

NOVA
IMS

**Information
Management
School**

MGI

Mestrado em Gestão de Informação
Master Program in Information Management

SEARCH ENGINE RANKING FACTORS ANALYSIS
Moz Digital Marketing Company Survey Study

Lucio Lanzarini de Carvalho

Project Work presented as requirement for obtaining the Master's degree in Information Management, with a specialization in Information Systems and Technologies Management

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

SEARCH ENGINE RANKING FACTORS ANALYSIS

Moz Digital Marketing Company Survey Study

by

Lucio Lanzarini de Carvalho

Project Work presented as requirement for obtaining the Master's degree in Information Management, with a specialization in Information Systems and Technologies Management

Advisor: Pedro Cabral, PhD

September 2018

ABSTRACT

The use of the Internet increases every year in the world for multiple purposes and at significant rates. In the same way, access to electronic business and personal pages allowing commercial transactions follows these high evolution rates. Many studies on this subject have pointed that it is important for most businesses to have a web presence. The key to be found by the right product or service target audience, at the right moment, according to most of authors, lies with search engines (SE) advent.

However, there had been frequently changes in search engines ranking website classification algorithms during the last years. To accomplish this model evolution, the Search Engine Optimization (SEO) professionals must to frequently adopt to constant changes regarding ranking classification strategies from SE schemas of work.

In this work the author explored a wide range of factors that may influence search engine result pages (SERP's) and examined recent aspects of user experience over a website that are increasing importance regarding the optimization to be done over the web pages, internal and external page links, and its technical components. In addition, it seems that the user action and involvement over the website are key factors that Google will probably continue to adopt to determine websites rank in SERP's.

As an empirical study, all efforts to discover the SE website promotion ranking factors are based on trial and error activities and there is no official knowledge base regarding these protected secrets kept by the major players of this valuable market. Due to the lack of published academic research works in this area, the present work has discovered and documented SE ranking factors based on survey data by a large quantity of companies in digital marketing segment. At the end of the project the author intends to present the state-of-the-art in this field of study as well as some market perception evolution of this subject based heavily on practical experiments and most recent literature in this area.

Moreover, it is growing the debate about the limits of digital marketing. Due the powerful influence of SE to market and people behavior, the presented study data and considerations raise an important forum of discussion now and in the future concerning ethics and socially acceptable limits and controls over personal information on the internet.

KEYWORDS

Digital Marketing; Internet Ranking; Internet Top Position; Search Engine Optimization; Web Crawling; Web Rank; Web Ranking Position

INDEX

1. Introduction.....	1
1.1 Background and Problem Definition	1
1.2 Study Objectives	2
2. Literature review	3
2.1 Search Engine Definition and Basic Features	3
2.2 Search Engines Optimization (SEO)	4
2.3 Moz SEO Company	5
3. Methodology	6
3.1 The Case Study Method.....	6
3.2 Moz Ranking Factors Biannual Survey.....	6
3.3 Sudy Approach and Methodology.....	7
3.3.1 Study Configuration	7
3.3.2 Study Variables	9
4. Results and Discussion.....	15
4.1 Results.....	15
4.2 Discussion	21
4.2.1 Click-Through Rate (CTR)	21
4.2.2 Time Spent on the Site (User Dwell Time and Bounce Rate)	22
4.2.3 Social Media.....	22
4.2.4 Voice Search.....	23
4.2.5 Technical Factors Still in the Game: The Mobile Paradigm	23
4.2.6 Combining Technology and Human Behavioral Patterns.....	23
5. Conclusions.....	25
Bibliography.....	26

LIST OF FIGURES

Figure 1 - Total US spending for SEO	4
Figure 2 - Cross Industry Standard Process for Data Mining	7
Figure 3 - View of Website Usability Attributes	8
Figure 4 - Search Engine Factors Categories and its Influence on Google Algorithm	9

LIST OF TABLES

Table 1 - Variables of General Google Search Ranking by Moz	10
Table 2 - Variables related to Domain-Level Link-Authority Group	11
Table 3 - Variables related to Page-Level Link Metrics Group	11
Table 4 - Variables related to Page-Level Keyword Features Metrics Group	12
Table 5 - Variables related to Page-Level Keyword-Agnostic Features Group	12
Table 6 - Variables related to Domain-Level Brand Metrics Group	13
Table 7 - Variables related to Domain-Level Keyword Usage Features Group	13
Table 8 - Variables related to Page-Level Keyword-Agnostic Features Group	14
Table 9 - Variables related to Page-Level Social Metrics Group	14
Table 10 - Google's Broad Search Ranking Algorithm respondent results	15
Table 11 - Domain-Level Link-Authority Features respondent results	15
Table 12 - Page-Level Link Metrics respondent results	16
Table 13 - Page-Level Keyword Features Metrics respondent results	16
Table 14 - Page-Level Keyword-Agnostic Features respondent results	17
Table 15 - Domain-Level Brand Metrics respondent results	18
Table 16 - Domain-Level Keyword Usage respondent results	18
Table 17 - Domain-Level Keyword-Agnostic Features respondent results	19
Table 18 - Page-Level Social Metrics respondent results	19

LIST OF ABBREVIATIONS AND ACRONYMS

B2B	Business to Business
B2C	Business to Consumer
CPC	Cost Per Click
CRM	Customer Relationship Management
CTR	Click-Through Rate
SE	Search Engine
SEO	Search Engine Optimization
SERP	Search Engine Result Page

1. INTRODUCTION

1.1 BACKGROUND AND PROBLEM DEFINITION

The ubiquity of information is one of the greatest characteristics of the globalized world. Also empowered by mobile technology, almost half of the world total population - around 3.8 billion individuals - had already gained access to the internet (International Telecommunication Union, 2017). It seems evident the use of the world wide web as a powerful tool for the search of information regarding individual, academic, business or from any other perspective of study. In addition, the growing technological integration between countries and cultures through internet network systems have influenced societies living standards in many different aspects. As part of this constant evolution, economic globalization has played the most relevant role since the industrial revolution (Coulibaly, Erbao & Mekongcho, 2018).

In this context, web digital marketing becomes a strongly important source of competitive advantage either in Business to Consumer (B2C) electronic transactions and Business to Business (B2B) market operations (Leefflang, Verhoef, Dahlstrom & Freundt, 2014). From the web user viewpoint, other strong reason for this large internet use growth is that most people prefer to use the web search facility regarding the low cost of information retrieval rather than the need to pay for any other source of knowledge. Additionally, the internet search tools include convenience of physical localization independence, easy-to-use access as well as the possibility to save the search results (Kritzinger & Weideman, 2013).

Considering this scenario, a specific kind of software has a key role and becomes highlighted: the web SE mechanism. Alexa Internet traffic evaluation Company pointed Google.com as the world most accessed website. By the end of 2011, Google was the first internet page to reach one billion visitors (unique) in a single month. It is important to note that SE's are primary used as a former point of entry to navigate the internet, making these tools a crucial part in linking content producer and final users (Baye, Santos & Wildenbeest, 2016). Borrell Associates Company (US) points SEO-related budget in United States will be around US\$ 80 billion, by 2020. It is important to note that these numbers are not about advertise on web (paid announcements) but projected spends on SE optimization initiatives in the United States (Borrell Associates Inc, 2016).

Moreover, in terms of marketing share percentage, we can point Google with over 77% of total market, followed by Baidu 8,13% and Bing 7,31% (netmarketshare.com, 2017). Despite of understand the fact that Google is by far the market leader, it is also interesting to note that Google's large market share is still increasing. In 2016, we had a 67% market share for Google, which indicates that Google has taken another 10% of the market from its competitors over the past year. Therefore, as can be seen in the next chapters the author will focus the study effort into Google mechanism SEO factors understanding, based strongly on its market leadership.

But, of course, the internet environment is not a static domain. The evolution of the web user navigation reflects in the challenge of the major SE algorithms to materialize this model transformation in their new search results pages, falling to SEO professionals understanding the factors that affect SE ranking classification, as well as the degree of influence of these factors. Until recently, these factors were just known as technical, with metrics related to page code and did not considering the experience of users. Recently, it has been noticed that major SE players are trying to improve the quality of the website content to the users as the major factor for a more precise internet page evaluation and to enhance SE ranking classification (Visser & Weideman, 2014).

However, as an empirical study, this context only can be evaluated under the prism of user experience and heavily based on website ranking factors user surveys. Furthermore, since SE's are not transparent regarding their ranking algorithms and sites classification, evidences about this kind of study are only extracted through their result pages analysis and by trying to interpret the page results behavior against each kind of search being done, being supported by multiple and varied data sources. In the same way, theoretical contributions of case study designs can be only evaluated in terms of understanding, theory-building, development, and theory testing (Ridder, 2017).

1.2 STUDY OBJECTIVES

The main objective of this work is to discover and document in a structured format, the SE ranking factors which cause internet sites to be rank promoted and better classified in the web environment.

As a practical approach, this project is based heavily on surveys data by a significative quantity of players in SE marketing segment also in some other case studies found on literature. At the end of the project the author intends to present the state-of-the-art in this area as well as the market perception evolution on this subject, during the last years.

To achieve the main study objectives the work will consider the most important and biannual survey performed by Moz SEO Company - one of the most prominent companies in SEO field (Mavridis & Symeonidis, 2015) - which considered 150 players in United States SEO marketing segment - as the primary source of data being collected. After analyzing and documenting these practical experiences the author also intends to make some brief forecast about future trends in this market.

2. LITERATURE REVIEW

2.1 SEARCH ENGINE DEFINITION AND BASIC FEATURES

it is important to define, on the first place, what is a Search Engine. SE can be designated as a particular software category that retrieves and stores data about public access websites. The referred data includes the Uniform Resource Locator (URL) of an internet page, tags and keywords that are related to the page content and the page codification structure itself including other external website links. This set of information are formatted in a specific type of index and then saved on a proprietary SE repository (Gudivada, Yalçin & Köse, 2014).

Based on the user search, the SE software compares the user performed search keywords with the contents of an index file that contains data about large quantity of sites. The search matches found are then returned to the user through the front-end navigator interface. The SE index is regularly updated by human editors or by automated programs. For a quick comprehension of how this work is done, in a first phase, the search engine discovers and validates the maximum number of public websites internal hyperlinks – which are links pointed to other public websites. In a second step, it builds the indexing of these linkages based on indicators of relevance and content characterization of each web page. At the final step, when a user performs a search, the SE compiles a ranking of positioning of all previously identified web pages for a specific keyword being searched, using complex, constantly updated algorithms, with the aim of returning the best results to the user, also considering all the used-defined search context (Kritzinger & Weideman, 2013).

In a more technical perspective, to guarantee a quick response of presentation results, minimizing requests to the SE database, some search engines save all or part of the pages and its contents in a cache mechanism. This cache also solves the problem in which search text has been already indexed and may be useful later when the content of the page was updated and the search terms are no longer stored in it. However, the cache of a search engine has the disadvantage of a link loss, that is, when the sites cease to exist or change their address, create results in pages that no longer exist and can lead the user to an unsatisfactory result. To solve this problem, the search engines has an indexing routine that demote all links that do not return a response validation of a server by canceling the indexing of those links in their database. To increase relevance of searches, these pages are kept in the SE database being strongly useful in caching, even with the fact that may retain data that may no longer be available in the internet (Killoran, 2013).

All the above scenario highlights the fundamental importance of the web pages being listed with the search engines. It explains the crucial strategy for any site owner to plan how visitors can find their way to their specific website. Furthermore, driven by constantly evolving information technologies, including mobile and analytics, Internet SE's have been either a powerful information seeking instrument and a suitable online marketing tool for different kinds of business activities and markets (Li, Lin,M., Lin,Z. & Xing, 2014).

2.2 SEARCH ENGINES OPTIMIZATION (SEO)

From the business viewpoint, search engines commercial role becomes clear. Appearing or not on user searches results pages can be the key factor to the success or decline of a firm. This seems to be the reality not only for companies which sell products directly in their e-commerce stores, but also for those business where customers use the internet to find their suppliers or service providers. At this point a new professional category raises importance and become more active – the *Search Engine Optimization (SEO)* workers which are responsible for the application of a set of techniques that have as major goal to make the websites more convenient for the SE's ranking algorithms helping them to achieve better classification in the SE's page results (Gudivada, Yalçın & Köse, 2014).

Regarding the players in SEO market, several organizations have defined their own metrics for the evaluation of web SE result pages from different purposes. It seems there is a good opportunity to try to consolidate information and start to build a solid common base of knowledge to be shared by this market stakeholders, which is one of the main purposes of this study (Mavridis & Symeonidis, 2014).

As quoted in the Introduction chapter, Borrell Associates Company (US) evaluation - shown in Figure 1 - points SEO-related budget in United States will be around US\$ 80 billion, by 2020. It is crucial to reinforce that the numbers graph below is not about advertise on web (paid announcements) but projected spends on SE optimization initiatives in the United States.

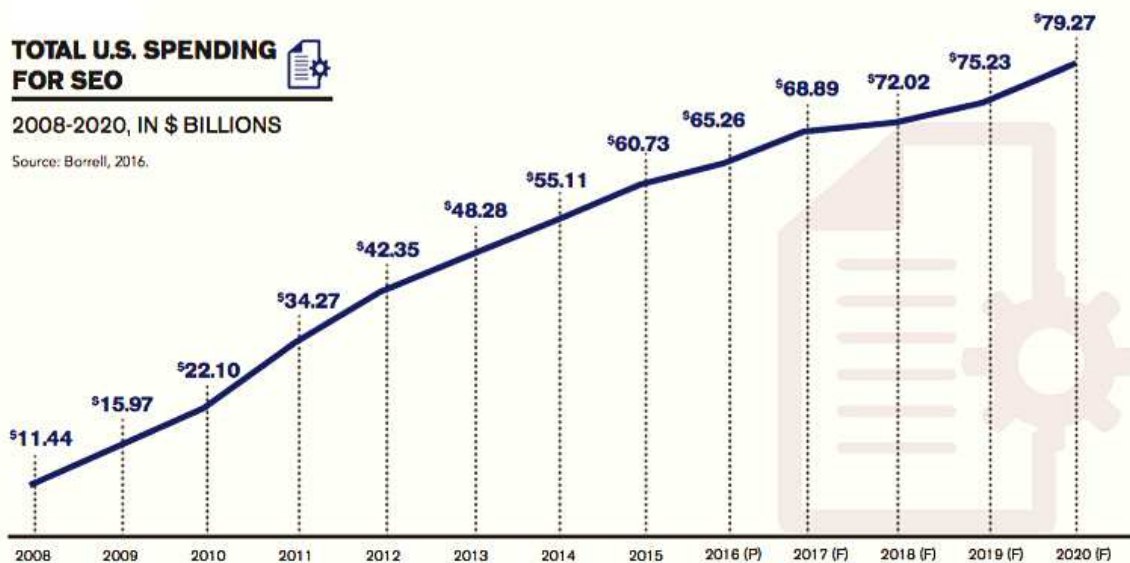


Figure 1 – Total US spending for SEO
Source: Borell, 2016

Despite of the great market values presented above, much of SEO efforts have limited chance of success. Usually, people tend to care only about the first few pages of search results. The low ranked pages in SE classification receive a small probability of attention. Therefore, companies had understood it is crucial to improve website ranking to increase website traffic flow using the SE optimization (SEO) techniques (Hui, Shigang, Jinhua & Jianli, 2012).

Unfortunately, while SEO is gaining much of the focus on online marketing industry, there have been few published academic researches works in this area. In fact, it is an arduous work to try to decipher this valuable secret kept by SE companies. The only way to get some conclusions is totally based on

exhaustive trial and error initiatives based on practical experiments or applying end user surveys data analysis as intended to be done in this practical case study (Li, Lin,M., Lin,Z. & Xing, 2014).

Finally, it is a common knowledge that the web experience has been continuously evolving. It reinforces exactly the major focus of this study which is to analyze and document information related to the billion-dollar giant market regarding SE ranking factors, intending to be a reference study for this kind of analysis and supporting the perception of this model evolution which is strongly based in practical use experiences (Hui, Shigang, Jinhua & Jianli, 2012).

2.3 Moz SEO COMPANY

One of the biggest players in SEO market - nominated Moz SEO Company - has defined its own metrics regarding the SE website ranking classification. These standard variables have been commonly used in the world as an instrument to evaluate the evolution of website page versions and to compare their SE result pages analysis (Mavridis & Symeonidis, 2014).

Moz SEO Company is a software as a service (SaaS) firm based in Seattle, Washington, U.S., that sells software subscriptions regarding inbound marketing and marketing analytics. Moz offers SEO tools which includes link building, keyword research, site audits, and page optimization insights to help companies to have a better overview of their web presence and how to improve their ranking in the search engines (Moz.com website, 2017).

The company also maintains a community of digital marketers and, every two years, surveys the opinions of several search marketers and performs correlation studies to better understand the search engine algorithms complexity.

3. METHODOLOGY

3.1 THE CASE STUDY METHOD

The fact that distinguishes the case study approach from the other research types is based on a method which focuses on the detailed and in-depth study, in its natural context, of a well-defined entity: the "case". The case study is the exploration of an in-time and depth limited system, through deep data collection involves multiple sources of rich information in a particular or specific context. It also reinforces the idea that case studies can be only developed based on theory building and testing (Ridder, 2017).

Another important characteristic to consider in the case study method is the case study generalization approach. This implies that If two or more case studies are intended to support the same supposition or situation being analyzed, replication can be applied. Moreover, the generalization advent lies in the fact that each single case may be considered as a unique observation. The higher the quantity of case studies that replicate the same experimentation being studied the greater will be the accuracy in which a new theory can be formalized. The case study is an empirical investigation task that studies a phenomenon in its natural environment particularly when the boundaries between the phenomenon and the context are not well defined and in which multiple sources of evidence are used. It is the most suggested research strategy when we want to know the " how "and the "why" of contemporary events on which the researcher has little or no control "(Yin, 2014).

Considering that all the above conditions are present in the current work, the author intends to apply the case study method as a useful approach to discover the SE ranking factors which allow websites to be better classified in the internet environment.

3.2 MOZ RANKING FACTORS BIENNIAL SURVEY

Moz's 2014/2015 Search Engine Ranking Factors Survey contains significant responses data about over a hundred marketing professionals of SEO market. This survey illustrates the most recent global Moz effort considering collected data of user searches performed in all USA states during 2015, May and its content provides quality information to the search marketing both now and in the future.

Furthermore, the data present in this study represents the respondent opinion concerning the various weighting of search engine ranking factors to be used (or not) in Google's search algorithm. The respondents rated the relative levels of influence of ranking factors on a scale of 1 (not influential) to 10 (highly influential). Moz also obtains this data to generate insights to the factors that may influence a website's visibility in search engines. In addition, Moz performs an extensive correlation study to determine which features of websites and webpages are associated with higher search rankings. "Moz surveyed over 150 leading search marketers who provided expert opinions on over 90 ranking factors" (Moz.com website, 2017).

It is important to note that these ranking factors are not "real proof" of what SE consider when ranking websites, but in a simple way they can point the web pages features that collaborate to higher ranking. Joining this known with both experience and understanding of search engine algorithms can help lead to better SEO practices (Mavridis & Symeonidis, 2014).

3.3 SUDY APPROACH AND METHODOLOGY

As an empirical study based strongly in survey application data over user experiences and perceptions, a data mining approach will be applied to collect, store, qualify and evaluate data about the SE ranking factors.

Data mining technology allows to extract implicit information from where people do not know in advance but potentially can be transformed in useful information. As referred in Introduction chapter, SE is some software category which collects data about websites. The related collected data is indexed and stored on a database. Furthermore, a production of software has a lot of affinity with data mining application techniques. In both scenarios, data analysis must be done to allow acquisition and automation of a new knowledge. Here, business domain may be understood to mean information that would facilitate new knowledge comprehension, in other words, it is not just documenting some existent knowledge being applied in the field of business, but also a process of knowledge discovery (Liu, Zhon & Zong, 2010).

3.3.1 Study Configuration

To achieve the targeted objectives using data mining techniques, four major phases (Figure 2) will be assigned to the method of this project, as follows:

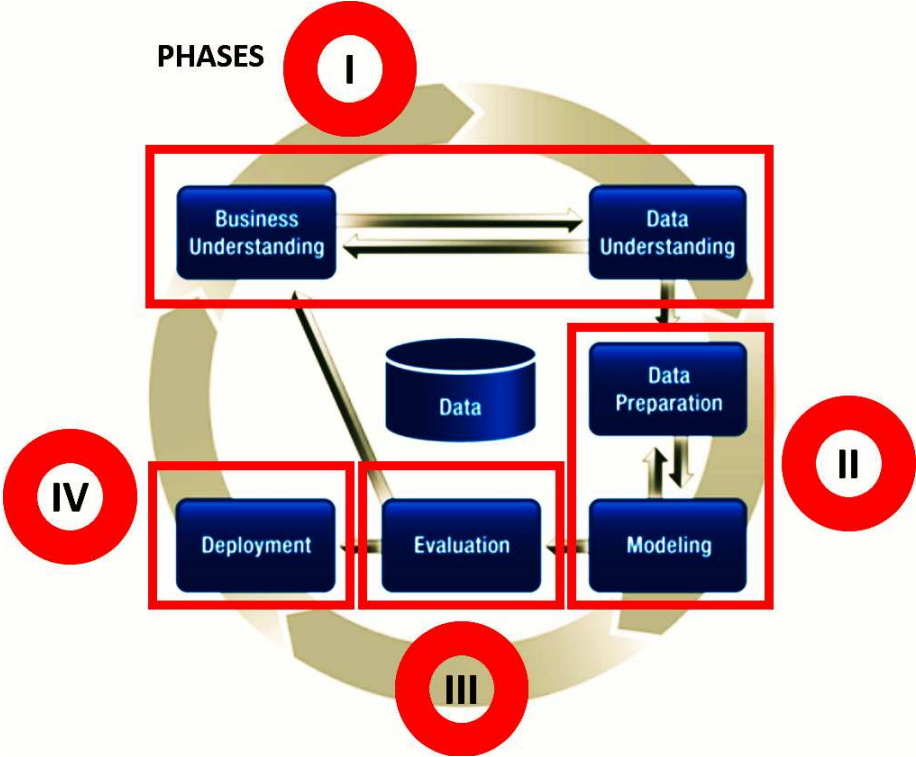


Figure 2 - Cross Industry Standard Process for Data Mining
Source: CRISP-DM Consortium, Aug 2006.

I) Understand SE business concepts and collect the most recent Google ranking factors data based on a large amount of responses regarding user experiences and registered in a biannual Survey performed by MOZ, Inc. (2014/2015 survey) and applied to 150 players of digital marketing segment.

II) Prepare and format the data to be loaded into a specific-purpose data model to allow the comparison between Google ranking factors attributes itself. The data model to be considered must have some variables to record the ranking factor attributes, as following:

a) The website identified ranking factor characteristics itself;

b) The Moz survey perceived importance of the ranking factor to the website internet positioning;

III) Evaluate and match the MOZ ranking factor variables against the recognized and established reference model of website usability (Visser & Weideman, 2011), presented below:

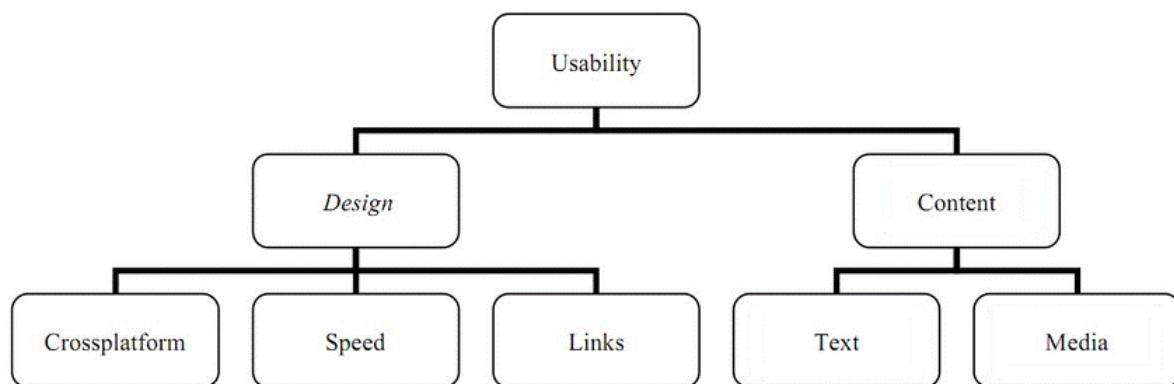


Figure 3 - View of website usability attributes
Source: Visser & Weideman, 2011

In the above presented model, a website usability is defined by the quality of the visitors' experiences using the website and their ability to interact successfully with the page information regarding the main purpose of their visit. The website usability is primarily supported by 2 subcomponents: *Design* which includes cross-platform technical features, website load speed and internal page links; *Content* which encompass text and media content features (Visser & Weideman, 2011).

The author will match the MOZ survey factors to the website usability classification (Design and Content) and will assign other complementary values, as "External" to the variables that are not part of internal website environment, being considered external-related to website boundaries. As part of this evaluation, some of these *External* factors are then assigned to a direct *Human* action or not. To better understand internal and external websites variables, it is important to consider the principle that nowadays new studies or methods are necessary to comprehend surfing behavior and the people habits on the internet (Egri & Bayrak, 2014). In addition, regarding external factors, the study will also evidence those elements which are related to human behavior over the website.

IV) Complement the analysis results with existent and dispersed literature articles and further studies in this area. Register and document the conclusions considering the state-of-the-art in this field also performing a brief forecast of the new approach to be used in this area by SE companies. It is

expected to verify obsolete ranking factors considerations also outline new trends about recently and increasing importance ranking factors being considered by SE algorithms.

3.3.2 Study Variables

This section lists the definition about the data that were collected by Moz Survey and its related category regarding SE specific features. The data are classified into 8 different master categories and distributed into particular tables as shown in Figure 4.

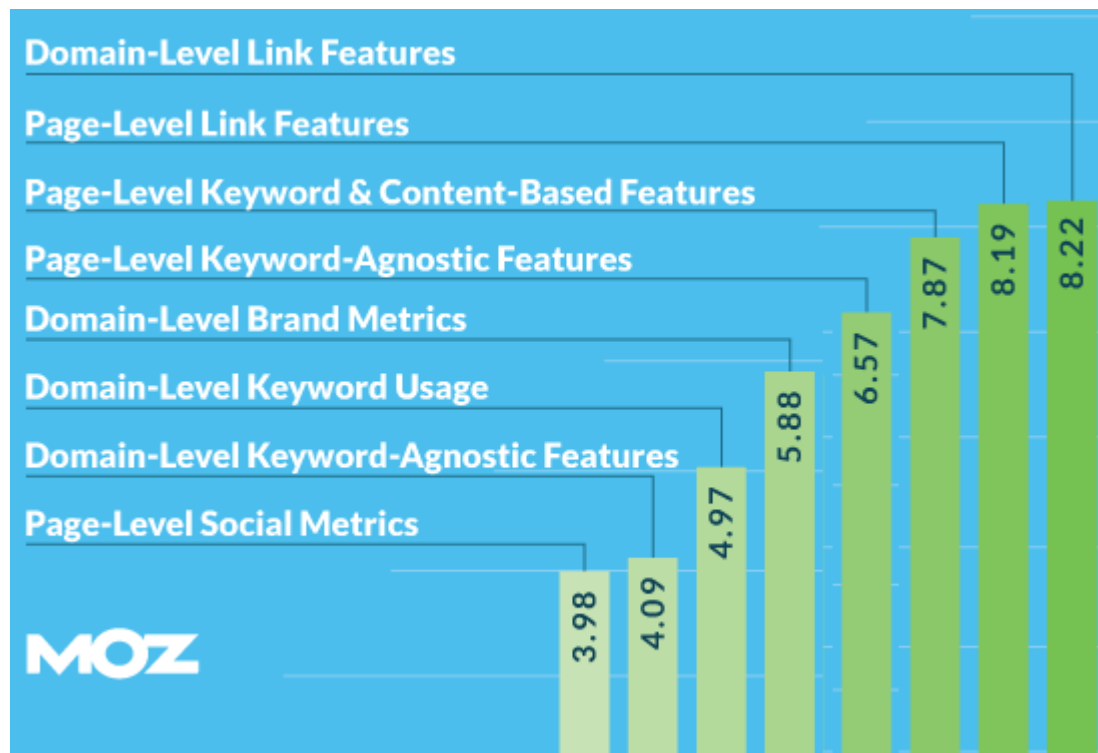


Figure 4 – Search Engine factors categories and its Influence on Google algorithm
Source: Moz Survey 2014/2015

The next pages show in more detail the explosion of each major category into its related subcomponents and its respective weight in ranking factor Google algorithm pointed by the respondents. The first category group was created by Moz just to allow an overview regarding the respondent's general perception of **Google's Broad Search Ranking Algorithm** and, further on, each category will be detailed in subsequent tables.

[1] Google's Broad Search Ranking Algorithm Data

Table 1 illustrates global variables – in a generic approach - that influence in the results of a global search query performed on Google. Each group is also categorized into related subcomponents that are listed in the succeeding tables below.

Variables of General Google Search Ranking by Moz	
Domain-Level, Link Authority Features	These metrics are based on link or referral citations such as quantity of quality links to this domain, also determinant to domain-level final page rank result
Page-Level Link Metrics	These elements include link metrics to the individual ranking page such as quantity of referral links and Page Rank. It is associated with quantity of external links to the page, anchor text distribution and quality of external links
Page-Level Keyword Features	These features describe use of the keyword term/phrase in particular parts of the HTML code on the page (title element, H1s, alt attributes)
Page-Level Keyword-Agnostic Features	These data relate to the search result page features despite of the explicit keyword being used on search and its semantical value. It includes page content length, use of https, use of structured data among others
Domain-Level Brand Metrics	These features describe elements that indicate qualities of global branding and brand metrics related to the root domain. Example: offline usage of brand/domain name, mentions of brand/domain in news/media/press, toolbar/browser data of usage about the site, entity association
Domain-Level Keyword Usage	These elements cover how keywords are used in the root or subdomain name, and how much impact this might have on search engine rankings to a particular query search
Domain Level Keyword-Agnostic:	These features relate to the entire root domain features despite of the explicit keyword being used on search and its semantical value. It includes domain age, responsive design and other general domain attributes
Page-Level Social Metrics	These features relate to third-party metrics from social media sources for the ranking page.

Table 1. Variables of General Google Search Ranking by Moz

Furthermore, for each group listed above the respondents evaluate each one of related subcomponents separately.

[2] Domain-Level Link-Authority Features

Table 2 shows page variables that are based on link or citation metrics such as quantity of quality links to this domain, which determine the page rank. In other words, it's relevant that your site has quality links coming into all of its pages not just the page that is the particular result of the search algorithm. Furthermore, Google will prioritize sites that are considered an "authority" in related market segment, and authority sites frequently have lots of "citations" to several pages in the same domain.

Variables related to Domain-Level Link-Authority Group
Quantity of unique linking domains to the domain
Topical relevance of linking domains
Raw popularity of the domain as measured by Moz
Trust of the domain from as measured by Moz
Distribution of linking domains' authorities/relative importance/popularity
Backlinks from sites of your own geotargeted area or language
Percentage of links with brand terms in the anchor text
Velocity of link acquisition to the domain
Sentiment of the external links pointing to the site

Table 2. Variables related to Domain-Level Link-Authority Group

[3] Page-Level Link Metrics

These features describe link metrics to the individual ranked page such as quantity of referral links and the particular page authority rank. In this case, Google will not consider the whole domain popularity, it will put the focus on the individual page citations over the internet. See Table 3 for a better comprehension about Page Level Link aspects.

Variables related to Page-Level Link Metrics Group
Raw quantity of links from high-authority sites
Topical relevance of linking pages
Topical relevance of linking domains
Diversity of link anchor text to the page
Raw quantity of links from known brands/entities to page
Raw quantity of unique linking domains to the page
Trust as measured by the distance from a trusted seed set of pages/sites
Position/context of inbound link
Popularity of the page as measured by algorithms (PageRank)
Link velocity of the page
Raw quantity of links that employ the keyword as partial-match anchor text
Raw quantity of links that employ the keyword as the exact-match anchor text

Table 3. Variables related to Page-Level Link Metrics Group

[4] Page-Level Keyword Features Metrics

Table 4 lists variables which describes use of the keyword term or phrase in parts of the HTML page code (title of the page, H1 elements, image alt attributes). Google rank websites not only based on their link popularity, in addition, Google have a kind of synonyms correlation and to rank well it is important to create topical content for pages not targeting only a single keyword but using related content in association with major keywords being searched.

Variables related to Page-Level Keyword Features Metrics Group
Keyword present in the title element
Keyword appears in the main content area of the page
Page contains close variants and/or synonyms of the keyword
Degree of optimization for a topic modelling algorithm
Keyword present in the page's URL
Keyword present in the anchor text of an on-page link
Keyword present in the alt attribute of an image on the page
Keyword present in specific HTML elements (bold/italic)

Table 4. Variables related to Page-Level Keyword Features Metrics Group

[5] Page-Level Keyword-Agnostic Features

These features relate to page attributes itself despite of the explicit keyword being used on search and its semantical value. As can be verified on Table 5, it includes page content length, Open Graph markup, https among others.

Variables related to Page-Level Keyword-Agnostic Features Group
Uniqueness of the content on the page (evaluated by Google)
Page is mobile friendly (for mobile rankings)
Relative CTR from Google SERPs (to the page for the keyword)
Page Load Speed
Quantity of searches for this keyword (+specific brand name, URL, or domain name)
Pure bounce rate of the page
Overall design and/or user experience
Long click metrics or dwell time (several clicks for the same user in the session reflected in the user time spent on the page)
Freshness of the content on the page
Return visits to this page after initial query/click

Length of content on the page
The age of the page
Average browse rate (after a click on this page in the SERPs)
Quality of supplemental content on page
Page contains Schema.org (or other structured data)
Reading level of the content on the page
Use of images on the page
Use of rich media (video, slides, etc.)
Page supports HTTPS / SSL

Table 5. Variables related to Page-Level Keyword-Agnostic Features Group

[6] Domain-Level Brand Metrics

These features describe elements that indicate qualities of internet global branding related to the root domain of a search result. See Table 6 for more detailed information.

Variables related to Domain-Level Brand Metrics Group
Search volume for the brand/domain
Existence/quality of verified real-world business info
Quantity of citations for the domain name across the web
Quantity of co-occurrence keyword + brand across web
Quantity of mentions of the brand/domain on social sites
Popularity of business's official social media profiles

Table 6. Variables related to Domain-Level Brand Metrics Group

[7] Domain-Level Keyword Usage Features

These features cover how keywords are used in the root or subdomain name, and how much impact this might have on search engine rankings to a particular query search. For a quick comprehension and according Table 7, Domain-Level Keyword Usage is simply whether or not your domain name is an exact match or partially contains your primary keywords.

Variables related to Domain-Level Keyword Usage Features
Keyword is the exact match root domain name
Keyword is present in root domain name
Keyword is closely related to domain name through entity association
Keyword is the subdomain name
Keyword is the domain extension

Table 7. Variables related to Domain-Level Keyword Usage Features Group

[8] Domain-Level Keyword-Agnostic Features

These features relate to the entire root domain despite of the explicit keyword being used on search and its semantical value. As shown on Table 8, it includes domain age, responsive design and other general domain attributes.

Variables related to Domain-Level Keyword-Agnostic Features Group
Uniqueness of content across the whole site (evaluated by Google internal algorithm)
Use of responsive design and/or mobile-optimized
Aggregated CTR from Google SERPs for the domain
Freshness of content on the site
Aggregated page load speed for pages of domain
Dwell Time or Long click metrics for domain (several clicks for the same user in the session reflected in the user time spent on the domain pages)
Quantity of error pages crawled on the site
Age of domain
Domain is associated with high-authority authors
Domain contains trust signal pages
Domain lists contact information
Quality of other sites hosted on the same block of IP addresses

Table 8. Variables related to Page-Level Keyword-Agnostic Features Group

[9] Page-Level Social Metrics

These features relate to metrics from social media sources for the ranking page. In other words, Page-Level Social Features consist of the interactions that occurs for that page across the internet major social media sites. In example, people likes and shares to the page in social posts, among others aspects presented on Table 9.

Variables related to Page-Level Keyword Features Metrics Group
Engagement with content/URL on social networks
Raw count of Google+ shares and +1s associated with URL
Raw count of Tweets associated with URL
Raw count of Facebook likes and shares associated with URL
Comments about the page on social sites
Sentiment of social links and citations referring to the page
Raw count of Pinterest pins associated with URL

Table 9. Variables related to Page-Level Social Metrics Group

4. RESULTS AND DISCUSSION

4.1 RESULTS

The data outlined below represent the opinions of marketing SEO professionals, respondents of the various weighting of factors thought to be used in Google's search algorithm. The respondents rated the relative levels of influence exerted by areas of ranking factors on a scale of 1 (not influential) to 10 (highly influential). After that, the respondents evaluate each one of components separately.

1) Google's Broad Search Ranking Algorithm Results

1. Google's Broad Search Ranking Algorithm	Value	Content	Design	Ext	Human
1.1 Domain-Level, Link Authority Features	8.22			X	
1.2 Page-Level Link Metrics	8.19			X	
1.3 Page-Level Keyword Features	7.87	X	X		
1.4 Page-Level Keyword-Agnostic Features	6.57	X	X	X	X
1.5 Domain Level Brand Metrics	5.88			X	X
1.6 Domain Level Keyword Usage	4.97			X	
1.7 Domain-Level, Keyword-Agnostic Features	4.09	X	X	X	X
1.8 Page-Level Social Metrics	3.98			X	X
1. TOTAL SCORE		18.53	18.53	41.90	20.52

Table 10. Google's Broad Search Ranking Algorithm respondent results

Table 10 Total Score illustrates that based on marketing respondents' perception about Google general ranking factors it seems that the Google algorithm is considering external website factors - *most of them related to human behavior patterns* - rather than internal contents or technical aspects of the page.

2) Domain-Level Link-Authority Features Results

2. Domain-Level Link-Authority Features	Value	Content	Design	Ext	Human
2.1 Quantity of unique linking domains to the root domain	7,45			X	
2.2 Topical relevance of linking domains	7,36			X	
2.3 Raw popularity of the domain as measured by MozRank, PageRank, etc.	7,15			X	
2.4 Trust of the domain from as measured byTrustRank, MozTrust, etc.	7,01			X	
2.5 Distribution of linking domains' authorities/ relative importance/popularity	6,95			X	
2.6 Backlinks from sites of your own geotargeted area or language	6,69			X	

2.7 Percentage of links with brand terms in the anchor text to the domain	6,25		X
2.8 Velocity of link acquisition (growth) to the domain	6,07		X
2.9 Sentiment of the external links pointing to the site	3,91		X
2. TOTAL SCORE			58.84

Table 11. Domain-Level Link-Authority Features respondent results

Table 11 Total Score shows that the website **Domain-Level Link-Authority Features** which were considered the top ranked factor based on marketing respondents regarding Google algorithm are also totally based on external website boundaries aspects.

3) Page-Level Link Metrics Results

3. Page-Level Link Metrics	Value	Content	Design	Ext	Human
3.1 Raw quantity of links from high-authority sites	7.78			X	
3.2 Topical relevance of linking pages	7.40			X	
3.3 Topical relevance of linking domains	7.26			X	
3.4 Diversity of link anchor text to the page	6.94			X	
3.5 Raw quantity of links from known brands/entities to page	6.92			X	
3.6 Raw quantity of unique linking domains to the page	6.84			X	
3.7 Trust as measured by the distance from a trusted seed set of pages/sites	6.69			X	
3.8 Position/context of inbound link	6.43			X	
3.9 Popularity of the page as measured by algorithms like PageRank, etc.	6.39			X	
3.10 Link velocity of the page	6.14			X	
3.11 Raw quantity of links that employ the keyword as partial-match anchor text	5.85			X	
3.12 Raw quantity of links that employ the keyword as the exact-match anchor text	5.75			X	
3. TOTAL SCORE				80.39	

Table 12. Page-Level Link Metrics respondent results

Table 12 Total Score illustrates that the website **Page-Level Link Features** which were considered the second well ranked factor based on marketing respondents' point of view regarding Google algorithm are strongly based on external website boundaries circumstances.

4) Page-Level Keyword Features Metrics Results

4. Page-Level Keyword Features Metrics	Value	Content	Design	Ext	Human
4.1 Keyword present in the title element	8.34	X			
4.2 Keyword appears in the main content area of the page	7.80	X			
4.3 Page contains close variants and/or synonyms of the keyword	6.95	X			
4.4 Degree of optimization for a topic modeling algorithm	6.70		X		
4.5 Keyword present in the page's URL	5.90		X		
4.6 Keyword present in the anchor text of an on-page link	5.38		X		
4.7 Keyword present in the alt attribute of an image on the page	4.56		X		
4.8 Keyword present in specific HTML elements (bold/italic)	4.16		X		
4. TOTAL SCORE		23.09	26.70		

Table 13. Page-Level Keyword Features Metrics respondent results

Table 13 Total Score presents that the website **Page-Level Keyword Features** are typically independent of external website boundaries aspects and are essentially related to the design and content of a page.

5) Page-Level Keyword-Agnostic Features Results

5. Page-Level Keyword-Agnostic Features	Value	Content	Design	Ext	Human
5.1 Uniqueness of the content on the page (evaluated by Google)	7.85	X			
5.2 Page is mobile friendly (for mobile rankings)	7.77		X		
5.3 Relative CTR from Google SERPs (to the page for the keyword)	6.92			X	X
5.4 Page Load Speed	6.60		X		
5.5 Quantity of searches for this keyword (+specific brand name, URL, or domain name)	6.20			X	X
5.6 Pure bounce rate of the page	6.15			X	X
5.7 Overall design and/or user experience	6.02		X		
5.8 Long click metrics (several clicks for the same user in the session)	5.91			X	X
5.9 Freshness of the content on the page	5.86	X			

5.10 Return visits to this page after initial query/click	5.64	X		
5.11 Length of content on the page	5.56	X		
5.12 The age of the page	5.42		X	
5.13 Average browse rate (after a click on this page in the SERPs)	5.37		X	X
5.14 Quality of supplemental content on page	5.34	X		
5.15 Page contains Schema.org (or other structured data)	5.14		X	
5.16 Reading level of the content on the page	4.89		X	
5.17 Use of images on the page	4.87		X	
5.18 Use of rich media (video, slides, etc.)	4.84		X	
5.19 Page supports HTTPS / SSL	4.59			X
5. TOTAL SCORE		30.25	35.29	40.56 30.55

Table 14. Page-Level Keyword-Agnostic Features respondent results

Table 14 Total Score shows that the website **Page-Level Keyword-Agnostic Features** are evenly affected by internal and external website aspects with a small predominance of external variables over the page content and design.

6) Domain-Level Brand Metrics Results

6. Domain-Level Brand Metrics	Value	Content	Design	Ext	Human
6.1 Search volume for the brand/domain	6.54			X	X
6.2 Existence/quality of verified real-world business info	6.33			X	X
6.3 Quantity of citations for the domain name across the web	6.36			X	
6.4 Quantity of co-occurrence keyword + brand across web	6.15			X	
6.5 Quantity of mentions of the brand/domain on social sites	4.35			X	X
6.6 Popularity of business's official social media profiles	3,99			X	X
6. TOTAL SCORE				33.72	21.21

Table 15. Domain-Level Brand Metrics respondent results

Table 15 Total Score illustrates that the website **Domain-Level Brand Features** are heavily affected by external website aspects, commonly considering human interaction.

7) Domain-Level Keyword Usage Results

7. Domain-Level Keyword Usage	Value	Content	Design	Ext	Human
7.1 Keyword is the exact match root domain name	5.83			X	
7.2 Keyword is present in root domain name	5.22			X	
7.3 Keyword is closely related to domain name through entity association	4.08			X	
7.4 Keyword is the subdomain name	3.82			X	
7.5 Keyword is the domain extension	2.55			X	
7. TOTAL SCORE				21.50	

Table 16. Domain-Level Keyword Usage respondent results

Table 16 Total Score illustrates that the website **Domain-Level Keyword Usage Features** are primarily affected by external website aspects, although not considering human interaction.

8) Domain-Level Keyword-Agnostic Features Results

8. Domain-Level Keyword-Agnostic Features	Value	Content	Design	Ext	Human
8.1 Uniqueness of content across the whole site (evaluated by Google internal algorithm)	7.52			X	
8.2 Use of responsive design and/or mobile-optimized	6.33		X		
8.3 Aggregated CTR from Google SERPs for the domain	6.24			X	X
8.4 Freshness of content on the site	6.22	X			
8.5 Aggregated page load speed for pages of domain	5.91		X		
8.6 Long click metrics for domain (user spent time on website)	5.76			X	X
8.7 Quantity of error pages crawled on the site	5.59			X	
8.8 Age of domain	5.37			X	
8.9 Domain is associated with high-authority authors	4.89			X	
8.10 Domain contains trust signal pages	4.71			X	
8.11 Domain lists contact information	4.57			X	
8.12 Quality of other sites hosted on the same block of IP addresses	3.89			X	
5. TOTAL SCORE		6.22	12.24	48.54	12.00

Table 17. Domain-Level Keyword-Agnostic Features respondent results

Table 17 Total Score shows that the website **Domain-Level Keyword-Agnostic Features** are strongly affected by external website aspects eventually depending of human interaction events.

9) Page-Level Social Metrics Results

9. Page-Level Social Metrics	Value	Content	Design	Ext	Human
9.1 Engagement with content/URL on social networks	3.87			X	X
9.2 Raw count of Google+ shares and +1s associated with URL	3.83			X	X
9.3 Raw count of Tweets associated with URL	3.45			X	X
9.4 Raw count of Facebook likes and shares associated with URL	3.16			X	X
9.5 Comments about the page on social sites	2,92			X	X
9.6 Sentiment of social links and citations referring to the page	2,90			X	X

9.7 Raw count of Pinterest pins associated with URL	2.77		X	X
9.8 Upvotes for the page on social sites	2.70		X	X
9. TOTAL SCORE			25.60	25.60

Table 18. Page-Level Social Metrics respondent results

Table 18 Total Score illustrates that the website **Page-Level Social Features** are primarily affected by external website aspects and - as a social variable - totally dependent of human interaction.

4.2 DISCUSSION

Based on marketing respondents' point of view and also on their perception about general ranking factors (Table 10, *External* and *Human* final scores) it seems that Google ranking factors are evolving from an internal and technical website perspective to a more complex approach, considering external website factors and much probably human behavior patterns. The discussion about the use of behavioral factors in SE ranking raised a few years ago and remains up-to-date. According to Google's patent named "Modifying Search Result Ranking Based On Implicit User Feedback" (2014), Google points the search users behavior as a crucial ranking determinant factor, so that if the user select a particular search result, it is considered to be relevant, or at least more relevant than other search results presented alternatives.

For other SEO study authors, user experience factors will keep increasing importance regarding the optimization to be done over the ranked pages, internal and external page links, and all the complete set of technical aspects of search engine result pages. In addition, user experience over the website is a key factor that Google will probably use to determine websites rank in SERP's. Furthermore, from a more technical viewpoint, it is quite capable of digital platforms to identify users and to drive significant opportunities for digital marketers.

The points listed in the next pages show a correlation between Moz survey results analysis and other SEO market studies in order to answer to the main questions and objectives established in this work.

4.2.1 Click-Through Rate (CTR)

According to respondents (Table 14, item 5.3) one metric that is gaining importance to measure the user engagement to a website is the Click-through rate (CTR). A search engine results page CTR is known as the proportion of the number of times a search listing was clicked over the number of times it was presented to the user.

Google's former Search Quality Chief, Udi Manber, testified the following: "The ranking itself is affected by the click data. If we discover that, for a particular query, hypothetically, 80 percent of people click on Result No. 2 and only 10 percent click on Result No. 1, after a while we figure probably Result 2 is the one people want. So we'll switch it." (Ryan & Jones, 2016).

4.2.2 Time Spent on the Site (User Dwell Time and Bounce Rate)

Respondents have highlighted the content of the page as the top internal website rank factor (Table 14, item 5.1). It is consistent with recent field study initiatives that focus on understand which factors contribute to “keep a user on the site”. From this perspective the web page content is considered for many authors as the protagonist of user engagement over a website. It means that providing right content within a right time is the key role and it is gaining so much importance nowadays.

Recently studies try to explore new frameworks of customer engagement that integrates the elements of multiple content formats and realizing customer engagement toward superior results or innovations. These frameworks consider new media’s different information service and technologies as search engine, social recommender and social media (among others) that can be properly arranged to achieve a virtuous customer engagement circle (Yuan,Chou,Yang,Wu & Wuang, 2017).

As a result of the content-driven approach, and still according to the Google patent specifications, a more recent and gaining importance factor that Google can use to modify rankings of search results is nominated *dwell time*. In few words, dwell time is the total time that a user spends on a page after clicking on a specific search result and before returning to the original set of SERP’s. Once more, Google's patent on “Modifying Search Result Ranking Based On Implicit User Feedback” (2014) points the following user information may be used to rank pages:

"The information gathered for each click can include: (1) the query (Q) the user entered, (2) the document result (D) the user clicked on, (3) the time (T) on the document [...]. The time (T) can be measured as the time between the initial click through to the document result until the time the user comes back to the main page and clicks on another document result. In general, an assessment is made about the time (T) regarding whether this time indicates a longer view of the document result or a shorter view of the document result, since longer views are generally indicative of quality for the clicked through result."

4.2.3 Social Media

Although MOZ Survey respondents have not highlighted social media as a key factor on SE ranking algorithm (Table 10, item 1.8), it is possible to find several authors who defend its importance. Smith (2016) and Smith-Ditizio (2018) reinforce the practical implications of search engine technologies applied within a social media environment that has propitiated the development of a modern, user-driven internet experience to satisfy user needs and engagement. For those authors, individuals are increasingly taking advantage of some tools to better control their experiences and what they intend to share. Furthermore, the transition to newer forms of integrated marketing, the future for search engines as marketing tools by social media users appears to be strongly promising in adding contextual content within user homepage or blog.

For Pinsky (2018), social media interaction is important to search ranking. For this author, Google is focusing on certify users are getting the most relevant content possible and to assess relevance, Google may investigate how content is debated and shared in social. In addition, marketing people should connect social influencers and promote them to share the marketing target content on social media.

In fact, social media channels have been frequently considered an important pillar for most modern digital marketing strategies nowadays. During some time, digital marketing studies pointed that all businesses should be heavily present on most of the main channels notably as Twitter, Facebook,

Instagram and LinkedIn. Nonetheless, it is highly important to note that not all tools of social media are suitable for all business categories (Ahmed & Shabbir, 2013).

Recently, the general way of thinking is that companies and businesses should be more active where your customers really are. Much coupled with social media, it is important to realize the role of user reviews. Reviews not only help people make purchasing decisions at the right time, but positive reviews can also help with your search engine rankings. Nowadays, internet business experience allows customers to watch videos, read articles and blog posts, interact with website content, ask their friends on social media and read reviews before they figure out what product, service or business to experiment (Antoci & Bonelli, 2018).

4.2.4 Voice Search

Still according with respondents' perception of SEO human behavior patterns raising importance, some industries point the voice search as a new variable to be considered in this scenario. Although frequent user searches consist of one to four or five words in a phrase, voice searches tend to encompass full sentences. That new approach requires also a specific method of structuring content. It is recommended that emerging digital marketing strategies must focus on content that can satisfy those voice queries. It is crucial to discover how your target audience really communicate. Discover what words do they use and optimize for that language, that sound, and those words (Slivka,2018).

4.2.5 Technical Factors Still in the Game: The Mobile Paradigm

It is imperative to note that the website design features were still classified by the survey respondents as a high-influence factor on page ranking algorithms together with website page content (see Table 10, items 1.3 and 1.4 and Table 14, item 5.2 and 5.4).

Since April 2015 Google had publicly declared it would give preference in its result pages that were adapted to the mobile environment. Google intention is to promote the sites that provide the best user experience in the top places. In this respect, search engines, and most particularly Google, have made their position quite clear: responsive design represents the optimal solution (Codina & Pérez-Montor, 2017).

For many authors, if marketing professionals intend to maintain their target audience, they should make sure all the web content exists on the mobile page versions of their websites. More than that, it is totally advisable that websites with separate URLs for desktop and mobile experiences must have their desktop content mapped to their mobile URLs in a one-to-one correlation to potentialize successful initiatives (Slivka,2018).

4.2.6 Combining Technology and Human Behavioral Patterns

The responses of Table 10 and particularly the item 1.4 exposes the power of human and technological factors combination in achieving top ranking SE positions for a website.

In fact, the present context for understanding SEO ranking factors is becoming increasingly complex. Although it was not explicitly mentioned in the MOZ survey respondent results, technology improvements in Artificial Intelligence had allowed business firms to get significant more information regarding their users, and it is also transforming the way that SE performs. The recent implemented Google RankBrain Project is a machine-learning artificial intelligence system that allows Google

process its search results and the firm had already declared that the system is also used to help rank web pages. Among other functionalities, this technology was projected to learn and predict user behavior which helps Google to receive and respond to millions of queries daily. In a Practical way RankBrain helps Google to interpret both content and intent in some particularities. Although Google has given few information about how AI specifically works in this project, company announced that it has become a relevant ranking flag for a web page. This has placed an even greater emphasis on creating content that matches the user intent (Shantam, Kumar & Tiwary, 2017).

In addition, based on a complex set of business rules and constraints, prescriptive analytics can automatically improve prediction accuracy and best decision choice scenarios. In prescriptive analytics approach a set of techniques computationally identifies several alternative actions to be taken by management, given their complex objectives and limitations (Appelbaum, Kogan, Vasarhely & Yan, 2017). For different reasons such as complexity, protected patent rights and trade secrets, it may not be possible to completely understand what search algorithms are looking forward, how they are evolving, and the real intention supporting them. Thereby, it seems that significant reliance is placed in those who develop and monitor search engines. Assuming a position of unconditional acceptance of SE algorithms suggest some risks. Being awareness of biases may already exist now or in the future is at least prudent (Cleverley, 2017).

At this point we may think about all of these human behavioral patterns being monitored and also the ethics concerns about SE market methods and trends. Whether we want or not, each of us is already a part in this experiment context. It seems that SEO algorithms no longer just help us find what we know but can also suggest what we don't know. This is the fundamental of prescriptive analytics (Bertsimas and Kallus, 2014) which answers the question of what should be done given the descriptive and predictive analytics results.

Together with those ethics concerns, it is growing the debate about the limits of digital marketing. Therefore, this is the main reason of recent European Union implemented General Data Protection Regulation (GDPR) which is a regulation on data privacy and protection for all individuals within the European Union considering the aspects of privacy and safety regarding the use of collected personal data. It seems to be urgent to limit human liability for artificial intelligence. Although to stimulate innovation in artificial intelligence may prove to be beneficial, such improvements must be considered alongside legislation defining socially acceptable limits and controls over the application of artificial intelligence so that providing effective rights of redress for individuals and groups that may suffer harm (Butterworth, 2018).

5. CONCLUSIONS

Based on the presented survey results study it is possible to observe that Google ranking factors are evolving from an internal and technical website perspective to a more complex approach, considering external website factors and much probably human behavior patterns. Furthermore, according to Google's patent named "Modifying Search Result Ranking Based On Implicit User Feedback" (2014), Google points the search users behavior as a crucial ranking determinant factor, so that if the user select a particular search result, it is considered to be relevant, or at least more relevant than other search results presented alternatives.

In addition, global user experience over the website is a key factor that Google will probably use to determine websites rank in SERP's. Moreover, it was also important to note that for other SEO study authors, user experience factors will keep increasing importance regarding the optimization to be done over the ranked pages, internal and external page links, and all the complete set of technical aspects of search engine result pages.

Moreover, SEO algorithms no longer just help us find what we know but can also suggest what we don't know. Together with those ethics concerns, it is growing the debate about the limits of digital marketing. The author intends that due the lack of scientific works and documentation in this area together with the powerful influence of SE to business and society behaviors, the presented study data and considerations raise an important forum of discussion that affects whole society. Whether we want or not, each of us is already a part in this experiment context. Admitting to stimulate innovation in artificial intelligence may prove to be beneficial, such improvements must be considered alongside legislation defining socially acceptable limits and controls over the application of artificial intelligence.

Limitations and Recommendations for Future Works

Although the work has considered a significant sample of research respondents, the results analyzed here are limited to the opinion of the respondents of the study research. Due to lack of literature and scientific knowledge in this area, all presented work was based on empirical efforts. At this point, it is also important to emphasize that correlation is not causation. In such a way, new studies on the subject can even deepen the knowledge here developed.

As the market studied comprises figures of great business value, obviously Google does not publicly reveal its ranking secrets and search engines are updated frequently to adjust to the advancement of SEO professionals' capabilities and to the market. In addition of being a fully dynamic environment which is usually part of common human daily tasks, it seems extremely necessary, besides understanding the evolution of the SE technical aspects, also to develop ethical and moral studies regarding the accepted limit boundaries of search mechanisms performance, as well as considering recent data protection efforts and legislation.

BIBLIOGRAPHY

- Ahmed, I., Shahzad, R. K., Kashif-ur-Rehman, & Shabbir, J. (2013). Search engine optimisation: Evidence from Pakistan. *Asian Academy of Management Journal*, 18(2), 1–16.
- Amber A. Smith-Ditizio, Alan David Smith, Walter R. Kendall, (2018) "Integrating search engine capacity and gender preference within a social media captured authority locus of control", *Benchmarking: An International Journal*, Vol. 25 Issue: 2, pp.374-399.
- Antoci, Angelo & Bonelli, Laura & Paglieri, Fabio & Reggiani, Tommaso G. & Sabatini, Fabio, (2018), *Civility and Trust in Social Media*, IZA Discussion Papers 11290, Institute for the Study of Labor (IZA).
- Appelbaum, Deniz, Kogan, Alexander, Vasarhelyi, Miklos and Yan, Zhaokai, (2017), Impact of business analytics and enterprise systems on managerial accounting, *International Journal of Accounting Information Systems*, 25, issue C, p. 29-44.
- Baye, M. R., De los Santos, B., & Wildenbeest, M. R. (2016). Search Engine Optimization: What Drives Organic Traffic to Retail Sites? *Journal of Economics and Management Strategy*, 25(1), 6–31.
- Butterworth, Michael. (2018). The ICO and artificial intelligence: The role of fairness in the GDPR framework. *Computer Law & Security Review*. 34. 10.1016/j.clsr.2018.01.004.
- Codina, Lluís & Pérez-Montor, Mario. (2016). *Navigation Design and SEO for Content-Intensive Websites: A Guide for an Efficient Digital Communication*.
- Coulibaly, S. K., Erbao, C., & Mekongcho, T. M. (2018). Technological Forecasting & Social Change *Economic globalization , entrepreneurship , and development*, 127(October 2017), 271–280.
- Dick, M. (2011). Search Engine Optimisation in Uk News Production. *Journalism Practice*, 5(4), 462–477.
- Egri, G., & Bayrak, C. (2014). The role of search engine optimization on keeping the user on the site. *Procedia Computer Science*, 36(C), 335–342.
- Gudivada, V. N., Rao, D., & Paris, J. (2015). Understanding Search-Engine Optimization. *Computer*, 48(10), 43–52.
- Killoran, J. B. (2013). How to use search engine optimization techniques to increase website visibility. *IEEE Transactions on Professional Communication*, 56(1).
- Kritzinger, W. T., & Weideman, M. (2013). Search Engine Optimization and Pay-per-Click Marketing Strategies. *Journal of Organizational Computing and Electronic Commerce*, 23(3), 273–286.
- Leeflang, P. S. H., Verhoef, P. C., Dahlström, P., & Freundt, T. (2014). Challenges and solutions for marketing in a digital era. DOI: 10.1016/j.emj.2013.12.001.
- Li, K., Lin, M., Lin, Z., & Xing, B. (2014). Running and chasing - The competition between paid search marketing and search engine optimization. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 3110–3119.

- Liu P., Zhou Y. & Zong P., (2010). Data selection for statistical machine translation," Proceedings of the 6th International Conference on Natural Language Processing and Knowledge Engineering), pp. 1-5. doi: 10.1109/NLPKE.2010.5587827.
- Mavridis, T., & Symeonidis, A. L. (2014). Semantic analysis of web documents for the generation of optimal content. *Engineering Applications of Artificial Intelligence*, 35, 114–130.
- Mavridis, T., & Symeonidis, A. L. (2015). Identifying valid search engine ranking factors in a Web 2.0 and Web 3.0 context for building efficient SEO mechanisms. *Engineering Applications of Artificial Intelligence*, 41, 75–91.
- O'Connor, P. (2009). Pay-per-Click Search Engine Advertising: Are Hotel Trademarks Being Abused? *Cornell Hospitality Quarterly*, 50(2), 232–244.
- Ridder, H. (2017). The theory contribution of case study research designs. *Business Research*, 10(2), 281-305. doi:10.1007/s40685-017-0045-z.
- Ryan, D., & Jones, C. (2016). Understanding digital marketing: Marketing strategies for engaging the digital generation.
- Shantam, Kumar, H., & Tiwary, D.R. (2017). Analysis of Rankbrain Algorithm Using Machine Learning.
- Silva, N., & Aguiar, A. (2014). Optimização de sítios web para motores de busca: Um estudo empírico. *Iberian Conference on Information Systems and Technologies, CISTI*.
- Slivka, A. (2017). How Mobile And Voice Will Drive SEO Engagement In 2018. *Forbes*, 1.
- Soe-Tsyr Daphne Yuan, Szu-Yu Chou, Wei-Cheng Yang, Cheng-An Wu, Chih-Teng Huang, (2017) "Customer engagement within multiple new media and broader business ecosystem – a holistic perspective", *Kybernetes*, Vol. 46 Issue: 06, pp.1000-1020, <https://doi.org/10.1108/K-01-2017-0042>.
- Szetela, D., & Kerschbaum, J. (2010). Pay-Per-Click Search Engine Marketing - An Hour A Day. Optimization.
- Visser, E. B., & Weideman, M. (2011). Search engine optimisation versus Website usability-conflicting requirements? *Information Research*, 16(3).
- Visser, E. B., & Weideman, M. (2014). Fusing website usability and search engine optimisation. *SA Journal of Information Management*, 16(1), 9 pages.
- Wu, H. (2011). Search engine optimization of e-commerce websites. *International Conference on Management and Service Science, MASS 2011*.
- W. T. Kritzinger & M. Weideman (2013) Search Engine Optimization and Pay- per-Click Marketing Strategies, *Journal of Organizational Computing and Electronic Commerce*, 23:3, 273-286, DOI: 10.1080/10919392.2013.808124
- Yalçın, N., Köse, U., Gudivada, V. N., Rao, D., Paris, J., Egri, G., Weideman, M. (2014). Running and chasing - The competition between paid search marketing and search engine optimization. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 16(1), 3110–3119.

Yalçın, N., & Köse, U. (2010). What is search engine optimization: SEO? *Procedia - Social and Behavioral Sciences*, 9(July 2009), 487–493.

Yin R. (2014). *Case Study Research Design and Methods* (5th ed.). Thousand Oaks, CA: Sage. 282 pages.. *The Canadian Journal of Program Evaluation*. 10.3138/cjpe.30.1.108.

Zhou, H., Qin, S., Liu, J., & Chen, J. (2012). Study on website search engine optimization. *Proceedings - 2012 International Conference on Computer Science and Service System, CSSS 2012*, 930–933.

Z. Hui, Q. Shigang, L. Jinhua and C. Jianli, "Study on Website Search Engine Optimization," 2012 International Conference on Computer Science and Service System, Nanjing, 2012, pp. 930-933. doi: 10.1109/CSSS.2012.236.