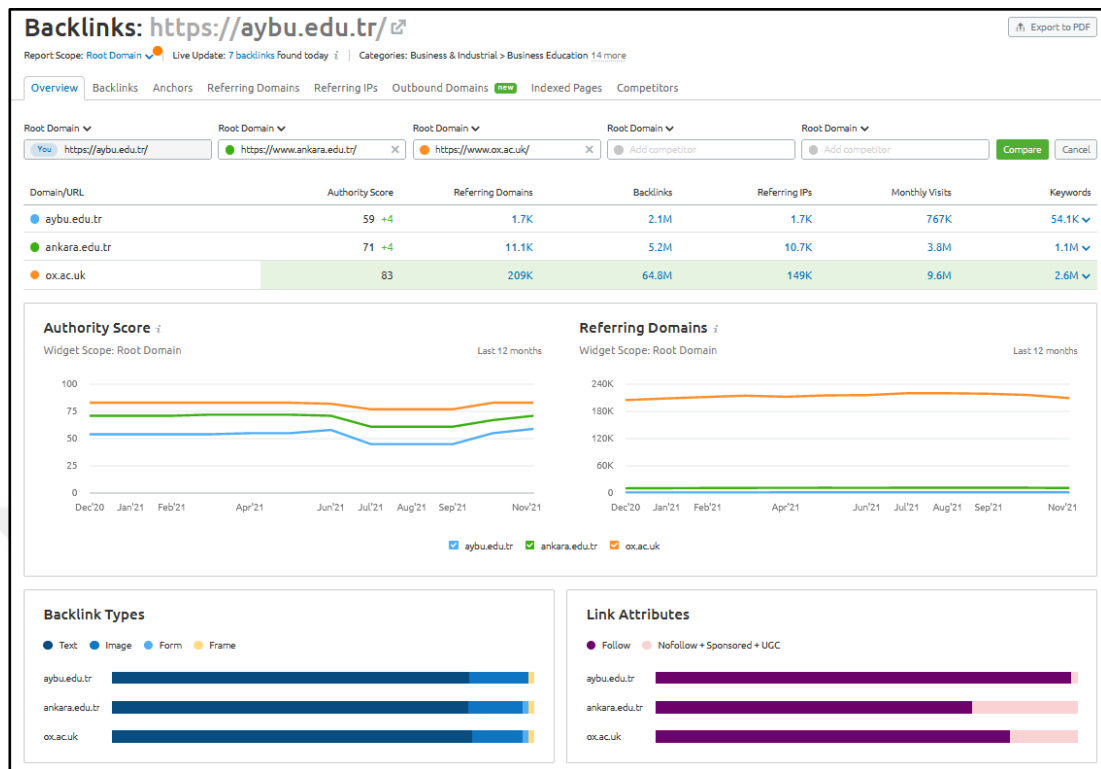


Figure 4.6 Backlink analysis comparison of university websites

4.2 Analysis of University Websites with Moz

The domain overview analysis of Ankara Yıldırım Beyazıt University's official website was made with the Moz tool. As seen in Figure 4.7, first the URL information of Ankara Yıldırım Beyazıt University is entered in the root domain section and the “Analyze” button is clicked. The generated analysis report includes domain authority, linking domains, inbound links and ranking keywords data. Moz online tool measured the domain authority of aybu.edu.tr as 38. While the number of linking domains is 2.1K, it has gained 170 of them in the last 60 days and lost 114 of them. While the number of inbound links is 252.4K, the number of ranking keywords is 156.

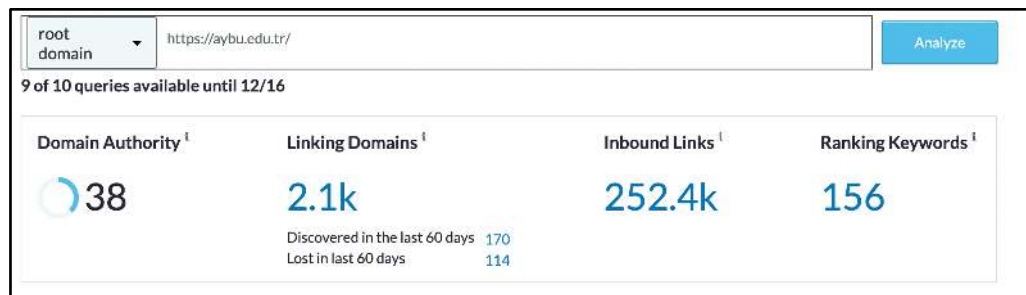


Figure 4.7 Domain overview of Ankara Yıldırım Beyazıt University website with Moz

The domain overview data of the Ankara University official website was also created with the Moz tool by entering the URL information in the same way. As seen in Figure 4.8 , the domain authority value was measured as 60. While the number of Linking domains is 13.3K, it gained 887 of them in the last 60 days and lost 379 of them. While the number of inbound links is 1.2M, the number of ranking keywords is 7.5K.



Figure 4.8 Domain overview of Ankara University website with Moz

Domain overview data of Oxford University official website is also shown in Figure 4.9. The domain authority value is measured as 93 and it is quite good. While the number of linking domains is 258.7K, the number of inbound links is 43.5M, and the number of ranking keywords is 277.9K.



Figure 4.9 Domain overview of Oxford University website with Moz

The analysis of Ankara Yıldırım Beyazıt University website regarding the discovered and lost connection areas in the last months was made with the Moz online tool. It is shown in Figure 4.10. While 170 linking were gained, it is observed that 114 linking were lost in the last 60 days. In recent days, the number of links discovered and lost in the site domain has been decreasing.

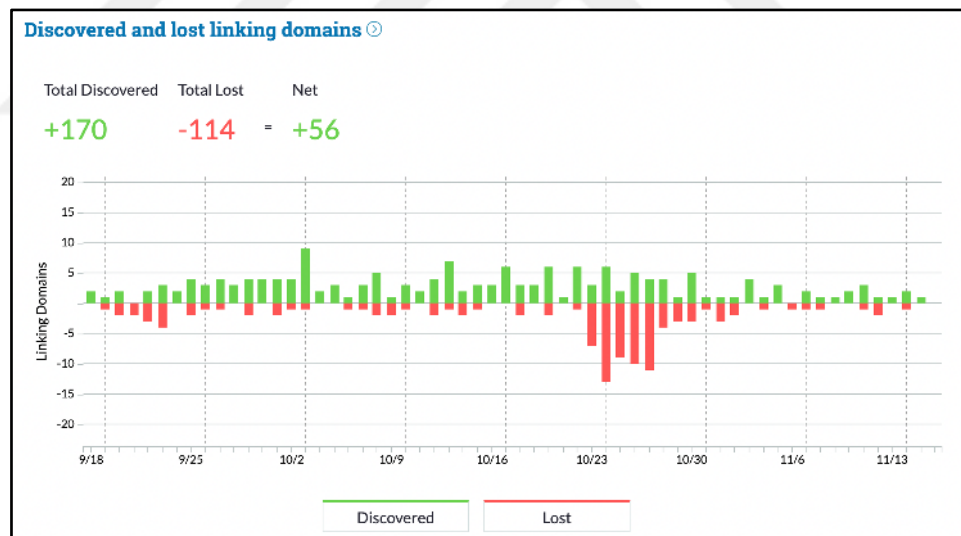


Figure 4.10 Discovered and lost linking domains of Ankara Yıldırım Beyazıt University website

Discovered and lost linking domains of Ankara University website in recently are shown in Figure 4.11. While 887 linking were gained in total, it is observed that 379 linking were lost. In mid-October, the number of links discovered reached the maximum level. However, it has been decreasing in recent days. The status of Ankara University website is slightly better in terms of number of domains.

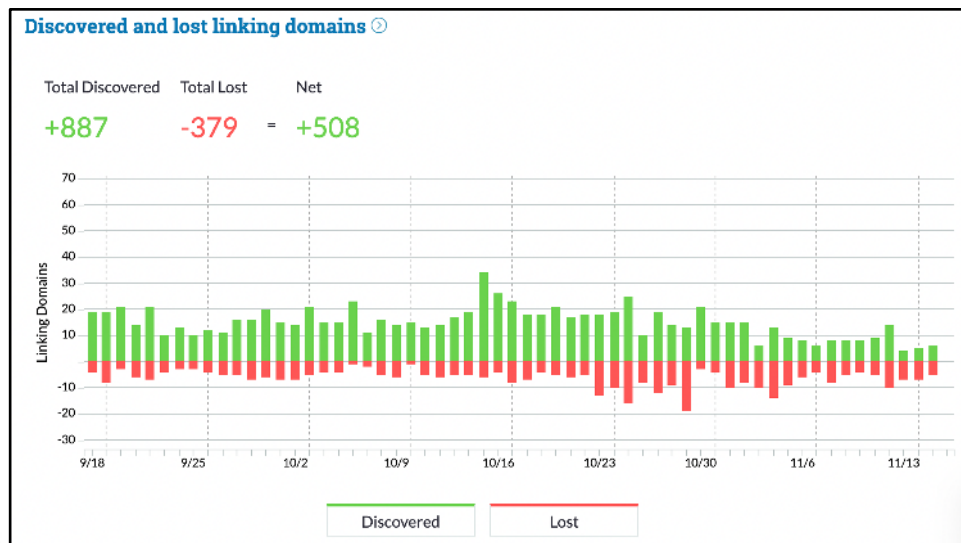


Figure 4.11 Discovered and lost linking domains of Ankara University website

Discovered and lost linking domains of Oxford University website in recent months are shown in Figure 4.12. While 17357 linking were discovered in total, it is observed that 5998 linking were lost. In general, the number of discovered links is always very high, while the number of lost links is low. The net link count is pretty good with 11359.

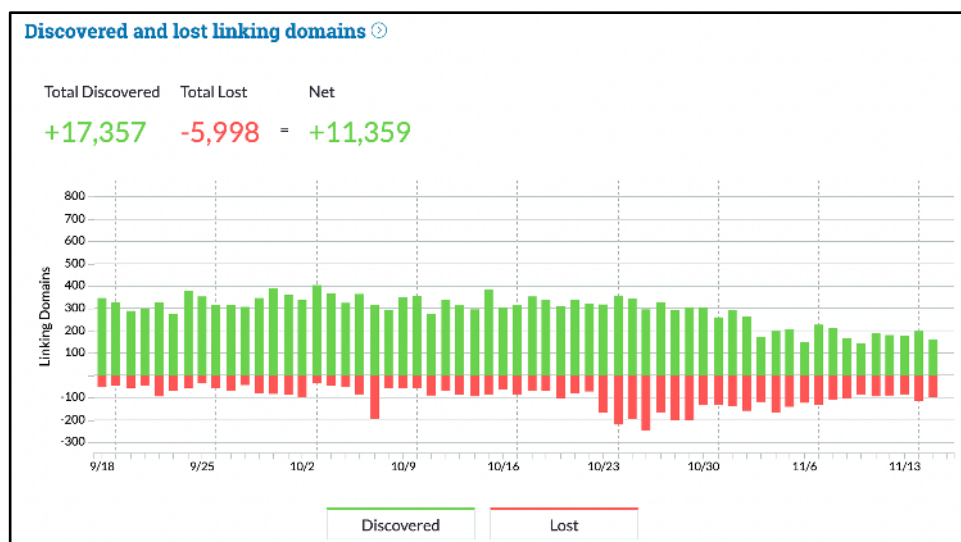


Figure 4.12 Discovered and lost linking domains of Oxford University website

4.3 Analysis of University Websites with Nibbler

Nibbler is not a well-known SEO tool. Although the actions are limited, the interface is simple, and the actions are useful. Therefore, several analysis results with this tool were examined.

Technical SEO report of Ankara Yıldırım Beyazıt University official website were created with Nibbler and are shown in Figure 4.13. The overall score was evaluated as 6.9 out of 10. It is rated 8.2 for accessibility, 7.7 for experience, 7.2 for marketing, and 6.6 for technology. On-page and off-page SEO criteria are also scored in the analysis. The internal links, server behavior and meta tags values of Ankara Yıldırım Beyazıt University's website were measured to be quite low. Headings, URL format and ratings on social networks are good.

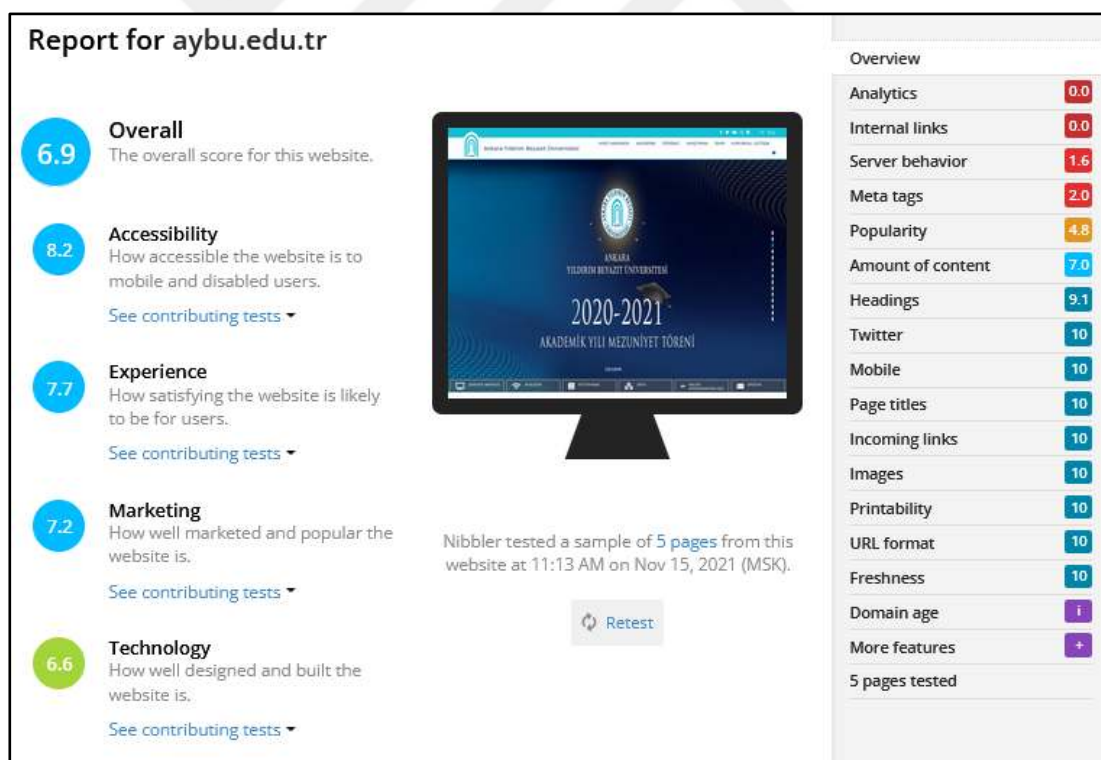


Figure 4.13 Technical SEO report of Ankara Yıldırım Beyazıt University website with Nibbler

Technical SEO report of Ankara University official website were created with Nibbler and are shown in Figure 4.14. The overall score was evaluated as 9.2 out of 10. It is rated 9.2 for accessibility, 9.0 for experience, 8.4 for marketing, and 8.8 for technology. Analyzed on-page and off-page SEO criteria are also displayed. The popularity of Ankara University's website was 7.1, the headings score was 6.0, and the server behavior was 7.0. Except for these low values, internal links, amount of content, URL format and meta tags values are above 9. Their scores are also high compared to other criteria.

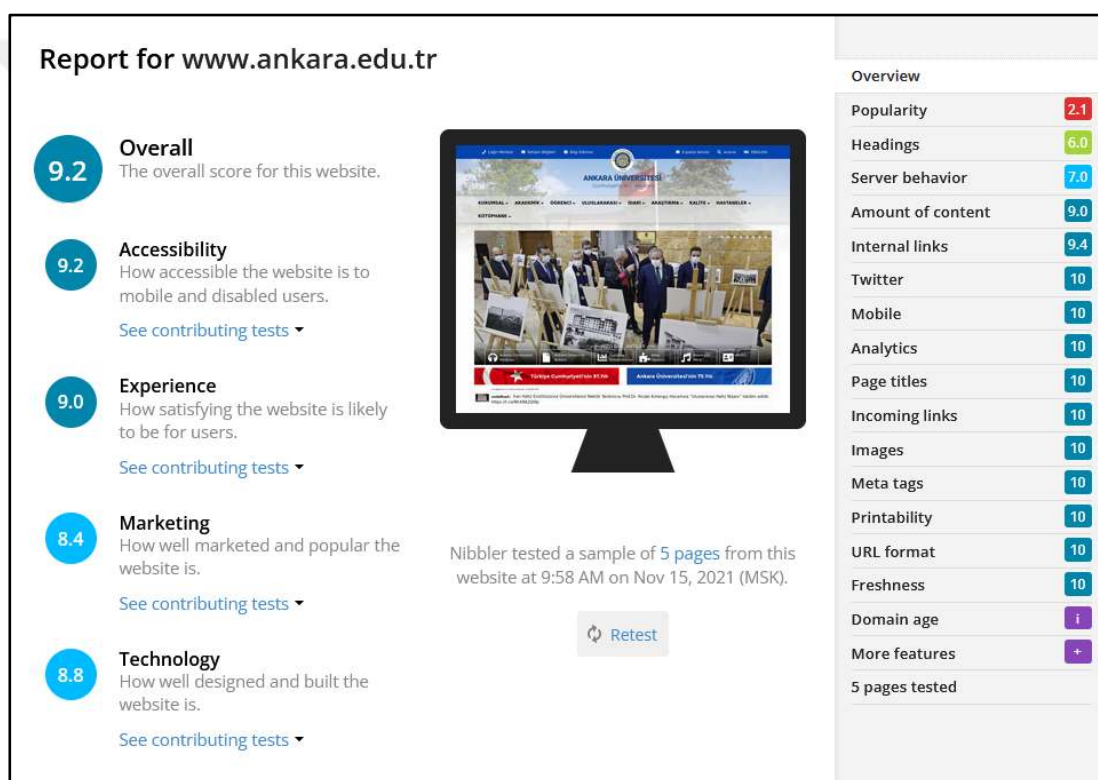


Figure 4.14 Technical SEO report of Ankara University website with Nibbler

With the Nibbler online tool, the technical SEO report of the Oxford University official website was also created. It is shown in Figure 4.15. The overall score was evaluated as 9.3 out of 10. It is rated 9.9 for accessibility, 9.1 for experience, 8.9 for marketing, and 8.8 for technology. Looking at the on-page and off-page SEO feature scores of the Oxford University website, server behavior, meta tags, amount of content values are

above average. Internal links, social networks, headings and page titles values are very good.

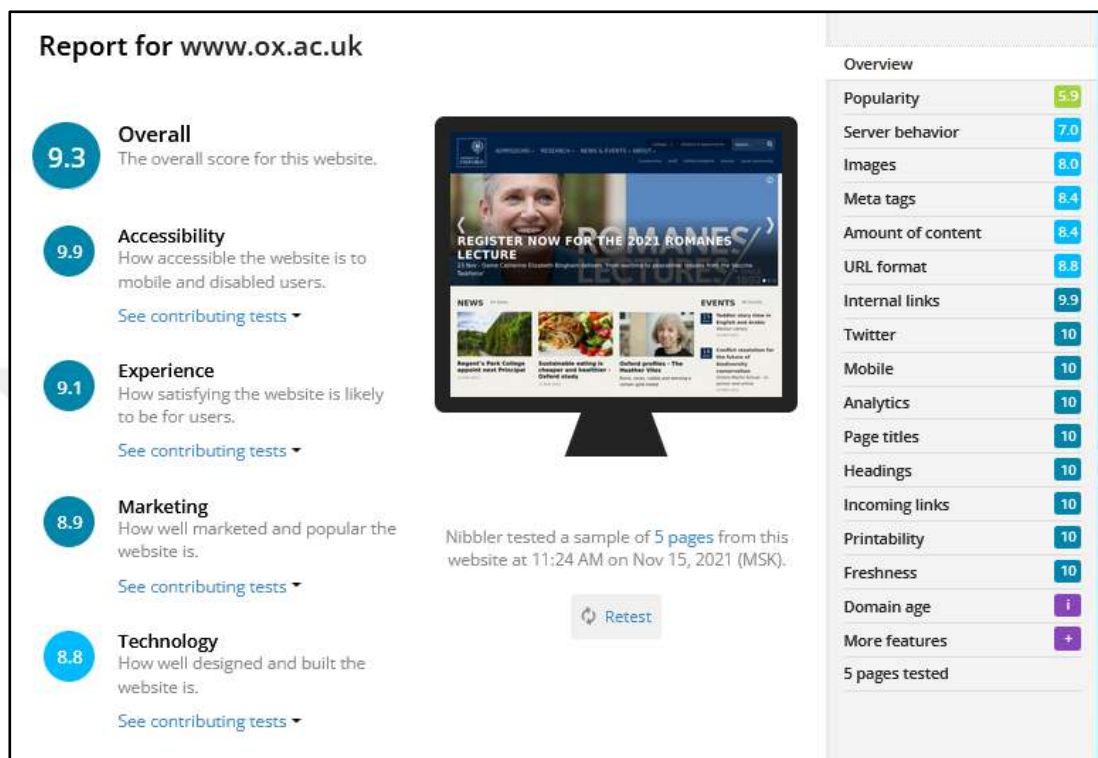


Figure 4.15 Technical SEO report of Oxford University website with Nibbler

CHAPTER 5

EXPERIMENTAL RESULTS

5.1 Findings

Search engine optimization is a long-term process, but it has great benefits in terms of returns. Search engine optimization and marketing has become a basic need for all businesses. The SEO industry has made great advances and become widespread in the last decade. Large-scale internet sites are constantly renewing themselves, depending on the new techniques developed for the search engine and the updating of the ranking and search algorithms of the major search engines.

Instead of traditional SEO techniques, Google has released more ethical and user-oriented algorithm updates and changed their SEO strategy significantly. It continues its updates depending on the returns of the new world order.

The answers to the research questions asked at the beginning of the study are as follows:

1. It has been determined that search engines apply different algorithms when ranking websites, but the most important improvement is achieved with the PageRank algorithm, and the PageRank value of the websites is effective in the ranking in the search results.
2. In search engine optimization, primarily on-page and off-page SEO techniques are applied. The criteria that are important when applying these techniques are explained in the study. After these processes, the technical SEO stages of crawling, indexing and ranking are performed.
3. While the algorithms used during the optimization processes differ according to the search engines, it has been learned that the contents of the algorithms are not shared.

4. It has been learned that there are many paid or free tools used in the SEO process. A few of these tools have been examined and comparatively analyzed.
5. In order to observe the performance of the websites in SEO tools, several university websites were selected, analyzes were carried out on these tools and the results were compared. It has been determined that the performance of competitor websites can also be compared with some of these tools.

Search engine optimization process has both advantages and disadvantages. According to their needs, areas and goals, website owners should decide to perform the search engine optimization process, considering these advantages and disadvantages.

5.1.1 Advantages of SEO

Some important advantages can be listed as follows:

- It increases the visibility of the website in search engines.
- It is cost effective.
- Provides long-term positioning.
- It creates a reliable image for users.
- The return on investment of SEO is much higher than paid advertising.
- With SEO, it is possible to reach the desired audience.
- It helps to increase brand awareness.
- SEO makes websites accessible globally.

5.1.2 Disadvantages of SEO

Some of the disadvantages of SEO are listed as follows:

- Search engine optimization is a process that takes time and patience.
- SEO takes a lot of time to show satisfactory results.
- SEO is a set of processes that require the correct application of tools and techniques and experience.
- It requires a lot of investment to be the best.

- SEO operations to be done may vary according to search engines. It requires constant updating of the data and information of the website with each update coming to the search engines.
- Since the details of Google algorithms are not shared, there is no guarantee that every SEO action will be effective.
- SEO is not very suitable for flash and graphic based websites.

5.2 SEO Recommendations

Before starting the SEO process, it is important to first select the targeted business area and industry. Accordingly, it is necessary to choose a good domain name. The second important point is to buy the website with a dedicated IP address. Then the SEO analysis can be started.

The first thing to do in SEO analysis is keyword research. It is very important to choose suitable keywords for the site in order to obtain high web traffic in search engines.

It is advantageous to do competitor analysis before starting website development. It is a process that can be done easily with existing SEO tools. Competitor websites are selected that rank high in the search engine. Seeing information such as the keywords they use and get good rankings, backlinks, outbound links, on-page SEO techniques is a guide when creating a website.

The URL information of the website is important. Choosing a Search Engine Friendly (SEF) URL is advantageous.

The page title tag should contain a keyword phrase consisting of two- or three-word combinations. The ideal length should be between 50-80 characters. Title content should be related to the page content.

The meta description tag should be a short and concise statement between 150-200 characters, consisting of two or more sentences, giving information about the website. It should be defined separately for each page depending on its content. It is very

important to use meta robots tags for indexing the website. It is necessary to use target-oriented keywords in the meta keywords tag.

Heading tags from H1 to H4 should be used for headings within the website. Spam sites don't use too many title tags.

For search engines to index the images on the website, alt tags should be added to the images and keywords should be used.

The website should have a sitemap and be updated whenever a new page is added. It is important for indexing by search engines.

Websites that have SSL certificates and use the HTTPS protocol are advantageous in search engine rankings.

Building backlinks of the website is an important point to have a good ranking. Having backlinks with high PageRank is an opportunity to increase a website's ranking quickly. At the same time, the opposite is also true. The content of the backlink, its relevance, its manual creation, and the IP address are also important.

It would be very helpful if backlinks to the web page could be provided from forums, social media networks and blogs.

5.3 Comparison of SEO Tools

The general features, strengths and weaknesses of the reviewed Ahrefs, Moz and Semrush SEO tools are compared in Table 5.1.

Among the SEO tools, Semrush is ahead of its competitors in terms of the wide range of features it offers. It has more advantages than other SEO tools. Keyword capacity, reporting capacity, ease of use and competitor analysis are among its strong features. In second place comes the Ahrefs SEO tool. While it's missing a few features offered by Semrush, it's still a great SEO toolkit. Finally, the Moz SEO tool is among the preferable SEO tools.

Table 5.1 SEO tools comparison

	Ahrefs	Moz	Semrush
Keyword Research	Offers lots of features and searches for different search engines.	Offers several features	Offers several features
Rank Tracking	Provides tons of information in a neat and organized dashboard	Offers a certain level of information	Offers a certain level of information
Site Audit	It is more difficult to use than its competitors	Easy to use	It provides more comprehensive data and is easy to use
On-Page SEO	Its features are slightly less	It has more features	It has more features
Backlink Analysis	It has a large index of live backlinks	It has a large index of live backlinks	Gives more accurate results
Competitive Benchmarking	Limited talents	Many talents	Too many talents
Content Marketing	There is a content toolkit	No special content tools	Advanced content toolkit
Broken Link Analysis	Easier	Harder	Harder
Ease of Use	Average	Least	Most
Report Generation	Limited	Limited	Too many reports to be generated daily
Trial	7\$ for 7 days	Free	Free

5.4 Comparison of Analysis Results on University Websites

The results obtained from the analyzes performed with the Semrush tool are summarized in Table 5.2.

The domain overview of Ankara Yıldırım Beyazıt University was examined with the Semrush tool. It has 59 authority score. Organic search traffic is 306.3K and has increased 9.8% recently, while paid search traffic is 132. Its peak traffic occurred in March 2021. 98% searched from Turkey.

The domain overview of Ankara University was also examined in the same way. It has 71 authority score. Organic search traffic is 4.1M and has decreased 7.2% recently, no paid search traffic data. Its peak traffic occurred in July 2021. 98% searched from Turkey.

The domain overview of Oxford University was examined with the Semrush tool. It has an authority score of 83. Organic search traffic is 5.2M and has increased 1.2% recently, while paid search traffic is 7.2K. Its peak traffic occurred in November 2021. Most searched from UK with 36%. Afterwards, the US, India and Turkey are ranked as the countries with the highest search rate.

When the backlink analysis reports are examined, aybu.edu.tr has 2.1 million, ankara.edu.tr has 5.2 million and oxford.ac.uk has more than 60 million backlinks. Looking at the backlink types of aybu.edu.tr, 85% consists of text, 14% is image and less than 1% is frame. When we look at the link properties, 99% of the follow links constitute, the number of nofollow backlinks is almost non-existent. Looking at the backlink types of ankara.edu.tr, 86% consists of text, 13% of images and less than 1% of frames and forms. Looking at the link properties, follow links make up 75%, while the number of nofollow backlinks is 25%. Looking at the backlink types of ox.ac.uk, 87% consists of text, 12% of images and less than 1% of frames and forms. Looking at the link properties, follow links make up 84%, while the number of nofollow backlinks is 14%.

The shaping of backlink types and features is also one of the optimization techniques that increase the visit rate. It has been observed in rival university data that if new domains increase and new backlinks are gained, the visit rate and authorization score will increase.

Table 5.2 Analysis results with Semrush Tool

	aybu.edu.tr	ankara.edu.tr	ox.ac.uk
Authority Score	59	71	83
Organic Search Traffic	306.3K (+ 9.8%)	4.1M (- 7.2%)	5.2M (+ 1.2%)
Paid Search Traffic	132	No data	7.2K
Backlinks	2.1M	5.2M	66.1M
Backlink Types	Text (85%) Image (14%) Frame (1%)	Text (86%) Image (13%) Frame and form (1%)	Text (87%) Image (12%) Frame and form (1%)
Most Searched Country	Turkey (98%)	Turkey (98%)	UK (36%)

In the analysis reports made with the Moz tool, it is shown that the link area, the number of inbound links and the number of keywords affect the search results of the website. The results of the analysis process are shown in Table 5.3. While the number of keywords for aybu.edu.tr is 156, it is 7.5K for ankara.edu.tr and 277.9K for ox.ac.uk. While the number of linking domains is 2.1K for aybu.edu.tr, it is 13.3K for ankara.edu.tr and 258.7K for ox.ac.uk. Accordingly, domain authorizations are listed as 38, 60 and 93.

The number of keywords and the number of linking domains increase the popularity of the page, and the opposite is true. The number of keywords and linking domains owned by the Oxford University website has been a great advantage.

Table 5.3 Analysis results with Moz Tool

	aybu.edu.tr	ankara.edu.tr	ox.ac.uk
Domain Authorization	38	60	93
Linking Domains	2.1K	13.3K	258.7K
Number of Keywords	156	7.5K	277.9K

As shown in Table 5.4, scoring was done according to the on-page and off-page SEO rules in the analysis reports made with the Nibbler tool. It has been observed that the internal link, server behavior and meta tag usage of aybu.edu.tr is very low. However, it scores well on headings and URL format. While ankara.edu.tr has good use of meta tags, it has 6 points in titles. Apart from these, it has been observed that the score is above 7 in terms of internal link, URL format, social media links. On the other hand, ox.ac.uk scores above 7 for all on-page SEO features. Based on all these, the overall website scores were measured as 6.9 for aybu.edu.tr, 9.2 for ankara.edu.tr and 9.3 for ox.ac.uk. The effect of on-page and off-page SEO techniques on website popularity has also been demonstrated.

Table 5.4 Analysis results with the Nibbler Tool

	aybu.edu.tr	ankara.edu.tr	ox.ac.uk
Overall	6.9	9.2	9.3
Internal Links	0.0	9.4	9.9
Server Behavior	1.6	7.0	7.0
Meta Tags	2.0	10	8.4
Headings	9.1	6.0	10

CHAPTER 6

CONCLUSION & FUTURE WORK

6.1 Conclusion

Competition between websites is increasing to rank higher in search results and be accessible to more users. The most important factor that will provide an advantage in this competition is search engine optimization studies.

The purpose of this study is to contribute to the development of a new search engine by providing basic information to the people who will do research on this subject. Another goal is to provide people and organizations who will create new websites with information that they need to pay attention to when building their pages from start to finish. In this way, they will ensure that their pages rank higher in the search results.

In this research, firstly, technical and theoretical issues related to search engine optimization were investigated. The history of the search engine was examined, and the types of search engines were mentioned. Optimization techniques to be followed while developing a search engine are explained. Knowing the basic principles of crawling, indexing and ranking is extremely important for website owners to make it easier for search engines to read and understand and direct them to the right search results. The features of the algorithms used for sorting are examined and their different aspects are emphasized. The algorithm updates made by Google were examined one by one, their successful and unsuccessful aspects and their effects on search results were compared. Then, a few important search engine optimization tools were examined. The features and strengths of these tools are highlighted. Optimization analyzes of university websites were made with Semrush, Moz and Nibbler, which are search engine optimization tools. These results are compared, and it is shown which features need to be developed in order to rank high in search engine results.

During the preparation and development of this thesis, many studies were examined, and a literature review was made. This study differs from other studies in that it examines search engine optimization from many different perspectives and makes comprehensive analyzes with different search engine tools that are not in the literature. When a new search engine is desired to be developed, the process should be managed according to the needs of the user, considering all these details. It can be a source for studies on search engine optimization.

Since the web has a dynamic structure, search engines also need to keep themselves updated. Accordingly, they need to renew their algorithms according to today's needs. As a result of the research, many optimization factors affecting the search results were examined. The number and quality of the number of backlinks and the selection of keywords in accordance with the content and the high number of them are important search engine optimization factors. Search engines love pages that are regularly updated, have uniquely optimized content and quality backlinks.

6.2 Future Work

It is recommended to give more weight to algorithms that search through dialogue rather than word-based search algorithms. T It would be beneficial to shape the search engines to be developed and the algorithms to be used accordingly. Research can be conducted on the concept of 'Semantic Web', which emerges when the user reaches the information he really wants to reach. It has been evaluated that there is a need for a transition from the keyword-oriented search model to the semantic and target-based search model. The development of search engine optimization tools can be analyzed day by day by creating a new website and applying successful SEO techniques to its content.

REFERENCES

- [1] "Statcounter Global Stats," [Online]. Available: <https://gs.statcounter.com/>. [Accessed 20 08 2021].
- [2] C. Razbonyalı, "Dikey Arama Motorlarının İncelenmesi Ve Bir Dikey Arama Motoru Uygulaması," 2011.
- [3] "Site İçi (On Page) SEO Nedir?," Optimist Hub, 24 01 2020. [Online]. Available: <https://www.optimisthub.com/site-ici-on-page-seo-nedir.html>. [Accessed 24 08 2021].
- [4] B. Schwartz, "Google Starts Giving A Ranking Boost To Secure HTTPS/SSL Sites," [Online]. Available: <https://searchengineland.com/google-starts-giving-ranking-boost-secure-httpsssl-sites-199446>. [Accessed 21 10 2021].
- [5] "Site Dışı (Off Page) SEO Nedir?," Optimist Hub, 23 09 2020. [Online]. Available: <https://www.optimisthub.com/site-disi-off-page-seo-nedir.html>. [Accessed 24 08 2021].
- [6] "Sitechecker," [Online]. Available: <https://sitechecker.pro>. [Accessed 10 11 2021].
- [7] "Semrush," [Online]. Available: <https://www.semrush.com/projects/>. [Accessed 13 11 2021].
- [8] "W3C Link Checker," [Online]. Available: <https://validator.w3.org/checklink>. [Accessed 10 11 2021].
- [9] "Dead Link Checker," [Online]. Available: <https://www.deadlinkchecker.com/website-dead-link-checker.asp>. [Accessed 10 11 2021].

- [10] B. Muller, "How Search Engines Work: Crawling, Indexing, and Ranking," Moz Pro, [Online]. Available: <https://moz.com/beginners-guide-to-seo/how-search-engines-operate>. [Accessed 16 08 2021].
- [11] G. Eğri, "Arama Motoru Optimizasyonu Teknikleri," 2013.
- [12] "PageRank," Wikipedia, 01 11 2021. [Online]. Available: <https://en.wikipedia.org/wiki/PageRank>. [Accessed 11 08 2021].
- [13] "TrustRank," Wikipedia, 21 02 2021. [Online]. Available: <https://en.wikipedia.org/wiki/TrustRank>. [Accessed 01 08 2021].
- [14] Z. Gyongyi, H. Garcia-Molina and J. Pedersen, "Combating Web Spam with TrustRank," in Proceedings of the 30th International Conference on Very Large Data Bases (VLDB), Toronto, 2004.
- [15] R. Montti, "Google's Hilltop Algorithm – A Foundation for Modern SEO," Search Engine Journal, 08 05 2018. [Online]. Available: <https://www.searchenginejournal.com/hilltop-algorithm/253893/>. [Accessed 15 08 2021].
- [16] "Google Panda," [Online]. Available: <https://moz.com/learn/seo/google-panda>. [Accessed 18 08 2021].
- [17] "What is the Google Panda SEO Update?," [Online]. Available: <https://www.brightedge.com/glossary/google-panda-update>. [Accessed 20 08 2021].
- [18] "Google Penguin Update," [Online]. Available: <https://www.searchmetrics.com/glossary/google-penguin-update/>. [Accessed 25 08 2021].
- [19] "What is the Google Penguin Update?," [Online]. Available: <https://www.brightedge.com/glossary/google-penguin-update>. [Accessed 01 09 2021].

- [20] "Google Penalty Recovery," 18 05 2020. [Online]. Available: <https://sites.google.com/site/recoverfrompenalty/>. [Accessed 25 08 2021].
- [21] K. Szastok, "Life after Google's Hummingbird algorithm update – what does it all mean?," 27 11 2013. [Online]. Available: <https://blog.searchmetrics.com/us/life-after-googles-hummingbird-algorithm-update-what-does-it-all-mean/>. [Accessed 03 09 2021].
- [22] Ö. Yoluk, "Google Hummingbird Algoritması," [Online]. Available: <https://omeryoluk.com/google-hummingbird-nedir/>. [Accessed 03 09 2021].
- [23] "Google Hummingbird," [Online]. Available: <https://moz.com/learn/seo/google-hummingbird>. [Accessed 02 09 2021].
- [24] "What is Google Pigeon Algorithm?," 08 12 2020. [Online]. Available: https://www.dotnek.com/Blog/Web/what-is-google-pigeon-algorithm#google_vignette. [Accessed 03 09 2021].
- [25] "Google Pigeon," [Online]. Available: <https://moz.com/learn/seo/google-pigeon>. [Accessed 10 09 2021].
- [26] "Google Pigeon Update," [Online]. Available: https://newsronic.com/google-pigeon-update-google-algorithm/#Advantages_of_Googles_Pigeon_update. [Accessed 09 09 2021].
- [27] "Mobilegeddon: The Effects of Google's Recent Mobile Algorithm Update," 23 07 2015. [Online]. Available: <https://www.sovrn.com/blog/mobilegeddon-the-effects-of-googles-recent-mobile-algorithm-update/>. [Accessed 14 09 2021].
- [28] S. Sanders, "Google's New Search Algorithm Stokes Fears Of 'Mobilegeddon'," 21 04 2015. [Online]. Available: <https://www.npr.org/sections/alltechconsidered/2015/04/21/401269739/googles-new-search-algorithm-stokes-fears-of-mobilegeddon>. [Accessed 20 07 2021].

- [29] "Google Mobile Updates," [Online]. Available: <https://www.searchmetrics.com/glossary/google-mobile-updates/>. [Accessed 25 08 2021].
- [30] E. Enge, "Mobilegeddon: Nearly 50% of Non-Mobile Friendly URLs Dropped in Rank," 03 07 2015. [Online]. Available: <https://blogs.perficient.com/2015/06/03/mobilegeddon-may-have-been-bigger-than-we-thought>. [Accessed 27 08 2021].
- [31] "Google Rankbrain: The Definitive Guide," [Online]. Available: <https://backlinko.com/google-rankbrain-seo>. [Accessed 11 09 2021].
- [32] "What is the RankBrain Algorithm?," [Online]. Available: <https://www.brightedge.com/glossary/rankbrain>. [Accessed 12 09 2021].
- [33] "What is Google Medic?," [Online]. Available: <https://www.bluecorona.com/faq/what-is-google-medic/>. [Accessed 09 09 2021].
- [34] "What is YMYL? SEO for YMYL Websites and Google," [Online]. Available: <https://raddinteractive.com/what-is-ymyl-seo-for-ymyl-websites-and-google/>. [Accessed 09 09 2021].
- [35] B. Schwartz, "Google Medic Update: Google's Core Search Update Had Big Impact On Health/Medical Sites," 08 08 2018. [Online]. Available: <https://www.seroundtable.com/google-medic-update-26177.html>. [Accessed 09 09 2021].
- [36] M. Cadogan, "Google Medic Update – EAT protocol," 03 11 2020. [Online]. Available: <https://litfl.com/google-medic-update-eat-protocol/>. [Accessed 09 09 2021].
- [37] Y. Korkmaz, "SEO Araçları," Crm Medya, [Online]. Available: <https://www.crmmedya.com/seo-araclari/>. [Accessed 15 08 2021].

- [38] "Ahrefs," [Online]. Available: <https://ahrefs.com/tr/site-explorer>. [Accessed 10 11 2021].
- [39] "Moz," [Online]. Available: <https://moz.com/>. [Accessed 12 11 2021].
- [40] A. Kılıç, "Türkçe Dokümanlar İçin Özelleştirilebilir Web Tabanlı Dikey Arama Motoru," 2008.
- [41] P. S. P., P. B. V. and P. A. S., "Search Engine Optimization: A Study," *Research Journal of Computer and Information Technology Sciences*, vol. 1, no. 1, pp. 10-13, 2013.
- [42] F. Horasan, "Genetik Algoritma İle Arama Motorları İçin Web Site Optimizasyonu," 2014.
- [43] N. Abdulwahid, "Crawling The Web Using Apache Nutch And Lucene)," 2014.
- [44] A. Kayakökü, "Mobil Tabanlı Bir Patent Arama Motoru Geliştirilmesi," 2017.
- [45] Y. Yuniarthe, "Application of Artificial Intelligence (AI) in Search Engine Optimization (SEO)," in *2017 International Conference on Soft Computing, Intelligent System and Information Technology (ICSIIT)*, 2017.
- [46] P. Ankalkoti, "Survey on Search Engine Optimization Tools & Techniques," *Imperial Journal of Interdisciplinary Research (IJIR)* , vol. 3, no. 5, pp. 40-43, 2017.
- [47] S. Türeç, "Türkiye'deki Özel Alışveriş Kulüplerinin Motorlarında Etkin Sıralanması," 2017.
- [48] A. Ulah, N. M. Nawı, E. Sutoyo, A. Shahzad, S. N. Khan and M. Aamir, "Search Engine Optimization Algorithms for Page Ranking: Comparative Study," *International Journal of Integrated Engineering*, vol. 10, no. 6, pp. 19-25, 2018.

- [49] K. S. Dhruva S. Patel, "Search Engine Optimization Page Rank Algorithm," *International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR)* , vol. 8, no. 3, pp. 17-24, 2018.
- [50] M. Karlık, "Arama Motoru Mimarisi Ve Uygulaması," 2018.
- [51] M. A. Karabıyık, "Akademik Yayınlar için Makine Öğrenmesi Tabanlı Arama Motoru Tasarlanması ve Uygulanması," 2018.
- [52] H. Koç, "Site içi Arama Optimizasyonunun Bir Web Sitesi Trafğinde Olduğu Değişimlerin Ölçümlenmesi," 2019.
- [53] E. G. Vuran, "Arama Motoru Optimizasyonu," 2019.

CIRRICULUM VITAE

PERSONAL INFORMATION

Name Surname : Melike SARIKAYA
Date of Birth :
Phone :
E-mail :

EDUCATION

Master's degree : Ankara Yildirim Beyazıt University,
Computer Engineering (2017-2021),
Ankara University,
Computer Engineering (2015-2017)
Bachelor : Ankara University,
Computer Engineering (2010-2015)
High School : TOKI Anatolian High School (2006-2010)

WORK EXPERIENCE

Assistant Specialist - Computer Engineer : April 2021 - Currently
TCMB, Ankara

Scientific Programs Specialist : June 2016 – April 2021
TÜBİTAK, Ankara

Intern : August 2014 – September 2014
Basarsoft Software Technologies Inc., Ankara

Intern : August 2013 – September 2013
AGMLab Software Technologies Inc., Ankara

TECHNICAL SKILLS

Programming Languages: Java Technologies (Java, Spring Framework, JPA, Maven, JSF, JSP), C, C++

UI Frameworks: Angular, React, Primefaces

Databases: Oracle, PostgreSQL

Products: Eclipse, Netbeans, Spring Tool Suite, IntelliJ Idea, Visual Studio Code, Docker Container, Atlassian Products (JIRA, Confluence, Bitbucket, Bamboo), SourceTree

Version Control Systems: Git, SVN

LANGUAGE SKILLS

English: Intermediate oral and writing, reading skills

HOBBIES

Travelling, sport, swimming

CONFERENCE PRESENTATIONS

Sarikaya M., Yeniad M. (2021) Search Engine Optimization, *11th International Scientific Research Congress (UBAK), Science and Engineering Sciences 2021, Ankara*.

PUBLICATIONS

Efe A., Sarikaya M., Altinbas M. (2019) The Security Vulnerabilities on Internet-Enabled Embedded Systems, *Journal of Erzincan University Institute of Science*, 12 (3), 1468-1484, DOI: 10.18185/erzifbed.565044.