IMPROVE SEARCH ENGINE OPTIMIZATION

Case study: Production Software

Lahti University of Applied Sciences
Degree programme in Business Information Technology

NGUYEN THI THUY LINH:　　　　　Improve Search Engine Optimization
　　　　　　　　　　　　　　　　　Case study: Production Software

ABSTRACT

Search Engine Optimization (SEO) is a collection of strategies, techniques used to optimize the web pages or whole site to make it friendlier with search engine and improve the ranking of this site in results pages. A research of Conductor (2014) found that 70% of internet users use search engines to find and access websites. Therefore, implementing SEO factors and methods in building the website of the company helps to increase the online presence of the firm to potential clients.

The purpose of this study was to describe how to use analyzed data to improve SEO in order to attract internet traffic and increase the ranking of a website. The key SEO factors and methodologies often used in building websites also are described in this study.

Both qualitative and quantitative research method were used to analyze the case study. Data was collected by observing the case company website. After that, the data was analyzed it to find out the effect of SEO on the website and give some suggestion in order to improve the website ranking.

According to the analysis results, the researcher concluded that the case company applied many SEO factors and methods to increase its ranking in search result page. Some SEO factors should be modified and other techniques should be applied to push its website ranking on the top of the SERPs.

Keywords: online marketing, search engine, On-Page SEO, Off-Page SEO, White Hat SEO, Black Hat SEO, Grey Hat SEO

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

HTML	Hypertext Markup Language
URL	Uniform Resource Locator
SEO	Search Engine Optimization
SEOs	Search Engine Optimizers
SERP	Search Engine Result Page
SEM	Search Engine Marketing
ROI	Return on Investment
DMCA	Digital Millennium Copyright Act

1 INTRODUCTION

Rajendran V. (2001) claimed that owning a website brought invaluable benefits to the business. The customer can find companies easily when searching for products and services they need on the Internet instead of reading any print advertisements or watching commercials. Websites help companies saving a huge cost on print brochures, coupons, and newsletters. Moreover, it is quick and efficient to update real time information on the website. Customers can also receive up-to-date information about your products and services. Another advantage is the website is available to the visitors anytime, anywhere, and any place. Through the website, the company can market its brand and products outside its located country no matter the size of the business.

In addition, almost people search information about products and services by surfing on the major search engines such as Google. The report in January 2015 of Statista showed that Google has dominated the search engine market with 88.2 percent market share. If companies want to succeed in the online business, their websites need to be found in the major search engines easily. In other words, in order to have an advantage competitive, companies need to optimize their websites so that the online presence of companies can be enhanced.

1.1 Statement of the problem

Internet Live Stats researched and reported that the number of worldwide internet users as of July 2014 was nearly 3 billion (around 40% of the world population), increased 7.9% compared to the previous year. In 1995, only 0.78% of the world population had internet connection. This number had increased five times during the previous decade. The internet has been becoming an integral part of human life. (Bargh, J. and McKenna, K., 2004)

Millions of people have used the search engine to find products, services and information every day. A study of the Chitika team in 2013 showed that Page 1 in

the SERP of Google contributed 92% of all traffic from the average search, and the first position receives 33% of the traffic. The clients only focus on Page 1 of search results while millions of results are returned when a search queries. The Search Engine User Behavior Study of 2639 users of iProspect (2006) revealed that 88% of users never look past first three pages of SERPs. As the results, the study claimed that if the website is not in the first three pages, customers might not even know that this site exists.

There are millions of websites on the Internet today, and this number has still continued increasing over time. In order to have an advantage competitive, a company has to ensure that potential clients can find its website before going to competitors' websites and its advertising has reached enough potential customers. The website is considered to be important and beneficial for companies. (Rodgers, 2002)

Search Engine Optimization (SEO) is considered as an effective methodology to help website increase its ranking in the SERPs of the search engines so that potential customers can find the website easily and quickly. It is a collection of strategies, techniques used to optimize the website to make it friendlier with search engine and increase the traffic.

Advertisement and sales via the website are the best choice for the small and medium business. It has many advantages compared to the traditional advertising methods such as low cost, high effective, and flexibility. SEO is a necessary activity with any websites to help them become well-known and reach the potential customers.

However, the algorithms of search engines are constantly updated. Moz (2015) revealed that Google updated algorithms of the search engine around 500–600 times each year. Some changes can impact the rank in the SERPs significantly, for example, Google Panda, Google Penguin, and Google Hummingbird. Thus, SEO is also needed to modify and change in order to the websites still maintain on the top lists of the SERPs. The Statistical SEO Tools help the webmaster have a gen-

eral view of the current SEO health to adjust it correctly. So, the question is: "How to use measured data to improve SEO of the website?"

1.2   Thesis structure

This thesis consists of nine sections. The first section describes the research problem and discusses the contextual background and the solution history of the problem briefly. Several basic concepts of SEO are also included in this chapters. Followed this, the objective of this study and the structure of this study are determined.

In chapter two, the researcher delineates the research method used in this study to approach these goals. The research question is identified, and every stage of the research process is described, including establishing a research model, collecting data and analyzing the collected data.

Chapter three explores the search engine, including the structure of the search engine and how it works. The author also introduces the page rank system via two popular page ranks: Google PageRank and Alexa rank. Two rank systems are important measures for the estimation of the website.

The literature review of SEO is outlined in chapter four. The definition of SEO is determined, and the significance SEO factors, methods, and technique are mentioned. Two SEO methodologies "White Hat" and "Black Hat" are described. Some tactics in each category are introduced.

After that, the researcher applies the knowledge to optimize the website of the case company, Production Software. Chapter five introduces briefly about the case company and the achieved goal. The common SEO factors and methods are implemented in order to attract internet traffic and place this website on top ten of Google results pages with keyword "Valmistuksenohjausjärjestelmät" and "Valmistuksenohjaus".

After implementing SEO, the research collect data by using online analysis tools. The collected data is provided in chapter six.

In chapter seven, the collected data is analyzed in details to find out the success and failure of the current SEO. The researcher also proposes some suggestion to improve SEO of the website.

Followed that, chapter eight concludes the works and the results achieved in the previous sections.

The final chapter discusses the limitations when conducting the research. Reliability and validity of this study are also considered.

2 RESEARCH METHOD

2.1 Research question

This thesis aims to suggest some idea to modify the used SEO factors, methods in order to increase the website ranking of the case company on SERPs. Through improving the website rank, the efficiency of online marketing is enhanced, and the enterprise can reach more potential customers. Depend on this objective, the research question of this study is: How to use measured data to improve the SEO of the website so that it will be maintained on the top lists of the SERPs?

2.2 Research approach

Two distinct methods of research approaches often referred is the deductive and inductive approaches.

Depend on Trochim W. (2006), deductive reasoning starts with a general theory, applies that theory in a specific case to gain evidence and draws a specific conclu-

sion about this case from those premises. This method is informally called a "top-down" or "waterfall" approach. The inductive approach works in the opposite way. It generates a broader generalizations and theories based on specific observations. This method is also called a "top-down" or ""hill-climbing" approach. Figure 1 describes briefly how deduction and induction work between the general and the specific.
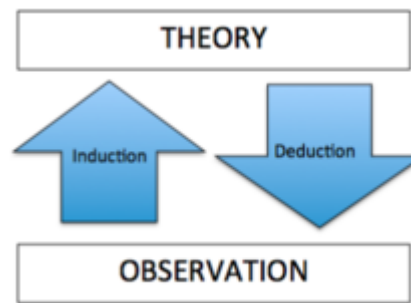
Figure 1: Deduction and induction (Trochim W., 2006)

In this study, the general key SEO factors and methods are identified through researching the resources. Afterward, they are used to compared to the currect SEO situation of the case company. Therefore, the research approach of this study is deduction, and the nature setting is the explanation.

Research methods in social sciences are often divided into two main approach: quantitative and qualitative methods.

Quantitative approach is described as a method explaining phenomena by collecting numerical data that are analyzed using mathematically based methods (in particular statistics) (Aliaga and Gunderson, 2000)).

According to Denzin N. and Lincoln Y. (1994), qualitative research is defined as followed:

> *"Qualitative research is multimethod in focus, involving an interpretive, naturalistic approach to its subject matter. This means that qualitative researchers study things in their natural*

settings, attempting to make sense of, or interpret, phenomena
in terms of the meanings people bring to them.”

Qualitative research method is a method of inquiry that is used in many different disciplines. Qualitative research method aims to seek an in-depth understanding of reasons and motivations, to explore and describe the studies and develop the theories. This method investigates the 'why' and 'how' questions, not just 'what', 'where' and 'when' questions. It is used when the researcher does not know what to expect or want to go deeper into issues of interest. The sample using this method is usually a small number of non-representative cases.

The quantitative approach provides more convincing evidence for the effectiveness and impact of the phenomena to the environment. Meanwhile, the qualitative method is often considered as more flexible, more thoroughly when exploring a topic and more appropriate to collect data to answer these questions.

A mixed-method approach is a method integrating a mix of both qualitative and quantitative research techniques. It provides more coherent, logical and reliable evidence by comparing results of different methods, complementing the strength of each technique and avoiding bias. (Heyse L., Zwitter A., Wittek R., Herman J., 2014)

According to the research objectives, this study uses the mix method to evaluate the theoretical foundation of the study. Qualitative technique is used to understand the background, identify topics, and analyze theoretically the collected data from the case study involving the investigation of a particular phenomenon, to understand deeply into the SEO elements. In addition, quantitative method is used to interpret empirical data collected via statistical analysis tools to convince the effectiveness and impact of SEO factors and methods onto the website.

Qualitative data is collected from many sources, for example, interviews and questionnaires, documents and texts, group discussions, observation and reflection field notes, pictures, and other materials. In this method, data can be collected by

a range of approaches. Methods for collecting data which are used widely are Participant Observation, Non-participant Observation, Field Notes, Reflexive Journals, Structured Interview, Semi-structured Interview, Unstructured Interview, and Analysis of documents and materials (Marshall, Catherine & Rossman, Gretchen B., 1998). In participant observation method, the researcher spends time with the group in their environment, or in the environment being studied to gain familiarity with the research subject.

> *"Qualitative researchers accomplish this through observation alone or by both observing and participating, to varying degrees, in the study community's daily activities. Participant observation always takes place in community settings, in locations believed to have some relevance to the research questions."*
>
> *(Mack N., Woodsong C., Macqueen K., Guest G., Namey E., 2005)*

Gill J. and Johnson P. (2010) develop a fourfold categorization of the role the participant observer can adopt as following:
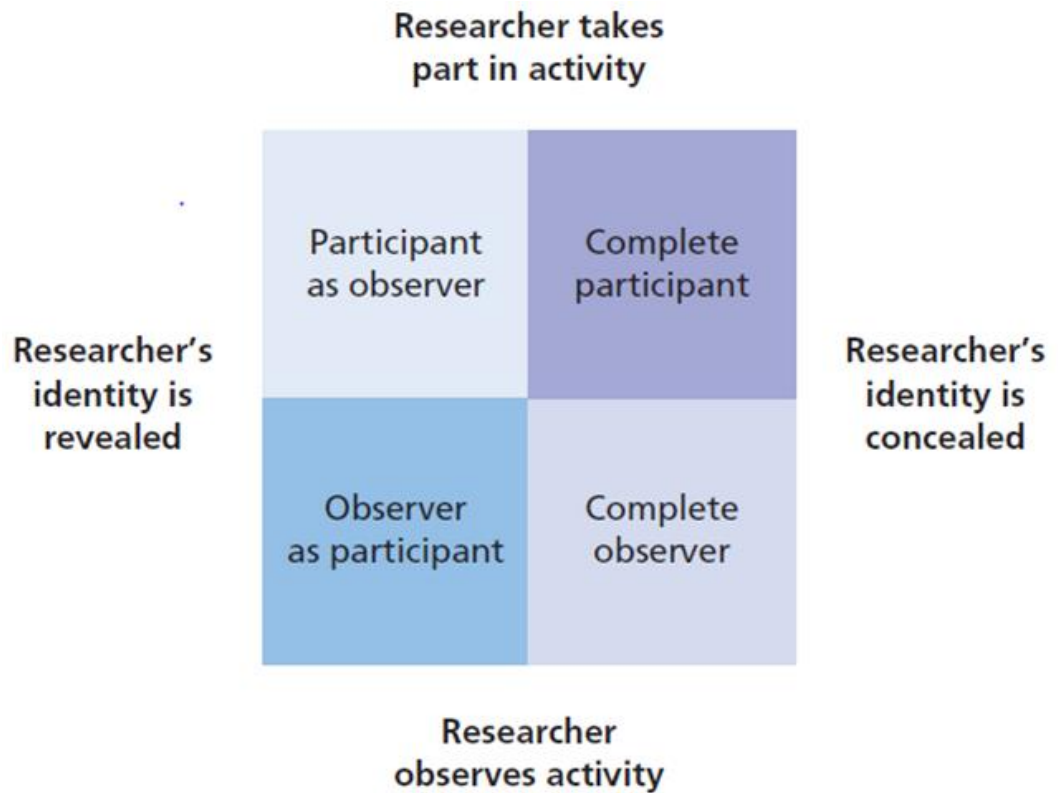
Figure 2: Participant observation research roles (Gill J. & Johnson P., 2010)

In this research, the researcher collected data during the building website process for the case company. The purpose of this study is revealed to those participants in the research setting. Therefore, depend on the model of Gill and Johnson, the role of the researcher is participant as observer.

> *"Data analysis is the process of bringing order, structure and meaning to the mass of collected data"*
>
> *(Marshall, Catherine & Rossman, Gretchen B., 1998)*

When conducting a research study, the major parts are data analysis, interpretation and drawing conclusions. In this study, the researcher interprets and explains the experience gained during the building website process. Online SEO analysis tools include Majestic SEO, WooRank, and SeoSiteCheckup are also used to assess and attain more details about the website of the case company. In the end, the re-

searcher concludes the result and gives the suggestion for the further development of the website.

Therefore, concerning the analysis method, Interpretative Phenomenological Analysis, the most common analysis of qualitative data is suitable to interpret this study.

3   SEARCH ENGINE

Search Engine is defined in Dictionary of the English Language of the American Heritage (2013) as a program designed to help internet users find out the information which contains or is related to their targeted search phrases and keywords. Users can search information by keywords, images, video and location. When receiving the command query, the search engine will analyze the request, reviews, ranks and returns the most relevant results. The search results are displayed in a line of results called search engine results pages (SERPs). This chapter provides basic knowledge about search engine including types, structure, operation, and rank algorithms within search engine.

3.1   Search engine types

Depend on how information on the websites is entered into the index, Ledford J. (2009) divided search engines into three main types: Crawler-based search engine, human-powered engine, and Hybrid engine.

The crawler-based search engine indexes the websites automatically by using a software agent called a crawler or spider. The crawlers visit the targeted website to collect information. The collected data is returned, analyzed, converted into the entries and stored into a central repository during indexing progress. The search engine results are filtered from these entries. The crawlers are re-sent to the websites periodically and automatically to update information. Therefore, modifying the web pages may affect the website ranking in the SERPs. Today, the majority of active search engines are in this category, for example, Google, Ask Jeeves, and DuckDuckGo.

A human-powered engine or web directory uses human participation to create its listings. Typically, a brief manually edited description of the website is submitted to the directory and search engine only searches within the description store. These search engines in this category often return a limited but relevant number of

results. The human-powered directory helps users narrow the search area and re-fine results. Therefore, it is efficient in finding information for a general topic. Examples of the human-powered directory are Wikia Search, Mahalo.com, and DMOZ.

Hybrid engine or mixed result is a combination of the two above search engines. It uses different types of data submitted by web crawler and human submission and provides both crawler-based results and human-powered results. However, a hy-brid engine still uses one type of listings as its main results type. Yahoo and Bing are search engines which have used hybrid-based model.

Furthermore, there is another model of search engine called meta-search engines. Berger S. (2005) wrote that this search engine generates its owned results by inte-grating the other search engines results, eliminating the duplicates, adding some features and combining into one large listing. Examples of Meta search engines include Dogpile, Mamma, and Metacrawler.

3.2   The anatomy of the search engine

In an article in 2006, New Idea Engineering described a basic search engine as a coordinated set of three components including spider, fulltext index, and engine (Figure 3).
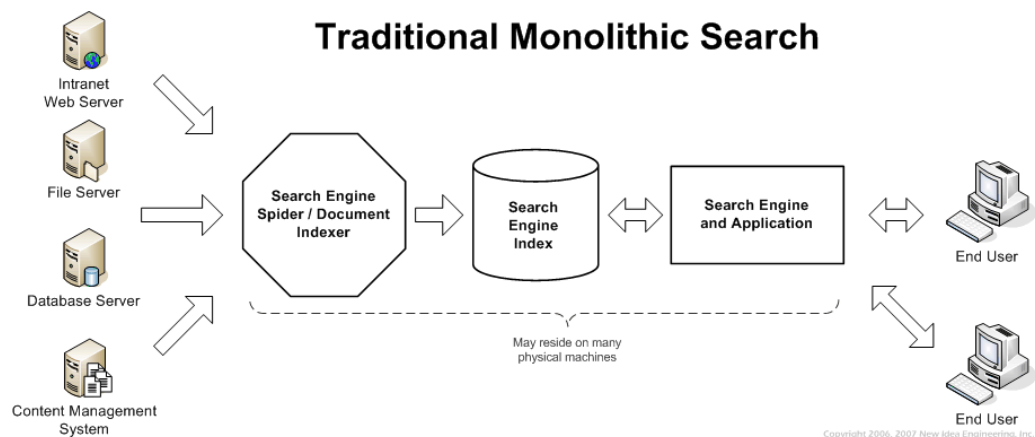
Figure 3: Traditional monolithic search engine (New Idea Engineering, 2006)

A spider (also called a "crawler", "bot", or a "data prep") is sent to every website to collect data. It goes to every page which is allowed to be crawled, parse these pages, extract links, and keywords, and return data to the repository.

A binary Fulltext Index (also called "the index" or a "catalog") indexes pages which will be provided to SERPs. It analyzes the collected data and converts them into the entries. In some Search Engines, this program also creates a list of URLs on each site.

The engine is the program that receives command queries, compares them to the entries in the index, and finally returns the most relevant results to users.

The basic search engine structure is simple, but there is a collection of complex algorithms inside each program. Modern search engines have subdivided their components into subsystems to improve their modularity and the relevance of the results. For example, in the article The Anatomy of a Large-Scale Hypertextual Web Search Engine in 1998, Page L. and Brin S. described in depth the architecture of Google as pictured in Figure 4. Since 1998, this system has still been developed and become complex increasingly over time.
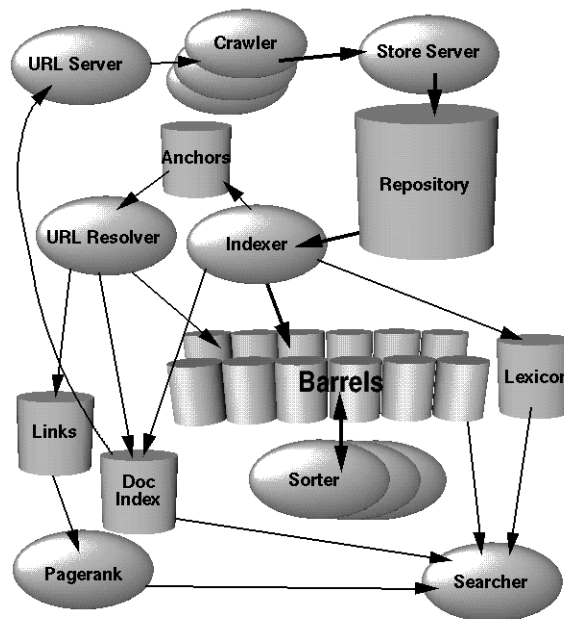
Figure 4: High level Google architecture (Page L. and Brin S, 1998)

3.3   Search engine operation

According to Google (2012), the operation of the Crawler-based Search Engine includes four processes: Crawl, Index, Analysis, and Result as figure 5.

Figure 5: Search engine operation (Google, 2012)

Crawl: When a site is uploaded, the Search Engine will send Web Crawlers (or spiders) to retrieve this site. These web crawlers are programmed to track automatically the links to go to every page in this site, collect and evaluate information on the website before going other pages. However, the owners of the website also choose which parts in their sites are crawled or hidden.

Index: After receiving the information from the Web Crawler, Search Engine stores them. With infinite storage capacity, search engines may contain millions of relevant results. Google (2012) declared that over 100 million gigabytes were stored in their central repository. A website can be indexed fast or slowly depend on crawling speed, the trustability of website and many other elements.

Analysis: in this stage, the search engine works on the collected data and calculate the relevance of the data with the targeted keywords of users. Algorithms and

formulas are utilized to deliver the most relevant results as possible. Different search engines use different algorithms to analysis and algorithms are also changing constantly. It leads to the difference in results between search engines. After that, the results is ranked by using over hundreds of factors.

Results: The results which are relevant to the targeted keywords of users are displayed in this phase. Google shows their result pages after around 1/8th of a second. The most relevant results appear first. However, the results do not always reach the requirement of searchers, spam links often appear in the SERPs.

Up to this time, Search Engines have prevented and fight continuously spam to keep results relevant. The spam is removed both automatically and manually. If the spam is found out, the website owners will be notified, and these sites may eventually be banned. The webmaster must fix their sites to be restored to Search Engines' list. For example, on 6th February 2006, Google is reported to have removed the websites of BMW Germany and Ricoh Germany because of the deceptive practices to raise theirs website ratings. Both companies fixed the offending pages in order to appear in the result pages again. (CNET News, 2006)

3.4   Popular search engines

Currently, there are many active search engines. However, almost of the search market is dominated by several major search engines such as Google, Baidu, Bing, and Yahoo.

Figure 6 and 7 show the newest statistic of NetMarketShare about global search engines market share in desktop and mobile device.
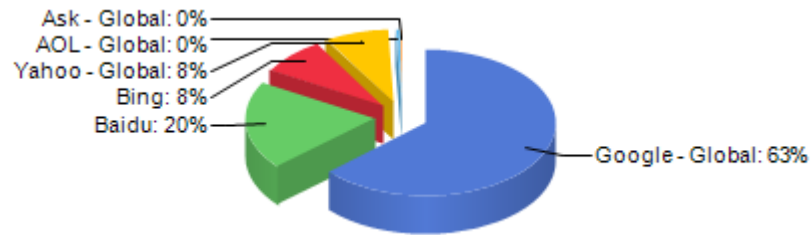
Figure 6: Desktop search engines market share (NetMarketShare, 2015)
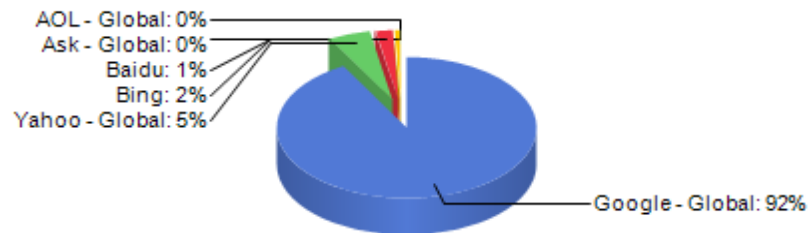
Figure 7: Mobile/tablet search engines market share (NetMarketShare, 2015)

This statistic reported that Google globally is the most widely used search engine, with 63% of the desktop market and 92% of mobile/tablet market. Therefore, SEO on Google is the top priority for the websites.

However, this above statistic did not include local search engine. There are several places where Google is not the leading search engine especially in East Asia and Russia. Global Search Engine Marketing (2010) reported that Yandex is the largest search engine in Russia with a market share of 68.06% in that country while Baidu is the most popular search engine in China with close to 60.74% of the search market. The most popular search engine in South Korea is the local engine Naver, which is used for 47.32% online searches.

As a result, the webmaster should optimize the website on the suitable search engine according to the geographically limited scope of the visitors.

3.5   Rank system

Rank system is the algorithms search engines use to estimate and rank the website in the SERPs. An algorithm contains a lot of factors, and a search engine can use one or more algorithms. Google have used search algorithms PageRank, Panda, Penguin, and Hummingbird with more than 200 factors (Brian D, 2015). Rank system is not only used by search engines; it is also developed by a particular web-site, for instance: Alexa, to evaluate the traffic of the websites.

3.5.1   PageRank

PageRank is a ranking algorithm that Google uses to determine the popularity, the relevance and the importance of the website. Sullivan D. (2007) wrote that PageRank was developed as a system for ranking web pages in a research project about search engine by Larry Page and Sergey Brin at Stanford University in 1996. Today, Google has still considered PageRank as an important part of their web search tools.

According to Google (2011), PageRank counts the number and quality of links to a page to estimate roughly the importance of this page. However, it is not the sole factor used to decide the rank of the website. This estimation are analyzed with other important elements, for example: the usefulness of the information, the popular of the website and the quality of techniques used to building it, to determine the position of the website in SERPs.

Depend on Google Expert Karch M. (2015), PageRank is divided into ten levels and assesses individual web pages, not the entire website. The value of the page ascends from 1 to 10. PageRank is considered as a reliable indicator to assess the importance of each website. A website with high PageRank is believed as trustworthiness to visitors and easy to appear in the top list of Google SERPs.

To increase the PageRank of the website, the webmaster need to increase its backlinks. Linking to pages within the website is a simple way but it should not be

abused. Karch M. (2015) supposed "*The best way to increase your PageRank is to have quality content that other people want to link.*" The website should not contain link farms or HTML errors with coherent sitemaps and useful abundant content.

3.5.2   Alexa rank

Alexa (2009) defined their ranking system as a tool measuring the traffic data and the popularity of the website. The rank of every website is combined with two elements: the number of users visiting this website (Reach) and the number of times an individual page is viewed (Page view) on a daily basis and is calculated by performing the geometric mean of these two quantities averaged over three months. However, this traffic data only records users who install the Alexa toolbar.

Avangate (2015) claimed that Alexa Traffic could be used as a competitive advantage tool. A website with high Alexa ranking index means it is credibility and has a lot of visitors. In the online advertising market, Alexa Rank can be used to evaluate the marketing potential of the website.

Avangate also suggested some advice to boost the Alexa traffic ranking including: the website owner should provide useful quality content to attract users come and back to the website frequently, the layout and the structure of the website should be designed logically so that the website can obtain a lot of clicks from the visitors.

4   SEO FACTORS AND METHODS

The literature review of SEO is outlined in this chapter. The significance SEO factors and methods are introduced thoroughly. Moreover, the history of SEO and its benefit are also mentioned.

4.1   SEO overview

SEO is the abbreviation of Search Engine Optimization. It is a collection of factors and methods used to affect the position of a website in the SERPs. (Search Engine Land, 2011) When users raise a query, search engines return several lists on the search results page including paid lists, advertising lists, pay-per-click lists and natural (or 'organic') list. The goal of SEO is to improve the ranking of the website in the 'organic' lists to promote the number of visitors' access to this site. SEO can be considered a subfield of marketing through search engines.

SEO helps the website become friendly with search engines by improving the structure and the source code, increasing the quantity of links, and optimizing the content of the website. A good SEO website always appears in the top results when users use search engines to find the needed information. SEO may include different types of targets when searching such as images, videos, and local.

The term SEO can also be an acronym for search engine optimizers, who work as the consultant providing SEO strategies for the clients' websites.

4.2   The history of SEO

Pabitha. C (2015) declared that from the early 1990's, with the development of search engines and the World Wide Web, the idea of having high position in the SERPs would bring significant value to business also grew up. Higher rank could bring more traffic to the website. The website owners started to recognize the value of having better position in search engine results. Because of that, the industry of SEO was launched.

In the middle of 1990s, webmasters began optimizing websites. However, at this time, they only sent URL link of a website to search engines. A 'spider' would send back to this site to retrieve it, track the links of every page in this website, and collect data. After data is returned, this website would be indexed on the server of search engines. The information on this website, for example, keywords, links, and contents, is stored. The spider would repeat the indexed process as scheduled in search engines algorithms. According to analyst Sullivan D. (2004), the phrase "Search Engine Optimization" was probably used since 1997.

Initially, search algorithms relied on the information provided by the webmaster, for example, the keyword Meta tag, index files. Meta tags offers a brief guide to the content of each page, normally is a collection of words and phrases. However, Meta tags were considered to be less reliable because the accuracy of a website's content stored on search engines' server depended on the chosen keyword meta tags. If Meta tags were inaccurate, incomplete, and inconsistent, the page could be ranked for irrelevant searches. Consequently, the optimized information and keywords of the website need to be provided accurately to attain a high position on result pages.

Because of relying on factors controlled by webmasters, for instance: keywords, early search engines were used to manipulate rankings. For the sake of searchers, search engines had to ensure that the results pages only showed the relevant search results instead of unrelated trash pages. Search engines have developed ranking algorithms which are more complex, have additional factors to cause difficulties for webmasters when manipulating.

4.3   The benefit of SEO

SEO has become an important online marketing strategy during the last few years. A logical SEO strategy can deliver a significant benefit to the business. Machin E. (2014) recorded five main reasons why the business should invest into SEO.

Increased Traffic: Chitika (2013) reported that Page 1 in the Google SERPs contributed 92% of all traffic, and the first position is responsible for 33% of the traffic. Therefore, high positions on the SERPs receive a large of the attention and clicks from internet users who are impatient to scroll through the page or go to the next pages to look for more relevant information.

Cost effectiveness: Compared to the other types of internet marketing, the primary advantage of SEO is cost effectiveness because the organic results are free. There is no payment for search engines to be displayed in organic listings. If the website is designed and optimized well, it will be listed on the top, and the owner does not need to waste money for advertising the website or pay per click. As the result, SEO becomes one of the most cost-effective marketing strategies.

ROI: One of the major advantages of SEO is to bring better Return on investment (ROI) than normal paid advertising such as Adwords, PPC. Enquisite (2008) conducted a research to study visitor behavior on both organic search listings (SEO) and pay per click search listings (PPC). The statistics revealed that with the same keyword, searchers clicked on the organic results eight times more than the paid one. It means for every dollar spent on SEO; the owner has to spend eight dollars on PPC to receive the same value. Furthermore, the possibility of users clicking on organic search is 5.66 times higher than paid search. As the results, the value of SEO is 45 times more than paid search. Due to the low cost and high conversion rates, it is concluded that SEO offered the better value in internet marketing than paid advertising.

Increased site usability: SEO helps the website become friendlier, easier to be crawled and indexed by improving its structure and source code, arranging the links within the website, and optimizing the content of the website. These changes

not only help the crawlers index the website easier, but also help searchers find information quicker and feel comfortable on this website.

Brand Awareness: Being ranked in high positions of the SERPs makes a strong impression on the potential customers, creates the reputation to the company, and can make the website become a brand name. Moreover, if the website is on the top list, the customers can associate it with the keywords, phrases leading to this result. The more pages or links to the website are placed on the top lists of the SERPs, the more trustworthy and credibility this website is perceived. It opens higher opportunities for users to visit the website and become faithful customers.

For these reasons, SEO is the wiser long term investment for every company wanting to improve the business.

4.4   SEO Factors

Search engines rank the websites by relying on unique "signals" – the ranking factors. SEO ensures that the website is suitable and spawns the correct "signals". Search Engine Land (2013) researched and designed the Periodic Table Of SEO Success Factors (Figure 8) to describe the major SEO factors that can affect the website ranking in the SERPs.
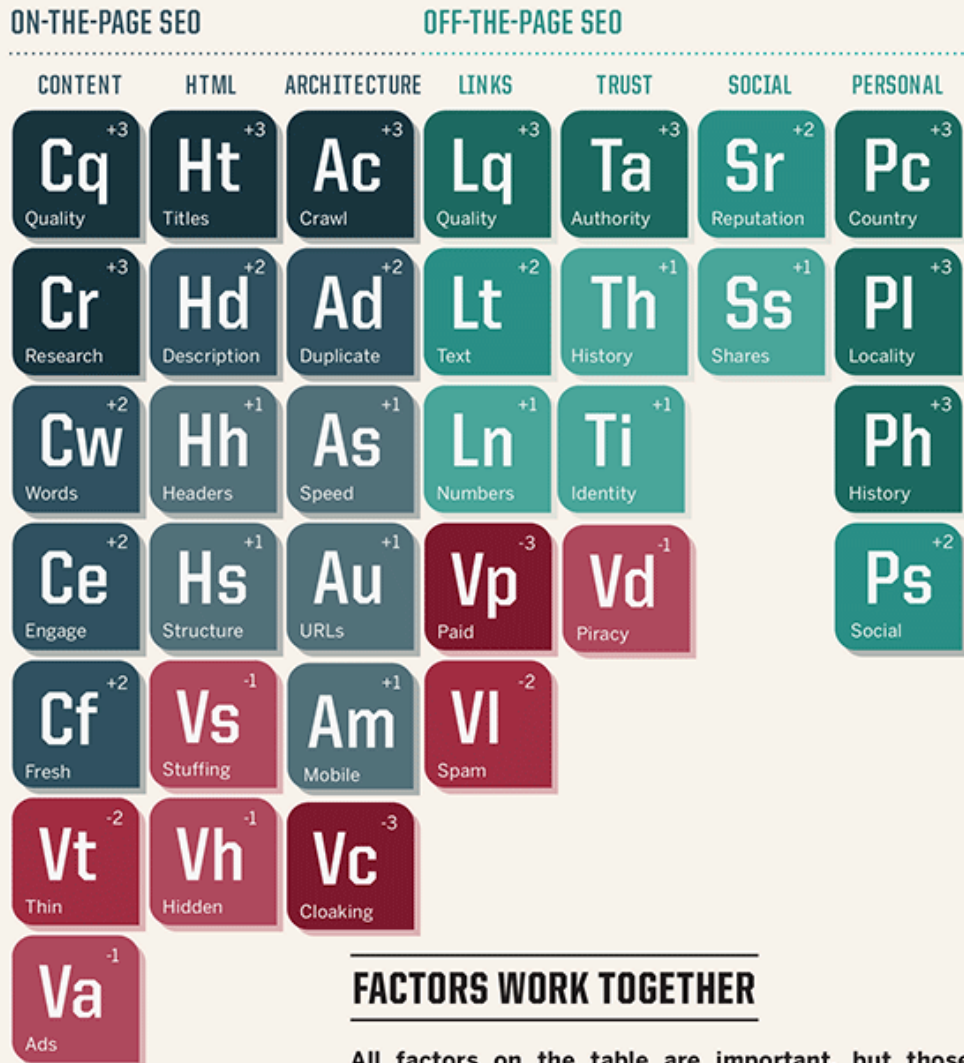
Figure 8: The periodic table of SEO success factors (Search Engine Land, 2013)

This chart includes three major groups: On-page SEO, Off-page SEO, and Violations. Each group consists of several subgroups which contain one or some individual SEO "elements". The symbol of the SEO "element" is created by combining the first letter of the subgroup it belongs to and the first letter of this SEO factor's name. In the top right corner, there is a number which is the representative of the importance of this element, measured on a scale of one to three. A factor weighed of three has the bigger influence on the visibility of the website than others. However, there are also factors with negative numbers. These factors can cause adverse impacts on the online presence of the website. The smaller its number is, the worse effect it can grow.

These SEO elements listed in this table are assessed to be important in the increasing of the rankings. However, these factors are only effective when they are combined with the others, and the negative factors should be avoided.

4.4.1   On-Page factors

These factors in On-Page group are the factors controlled entirely by the webmaster. They are divided into three subgroups: the content, the HTML code, and the structure of the website.
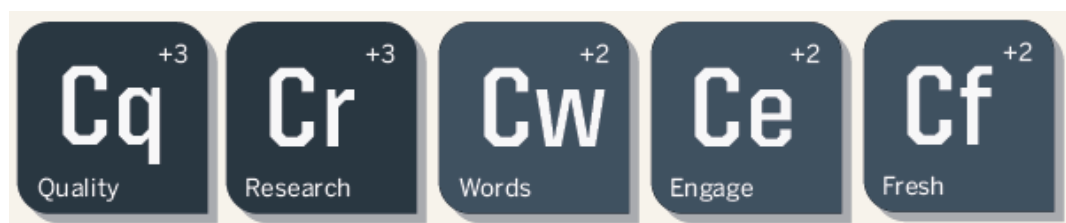

Figure 9: Content factors (Search Engine Land, 2013)

Depend on the table in figure 8, content is the most important part to make SEO success because it involves these elements having high weighting point. Content is said to be the cornerstone for other factors. The owner should consider the quality, the keyword, the engagement, and the freshness of the content (Figure 9).

Quality: High-quality content improve the SEO effects significantly. The content of the website is assessed as quality if it is unique, contains useful information, and cannot be found on other sites. If that is a selling website, the owner should ensure the goods are real, guarantee the quality of the products, offer the marketing campaigns to the customers.

Keyword: This factor can be separated into two sections: keyword research and using the keyword in the website. Today, there are a lot of keyword research tools which help the websites' owners target groups of related keywords, offers further suggestions, and even provides statistics about average monthly searches, competition level of these keywords. For example: WordStream's Keyword Tool, Fresh-Key, Ubersuggest, and Google Keyword Planner. The owner discovers the common keywords leading to the website which users may use. Based on those keywords, the owner can create or modify the content of the website in order to produce contents which can satisfy searchers when raising that query. Dean B. (2015) showed that in Google's ranking factors, there are about twenty factors concerned to keywords and around ten factors collecting data about keywords on page-level. However, the targeted keywords and the related words should be used naturally to the content for the best results.

Engagement: Having high-quality content is not enough, this content should be attractive and satisfying the visitors to keep them stay on the website. This interaction – engagement – is measured to assess the value of the website. There is a variety of ways for search engines to measure the engagement. Two of them are 'Pogo-sticking' metric and 'time on site' metric. 'Pogo-sticking' refers to the searchers' behaviors when they click through to the website, back to the SERPs and click through to other pages (Berg A., 2015). If the pogo-sticking rate is high, it means the website content may be not engaging. Conversely, when the visitors spend longer time on the particular website than other sites, it is called "time on site" or "long click". Moreover, the social gestures, for instance: comments, shares, and likes, are used as an engagement metric. However, the website owners cannot measure engagement themselves, engagement metrics are collected secretly and used to evaluate the website.

Freshness: Search engines pay special attention to the websites updated frequently. Google developed an algorithm named Query Deserves Freshness (QDF) to query if there is any news on the website (Willoughby S, 2009). If the website is often updated with new information, news, and stories, it has remained high position or increased ranking on the SERPs. However, the updated information should be relevant to the old content.
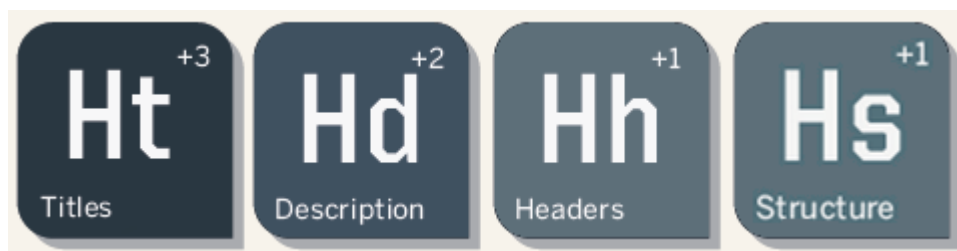

Figure 10: HTML factors (Search Engine Land, 2013)

HTML code is an important part to achieve SEO success. Search engines use several specific HTML elements as the clues to assess the websites. These elements are Title tag, Meta description tag, header tag, and the structured data of the website (Figure 10).

Title tag: Title tag is a strong on-page SEO signal because search engines understand the purpose, the target of the website through it. If every page on the website has the same title, the search engine is confused. Therefore, the keyword should be added to the title tag, and each page of the website should have a unique, descriptive, and related-to-keywords titles.

Meta description tag: According to Moz (2015), Meta descriptions are HTML elements that provide the brief description of the contents of the websites. Meta descriptions are used on the SERPs to provide searchers a preview about a specific page. In 2009, Google announced that Meta description tag is not a "ranking factor", it means this tag is not influenced the ranking results. However, Moz claimed that it was an extremely important part of online marketing to attract searchers to a website from the SERP. Because search engines only display 160 characters on the snippets, the Meta description should be long enough, from 150

to 160 characters and contains the keywords. In addition, the owner should not duplicate Meta description in the pages of the website.

Header tag: Similar to the title tag, search engines use header tag as the signal to understand the purpose of a particular page on the website. The owner should use the keywords in header tag to increase the chance of displaying on the SERPs for these keywords.

Structured data: Structured data helps search engines index the website easier and understand it completely. Yerian N. (2013) informed that structured data is used to create rich snippets, which are the extra information displayed beside the title, the Meta description, and the URLs of the website on the SERPs, for example: image, address, and map. Similar to Meta description tag, it is a "display factor" used to attract searchers' eyes.
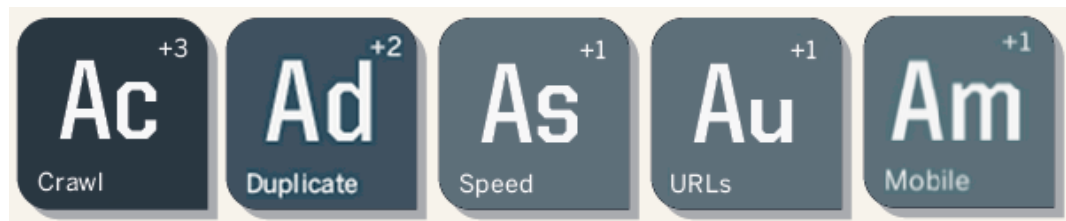
Figure 11: Site architecture factors (Search Engine Land, 2013)

Site architecture affects how search engines index the website. A well put-together architecture can boost SEO efforts while the wrong one can harm them. (Search Engine Land, 2013). The issues concerned to website architecture include: the crawl ability, the duplication, the speed, the URLs, and the Mobile Search. (Figure 11)

Crawl ability: In order to appear on the SERPs, the websites need to be indexed. To be indexed, they have to be crawled by search engines. Based on the authority and the trustability of each website, search engines set an individual crawl budget - the approximate amount of pages or time this search engine crawls periodically. To improve the crawl efficiency of the website, the owner can apply some meth-

od, for example: using robots.txt, building smart site architecture, creating the authority for the site, and creating a sitemap. (Lurie I, 2011)

Duplication: If the website has many version indexed, for example: with www and without www or pages with the same content in the central repository, search engines can estimate this site incorrectly. The owner should remove or fix duplicated content in order to waste crawl budget, and the crawling process can progress easily and quickly. Implementing 301 redirects, 'rel=canonical' tags, and managing URL parameters also helps the website avoid duplication.

Speed: Cutts M. (2011) claimed that are the site speed ranking factor only impacts 1 out of 100 queries that make it become a minor factor. However, the speed can support the other factors, for example, engagement and social media. Therefore, the owner should speed up the website to appreciate search engines and searchers.

URLs: Search engines algorithm also considers to keywords found within the domain name and page URLs or not. Therefore, adding keywords to these places can help the ranking prospects of the website. Moreover, Smarty A. (2009) reported that the URLs should be not so long to catch users' eyes.

Mobile Search Factors: In 2015, Google updated their mobile algorithm. Depend on that, the position on the SERPs of a page may be different when users raise queries on the desktop and mobile device. If the page is assessed as mobile-friendly, for example: having mobile viewport and big text, it could rank higher than a non-mobile-friendly page.

4.4.2   Off-Page factors

Off-Page factors are the elements which are not controlled directly by the owner and affected by visitors and other publishers. This group contains four subgroups: Links to the website, trustability, social media and personalization.
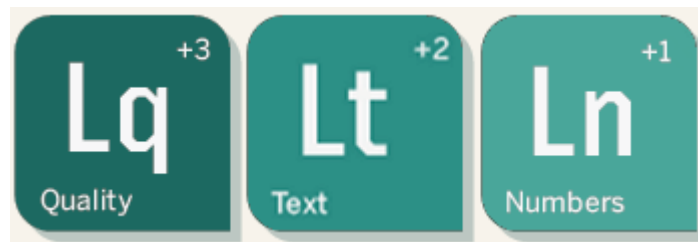

Figure 12: Link factors (Search Engine Land, 2013)

Backlinks include both external links and internal links to the website. According to Dean B. (2015), Google use about 40 ranking factors concerned to backlink in their algorithms. PageRank is an example for the algorithm which relies on backlink to evaluate the relevancy of the website. Therefore, backlink is considered as the letter of recommendation for the website and the most important external signal for search rankings (Search Engine Land, 2013). However, the value of backlinks is not equal. The aspect of the link building need to be considered when optimizing off-page are the anchor text, the quality and the number of backlinks. (Figure 12)

Quality: Among link building issues, the link quality is the most important one. Search engine counts every link to the website. However, the links considered as high-quality ones are more weighted. These links are from any large website, prestige pages, or reputable blogs. Moreover, links from local websites which are relevant to the website may also be assessed higher.

Anchor Text: Anchor text, also called link text, is the visible, clickable, and underlined text in a hyperlink (Moz, 2015). It represents for how the writers describe the website linked to and cannot be controlled by the website's owner.

Number of links: Obtaining a lot of links can improve the ranking of the website on the SERPs. However, if the links come from a single website, they are evaluated less value than the links come from different websites. Websites can use automated link-building software to generate link spams. This method affects both positively and negatively on website ranking, and the search engines have developed their algorithms to reduce its effects.
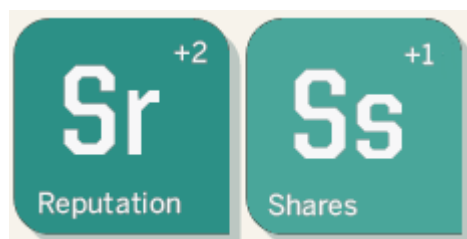
Figure 13: Social media factors (Search Engine Land, 2013)

Cutts M. (2014) announced that social media signals, for example, Twitter and Facebook were not used to rank websites. However, social signals are considered as ranking factors because they still affect SEO indirectly. The Internet users can visit the website by clicking links shared on the social networks. Figure 13 showed two factor concerned to social media signals: reputation and shares.

Social Reputation: If a respected blog or social account shares link or reviews about a particular website, the amount of people knowing this website increases. Social networks can become a communication channel to connect with the users and reach the potential customers for the business websites. As a result, having an own social account is also important to the website.

Social Shares: Sharing on the social networks is similar to the link. It helps the web-site spread the content, reach more users and obtain more visitors. Therefore, participating in social networks is important.

Figure 14: Trustability factors (Search Engine Land, 2013)

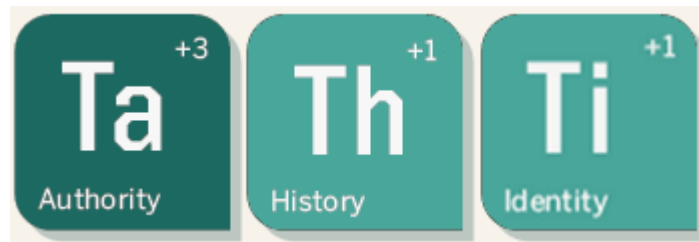Trust is believed that it takes an important part to achieve the SEO success (Search Engine Land, 2013). Authority, history, identity are subjects search engines examine when evaluating the trustability of the website (figure 14).

Authority: An authority website is the high-quality website that recognized widely in its area. How the authorities of the websites are assessed is still a mystery because the search engines have not revealed their algorithms about it. Google (2011) only announced that the authoritative issue is evaluated when assessing the quality of a page or content of the website.

History: The history of the website is considered by search engines when ranking the website. A website which was discovered to violate guidelines and use deceptive practices is more difficult to reach the high position on the SERPs than the others sites with a good record. Furthermore, an old, respected website is assessed higher on the organic result than a new website in the same industry.

Identity: The importance of the identity has increased because search engines want to ensure that they provide to searchers the 'real' results, not the fake ones. Using the real profile, building the brand and relationship with respected websites are the ways to create identity signals.
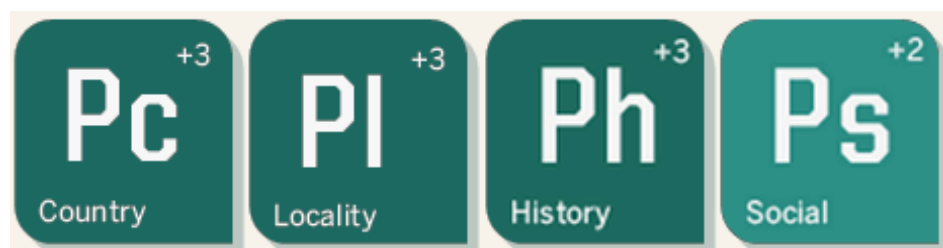

Figure 15: Personalization factors (Search Engine Land, 2013)

Today Google and Bing offer the private results to their users. It means the SERPs of their users may be different with general results and others when searching for the same keyword. Depend on figure 15, search engines depend on the geographical location, searched history, and the social connection of users to provide the most relevant results. Because it is users' behavior, the webmaster cannot control it. However, the owner can modify the website so that it is more relevant to users.

Country: This ranking factor is one of the earliest factors concerned to personalization. Search engines prioritize to websites having the same country server IP with the users

Locality: Search engines locate users' area through location signals, for example, the users' profile or IP address. Algorithms have been developed to deliver results in the same area or near users' location. The owner should ensure that the website is relevant to the locality in order to appear in local search results.

Personal history: Personal history of the user impacts significantly search results when delivering the private results. Both browsing history and search history of users are considered when the query is raised. These websites that users visit frequently are prioritized to place on the top.

Social connection: Similar to personal history, the social connection of users can affect their private results. If the website is shared by someone, search engines supposed that this user trusts this website, and it may place higher in the later private SERPs of this user.

4.4.3   Violation factors

Figure 16: Violation factors (Search Engine Land, 2013)

These positive factors can help the website improve its ranking on the SERPs, while there are negative factors which presence can harm the website. If worse, this website can receive a penalty or be banned from the SERPs. Figure 16 displayed these signals including:

"Thin" or "Shallow" Content: In 2011, Google released Google Panda algorithm to reduce rankings for low-quality sites that are described as the lack of content, useless information and low-value. If the website has empty content or redundant articles, it can receive a Panda penalty.

Ads / Top Heavy Layout: High ads ratio and heavy template footprint can return a warning for the website. Search engines want to ensure that when searchers visit the website, they find some useful information not only a bunch of advertising.

Keyword Stuffing: Adding the keyword in the content would increase the chance of appearing in the top result. However, repeating keyword many times is considered as a spam method. High keyword density makes the website become low quality in the search engines' view. Some online SEO Tools, for example, Ranks NL and SEOBook, offer the keyword density tool to calculate the percentage of a keyword over the total number of words on the website.

Hidden Text: if the webmaster decides to use keyword stuffing, the hidden text method often use together to hide the stuffed keywords. Two most common ways are changing the color of the text to be the same as the color of the background and using display: none. However, search engines want to see all the content of the website. Therefore, hidden text can bring a minus point for the website.

Cloaking: Cloaking is referred to the action of webmasters when they provide to search engines' crawlers a different version of the website compared to what searchers see. Searchers can see a website which is not related to their search keyword but still ranked high on the SERPs. While spam sometimes can be considered as an accident, cloaking is estimated as a deliberate action. A website will receive a heavy penalty or banned if it is discovered to use cloaking.

Paid Links: An example for paid link penalty is Google Japan was detected to pay to receive positive reviews (Schwartz B, 2009). After that, Google penalized Google Japan by drop its PageRank point from 9 to 5. Search engines want to eliminate the paid links because they can impress the ranking significantly.

Link Spam: Link spam is considered as one of the signals of a "thin" site. It appears everywhere from forums to blogs with the attractive anchor text. The algorithms of search engines, for example, Google Penguin, have been constantly developed to remove web spams.

Piracy / DMCA Takedowns: This factor have added to Google ranking factor recently (Sullivan D., 2012). If the website infringes on copyright law, search engines may penalize this site. It is not the problem of every website, but the webmaster should handle carefully if a DMCA takedown request appears.

4.5 SEO methods and techniques

SEO techniques can be divided into two broad categories: White Hat and Black Hat. White Hat SEO uses techniques which are approved by search engines and does not involve deception. Black Hat SEO does not follow guidelines, used those techniques search engines do not approve and deceptive practices. Despite the high-risk level, Black Hat SEO still is accepted and used because it can help the website increase its ranking quickly. In addition, another techniques sometimes mentioned is Grey Hat SEO, the combination of both White Hat SEO and Black Hat SEO. Websites can be banned or penalized by search engines if they are discovered using black hat methods to affect the website ranking.

The writer introduces some tactics in three categories which are used widely in SEO industry.

4.5.1 White Hat SEO techniques

Link Baiting: Link Baiting is defined as the creation of compelling content in order to attract people link to the website or share this content. (SubmitExpress, 2015). This method aims to increase the number of links and improve the PageRank of the website. The article should be easy to understand, contain useful information, and mention to the controversial or sensationalistic topic.

Quality Content: It is defined as the creation of an informative and valuable content. This method can be time-consuming. However these efforts bring great benefit after the long period.

Internal Linking: Internal links are links pointing to pages within the website. They are usually used in the main menu to navigate to the website. Moreover, internal links help to build site architecture, increase the number of links and help the crawler reach every page of the website. (Moz, 2015)

Site Optimization: This tactic is the essence of SEO. It includes editing the content, modifying the site architecture, and fixing the HTML code. This method affects almost On-Page SEO factors.

4.5.2 Black Hat SEO techniques

Black Hat SEO is also called spamdexing. In general, almost tactics in this technique impact the link factors by boosting the quantity of links to the website from different sources. Several tactics are introduced in Section 4.4.3: Violations factors. They are cloaking, paid links, link farms, keyword stuffing, and hidden text. There are still three Black Hat SEO methods often used: Scraping, Doorway Pages, and Parasite Hosting.

Scraping: Scraping is a Black Hat SEO spam technique. The spammers copy the popular contents or the articles of other websites and paste them into in their sites. This method is often used to increase the view, to sell advertising slots on the website. (Gavrilas R., 2012)

Doorway Pages: Doorway page is often called by the others names, for example, bridge page, portal page, jump page, gateway page, and entry page. A doorway page is considered as a form of cloaking. It is created to achieve high rank in the search results with a particular phrases or keywords. If the searchers click on its link in the SERPs, it will send them to the different page called landing page. (SubmitExpress, 2015)

Parasite Hosting: Parasite Hosting is explained as the process of creating blogs or topics on a highly respected and high-rank website and links to the targeted website. This method helps to create the weighted backlinks to the website because of the high rank of the host domains. (Techopedia, 2015)

4.5.3 Grey Hat SEO techniques

Three-way links exchange: Three-way links is described as to link three websites together. The website post links to the two other sites. This method is used to build link popularity for new or small websites. (Techopedia, 2015)

Article spinning: It is similar to scraping tactic. However, the spammers rewrite popular existing contents to release the different version of these articles on their websites. (Gavrilas R., 2012)

Buying old domains: The old, authorized and respected domains are bought in order to create the weighted backlinks to the targeted websites to achieve the high rank on the SERPs. In some way, it is similar to the Parasite Hosting tactic of Black Hat SEO.

5   RESEARCH DESCRIPTION

The case study is a qualitative research method which offers an opportunity to study a particular subject. The researcher takes part in the organization to gather and analyze information.

Production Software Company is a Finnish company located in Lahti, Finland. This company offers consulting services and software solutions to improve the IT systems of industrial manufacturing companies. The researcher joined the team building the website and creating SEO strategy for this company. Therefore, it was selected to be the case study in this study.

Global Search Engine Marketing (2010) reported that the market share of Google in Finland is 98.12%. Therefore, the targeted search engine for Production Software Company is Google.

The information is collected by observing the website. Online analysis tools contain Majestic SEO, Woorank, and SeoSiteCheckup are used to analyze SEO Strategy. Majestic SEO analyzed the backlinks of the website while Woorank and SeoSiteCheckup are used to evaluate the other SEO factors. These analysis reports were attached on the appendices section.

Based on the analysis results, the researcher suggests some advice for both On-Page and Off-Page SEO factors to improve website ranking.

6   RESEARCH DATA

After applying SEO strategy on the Production Software website, the researcher received the below results.

When raising the query with keyword "Valmistuksenohjausjärjestelmät" in Google, website of Production Software is ranked at the third position on Page 1 of the SERPs as figure 17.
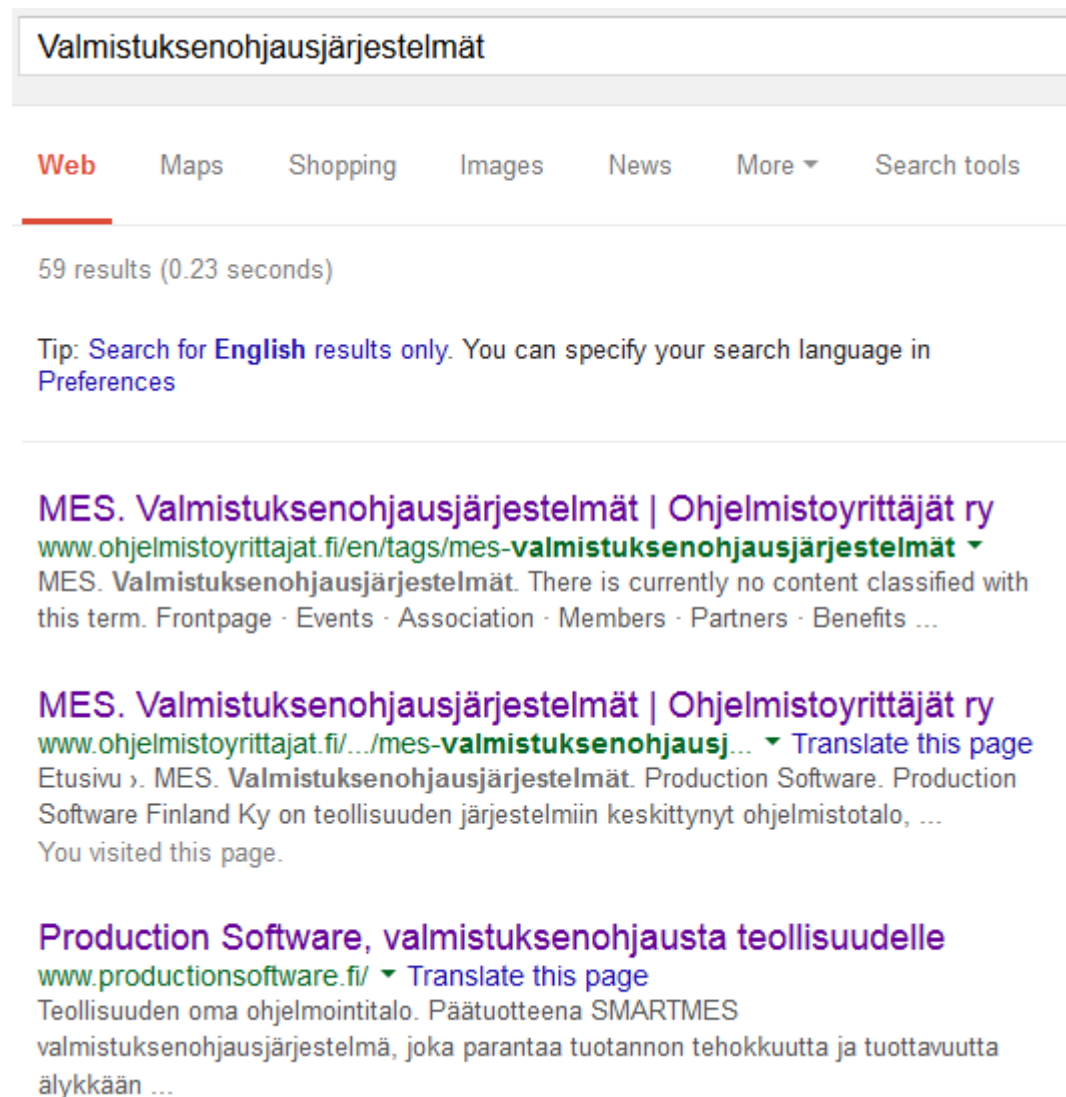

Figure 17: Google SERPs for keyword "Valmistuksenohjausjärjestelmät"

Figure 18 revealed that the rank of this website is the second of the organic results with keywords "Valmistuksenohjaus".
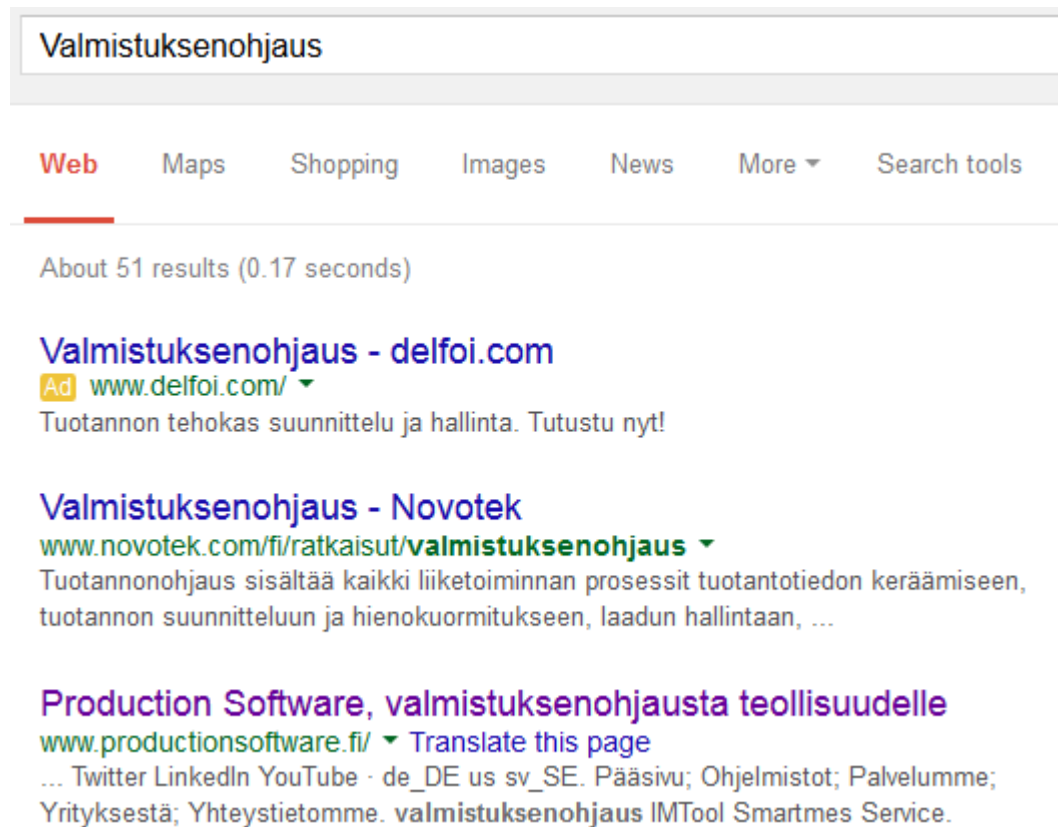
Figure 18: Google SERPs for keyword "Valmistuksenohjaus"

With regard to website traffic, table 1 showed the estimation of Statshow (2015) about the number of visitors accessing Production Software website and total amount of pageviews earned daily, monthly, and yearly.

	Daily	Monthly	Yearly
Pageviews	108	3240	39420
Visitors	49	1470	17885

Table 1: Average website traffic of Production Software by Statshow (2015)

Majestic reported that, Production Software website has 271 external links, come from 46 different domains this year. This number in the last five years are 630 and 73 respectively (see Table 3). Majestic also showed the geographic map of the referring domains represented by circles as figure 19. The larger and darker red circle is, the more links from this area to the website are. According to the statis-

tic, Production Software website is recommended most in Finland, following by US, UK, and China.

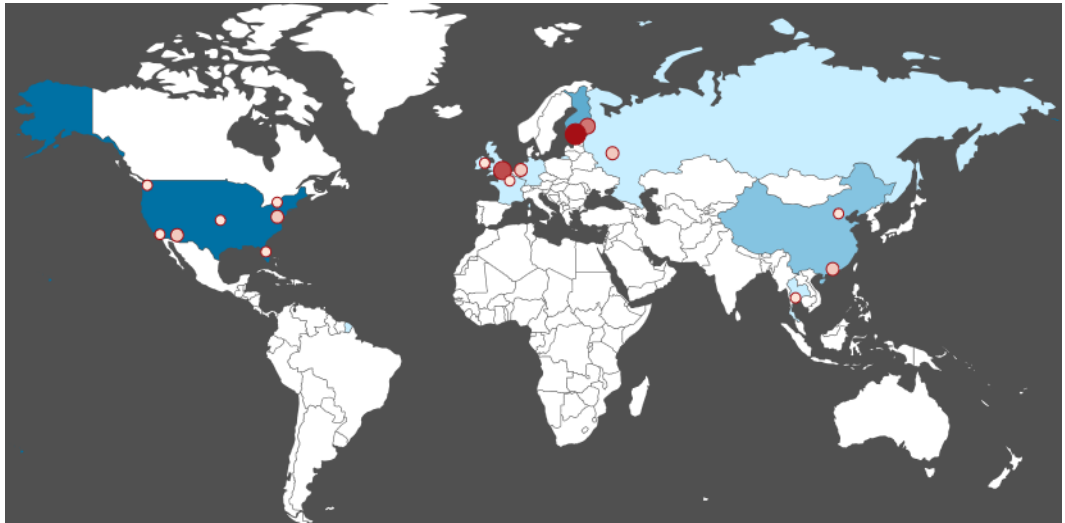*Figure 19: Geographic map of the referring domains (Majestic, 2015)*
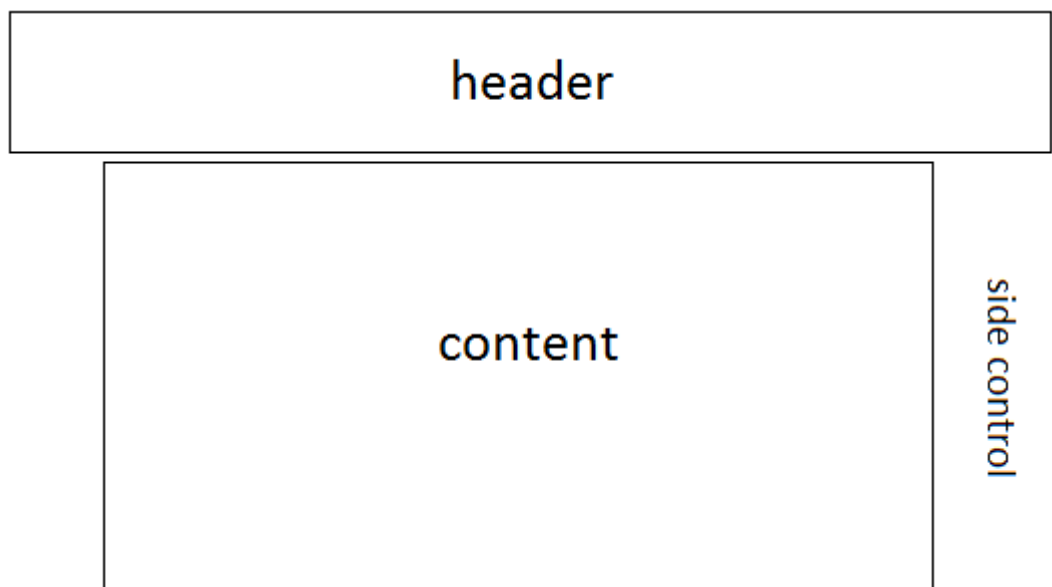
6.1 Website template

Figure 20: The layout of Production Software website.

Production Software chose WordPress as platform to building the website. The template of this website is one page themes which load all the content in a single page (figure 20). A side-control bar is placed in the right hand for the visitors to

change the section. The users can move to the other parts by scrolling mouse or click on the part name in menu bar. In addition, Production Software website does not include ads, only contains header, body part and do not have footer in its website.

6.2 Content

The content of the website is also reported that it had been download without errors.
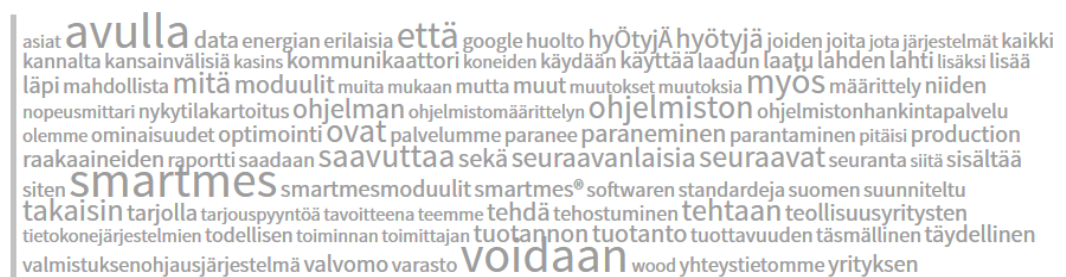
Figure 21: Common word cloud (Seositecheckup, 2015)

Seositecheckup concluded the most common words in the website and its usages as figure 21. In the word cloud, the words which is bigger than the other words are repeated many times.

According to the statistic, the word displayed most is Smartmes with the density of 3.09% of the website content. It's followed by 'voidaan', 'avulla', 'hyötyjä' and 'että' with the range of 0.88-1.68%. (figure 22)

Figure 22: Word density (Seositecheckup, 2015)

6.3   HTML

Text/HTML ratio is a signal which help search engine estimate a website as a spam site or not. The ideal ratio is above 15% (Seositecheckup, 2015). If the Text/HTML ratio is out of range, this website can be considered as spam. This ratio can be improved by adding more text into the content.

Seositecheckup informed that the HTML file size of Production Software website is 74.09 KB while the content text size is 12.46 KB. The content contributes 16.82% to the source code and its Text/HTML ratio is assessed as good.

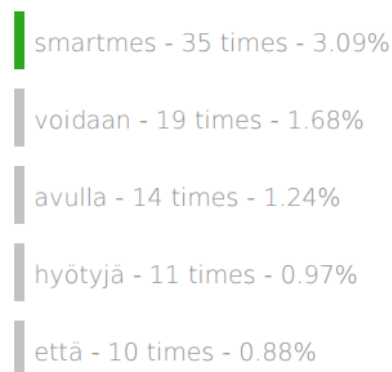Production Software, valmistuksenohjausta teollisuudelle

Figure 23: The title of Production Software website

Figure 23 showed the title of Production Software website. The website title should be unique, explicit and contains the keywords of this website. The title should include from 10 to 70 characters because search engines will truncate titles if its length is more than 70 characters. (Seositecheckup, 2015). The title of Production Software website has a length of 56 characters and in the safe range.

Teollisuuden oma ohjelmointitalo. Päätuotteena SMARTMES valmistuksenohjausjärjestelmä, joka parantaa tuotannon tehokkuutta ja tuottavuutta älykkään optimoinnin avulla.

Figure 24: Meta description of Production Software website

Similar to the title, the Meta description of the website should be unique, explicit, contains the keywords, and summarizes briefly about the website content. Almost search engines will truncate Meta description if it contains more than 160 characters. (Seositecheckup, 2015). The Meta description of Production Software website display in figure 24 has a length of 174 characters.

<H1>	<H2>	<H3>	<H4>	<H5>
16	25			

Figure 25: Quantity of header tags (Woorank, 2015)

The heading tags are used as the keywords for the indexing progress and should be not duplicated with the title tag (See section 4.4.1: On-Page factors). Depend on figure 25, Production Software website includes 41 heading tags with 16 tags in the first level (<H1>) and 25 tags in the second level (<H2>).

7 ALT attributes are empty or missing.

http://www.productionsoftware.fi/wordpress/wp-content/uploads/fi/... /smartmes_100.jpg

http://www.productionsoftware.fi/wordpress/wp-content/uploads/fi/b... /tuotanto_100.jpg

http://www.productionsoftware.fi/wordpress/wp-content/uploads/fi/bro... /huolto_100.jpg

http://www.productionsoftware.fi/wordpress/wp-content/uploads/fi/broc... /laatu_100.jpg

http://www.productionsoftware.fi/wordpress/wp-content/uploads/fi/br... /varasto_100.jpg

http://www.productionsoftware.fi/wordpress/wp-content/uploads/fi/.../ohjelmisto_100.jpg

http://www.productionsoftware.fi/wordpress/wp-content/uplo... /nykytilakartoitus_100.jpg

Figure 26: List of images missing 'alt' attribute (Woorank, 2015)

Because search engines cannot see the images like human, alt attributes help them understand clearly about the images. (WooRank, 2015) Production Software website is reported to have 21 'img' tags. However, only 14 tags have the 'alt' attributes, seven left attributes are empty or missing. The image missing 'alt' attribute are showed in figure 26.

With regard to CSS and JavaScript files, moving all the inline CSS and JavaScript code into external files makes the website lighter in weight and increasing the Text/HTML ratio. Production Software website is informed that there is still 179 inline CSS styles in the HTML code. The external JavaScript and CSS files should be minified to decrease the number of HTTP requests. This website have four JavaScript files and three CSS files.

6.4   Site architecture

Robots.txt file and sitemaps are used to inform search engines which pages on the website are and are not available for crawling. Production Software placed both robots.txt and Sitemap files in the root directory of the web server. The sitemap of Production Software website is an XML file containing the list of URLs, which crawlers are allowed to crawl. Figure 27 revealed the content of the Robots.txt file. Robots.txt file consists of three parts: 'User-agent', 'Disallow', and sitemap parameter.

Figure 27: Robots.txt file of Production Software

'User-agent' is represented for one or more specific search engines. The asterisk character (*) means all search engine indexing this website. 'Disallow' blocks the crawler from a specific folder and Sitemap parameter shows link to the sitemap of this website.

In terms of canonicalization, http://productionsoftware.fi/ and http://www.productionsoftware.fi/ are reported to the same URL. However, the IP address 81.90.71.19 does not redirect to the domain name of the website.

Seositecheckup informed that there are several errors in the URLs in the website. Some links in Production Software website do not follow the SEO Friendly URL Structure as Figure 28. For example:
http://www.productionsoftware.fi/wordpress/se/index.php?data=0

Figure 28: SEO friendly URL structure (Vertommen, 2012)

Moreover, this tool discovered some broken links, for example:
http://www.hamechamber.fi/en, http://www.productionsoftware.fi/wordpress/wp-content/uploads/fi/brochure/huolto-A4.pdf, and links which contain underscores, for example: http://www.productionsoftware.fi/wordpress/wp-content/uploads/fi/brochure/Production_Software_ESITE.pdf.
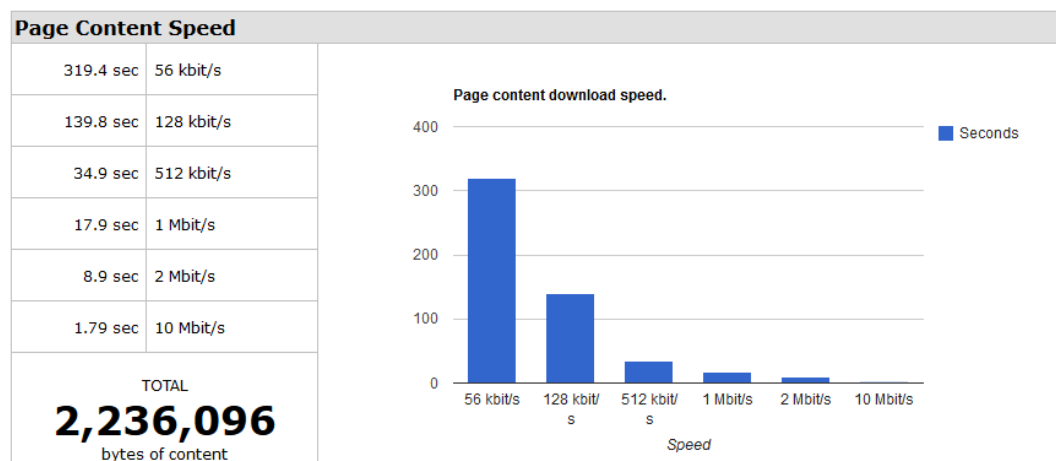
Table 2: Site loading speed

Table 2 shows the total time to load all Production Software website including HTML, CSS, JavaScript files, and images. The average loading time with speed of 10Mbit/s is around 1.79 seconds.

Figure 29: Website displays in mobile device

Figure 29 showed how website is rendered in mobile device. Production Software website is considered as non-mobile friendly website because it is not using media queries to implement responsive design and the Mobile viewport is not set. (WooRank, 2015) When loading the website in the mobile device, the text content is too small to read and the buttons/links is not large enough to be tapped easily. Moreover, the loading speed in mobile device is reported as slow.

6.5   Backlinks

External Backlinks	Referring Domains	Referring IPs	Referring Subnets
270	46	38	34
In the last 5 years			
630	73	64	61

Table 3: Quantity of backlinks and referring domains (Majestic, 2015)

Depend on table 3, Production Software website has had 271 external links which come from 46 different domains since the beginning of 2015. In the last five years, there are 630 backlinks and 73 domain link to the website. Furthermore, backlinks and referring domains charts in figure 30 revealed that the majority of

links which Production Software has received are in two early months of 2015. There are a little number of links reviewed on March and April, 2015.
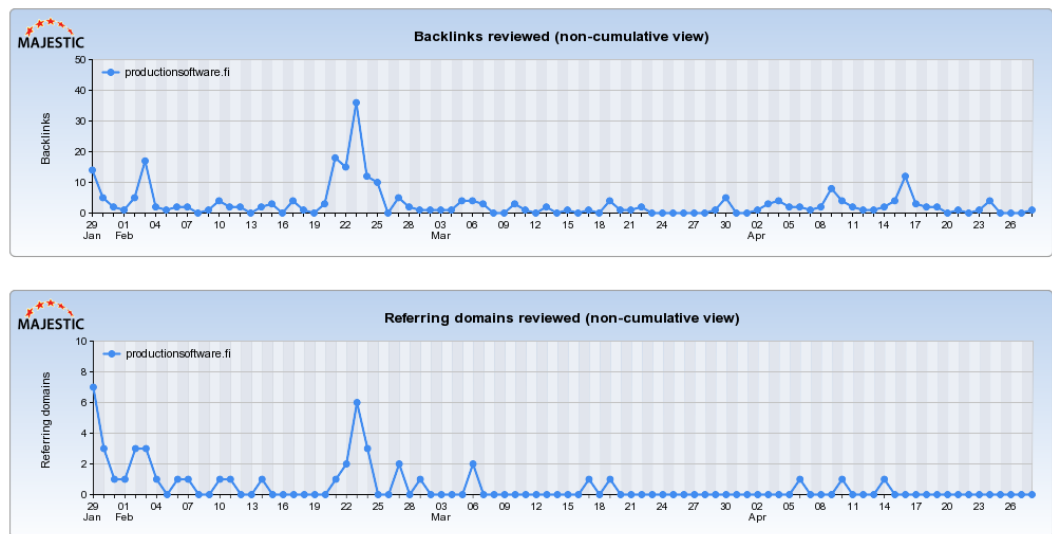
Figure 30: Backlinks and referring domains charts in 2015 (Majestic, 2015)

6.6   Social network

Production Software Company has Google+, Facebook, Twitter and LinkedIn account. However, the link to the Google+ account is missing in the website. According to the social shareability statistic in table 4, the website is only shared 5 times on Facebook and 1 times on Google+. It has not received any like or comment on Facebook or Twitter tweet. Moreover, there are no shared or liked button on the website.

Social Media Signals	
0	Twitter Tweets
5	Facebook Shares
0	Facebook Likes
0	Facebook Comments
1	Google+ Pluses
0	LinkedIn Shares
0	StumbleUpOn Views
0	Pinterest Pins

Table 4: Social shareability of Production Software website

6.7   Trustability

Figure 31: Trust indicators provided by WOT™ (WooRank, 2015)

Figure 31 showed the trust indicators provided by Web of Trust, a website reputation in rating and review provided by mywot.com. According to this, the trustworthiness score of Production Software website is 60 (out of 100) and the website is declared as a trustworthy website.

7 DATA ANALYSIS

The purpose of this study is using online analysis tools to find out the success and failure of the current SEO strategy and suggest some ideas to improve website r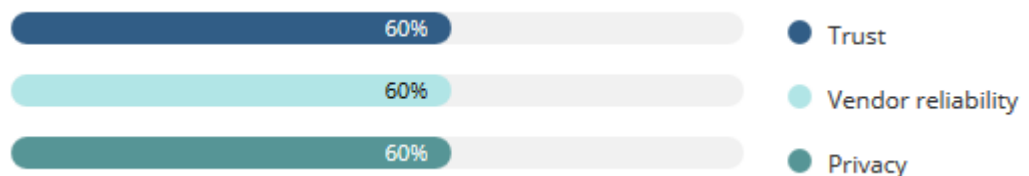anking. The researcher categorizes and analyzes the detail and implication of collected data. The researcher also give the advice to make the website become friendlier with the search engines.

Depend on the search results, Production Software website is placed on the top three of the SERPs when searching by its targeted keywords. However, this rank can be changed when the algorithms of the search engine are updated.

Production Software website is reviewed from many areas but the current website traffic is low. To improve the website traffic, the webmaster should modify SEO factors and apply other methods.

7.1 Website template

WordPress is recommended when building a small business website. (Escobanas J, 2013) The webmaster can find many plugins in this platform to manage SEO and perform marketing strategy. Single page theme is simple, easy to navigate the website. The users can read through all the content easily. However, the owner needs to manage content in order to provide the most valuable and useful information to the visitors.

7.2 Content

There are no problems encountered to retrieve all web page content files. All content files in the website have been download successfully.

The problem in the content is the keyword density. The keyword density is a signal for the relevance of keywords with the content. The keywords should be used

consistently with the content and displayed frequently in the text. SEOreporters (2014) claimed that the safe keyword density range between 2-4% for Google search engines. It means keyword should appear 2-4 times for each 100 words.

However, the most important targeted keywords do not appear in the common words list of Production Software website. Therefore, the relevance of the keyword to website content is low. The owner should modify the content to add more the keywords into the website, for example: header tags, image alt attributes, and internal link anchor text, to improve the chance of ranking high in the SERPs for the targeted keywords.

The owner can create a new section for blog or posting news to create a rich-content page and remain the freshness of the website.

7.3   HTML

The current Text/HTML ratio is considered as safe but only slightly above the limitation. However, if the owner want to add more source code, the content also needs to be added more text to keep the Text/HTML ratio in the safe range.

The title and Meta descriptions of the website are optimized well. Both of them are unique, explicit, contains the keywords. However, the Meta description has a length of 174 characters which is greater than 160 characters. The webmaster should use a shorter description.

Header tags can help emphasize important topics and keywords within a website. H1 tags in Production Software website are long enough and different to each other. In addition, the website also use H2 tag to describe the sub-topics.

Production Software website has 21 images but only 14 images have the required 'alt' attributes. Moreover, the website contains inline CSS styles and has more than one JavaScript file and CSS file. These HTML errors can decrease the web-

site ranking. The webmaster should extract all the inline CSS and combine JavaScript files/CSS files into one file.

7.4   Site architecture

The webmaster blocks the crawler from folder wp-admin and wp-includes. However, to increase the indexing speed and optimize the content, some other folders should be blocked, for example, wp-content/plugins and wp-content/themes. Moreover, the document files should be blocked by adding the code: Disallow: /*.[file extension]$. For example: Disallow: /*.pdf$

There are several errors in the URLs in the website including broken links, links containing underscores, and non-SEO friendly URL. The owner should remove/replace the broken links, use hyphens instead of underscores, and fix or remove non-SEO friendly URLs from the indexed links list.

The average loading time of Production Software website with 10Mbit connection is around 1.79 seconds. The site loading speed is fast, but it can still be improved by reducing the images size in the website.

According to the collected data, Production Software website is not optimized for mobile visitors. It is important to have a responsive website for visitors accessing the website on mobile devices because the mobile algorithm of the search engine has been developed to display the mobile-friendly website on the top of mobile search results. The web owner should use media queries to create a mobile viewport. For example:

```
@media screen and (min-width: 480px) and (max-width: 960px) {
   #header, img {display: none ;}
   body {text-size: 20px ;}
   .button {width: 20em; height: 2em ;}
}
```

7.5   Backlinks

The current quantity of backlinks and referring domains to the Production Software website is too little. The company should use Link Baiting and Internal Link method to increase number of link to the website.

The content should be updated/refreshed to attract visitors come to the website. The webmaster has to ensure that the website does not contain link spam.

7.6   Social networking

WooRank informed that this website is not popular on social platforms. The social media presence of Production Software Company is too low. This website is rarely shared in the social network. The owner should increase the social media activity by promoting the website more in the social network accounts of the company. In addition, the shared and liked button should be added into each article in the website.

7.7   Trustability

The worthiness score provided by WOT of Production Software website is 60. It means this website is trustworthy. Trust Indicators can be improved by building authority and identity. Using HTML5 microdata can help search engines understand the website content clearly. Here is an example of microdata:

```
<div itemscope itemtype="http://yourwebsite.com">
   <span itemprop="name">the owner name</span>
   <span itemprop="company">the company</span>
   <span itemprop="tel">telephone number</span>
   <a itemprop="email" href="mailto:exampleweb@example.com"> exampleweb@example.com</a>
   </div>
```

8 CONCLUSION

The theoretical foundation section of this research provides the basic knowledge about search engine and SEO factors and method. There are four types of search engines: Crawler-based, human-powered, hybrid, and meta-search engine. Almost the active search engines are Crawler-based engines. As the result, SEO factors and methods focus on this type.

The SEO factors can be divided into two categories: On-Page and Off-Page. On-Page SEO are the factors controlled entirely by the owner and can be divided into three subgroups: the content, the HTML code, and the structure of the website. Off-Page factors are affected by visitors and the owner cannot controlled directly. It includes: backlink, trustability, social media and personalization

SEO techniques contain three categories: White Hat, Black Hat, and Grey Hat. Implementing SEO takes a lot of time and effort if the webmaster utilizes the White Hat techniques. The website can achieve high ranking after the long time but does not involved any risk. Conversely, the Black Hat techniques can help the website appear in the top list of search result quickly with high risk level. Therefore, Black Hat techniques is only suitable for short term strategies. The website owner should research and analyze carefully to choose right methods for the website.

The algorithm of search engines are updated constantly and impact the website ranking significantly. SEO element are also modified to be compatible with them. The webmaster can use online analysis tools to audit SEO effect. These tools can help companies evaluate effectively the SEO strategy of these companies. Depend on the collected data and the analysis report, companies can modify the content and adjust SEO factors promptly and reasonably to remain on the top list of the SERPs.

After analyzing the data, some advice are suggested. Depend on this, the case company should modify the content by adding more the keywords into the article, creating blog or posting news in order to develop a rich-content page and remain

the freshness of the website. Furthermore, HTML errors need to be fixed. The webmaster also has to remove inline CSS styles to external CSS file and minify the JavaScript and CSS file. Two problems that webmaster should solve immediately are URLs errors and mobile viewport. Some URLs are reported as broken links or non-SEO friendly URLs. The mobile viewport should be created to attract mobile visitors. In addition, the social media should be focused more. The case company should increase the social media activity by sharing frequently link to the website in the social networks. The sharing and liking button should be added in each article on the website. Finally, the case company should create the profile using HTML5 microdata to improve its trustability and attract the search engines.

9   DISCUSSION

9.1   Scope and limitations

The purpose of this study is analyzing the current SEO situation to propose some suggestion to improve website ranking. The other companies can use this results as reference in order to improve SEO on their website.

However, this research still has some limitations. Firstly, this study only targets one search engine - Google. There are many active search engines, and each engine uses different algorithms and rules. It leads to the difference in optimizing the website to achieve the high position in the SERPs. Almost Finnish internet users use Google as their main search engine. Thus, this study only focuses on Google. Secondly, the research uses the White Hat SEO techniques and ignores the other method because they involve no risk. However, other technique can achieve the goal in the short term. Thirdly, the specific keywords are rarely used. There are under 100 results which are relevant to them. Therefore, the case company does not need to compete fiercely with competitors.

9.2   Reliability and validity

The reliability of this thesis is measured by the search result. In addition, suggestions to improve SEO are provided after comparing the literature review of SEO factors and the current SEO situation on the case company website thoroughly. Therefore, it is considered as reliable but the reality results need to be observed further. Moreover, because of the limitation of the case study which mentions in the previous section, the different websites with different keywords can reach different ranks.

9.3   Suggestions for further research

For the future research, the research can extend the target search engines, involve Yahoo and Bing in the research. The research also tries to use the other techniques besides the White Hat method to verify the effect of this method.

In the future, companies should pay more attention to the mobile viewport. It is important to have a responsive and friendly website for visitors accessing the website on mobile devices. The social media activities need to be increased to improve the social media presence of companies.

REFERENCES

Published References

Aliaga, M., & Gunderson, B. (2000). *Interactive Statistics*. Saddle River, p3-15

Bargh, J. and McKenna, K., (2004), *Internet and Social Life*. New York: New York University

Berger S. (2005). *Sandy Berger's Great Age Guide to the Internet.* Que Publishing

Denzin N. and Lincoln Y. (Eds) (1994). *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage.

Gill, J., & Johnson, P. (2010). *Research Methods for Managers.* London: Sage.

Mack N., Woodsong C., Macqueen K., Guest G., Namey E., (2005) *Qualitative Research Methods: A Data Collector's Field Guide*. Research Triangle Park, North Carolina: Family Health International

Heyse L., Zwitter A., Wittek R., Herman J., (2014). *Humanitarian Crises, Intervention and Security: A Framework for Evidence-Based Programming.* Routledge. P182-184

Marshall, Catherine & Rossman, Gretchen B. (1998). *Designing Qualitative Research*. Thousand Oaks, CA: Sage.

Pabitha C. (2015). *Search Engine Optimization by Eliminating Duplicate Links*. International Journal of Emerging Research in Management & Technology Volume 4 Issue 1. P53-57

Rodgers, J. A., Yen, D. C. & Chou, D. C. (2002). *Developing e-business: a strategic approach. Information Management & Computer Security*, 10(4), 184-192.

The American Heritage (2013). *Dictionary of the English Language, 5th edition.* Houghton Mifflin Harcourt Publishing Company.

Trochim W. (2006). *Research Methods Knowledge Base, 3rd Edition*. Mason, Ohio: Atomic Dog.

Electronic References

Alexa (2009). *About us*. Retrieved from: http://www.alexa.com/about

Berg A. (2015). *Pogo-Sticking, Bounce Rate – Ranking Factors?* Retrieved from: http://www.yourseoplan.com/pogo-sticking-bounce-rate-ranking-factors/

Chitika (2013). *The Value of Google Result Positioning*. Retrieved from: http://chitika.com/google-positioning-value

CNET News (2006). *Google blacklists BMW.de*. Retrieved from: http://news.cnet.com/Google-blacklists-BMW.de/2100-1024_3-6035412.html

Conductor (2014). *Organic Search Is Actually Responsible for 64% of Your Web Traffic*. Retrieved from: http://www.conductor.com/blog/2014/07/update-organic-search-actually-responsible-64-web-traffic/

Cutts M. (2011). *How does Google determine page speed?* Retrieved from: https://www.youtube.com/watch?v=SO4YuDAkplU

Cutts M. (2014). *Are pages from social media sites ranked differently?* Retrieved from: https://www.youtube.com/watch?v=udqtSM-6QbQ

Dean B. (2015). *Google's 200 Ranking Factors: The Complete List.* Retrieved from: http://backlinko.com/google-ranking-factors

Enquisite (2008). *Via Enquisite: PPC Agencies Make 45X What SEOs Do for the Same Value.* Retrieved from: http://moz.com/blog/via-enquisite-ppc-agencies-make-45x-what-seos-do-for-the-same-value

Escobanas J. (2013). *Top 5 CMS Website Platforms for Small Business*. Retrieved from: http://www.business2community.com/small-business/top-5-cms-website-platforms-for-small-business-0556849

Gavrilas R. (2011). *Black Hat vs White Hat SEO – Infographic*. Retrieved from: http://cognitiveseo.com/blog/229/black-hat-vs-white-hat-seo-infographic/

Global Search Engine Marketing (2010). *Internet and Search Engine Usage by Country*. Retrieved from: http://ptgmedia.pearsoncmg.com/images/9780789747884/supplements/9780789747884_appC.pdf

Google (2009). *Google does not use the keywords Meta tag in web ranking*. Retrieved from: http://googlewebmastercentral.blogspot.fi/2009/09/google-does-not-use-keywords-meta-tag.html

Google (2011). *Facts about Google and Competition*. Retrieved from:

http://www.google.com/competition/howgooglesearchworks.html

Google (2011). *More guidance on building high-quality sites*. Retrieved from:

http://googlewebmastercentral.blogspot.fi/2011/05/more-guidance-on-building-high-quality.html

Google (2011). *Finding more high-quality sites in search*. Retrieved from:

http://googleblog.blogspot.fi/2011/02/finding-more-high-quality-sites-in.html

Google (2012). *Algorithms*. Retrieved from:

https://www.google.com/insidesearch/howsearchworks/algorithms.html

Google (2012). *How Search Works*. Retrieved from:

http://www.google.com/insidesearch/howsearchworks/thestory/

Google (2015). *Rolling out the mobile-friendly update*. Retrieved from:

http://googlewebmastercentral.blogspot.fi/2015/04/rolling-out-mobile-friendly-update.html

Internet Live Stats (2014). *Number of Internet Users (2014)*. Retrieved from:

http://www.internetlivestats.com/internet-users/

IProspect (2006). *The Search Engine User Behavior*. Retrieved from:

www.iprospect.com

Karch M. (2015). *What Is PageRank and How Do I Use It?* Retrieved from:

http://google.about.com/od/searchengineoptimization/a/pagerankexplain.htm

Ledford J. (2009). *Search Engine Optimization Bible*. Hoboken: Wiley, p34-35

Lurie I. (2011). *How (I Think) Crawl Budget Works (Sort Of)*. Retrieved from:

http://searchengineland.com/how-i-think-crawl-budget-works-sort-of-59768

Machin E. (2014). *Top 5 Benefits of SEO*. Retrieved from: https://www.titan-seo.com/NewsArticles/5-seo-benefits.html

Moz (2015). *Google Algorithm Change History*. Retrieved from:

http://moz.com/google-algorithm-change

Moz (2015). *Meta description*. Retrieved from: http://moz.com/learn/seo/meta-description

Moz (2015). *Internal Link*. Retrieved from: http://moz.com/learn/seo/internal-link

Moz (2015). *Anchor Text*. Retrieved from: http://moz.com/learn/seo/anchor-text

NetMarketShare (2015). *Search Engine Market Share*. Retrieved from:

https://www.netmarketshare.com/search-engine-market-share.aspx?qprid=4&qpcustomd=0

Page L. and Brin S. (1998). *The Anatomy of a Large-Scale Hypertextual Web*

Search Engine. Retrieved from: http://infolab.stanford.edu/~backrub/google.html

Rajendran V. (2001). *How Can A Website Benefit Your Business*? Retrieved from: http://www.articlecity.com/articles/computers_and_internet/article_4848.shtml

Schwartz B. (2009). *Google Penalizes Google Japan For Buying Links*. Retrieved from: http://searchengineland.com/google-penalizes-google-japan-16541

Search Engine Land (2011). *What is SEO?* Retrieved from: http://searchengineland.com/guide/what-is-seo

Search Engine Land (2013). *The Periodic Table Of SEO Success Factors*. Retrieved from: http://searchengineland.com/seotable

Search Engine Land (2013). *Content & Search Engine Success Factors*. Retrieved from: http://searchengineland.com/guide/seo/content-search-engine-ranking

Search Engine Land (2013). *HTML Code & Search Engine Success Factors*. Retrieved from: http://searchengineland.com/guide/seo/html-code-search-engine-ranking

Search Engine Land (2013). *Site Architecture & Search Engine Success Factors*. Retrieved from: http://searchengineland.com/guide/seo/site-architecture-search-engine-ranking

Search Engine Land (2013). *Link Building & Ranking In Search Engines*. Retrieved from: http://searchengineland.com/guide/seo/link-building-ranking-search-engines

Search Engine Land (2013). *Social Media & Ranking In Search Results*. Retrieved from: http://searchengineland.com/guide/seo/social-media-ranking-search-results

Search Engine Land (2013). *Trust, Authority, Identity & Search Rankings*. Retrieved from: http://searchengineland.com/guide/seo/trust-authority-search-rankings

Search Engine Land (2013). *Personalization & Search Engine Rankings*. Retrieved from: http://searchengineland.com/guide/seo/personalization-search-engine-rankings

Search Engine Land (2013). *Violations & Search Engine Spam Penalties*. Retrieved from: http://searchengineland.com/guide/seo/violations-search-engine-spam-penalties

SEOReporters (2014). *Optimal Keyword Density SEO Best Practices.* Retrieved from: http://www.seoreporters.info/learn-seo/keyword-density

Seositecheckup (2015). Search Engine Optimization (SEO) Tools.
http://seositecheckup.com/checkup/result/43d7240a09974685f42b1c2083275a7b?url=http://www.productionsoftware.fi

Smarty A. (2009). *URL Structure: The Closer to the Root – the Better?* Retrieved from: http://www.searchenginejournal.com/url-structure-seo/11801/

Statista (2015). *Global market share of search engines as of January 2015*. Retrieved from: http://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/

Statshow (2015). *Worth & Traffic Estimation*. Retrieved from: http://www.statshow.com/www/productionsoftware.fi

SubmitExpress (2015). *SEOdictionary*. Retrieved from: http://www.submitexpress.com/seodictionary.html

Sullivan D. (2007). *What Is Google PageRank? A Guide for Searchers & Webmasters.* Retrieved from: http://searchengineland.com/what-is-google-pagerank-a-guide-for-searchers-webmasters-11068

Sullivan D. (2004). *Who Invented the Term "Search Engine Optimization"?* Retrieved from: http://www.searchenginejournal.com/who-invented-the-term-search-engine-optimization-part-deux/3925/

Sullivan D. (2012). *The Pirate Update: Google Will Penalize Sites Repeatedly Accused Of Copyright Infringement.* Retrieved from: http://searchengineland.com/dmca-requests-now-used-in-googles-ranking-algorithm-130118

Techopedia (2015). *Parasite Hosting*. Retrieved from: http://www.techopedia.com/definition/27202/parasite-hosting

Vertommen, K. (2012). *How to Create an SEO-Friendly URL Structure*. Retrieved from: http://webdesign.tutsplus.com/articles/how-to-create-an-seo-friendly-url-structure- -webdesign-9569

Willoughby S. (2009). *Query Deserves Freshness*. Retrieved from: http://moz.com/blog/whiteboard-friday-query-deserves-freshness

WooRank (2015). *Website Review for productionsoftware.fi.* https://www.woorank.com/en/www/productionsoftware.fi

Yerian N. (2013). *What is Structured Data and Why Does it Matter for SEO?* Retrieved from: http://www.adherecreative.com/blog/bid/153475/What-is-Structured-Data-and-Why-Does-it-Matter-for-SEO

APPENDICES

Majestic SEO analysis report

Majestic-Site-Explor
er-Results.pdf

Seositecheckup analysis report

Seositecheckup-An
alysis-Report.pdf

WooRank analysis report

Woorank-Analysis-
Report.pdf