

**DELIVERING USER-FRIENDLY EXPERIENCES BY
SEARCH ENGINE OPTIMIZATION**

Lian Heng

Bachelor's Thesis
School of Business and Culture
Degree Programme in Business Information Technology
Bachelor of Business Administration

2014

School of Business and Culture
Degree Programme in Business
Information Technology

Author	Heng Lian	Year	2014
Supervisor	Tuomo Lindholm		
Commissioned by	N/A		
Title of thesis	Delivering User-Friendly Experiences by Search Engine Optimization		
No. of pages + app.	60		

The objective of this research is to study different and recent Google algorithms to encourage the uses of White Hat search engine optimization in order to create a user-friendly experience on websites. Moreover, reasons for not using Black Hat search engine optimization are justified.

This research focuses on Google search engine, specifically on the search engine optimization principles and its benefits to deliver user-friendly experiences.

The exploratory research approach was adopted in this thesis research due to the formulation of the research questions. Descriptive method was also utilized as a secondary research method to enhance the findings. To ensure the validity of this thesis research, the sources drawn from are more recent than from year 2007.

On the basis of research analyses and findings, this thesis research proposes different techniques and principles to be applied to webmaster's websites in order to create websites acceptable by Google standards, more importantly user-friendly experiences.

Keywords: user-friendly experience, SEO, Google algorithm, Google Panda, PageRank, Google Penguin, Page Layout,

CONTENT

ABSTRACT

FIGURES

1	INTRODUCTION	5
1.1	Motivation and background	5
1.2	Objectives	6
1.3	Structure of the thesis	6
2	TOPIC, RESEARCH QUESTIONS AND METHODOLOGY	8
2.1	Research questions	8
2.2	Research methodology	9
3	GOOGLE SEARCH ENGINE	11
3.1	How does a general search engine work?	11
3.2	How does Google search engine work?	13
4	GOOGLE ALGORITHMS	15
4.1	PageRank	15
4.2	Google Panda	18
4.3	Google Page Layout	20
4.4	Google Penguin	23
4.4.1	Keyword stuffing	24
4.4.2	Hidden texts or links	25
4.4.3	Links Schemes	25
4.4.4	Cloaking	27
4.4.5	Duplicate Content	27
5	OPTIMIZATION	31
5.1	Quality Content	32
5.2	Keyword	33
5.2.1	Keyword selection	34
5.2.2	Keywords Placement	36
5.3	Linking	43
5.4	Other optimizations	45
5.5	Case study	45
6	CONCLUSIONS	52
	REFERENCES	54

FIGURES

Figure 1. Indexing and Query Process (a combination of Fig. 2.1 & 2.2 from Croft & Metzler & Strohman 2009)	12
Figure 2. Googlebot (Lievonen 2013, 14)	13
Figure 3. PageRank Example (Cutts 2009)	16
Figure 4. Outline of good quality content (Lievonen 2013, 20).....	19
Figure 5. Page Layout 1 (Google AdSense 2012)	21
Figure 6. Page Layout 2 (Google AdSense 2012)	22
Figure 7. Keywords Stuffing.....	24
Figure 8. Title tag 1	37
Figure 9. Title tag 2	38
Figure 10. Title tag 3	39
Figure 11. Meta Description.....	40
Figure 12. Keyword density curve (Hill 2014)	42
Figure 13. Backlink commenting (Krishna 2013).....	47
Figure 14. Content for the case website (Siteliner 2014)	49

1 INTRODUCTION

The motivation and background of this thesis topic are discussed in this chapter. Likewise, the objectives of the research are illustrated. Moreover, the general structure of this thesis work is described.

1.1 Motivation and background

Due to the rapid development of the Internet, more people use it, and they are also increasingly dependent on it. Search engine has become an important and inseparable tool for anyone who uses the Internet. For this reason, most businesses are willing to acquire more targeted customers through search engines, the webmasters use search engine optimization (hereinafter SEO) techniques to get a high ranking on the search result page. In consideration of attempting to gain a high ranking, many SEO tricks are against Google's quality guidelines, and therefore, Google designed many algorithms to fight with the website that violates Google's quality guidelines.

From the point of view how this research was conducted, the thesis research is mainly intended for the webmasters who are beginners in SEO. Furthermore, beginners who are willing to get a high rank in Google search result page should use ethical methods, i.e. White Hat SEO.

There are two conceptions that need to be differentiated, which are SEO and SEM (stands for Search Engine Marketing). SEO is different from SEM, in fact, SEO is one of the methods in SEM. There is pay-per-click advertising (hereinafter PPC) in SEM as well. SEM can be defined as a practice using a variety of methods to achieve the highest possible website ranking in search engine results. (Ledford 2009, xvi.) However, on the basis of objectives of this research, the discussion of this thesis is narrowed down to only include SEO.

The reason for choosing the most sources from Matt Cutts is because he is the head of Google's Webspam team. In fact, Matt Cutts joined Google as a software engineer since January 2000 until now. (Cutts 2014e.)

Danny Sullivan and Vanessa Fox's article was also used in this thesis, in a reason of Matt Cutts recognized that they are the expert for SEO (Cutts 2014d).

1.2 Objectives

The objective of this research is to deliver a user-friendly experience on websites by using the White Hat SEO. Moreover, devaluing the utilization of the Black Hat SEO is justified, as the methods in the Black Hat SEO are not agreed upon by Google.

This objective is attained by finding out the factors that impact on the rank of Google search engine, based on Google algorithms and user-friendly experiences. To measure up this objective, studying, understanding and analyzing Google algorithms are demanded.

1.3 Structure of the thesis

This thesis is divided into six main chapters. Chapter 2 discusses the topic of the research, the research questions and research methodologies used in conducting the research. Chapter 3 begins by providing an introductory overview of Google search engine. Moreover, the basics of search engines in general need to be comprehended in Chapter 3 as well. The main topic is presented and discussed in Chapter 4 as the chapter is about algorithms used by Google for ranking websites. Additional, Chapter 4 gives reasons why webmasters should use the White Hat SEO instead of the Black Hat SEO. Chapter 5 recommends techniques for using, in order to create high-quality content websites which are liked and accepted by Google guidelines. Furthermore, a case website is analyzed in Chapter 5 based on the discussion in Chapter 3, Chapter 4 and Chapter 5. Chapter 6 concludes the research and presents the findings of this

research while suggesting users to concentrate on the White Hat SEO in order to create user-friendly experiences.

2 TOPIC, RESEARCH QUESTIONS AND METHODOLOGY

The topics of this research are concepts and technologies of Google, the White Hat SEO and algorithms used by Google for ranking websites. To discuss these ideas and techniques, this thesis research presents a brief introductory overview of Google and its ranking systems. Furthermore, to form a text comprehensive for beginners of SEO, a few examples for illustration are provided. On the basis of these discussions and examples, the outcome of the thesis research is to deliver user-friendly experiences.

2.1 Research questions

To achieve the objectives of the thesis research, three questions are defined and answered below.

1. How does a search engine work?

To answer the question above, Google is analyzed, explored and described from the point of view of the process of the search engine. Choosing Google as the ideal search engine for a reason that it is leading search engine today with about 76% of the whole search engine market (Sullivan 2013). The percentage converted into numbers is 1.17 billion unique searchers (Sullivan 2013). However, to understand the questions, the process of Google search engine and Google system need to be examined, in other words, Google search engine and Google algorithms need to be explored. The above question is answered in Chapter 3 and Chapter 4. In Chapter 3, Google search engine is touched upon. In Chapter 4, different algorithms are described.

2. How does Google determine the rank for a search query?

The question is answered in Chapter 4, and as Google has many different algorithms, most recent algorithms are chosen. To answer the question

above is necessary, as the objective of this thesis research is to encourage the uses of the White Hat SEO techniques on websites.

3. How can the rank be made higher in the search result?

The question needs to be answered in order to let users to build a high-quality content accepted by Google. Chapter 5 discusses different techniques for optimizing websites to rank higher. As the techniques discussed in the thesis research are the White Hat SEO techniques, they are legal and comply with the rules and guidelines of SEO.

2.2 Research methodology

This thesis research utilizes the exploratory research approach. This approach is appropriate for the research when the researcher has little or no knowledge about the topic. As the researcher had no knowledge of this topic, to understand the concepts and to help crystallize the definition of the research questions, the exploratory research approach was selected as suitable for this research (McDaniel & Gates 2010, 43).

Aside from the exploratory research approach, the descriptive approach is used. The descriptive approach is appropriate, as Google and SEO are not new concepts. However, further understanding and research are needed as both Google and SEO technologies are advancing. The objective of descriptive research is to describe characteristics that are being studied. Therefore, after gaining the necessary knowledge from the exploratory research approach, the analyses and results need to be described. Furthermore, decisive information for achieving the outcome of this thesis research needs to be delivered as well. Due to the nature of the thesis research is theoretical; the research is conducted relying on the official Google as a primary source of data. Secondary data sources were journals, articles, websites and printed literature sources.

The criteria for searching and choosing literature sources were complied with the Lapland University of Applied Sciences writing instructions. Therefore, sources that were used are relevant to the research, current and published by established authors. However, most of the book sources related to SEO were paraphrased from Chinese versions of original English books of the same title, as the original versions were not available at the time the researcher was deciding to purchase them. However, secondary sources were found by using a search engine, i.e. Google.

3 GOOGLE SEARCH ENGINE

Google was founded in 1998 by Larry Page and Sergey Brin (Google 2014a). The goal of Google is to provide searchers the most relevant answers as fast as possible when a search query has been delivered. To describe the 'perfect search engine', Larry Page, one of the co-founder and Chief Executive Officer (Hereinafter CEO) of Google clarifies as follows "Understands exactly what you mean and gives you back exactly what you want." (Google 2014k.) Due to the substance, Google concentrate on the quality of website and try to rank the best quality website with a higher ranking.

Google search engine has a leadership among all the search engines. "Google is the standout leader in search today. It has the most traffic by far, and it's the only search engine with its own entry in the dictionary." (Grappone & Couzin 2011). Therefore, if a company's website wants to acquire more targeted customers through Google search engine, a high ranking is needed. The website should strive to show up in the first three pages in the Google search result page because most of the searchers will not check the website which are ranked after three pages (Ledford 2009, 34). In other words, in order to increase the rank of a website, it is important to use SEO techniques.

3.1 How does a general search engine work?

Search engine components are mainly offering two functions that are indexing process and query process. Indexing process is used to create searchable data; query process using these data and searchers' queries to come up with a sorted result list. (Croft & Metzler & Strohman 2009.)

The components in the indexing process are text acquisition, text transformation and index creation; they are showing on the left side of Figure 1 below. The components in the query process are user interaction, ranking, evaluation, on the right side of Figure 1.

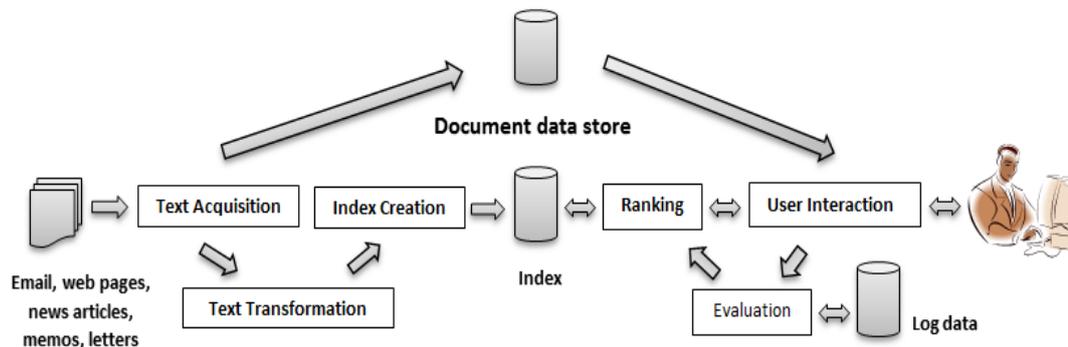


Figure 1. Indexing and Query Process (a combination of Fig. 2.1 & 2.2 from Croft & Metzler & Strohman 2009)

Figure 1 is a combination of the indexing process and query process. The indexing process can be seen from the left side of the Figure 1. Text acquisition usually uses for collecting text by crawling or scanning email, web pages, news, articles, memos, letters or other information sources. The text transformation component transfers the crawled text to be index term or feature, and through the index creation component to save into the index for future search uses. In addition, content and metadata of the crawled text will also be saved in document data store for showing in the user interaction when searchers search.

In the right side of Figure 1, the query process is shown. User interaction component is a connection between human user and search engine. There are two functions for user interaction. One of the functions is to receive users' queries and transfer them to be index term. Another function is to receive the sorted result list from search engine index by pass through ranking component. In addition, there are also some metadata from document data store showing to searchers by pass through user Interaction component. The evaluation component is an offline action used for evaluating and monitoring the result and efficiency of the system. The log data is used for record and analyze users' behavior.

3.2 How does Google search engine work?

When searchers search in Google search engine, they are searching in Google's index of the web. The whole procedure begins with crawling and indexing pages. Google's crawling robot is named "Googlebot". Googlebot arithmetically decides which sites need to crawl. Moreover, Googlebot regulates how often and how deep to crawl each website by following links from page to page on the web. (Lievonon 2013, 13.)

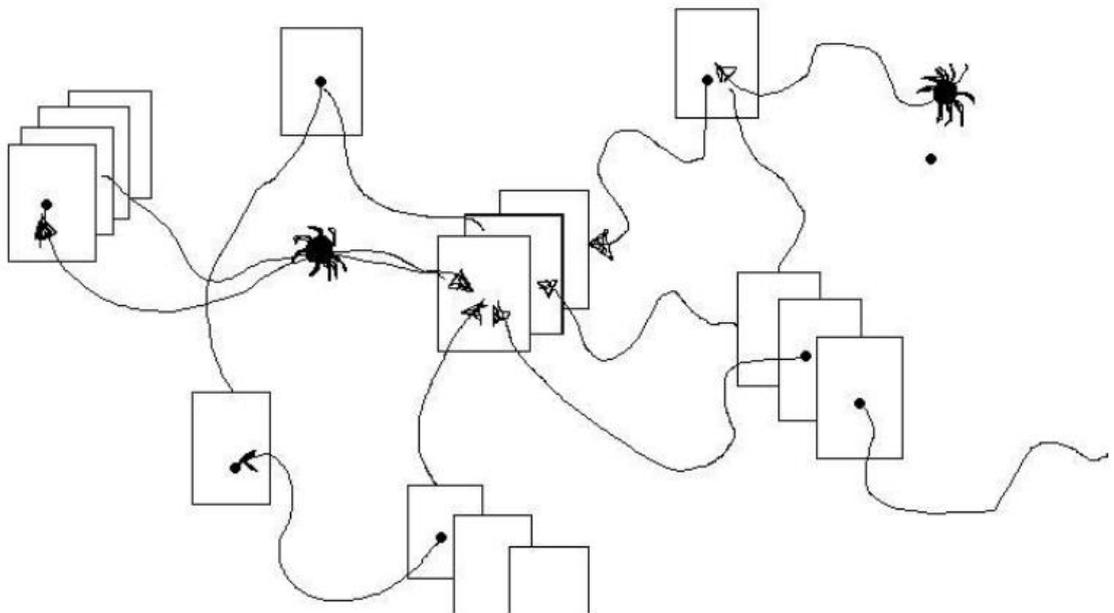


Figure 2. Googlebot (Lievonon 2013, 14)

Figure 2 shows the process of crawling. As can be seen, Googlebot starts to crawl by fetch a few web pages and repeat fetch and follow the links from extended pages. The process is crawling and fetching will not be ended until the final ranked websites are shown on the result page.

When Google search engine receives a search query from a user, in order to find every page that includes those search terms, Googlebot will search index as showed in Figure 1. To decide which documents suitable matching the search query, Google will consider the factors such as keywords density, the

placement of keywords and the quality of the webpage. Google uses PageRank algorithm to rate the importance of webpages, in order to make the final decision. PageRank depends on how many outside links pointed to the webpage and how important those links are, and combine all those factors together to produce an overall score for each webpage and send the ranked search result back to the searcher within half of a second. (Cutts 2010.) PageRank will be discussed in Chapter 4 of this thesis research.

Google encourage webmasters to use the White Hat SEO instead of the Black Hat SEO. The White Hat SEO and the Black Hat SEO will be discussed more in detail in Chapter 4 of this thesis research.

In addition, Google advocates webmasters to pay attention to users' experiences instead of solely consider the Googlebot. Google hopes webmasters to treat the Googlebot as same as the human users. (Google 2012a.) In other words, Google wants the webmasters making excellent websites for users, not just for algorithms.

4 GOOGLE ALGORITHMS

To determine the rank of a website in the search result, several algorithms were created by Google. In Google search engine, once the keywords are typed in the search query, there will be a significant amount of websites show in the search result page, which match this search query. However, there are about ten websites showing up on each search result page. In fact, most of the people will not look at the search result in ninth or tenth search result pages (Ledford 2009, 34). Why is there a difference in the rank of websites? In other words, what impact the rank in Google search engine?

According to Google, “Algorithms are computer programs that look for clues to give you back exactly what you want.” (Google 2014b.) Therefore, the goal of Google algorithms is to deliver the highest quality and relevant websites to users’ search queries. In order to help searchers find websites that provide a great user’s experience and fulfill their information needs, Google designed useful algorithms such as Google Panda, to fight with the Black Hat SEO techniques. In other words, the Black Hat SEO techniques are against Google’s quality guidelines.

In order to find the most useful information from thousands of search results, the computer program, Google algorithms rely on more than 200 unique signals or ‘clues’ to make it possible to find out the expected answer (Google 2014b). Understanding what the signals or ‘clues’ that Google algorithms care, helps with making websites showing up in the first couple of pages easier. On the other hand, high rank will also increase the amount of visitors.

4.1 PageRank

Google PageRank is one of the algorithms, the main purpose of which is to improve the quality of web search engines (Brin & Page 2014). PageRank is made by the two founders of Google, Larry Page and Sergey Brin, for the purpose of measuring the importance of websites (Karch 2014). In order to inter-

pret the importance of PageRank point and a way to improve the quality of web ranking, a research of PageRank is needed.

PageRank is basically about the number and importance of links pointing to a website (Cutts 2013c). Particularly, PageRank can be interpreted as voting as well, every link that directs to the website count as a vote. In other words, only getting more links is not enough for a high rank in Google search engine, and therefore, webmasters should also consider the quality of the linked websites. According to Sullivan, there are 11 points on the Google PageRank system, from 0 to 10 (Sullivan 2007). Point 10 indicates the highest PageRank value, and point 0 is the lowest (Sullivan 2007). In other words, the higher point the PageRank is, the higher possibility of ranking in the search result page. An example to interpret the idea of PageRank is illustrated in Figure 3.

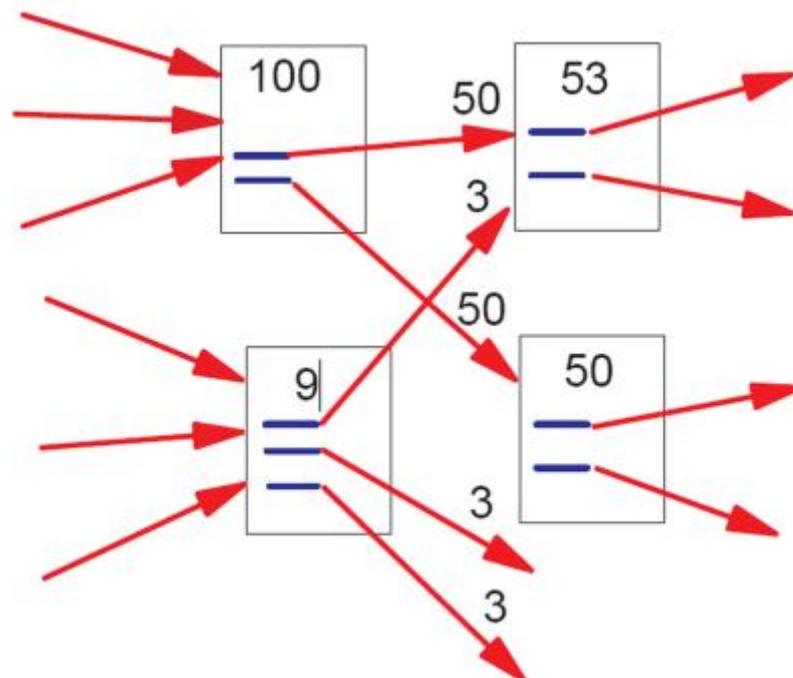


Figure 3. PageRank Example (Cutts 2009)

Figure 3 is an example explaining the basic idea of PageRank. In fact, the highest PageRank point is 10; however, the points in the example are only use for understanding easier. In Figure 3, there are two websites that have three in-

bound links for each website on the left of the figure. Additionally, these two websites are assumed as 100 PageRank points and 9 PageRank points, as in Figure 3. If a website on the left-top has two outbound links, one link goes to the right-top website, and the another link goes to the right-bottom website. Each of the website gets 50 points, because there are two links separated by 100 points. The website on the left-bottom has three outbound links from 9 points; hence there are 9 points goes to three different websites, each website gets 3 points. Therefore, the website on the right-top gets 53 points in total, which 50 points from the website on the left-top and 3 points from the website on the left-bottom. The website on the right-bottom only gets 50 points in total, which it is all from the website on the left-top.

To increase the value of the webpage in the search engine; in other words, to increase the points of Google PageRank, the best way is to add well-known and reputable incoming links. If there are more famous and reliable incoming links, the value of the website is higher which means that the quality of the incoming links is paramount.

In order to improve the value of PageRank, it is important to verify that links are only shared with relevant websites, and also to ensure that the shared links are not the 'link farm'. Otherwise, the result will be the opposite of what the website was aimed at. The link farm describes the type of websites that contains a high number of irrelevant links to other websites (Karch 2014). In addition, to increase the value of PageRank, webmasters can also register the website with high PageRank directories, i.e. the Open Directory Project (Karch 2014).

In short, Google PageRank is an algorithm measuring the value of the website through the quantity and quality of external links. Furthermore, Google ranking is about each webpage within a website, not the entity of the website. The most important indicator to improve the PageRank point is to share links to related websites that have high PageRank points. Moreover, webmasters can also register under high PageRank directories to increase the value of PageRank.

4.2 Google Panda

Google Panda algorithm is aimed to against the low-quality content providing websites and to downgrade the rank of these types of websites. This algorithm is launched in 2011, and after this algorithm was launched, there were 12% of webpages' ranking in natural search results were impacted. (Sullivan 2011.) Therefore, this algorithm received a lot of attention as it affected plenty of websites listed on the search engine.

The primary target of the Google Panda algorithm is to block the content farm from showing in Google search results. As Google announced on their official website, the Panda algorithm is designed for reducing rankings of low-quality websites, such as the websites with copied content and invaluable websites. (Google 2011b.) The content farm is one type of shallow and low-quality content with very little effort, or no any effort for searchers (Google 2011a).

Another target of Google panda algorithm is to act against the scraper websites (Google 2011b). The scraper websites are the websites that do not have original content, but pulling content from other websites. Scraper websites are using because webmasters trying to rank a website higher and taking the first page in Google in the search result (Sullivan 2011).

According to Google, it was inevitable that a major change which Google Panda brought was going to affect ranking for many websites. However, Google also states their motivation as it is important for high-quality websites to be ranked higher, because of the efforts these websites spend on creating quality content. "Google depends on the high-quality content created by wonderful websites around the world, and we do have a responsibility to encourage a healthy web ecosystem." (Google 2011b.) Therefore, creating a website with high-quality content not only attract more visitors, but also be appreciated by Google.

The improvement of the algorithm from Google was one step closer to offer users useful and informative content. Google Panda has been updated 27 times

since its first released by far (Moz 2014). Unless there is no more low-quality content, Google will continue updating and changing the algorithms to keep a healthy web ecosystem. Yet, for a reason that webmasters are wary of significant changes, and therefore Google makes small changes to its algorithms weekly or even daily, but the changes are so small most web users don't even notice them. (Google 2011b.)

On the basis of a discussion, there is a lesson to be learned. In SEO, whether for an entrepreneur or a webmaster, there is a need to write unique and high quality content. Google encourages creating high-quality contents, such as original and informational content. Furthermore, the contents should contribute to searchers instead of for a sole purpose of ranking higher in Google search engine. Moreover, branding and spreading the websites across social media platforms is even more important. The aspects that help in creating high-quality contents accepted by Google are outlined in Figure 4.

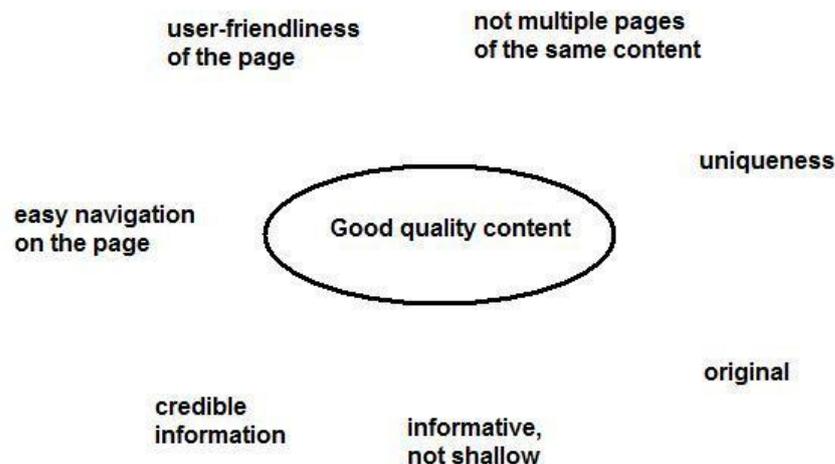


Figure 4. Outline of good quality content (Lievonon 2013, 20)

Figure 4 is a general overview of the needs for a website to be able to create good quality content. The important aspect is to create uniquely, original content while information should be credible. Information should be contributing and informational, in order to prevent Google panda labeling it as the 'thin content'

and ranking the website lower. In addition, there is a need to have a user-friendly layout and no repetition on multiple webpages.

More importantly, webmasters should keep in mind that the contents should be created for users. With this kind of mindset, the webmasters are able to create websites that users like to visit and also may recommend the website to other people. Consequently, having a healthy web ecosystem is what the Google panda algorithm aims for (Google 2011b).

4.3 Google Page Layout

Google Page Layout algorithm, also known as Top Heavy algorithm which was first launched in January 2012. This algorithm is used for checking the amount of content once users click on the result. The reason that Google makes this algorithm is because Google received complaints from users about unsatisfactory experiences during search. Users dissatisfied with the experiences of difficulty to find the actual content once they click on the result. However, "Rather than scrolling down the page past a slew of ads, users want to see content right away." (Google 2012b.)

Specifically downgrading of webpages also happens if the ads are deemed to be too distracting for using. This standpoint was confirmed by Google: "If you click on a website and the part of the website you see first either doesn't have a lot of visible content above-the-fold or dedicates a large fraction of the site's initial screen real estate to ads, that's not a very good user experience, such sites may not rank as highly going forward." (Google 2012b.)

However, Google announced that Page Layout algorithm does not mean that above-the-fold ads are fully not allowed. Therefore, a degree of how much ads to be used needs to be considered. (Google 2012b.) In addition, "pop-ups, pop-unders or overlay ads are not by this." The function of the Page Layout algorithm only applies to a static advertisement in a fixed position on the pages. (Sullivan 2012b.)

To avoid being penalized by Google Page Layout algorithm, it is important to consider the area of using ads, the size of the ads and to determine how much ads allowed in a webpage. In other words, it needs to be considered that how much of the visible content above-the-fold to visitors under various screen resolutions. In order to see how the website would look under different screen resolutions, Google Browser Size tool or similar tools are suggested to be used. (Google 2012b.)

Once when a website is penalized from Google, webmasters should make changes according to the algorithms, and consider whether the changes deliver a user-friendly experience. In order to change to be an acceptable website by Google Page Layout, the factors such as content, website structure, size and placement of ads need to be considered. In detail, webmasters should create original content and avoid copying or rewriting the same content from other websites. Moreover, structure and design of the website also valuable, it is encouraged to make navigation as easy and intuitive as possible. Figure 5 displays an example of the recommended page layout.

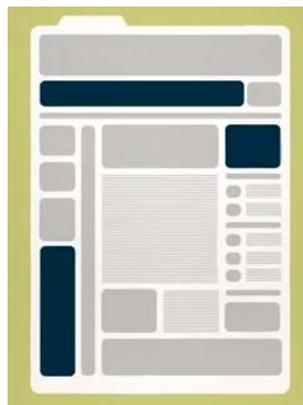


Figure 5. Page Layout 1 (Google AdSense 2012)

In Figure 5, the black squares are the recommended placement of ads example. In addition, according to Google, the recommend ads size for leaderboard, which is the black square on the top of the webpage, is 728x90. Size 300x250 is for medium rectangle that is the black square on the right-top of the webpage.

The size of wide skyscraper that is the black square on the left-bottom of the webpage recommend as 160x600. (Google AdSense 2012.) The figure above shows a valuable option of where to place the ads, and how big the ads should be. To the contrary, Figure 6 below has a high possibility to be punished by Google Page Layout algorithm, for a reason of delivering bad user experience.

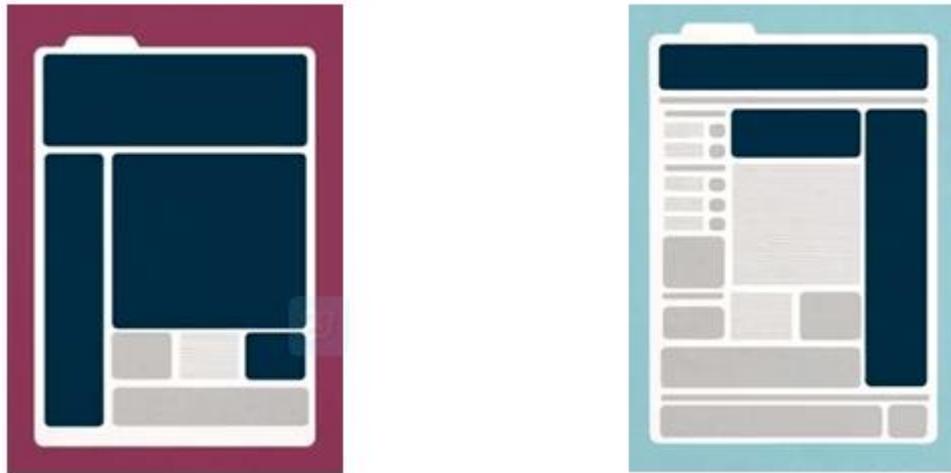


Figure 6. Page Layout 2 (Google AdSense 2012)

Figure 6 above are two webpage examples for placing ads in the wrong place or with a deprecated size. The webpage on the left side of Figure 6 has too big size of ads. Moreover, the webpage on the right side of Figure 6 is too concentrated on the top of the webpage, both of the w will make users to have a hard time in finding the valuable content.

The penalty does not instantly disappear right after changing the penalized website. Instead, Google makes note of the website when Googlebot visits the penalized website next time. A typical website needs to wait for a few weeks until the Page Layout algorithm automatically reflects the changes. Googlebot re-crawl and process enough pages from the penalized website to assess the changes. The duration of removing the penalty depends on several factors, such as the number of pages on the website and the level of efficiency Googlebot crawl the content. (Sullivan 2012b.)

In short, delivering the user-friendly experience is the main substance of Google. When creating a webpage with ads, the placement and amount of ads are important signals for the Google Page Layout algorithm. Too many ads in a webpage will be punished by Google with downgrade ranking. In order to improve the ranking, a reduction of ads above-the-fold is needed. The cancellation of penalization takes a few weeks after updating the changed page layout.

4.4 Google Penguin

The Google Penguin algorithm is designed to catch people who spam Google search results or aim to control Google rank on purpose. This algorithm was launched in April 2012 for the first time. The main goal of Google is to promote high-quality search results, and deliver high-quality search experiences to the user. (Sullivan 2012c.)

According to Google, Penguin algorithm affects all languages at the same time. A regular user might notice a total of 3.1% of queries in English to a degree. Moreover, there are roughly 3% of queries in languages such as German, Chinese and Arabic got affected. A language such as Polish received 5% noticeable effects because it is one of the heavily-spammed languages. (Google 2012a.)

Webspams are the techniques that for the sole reason to higher ranking or traffic. Moreover, the techniques of webspams do not benefit users. Moreover, the exploit aims to use shortcuts or loopholes to rank websites higher than they deserve to be ranked. (Google 2012a.) Google webspams are used to improve Google search results ranking typically by using the Black Hat SEO strategies such as keyword stuffing, hidden text or links, link scheme, cloaking, purposeful duplicate content. All these Black Hat SEO strategies against Google Publishers Guidelines and they violate users' search experience.

4.4.1 Keyword stuffing

Keyword stuffing refers to the practice of filling a webpage with keywords or numbers in an attempt of manipulating website ranking in Google search results (Google 2014c). For the reason that keyword stuffing delivers a negative experience to Google users, Google downgrades the website with keyword stuffing by using Google Penguin.

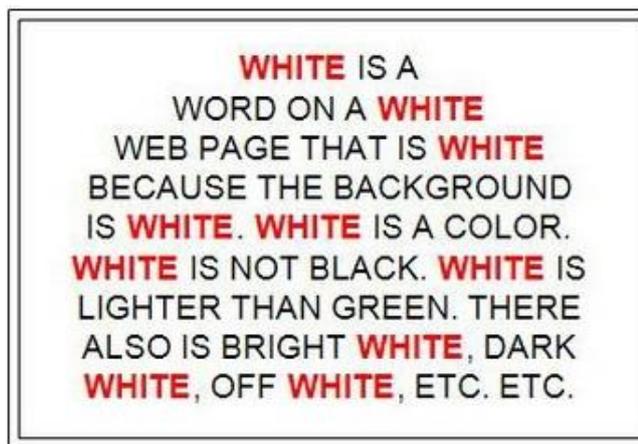


Figure 7. Keywords Stuffing

Figure 7 is an example of Keywords Stuffing; the keyword “WHITE” has been using ten times out of 42 words in total. The density is too high, and the repetitive keyword is not useful to searchers. The most of people do not want to read a paragraph with full of same words over and over unnaturally. The website as showing in Figure 7 is considered as keyword stuffing and has a high possibility to be penalized by Google Penguin algorithm.

In addition, Google has provided examples of keyword stuffing. For instant, “list of phone numbers without substantial added value; blocks of text listing cities and states a webpage is trying to rank for; repeating the same words or phrases so often that it sounds unnatural.” (Google 2014c.)

Keyword stuffing makes the website gross and disgusting. To avoid this unfriendly experience for the user, Google penalizes the website with keywords

stuffing. Moreover, the correct way of keyword using will be discussed in Chapter 5 of this thesis research.

4.4.2 Hidden texts or links

Hidden texts or links are another webspam. There are several ways in which keyword texts or links can be hidden on a website; such as camouflaging by using the same font and background color, hiding the keyword behind an image and using a 0pt font size.

The reason for webmasters using hidden texts or links is to try to manipulate Google ranking result, although these texts or links would not easily be viewed by visitors, it still can be read by Googlebot. Even this behavior does not violate to users' experience, though Google still penalize them since there should not be any texts or links solely for search engines rather than a human user (Google 2014d).

Google explains that not all hidden text is considered deceptive. Google suggests that using alternate descriptive text when dealing with images, JavaScript and flash plugins, animations or websites will dramatically improve the search ability of websites. (Google 2014d.)

4.4.3 Links Schemes

The purpose of Link Schemes is used for changing the search result ranking. Link Schemes are regarded as a violation of Google's webmaster guidelines. Link schemes include paid links, excessive link exchanges, keyword-rich anchor text links in articles or guest posting and comment spam. (Google 2014j.) Being tangled in a link scheming network might affect webpage badly and the downgrade the rankings. Google's primary goal is to deliver best relevant search result to searchers and to protect the searcher from bad experience. Therefore, Google recommends that the best way to get relevant links is to create high-

quality content (Google 2014j). Original, unique and relevant content will gain popularity on internet and the links will naturally follow.

Google reminds all webmasters that selling links that bypass PageRank against Google's quality guidelines, and Google does take actions for such violations. A clear and common paid link definition by far is a webmaster sells links to another webmaster, in exchange either for money or for anything that close to the value of money. (Cutts 2014b.)

Excessive link exchanges can be simply explained as 'Link to me and I will link to you' or partner pages completely for the sake of cross-linking. However, since there is no clear determine lines for excessive, webmasters have to 'wade across the stream by feeling the way' (an old Chinese proverb, translated by the researcher).

Comment spam, such as blog comment spam, spamming forum threads and spamming accounts on a free host. These spams comment only for the sole reason of trying to get a link from any platform. One type of the blog and forum comment spam is automatically generated content with links to sites that irrelevant with the article or forum content. Another type of comment spam involves seemingly friendly, but provide short and shallow comments with unrelated links to external websites in the poster's signature line, name, or profile page. (Huang 2012; Google 2014i.)

All advertorials and text advertisements bypassing PageRank are considered a violation of Google's guideline. Ignoring whether the advertisements and advertorials are paid or not. Again, PPC advertising links, which do not bypass PageRank, are subtle. Process of bypassing PageRank can be achieved by adding `rel="nofollow"` attribute to `<a>` tag or registering links to a page that has `robots.txt`. The latter process is for blocking search engine robots from crawl the webpage. (Sandy & Cutts 2013.)

When website penalized by Google, webmasters should either remove links or use “rel=nofollow” to not let these links pass through PageRank. Moreover, webmasters need to deliver as much as document that they can to Google to reconsider of the penalization. (Alex & Cutts 2013.)

4.4.4 Cloaking

Cloaking defined as showing different content to users than to Googlebot. Cloaking is considered as a webspam because it is a violation of Google’s Webmaster Guidelines, and it might deliver user’s awful experience. An example from Matt Cutts in the video, explained how does a cloaking website harm users’ experience: for instance, when Googlebot crawling for a webpage, the web server that was cloaking might return a page all about Disney cartoons, but when a user came and visited the webpage, the web server might return something else such as porn. This cloaking might bring a hugely sad experience to the user when a user does a search for Disney cartoons on Google, but get a porn website after clicking it. (Cutts 2011a.)

In addition, Matt Cutts reminds webmasters that cloaking is totally not acceptable in Google search engine. Webmasters should treat the crawler same as human users. (Cutts 2011a.)

4.4.5 Duplicate Content

The last but not the least webspam which listed on Danny Sullivan’s article is purposeful duplicate content, which includes syndicated content, affiliates and doorway websites. (Sullivan 2012c.)

Doorway pages are consisted of a vast number of low-quality webpages. Moreover, each of these doorway pages is optimized for a particular keyword or phrase. (Google 2014e.) Classic doorway page described the website with a ton of different webpages with very slightly different phrase or a different keyword in each page. It also explained in the video Webspam Content Violations by

Google,” Doorway sites are created purely to drive users to another website and often try to crowd the search results for a particular term or phrase.” (Wyszomierki 2012.) For example, a website with wonderful relevant content on the first page; however, after four or five pages the content is as same as the previous ones, the only difference might be some words in the content. More than one search result lead to the same website provides an unfortunate experience to users. Doorway websites violate Webmaster Guidelines; it can be negatively impacted in Google search results and can be even removed from Google search results (Cutts 2013d).

An affiliate website promotes products by providing users an eye-catching content, with the help of featuring product descriptions. Moreover, the affiliate websites appear on websites across affiliate. In fact, in order to attract people go purchase a product, a shopping website would like to link the product to a description webpage include original reviews and commentary. A high-quality affiliate website provides unique content and informational content. A low-quality affiliate website, also known as “thin affiliate”, would grab an affiliate feed from other websites. An affiliate feed is a description of different products that the merchant makes. Low-quality affiliate website just copies and paste the affiliate feed without any changes. (Google 2014f.)

To change the low-quality affiliate website, the website should add value and give the user a reason to visit the website. For instance, adding original insight, research, analysis, review and videos can be considered as adding value. “An example of a high-quality affiliate website would be a bicycle blog with original reviews and commentary. And an example of a low-quality affiliate website would be a bicycle blog that provides only copied product descriptions, accompanied by affiliate notes.” (Wyszomierki 2012.) A comparison of high-quality and low-quality affiliate sites could be in a contrast of blog contents. A blog, which provides original content, reviews and commentary, is seen as high-quality affiliate site. As for low quality affiliate sites would provide copied content, with affiliated notes.

Syndication content also known as “scrapped content”, it is the content that taken from another website High-quality syndication content which relevant for the topic and adding own value is accepted by Google. However, Syndication content with irrelevant content and without additional useful services or content will be considered as low-quality syndication content, also known as “thin syndication content”. Typically, in order to fill up webpages with a certain amount of words for a topic, webmasters grab low-quality syndication content from websites which offering free article syndication, such as low quality article bank or article marketing site. Since searchers do not respect this type of low-quality syndication content, Google considers it as webspam. (Google 2014g; Cutts 2013d; Wyszomierki 2012.)

Once the websites got penalized because of the thin content, webmasters can either remove the content that is scrapped content and without adding value or adding unique or addition informational content to make the site compelling. All these changes are designed to make user like the website that they will bookmark the website or introduce it to other users.

Hopefully, webmasters can look at a website from the perspective of searchers or regular users to understand why doorway pages, thin affiliates, low-quality syndication content are disliked by Google and considered as webspams.

In addition, Matt Cutts mentioned in his blog that webmasters, who use guest blogging solely to gain links, should probably stop in 2014, because over time this behavior is more and more considered as webspam. According to Matt Cutts, some of the guest blogs have low-quality articles, for example offering the same blog post multiple times, spinning the blog post and offering it to various outlets, should be stopped. However, not all the guest blogging should be arrested. In others words, the high-quality bloggers are encouraged by Google. (Cutts 2014f; Cutts 2012.)

In short, all algorithms are designed to provide a user-friendly experience to Google users. The simplest way to avoid by penalization from Google is to con-

sider of the user-friendly experience, always remember the primary goal from Google is to serve users with high-quality and relevant information.

5 OPTIMIZATION

In fact, Google advocates webmasters to use SEO, but only the White Hat SEO. This proposition can be confirmed with the Google official website, “Google also present that the White Hat SEO often improve the usability of a site, help create great content, or make sites faster, which is good for both users and search engines.” (Google 2012a.)

The difference between the Black Hat SEO and the White Hat SEO is about the techniques and strategies that used for trying to improve a website’s search ranking.

The uses of the Black Hat SEO techniques and strategies break Google’s rules. The Black Hat SEO solely focuses on search engine, but not so many focuses on human users. The Black Hat SEO is typically used by those websites which willing to get a quick return on their websites, rather than a long-term investment of the websites (Google 2012a). Some techniques are used in the Black Hat SEO such as keyword stuffing, doorway pages, hidden texts and links, and guest blog content spamming which were discussed in Chapter 4.4. In addition, the Black Hat SEO will be penalized by Google search engine and downgrade the rank of the webpage.

In contrast, the White Hat SEO follows Google’s rules, and it focuses on both search engine and human users. The White Hat SEO expects to make a long-term investment in the website. The White Hat SEO typically uses techniques such as using relevant keywords, writing a clear meta description and high-quality content. Google encourages webmasters to use the White Hat SEO to advance user-friendly experience (Google 2012a).

SEO can be positive and constructive. In fact, effective SEO is an essential factor in online marketing, and it can be the difference between the success and failure of a business, by making relevant pages easier to find thus the site more accessible and ranked higher in the pool of search results resulting in more

traffic. Google advises SEO experts, web developers and designers to utilize the White Hat SEO techniques that are compliant with search engine regulations, such as relevant meta tags versus the Black Hat SEO such as keyword stuffing, which could get a website owner penalized or banned. (Google 2012a.)

There are many factors impacting the page ranking in a Google search result page, such as keyword, links and content. In this chapter, the main exploration is on how to use these factors to improve the page ranking result. However, the factors and methods which will be mentioned in this chapter is not all SEO has to offer. The factors being explored in this chapter are significant and essential factors, including content, keyword, links and sitemap.

During optimize the website, webmasters should always remember to optimize not only for Google search engine, but also for human users. Webmasters should keep Google's goal in mind which is to show searchers the best with most relevant search result for a search query.

5.1 Quality Content

It is important to have high-quality content for a website. Content is the most important aspect of a website; it is as important as the head of the document (Cutts 2014c). High-quality content can bring more Google searchers to visit, and also high frequency visiting helps the webpage getting higher rank. This is because Googlebot considers the more visited website as a high-quality website, and Google likes high-quality websites. In contrast, low-quality content will be penalized by Google algorithm, such as Google Panda.

Google recommends original, engaging and valuable content. Webmasters who concentrate on link building limits themselves to a particular mindset, therefore, compelling content and marketing is more important than build links. The Google search engine is not the only reader that webmasters create a website

for, compelling content will bring more searchers to visit the website and to link to the webpage instead.

The key in building a quality website that performs well in Google search result is to focus on the website and its content to users. There are many factors that determine where a website displays in the search results. Some of these factors are explained subsequently. First and foremost, it is important to create content that is original. Therefore, webmasters should avoid using duplicate content and other non-value content. In addition, it is paramount that the website has a clear message or goal that it wants to convey to its users. Secondly, differentiate the website from other websites. This is done by building a website that is unique and stands out from other similar websites. Make searchers aware of the exclusive content that are offered and inform them about the content that may be available elsewhere. Another factor is to regularly update the website, and a good way to achieve this is to keep a blog. Keeping a blog keeps the site fresh; searchers will be engaged whilst always have new content to digest. Lastly, periodically update the website and archive any content that is no longer relevant. (Douvas 2012.)

In addition, according to Danny Sullivan, Google also considers ranking by searcher's reaction; such as if searchers stay on a page for a long time; if the searcher share the content; if searchers make links to the content from their own blogs. (Sullivan 2012a.) The signals above help Google to determine whether the website has high-quality content. To make searchers stay longer in a webpage, webmasters should create compelling and relevant content, which is useful and informative for searchers. Compelling content has a high possibility to be shared on searchers' blogs.

5.2 Keyword

The keyword is a core of all the SEO methods. In fact, every SEO method more or less related to keywords. In this chapter, the selection of keywords, the placement of keywords and the keyword density will be mentioned and ex-

plored. Furthermore, the keyword optimization which will be discussed in this chapter is all based on organic keywords. Organic keywords are those which are free of cost, and naturally will appear in search result page as well (Kim 2014).

Potential visitors will use the keywords to find relevant websites which are related to their search topic in the search engine. Selecting the appropriate keywords will attract more search engine users to the website. Based on research that most of the search engine users will not even check the websites that are ranked in the second search result page (Ledford 2009, 80). The research shows that the keyword optimization does not only determine the website ranking, but also decides if the search engine users are able to see the website.

5.2.1 Keyword selection

Keyword selection is the most essential parts of SEO; however, it is also the most important part. The quality of keyword will impact all other aspects of SEO. Keyword selection is important because the choice of keywords will determine the ranking of the websites on the search results page. (Ledford 2009, 92.)

Keyword can be divided into two : Brand keywords and Generic keywords. Brand keywords are used to define the keywords that are directly related to the company brand, such as Microsoft. In contrast, Generic keywords describing the type of keyword that has no direct relations to the company brand. Brand Keywords and Generic Keywords are both important. Despite Brand keyword itself is tightly certain up with the website; it is still important to use them as a keyword. (Ledford 2009, 92.) For instance, if the brand name was not being used in the official brand company's website, such as not in title, description and category on the website; there is an opportunity for another website which used this brand name as keyword to get higher rank. The users who search for this brand name as a keyword would not find the official website of the brand. Therefore, the official brand company's website would not show up on the

search result page; consequently the brand company may have lost potential customers.

The very first step of SEO is that to select keywords that are related to the content of the website. Otherwise, the irrelevant keywords will be considered as keyword stuffing, and Google Penguin will impact the rank of the website. (Google 2014c; Sullivan 2012c.) On the other hand, the penalized website will lose the chance to create good, credible business. For example, a website for shoes uses an irrelevant keyword 'rice' as keyword. When user who visit the website by search 'rice', they would not find the expected search result and they would turn to another website directly; this bad experience might cause these visitors refuse to visit this website anymore even when they need shoes, because they might consider this shoes website as not reliable website. In other words, the irrelevant keywords not only unable to bring any sale to this website, but also will lose the future buyers because of the bad experience for users.

Keywords should not be selected too broad. The broad keyword not only unable to bring any sale on the website, but also brings more pressure of competition. Website keyword selection is a process of gradual discovery. Most of the time, it is not enough to only pay attention to a certain keyword, but also should follow closely to the derivative keyword. (Ledford 2009, 92-93.) Therefore, when selecting a keyword, webmasters should consider what keywords that target audience will likely use in searching. On this basis, keywords that are with high interest and little competition are useful options.

Webmasters should make sure the webpage having the right keywords which people might type (Ledford 2009, 51). For example, a Finland travel agency website. If the webmaster of this website wants people to find it, he/she should not just use the keyword 'Finland', but to use the keywords phrase such as 'Travel in Finland' or 'where to visit in Finland', because if people plan to travel to Finland, they would type the keywords phrase such as 'Travel in Finland' or 'where to visit in Finland' instead of just type keyword 'Finland'.

Moreover, when choosing keywords, webmasters should avoid issues, such as selecting only one keyword for a website and using the same keyword in all webpages. In fact, there are specific topics for each webpage, and therefore the keyword of each webpage is also distinct

5.2.2 Keywords Placement

When the keywords are chosen where to place them need to be considered. Based on the regular pattern of Googlebot crawling, there are several important regions to be considered, such as title tag, heading tag and meta description.

Webmasters should put the keywords in title and description on the important webpages of a website, such as a home page. The reason is that the keywords in title will help searchers to find what the webpage is about.

In search engine, the title tag is one of the most important areas to place the keywords. When a keyword is typed in a search query, the title of the web page shows up in the first search result page, as being indicated in Figure 8.

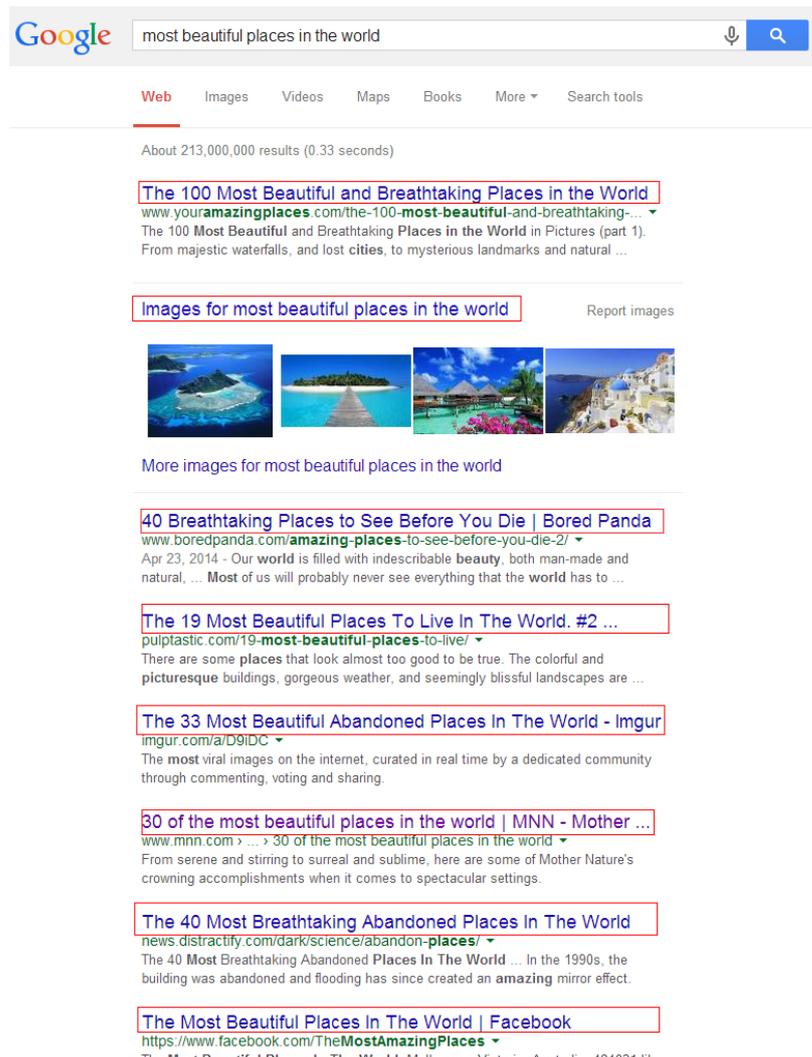


Figure 8. Title tag 1

As Figure 8 illustrates, when the search results appear, the text of the title tag is used as the main title for each web page. These main titles will link to the webpages by clicking. Thus, search engine users will decide which website to visit. After the keywords “the most beautiful places in the world” was the type in Google search query, the first page of the search result is shown in Figure 8. The title and link to each website are surrounded by a red square. By simple observation, all the results on this search result page include the keywords that are contained in the search query.

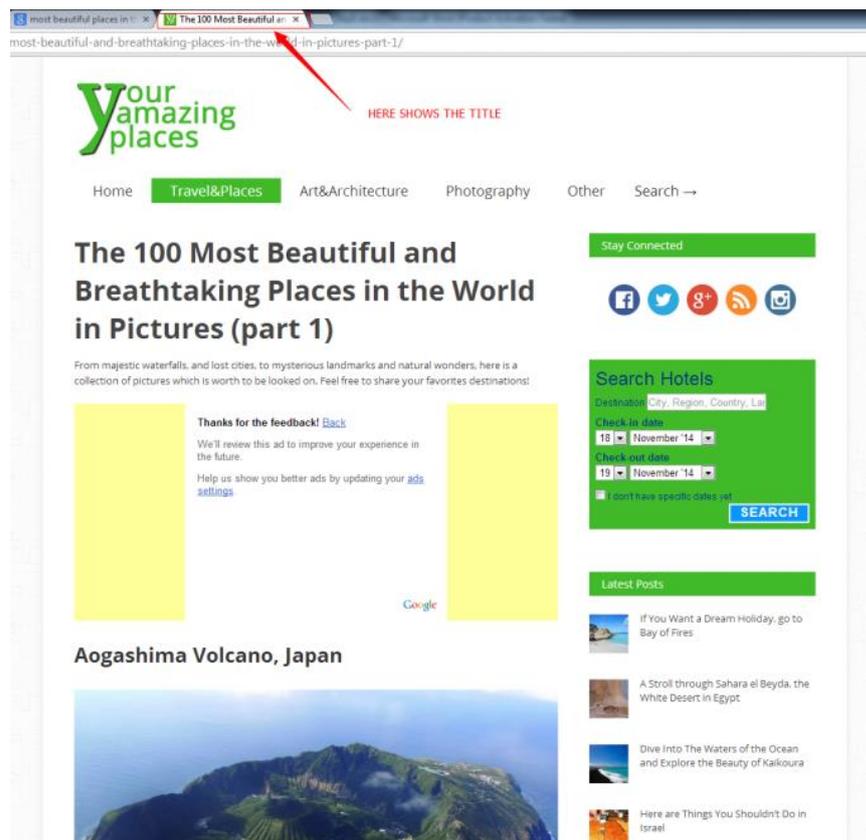
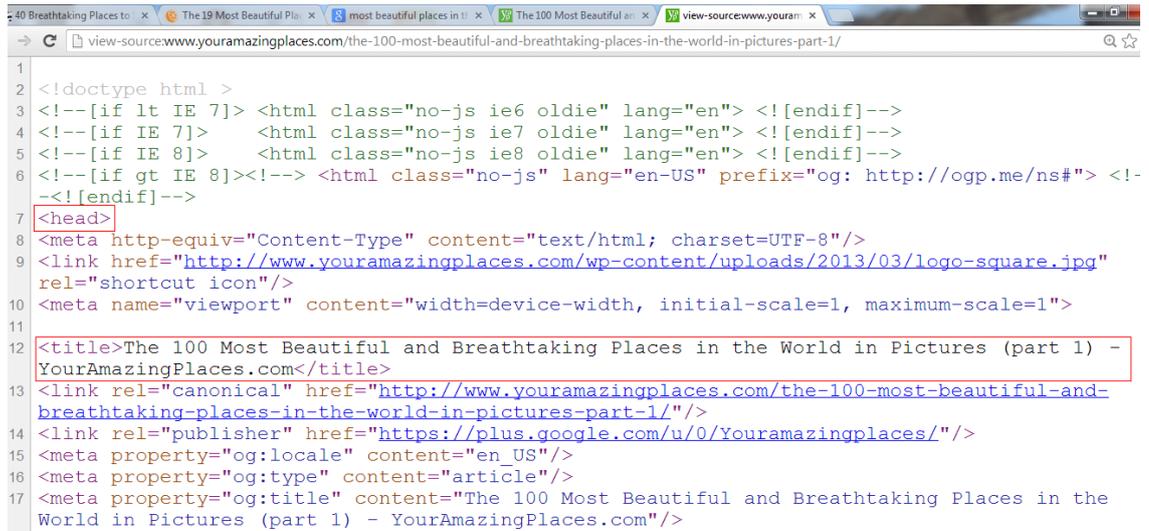


Figure 9. Title tag 2

Figure 9 is a website after clicking on the first title link from Figure 8, on the top of the page. The red square squares the title; as can be seen, it is not the whole title, but a part of it. This is because Google search engine has a title length limitation allowing users to read easier, especially when more websites are open at the same time. This rule has been confirmed in Welcome to Google's Search Engine Optimization Starter Guide, "Titles can be both short and informative. If the title is too long, Google will show only a portion of it in this search result." (Google 2010). Therefore, to avoid a situation where some keywords are not shown to users, webmasters should keep the titles clear, explicit, natural and put the keywords as in front as possible in the title tag. Moreover, webmasters should avoid keyword stuffing at all cost. The Figure 10 below is the source code of the website in Figure 9.



```

1 <!doctype html >
2 <!--[if lt IE 7]> <html class="no-js ie6 oldie" lang="en"> <![endif]-->
3 <!--[if IE 7]> <html class="no-js ie7 oldie" lang="en"> <![endif]-->
4 <!--[if IE 8]> <html class="no-js ie8 oldie" lang="en"> <![endif]-->
5 <!--[if gt IE 8]><!--> <html class="no-js" lang="en-US" prefix="og: http://ogp.me/ns#"> <!--
6 <![endif]-->
7 <head>
8 <meta http-equiv="Content-Type" content="text/html; charset=UTF-8"/>
9 <link href="http://www.youramazingplaces.com/wp-content/uploads/2013/03/logo-square.jpg"
10 rel="shortcut icon"/>
11 <meta name="viewport" content="width=device-width, initial-scale=1, maximum-scale=1">
12 <title>The 100 Most Beautiful and Breathtaking Places in the World in Pictures (part 1) -
13 YourAmazingPlaces.com</title>
14 <link rel="canonical" href="http://www.youramazingplaces.com/the-100-most-beautiful-and-
15 breathtaking-places-in-the-world-in-pictures-part-1/" />
16 <link rel="publisher" href="https://plus.google.com/u/0/Youramazingplaces/" />
17 <meta property="og:locale" content="en_US"/>
18 <meta property="og:type" content="article"/>
19 <meta property="og:title" content="The 100 Most Beautiful and Breathtaking Places in the
20 World in Pictures (part 1) - YourAmazingPlaces.com"/>

```

Figure 10. Title tag 3

Figure 10 is the code for Figure 9; the reason for the page length limitation is only a short part of the web page source is shown. The red circled part `<head>` is the beginning of the head tag. To finish and close the head tag, the mark `</head>` has to be included where the editor wants the header to end. In Figure 10, it is not shown up because of the source code limitation. For all tags in programming language, `<tag name>` is a starting of the tag, to close the tag is marked with `</tag name>`. “The `<title>` tag should be placed within the `<head>` tag of the HTML document.” announced in the Google SEO Guide (Google 2010). If the title tag is not placed within `<head>` tag, the website will fail to run. In Figure 10, the text between `<title>` and `</title>` is the main title of the website in Figure 9. This text is also shown as the first bold title in Figure 8.

A good title should contain keywords as this will help popularize the website with attractive and comfortable understanding. In addition, when creating the titles, the order of the words should be considered. By a reason of Google will show only a portion of the title in the search result, and therefore, the keyword should be placed as in front as possible. (Ledford 2009, 131-134.)

Likewise, the keywords in the meta description can sometimes determine if the searcher will be enticed into clicking or not. In fact, as there are a lot of popular but irrelevant keywords stuffed in meta tags are created by webmasters, who do

the Black Hat SEO, Google considers the meta descriptions with not that important for ranking (Ledford 2009, 134).

However, although meta tag does not impact the rank, meta description is still very useful. To help searchers when their intention is not clear for the search or different types of users have different motivations, Google shows a title and snippet for each website. The snippet, typically directly from the meta description tag, that is why the meta description tag is especially important. (Ledford 2009, 135.) Figure 11 displays the meta description on the search result page.

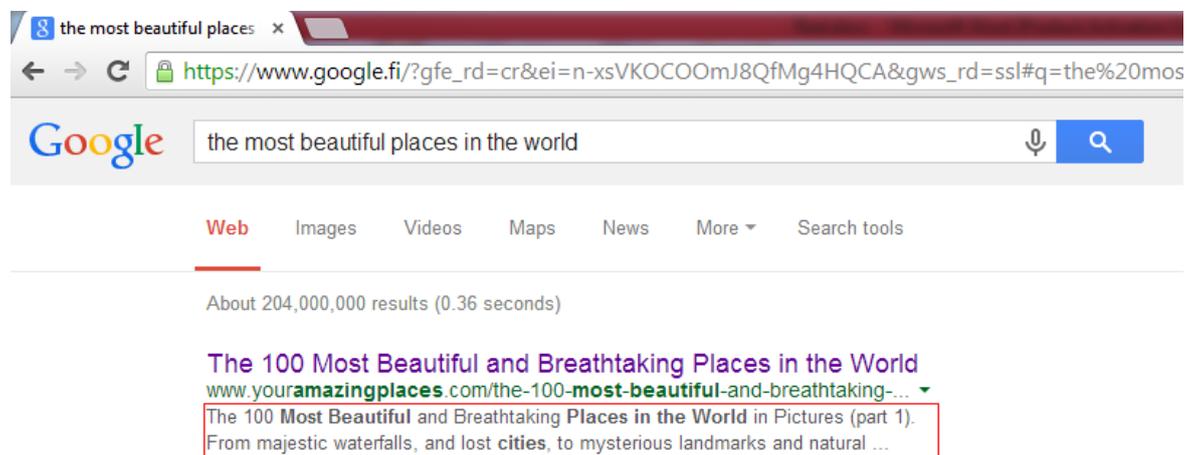


Figure 11. Meta Description

In Figure 11, the area that is marked with a red square is where the snippet shows on the Google search result page, in other words, it is where the meta description tag shows. It is better to keep the length of meta description tag clear and short, although the maximum of the meta description allows more length (Ledford 2009, 135). By cause of Google only displays is a snippet of the meta description on the search result page, as can be seen in Figure 11, the last will be truncated.

To distinguish whether the webpage is useful, Google highlight some of the places that query keywords occur on that page. As can be seen in Figure 11, the bold words in the circled area are the query keywords. Therefore it is important to place keywords in the meta description content.

For a very high-traffic website, a great experience and a real clicking from searcher are both significant. In order to attract searchers to click on the website, keep the description as relevant as possible will make the searcher know what the website is exactly about.

It is important to put keywords in <h1> heading tag. Heading tag is a common name for h1, h2, h3, h4, h5, h6 tags. In HTML, heading tag is used to identify the level of header in the file. In these six tags, h1 describes the main title; it is the most important one, and h6 is the smallest title. It is written as <h1> text </h1>. In each webpage, all levels of heading tags are accepted; however, h1 tag should be used once only in each web page, this is because using the h1 more than once can be judged as spamming by Google search engine. (Ledford 2009, 52-53; Enge & Spencer & Jessie & Fishki 2012, 219.)

As discussed previously, the h1 tag is the most important. Therefore, the content of h1 tag should contain relevant keywords because search engine values h1 tag. The search engine will pay more attention to the keyword in the h1 tag. (Ledford 2009, 53.) Moreover, the style of the heading tag will not impact the value of the website, which also means the size of the text does not matter for Googlebot (Enge et al. 2012, 219.)

However, too many keywords in a webpage will cost a problem of keyword density. According to Matt Cutts, there are no exact answers for this percentage. Google believes that the keywords should show up in a natural way and not in an artificial way. However, too many keywords will be penalized because Google thinks same keyword or keyword phrase showing too many times in a webpage will not help readers by reading the content, it can be considered as keyword stuffing as discussed in the previous chapter. (Cutts 2011b.) Figure 12 illustrates the curve of how Google considers the keyword density.

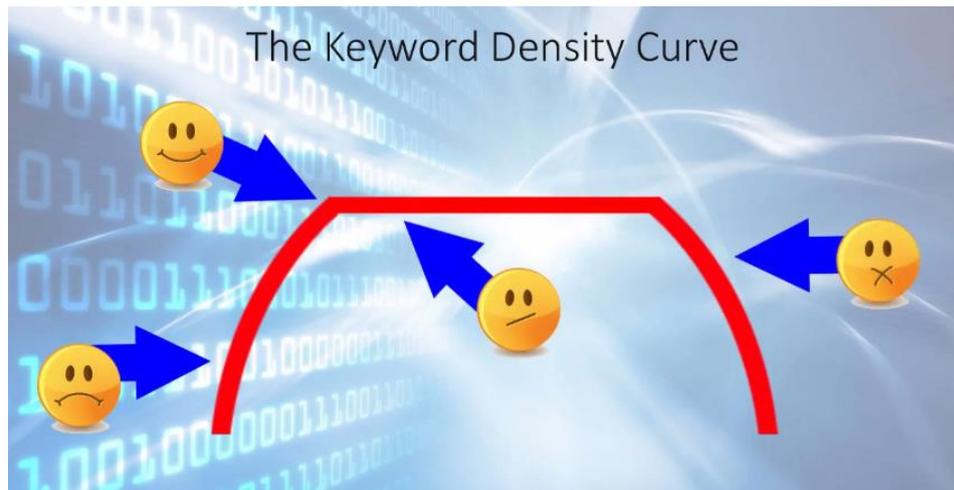


Figure 12. Keyword density curve (Hill 2014)

In Figure 12, the keyword density curve is presented. According to Google, the best way to obtain a good ranking is to write natural and relevant text which will gain popularity on the internet over time. The use of keywords is an important aspect because this is one of the criteria which are followed by Googlebot. In Figure 12, when a relevant keyword is found in the content, ranking algorithm will position the website in a higher position. If another keyword related to the first one is found, the ranking position will increase again. However, the algorithm is much more complex and by having a very high percentage of keywords will not guarantee a top position. Google will stop increasing the ranking or even decrease it if too many related keywords are found. However, there are not an exact percentage of the keywords that must be used. Therefore, there is no guarantee to get a webpage on top ranking if it will have the same percentage of keywords density as a top ranked page. (Cutts 2011b.)

To check whether the keywords was using naturally Matt Cutts recommends webmasters to read the articles which include keywords either aloud, or read it to someone else, or having someone else to read it, in order to check if the article sounds artificial or stilted. (Cutts 2011b.)

5.3 Linking

Linking is another important factor for SEO. The strategy of linking determines the role of a link in improving the search engine ranking. However, inappropriate linking strategy, such as a link farm, will also cause the website to be punished by Google search engine. One of the functions of linking is to link the website to other relevant websites to increase the website traffic. Another important function of linking is to recognize the value of the website by getting inbound links from other well-known and reputed websites. (Ledford 2009, 259-275.) Based on PageRank, the more inbound links from high-quality websites, the higher PageRank value can get. Therefore, website can get a better ranking in the search results.

However, not only inbound links are important, but also outbound links and internal links. The relevance of links as important as the quality of the links, therefore, webmasters should create links which relevant and valuable for users. In addition, the quantity of links is also important. (Ledford 2009, 268-269.) Furthermore, too many outbound links might be considered as a link farm by Google.

According to PageRank algorithm, the external links are important for a website. The quality of external links will affect website's PageRank value. External links include inbound links and outbound links (Ledford 2009, 43).

Inbound links also known as "backlinks", it is any links that received from other websites. For example, if the user A write a blog that user B finds interesting, and therefore write a relevant post on his/her own blog to link back to the user A's original post. As discussed previously, inbound links are equivalent to voting from a website to another website, and therefore it is important to get more backlinks. Backlinks help in terms of making sure that Google returns the best, most relevant, most topical set of search results. (Cutts 2014a; Google 2014I.)

Only have inbound links is not enough, to create more relevant outbound links is also important. Webmasters should collect high-quality relevant websites, and create outbound links to user-friendly experience, hence getting more repeat visitors. Linking to relevant links does not mean webmasters have to link to a website in the same industry, related websites are all can be linked. For instance, a shoe retail website can create an outbound link to fashion magazine website. The quality of outbound links needs to be considered carefully, linking with a high-quality website will deliver an advantage of high PageRank value. In contrast, linking to low-quality websites will provide a disadvantage of high PageRank value. Therefore, webmasters should try to avoid linking to low-quality websites such as a link farm websites; it has a great possibility to be penalized by Google. In addition, webmasters should not create a website with outbound links only; it can be considered as webspam. Moreover, webmaster should check the availability of outbound links often, and repair the invalid links. For the reason that search engines believe that a website with too many invalid links will deliver a bad experience to users, hence downgrade the ranking. (Ledford 2009, 268-269.)

Internal links are used to direct users to different webpages within the website. Moreover, internal links also help Googlebot to go deeper to the websites; it helps with PageRank. Webmasters should link all relevant pages together; in order to avoid users use “back button” too many times, for a reason that Googlebot might consider that users use too many “back button” because the website is not friendly. (Ledford 2009, 273-274; Cutts 2013a.)

An internal link can be placed either close to the article or keep it on the bottom, both ways can be accepted by Google PageRank as long as the article is original content. However, Matt Cutts personally suggests in his video that the anchor text of the link should be placed near by the article, for the reason that the links which leave on the bottom is not convenient for users to check. (Cutts 2013b.)

Google's primary purpose is to ensure that Google searchers get the best search results and to protect users from having a bad search experience. Therefore, all kinds of the link spams should be avoided. The link spams include paid links, comment spam, excessive link exchanges, link farm and other unnatural links as discussed in the previous chapter.

In conclusion, webmaster should ensure if the website is able to crawl. A website has to make sure there is no good content hidden away somewhere hard for people to find. To avoid this problem, webmaster should click around on the website to ensure that the pages on the website can be reachable by clicking on regular links, preferably in a text browser.

5.4 Other optimizations

The sitemap is used to help Google finding all of website's hosted content. Typically, a sitemap is a list of all webpages on a website, it is designed to help search engines discover and understand contents better and improve future crawling by Googlebot. (Lee 2012.) Webmasters create sitemaps to tell Googlebot, which content on the website can be indexed. The sitemap helps Google to understand where the content from website lives and what it is about. Therefore, Google can serve the content to users for relevant searches.

To create sitemaps, webmasters can either create by writing an XML sitemap file or use a third-party generator to create a sitemap file. When the sitemap file has been done, webmasters can upload it to Webmaster Tools. (Ledford 2009, 304-308; Lee 2012; Google 2014h.)

5.5 Case study

In this sub-chapter, a simple case study on the website is done. Justification of why a particular website was used for analyzing is presented. The optimization techniques and methods for analyzing were discussed in both Chapter 4 and Chapter 5 in this thesis. The case website will be analyzed by checking the con-

tent, keyword, links, and sitemap. Moreover, whether the case website has the possibility of being penalized by any of Google algorithms has to be taken into consideration. The final result is recommendation practices the case website should use, in order to rank higher in Google search result.

The case website, Technoupdates.org was found by searching low PageRank websites on Google search engine, and there was a list of relatively low PageRank websites displayed on '9to5blogger' (Krishna 2013). The case website, Technoupdate.org was selected because first impression of the website is good. However, by analyzing the case website from the perspective of user friendliness and SEO optimization, there are a lot of problems which could be enhanced through SEO optimization.

Even though Technoupdate.org was chosen for the case study, the integrity of information had to be confirmed. Therefore, different tools and services offered by various websites were used to verify the quality of the website. Firstly, www.9to5blogger.com displayed the PageRank of the website to be PageRank value of three, however after confirming a rank of the website on the website 'prchecker', the PageRank value was found to be one instead of three (Page Rank Checker 2014). However, as the aim is to find a low PageRank value website as the case website, therefore, Technoupdate.org still fit the criteria of being analyzed.

Reason of why the case website has very low PageRank has to be disclosed since PageRank of the case website is available. As discussed in Chapter 4.1 in this thesis that PageRank is related to the quality and quantity of external links. For this reason, the backlinks of the case website were checked from the website 'Backlink Watch' (Backlink Watch 2014). The number of backlinks to the case website seems decent, which about 1000 in total. As discussed in Chapter 4.3.3 in this thesis, link attributes without 'rel=nofollow' will all pass through PageRank, in other words, Googlebot is able to crawl these links. However after clicking some of the backlinks that Googlebot can crawl, there seems to be too many low-quality backlinks. Moreover, the most of the backlinks are unavailable

and gives an error, such as <http://www.eduways.in/p/disclaimer.html> and <http://gktricks.info/212-high-pr-web-directories-to-submit-your-website/>. In addition, among all these backlinks there are even 'link farm' websites which are not appreciated by Google, such as http://webnamelist.com/alexa/Alexa_6.html.

Furthermore, comment spam also used in this case website. Figure 13 below presents an example where author of the case website tries to gain backlinks, by commenting other people's blog for referring users to click on the case website.



Figure 13. Backlink commenting (Krishna 2013)

Figure 13 can be considered as comment spam by Google. Based on the discussion in Chapter 4.4.3 of this thesis, the comment in the figure is general, and anyone who does not read the whole blog would be able to write the same type of comment. In fact, the comment in the figure is short and shallow comment which only for the sole reason to try to get a backlink.

As Google has proven quality of backlinks thrives over quantity of backlinks, these backlinks not only fail to bring advantages to the case website, but also brings disadvantages for the case website, such as downgrade the rank of the case website.

As the case website was analyzed further, there are a lot of problems from the perspective of usability, user-friendliness and SEO optimization. It was inevitable that the website has low PageRank because of the quantity of low-

quality backlinks. However, the home page of the website does not even have a navigation function and searching function. Moreover, the contents are displayed without categorization and in generic formation. Furthermore, the only way users can use the case website is through pagination presented in the bottom of the page. Only use pagination is not convenient for users because user cannot find what they want immediately. In fact, users already want to quit the website after go through about three to four pages.

Changes have to be made in navigation function, for example creating a navigation bar. The navigation bar allows users to navigate to different webpages on the case website, and this makes the users more convenient. Moreover, search function enables users to search specific articles or content, which makes the use of the case website more convenient as well. Furthermore, although using pagination presented on the website is acceptable, how the pagination presented in the case website needs to be considered. As in the case website, it is presented without real motivation, meaning even if the users were to click next page, it would make no sense, because users do not know what is on the next page. Therefore categorizing the contents is a must as users may be able to present another interesting article. Specifically pagination would make sense after classifying the contents. These changes discussed have to be made if the webmaster of the case website wants to create a user-friendly website, which will turn into a successful website with a lot of visitors.

Based on Chapter 4.2 and Chapter 5.1 in this thesis, the quality of content in this case website need to be checked. For checking the key issues of the website, Siteliner was used (Siteliner 2014). The figure below is a result of the content on the case website from the website Siteliner.

Your Duplicate Content - Click below to see your duplicate content:

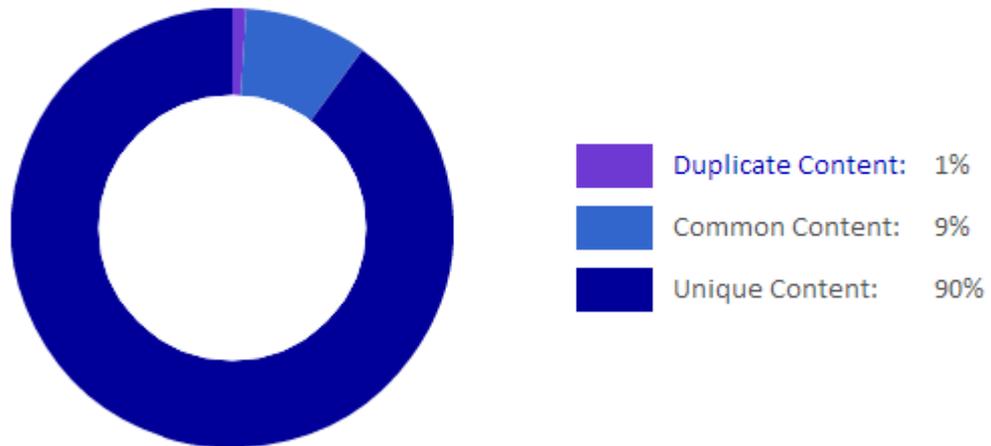


Figure 14. Content for the case website (Siteliner 2014)

Figure 14 illustrates the duplicate content percentage of the case website (Siteliner 2014). However, as a result in Siteliner, 133 pages were skipped as the free service provided by Siteliner offers maximum 250 pages to be inspected. Therefore, the result of duplicate content is estimation. Based on Figure 14, there are only 1% of duplicate contents and 90% unique content. Analyzing based on Chapter 4.2 and Chapter 5.1 in this thesis, to create quality content, Google recommends having original, engaging, valuable content, and should avoid using too much duplicate content. This result is showing that the most of contents using on the case website are original and avoided to use duplicate content, which is appreciated by Google.

However, after reading some articles on the website, the quality of the content seems to be in a common sense. Although the contents provided are unique, the content is not engaging. This unattractive content would not be expected by users who read the title of the webpage. In addition, some of the articles do not even have keywords; and it is bad if the webmaster aims for showing the website on Google search result page.

The website was created by using WordPress. WordPress is content management system which allowing users to add/edit/delete contents at ease. Therefore, the source code was analyzed in order to find problems. The source code for the home page and some of the webpages were analyzed as the case website is built based on them.

Firstly, keywords presented in each article are too broad and in generic level, such as the article '5 Reasons why Mac Users should buy upcoming Leap Motion', uses keywords "Mac Users" and "Leap Motion" are used for this article (Chopra 2012). As discussed in Chapter 5.2 in this thesis, keywords that are too broad and generic lead to having too much competition. Therefore, to rank in the first three pages of Google search result would be impossible. Another example for broad and general keywords is the keywords for the case website's home page, keywords for this webpage are Technoupdates, Techno Blog, Technology Blog, SEO, Blogging Tips and Wordpress. By searching these keywords in Google search engine, I found that these keywords are too general and with high competition. The case website should choose more concrete and low competition keywords or keywords phrases.

Secondly, by checking the source code of some webpages, there are not even keywords available, such as the article 'Ways that Technology and Marketing Are Symbiotic' (Chopra 2014). Based on Chapter 5.2 of this thesis, in order for potential visitors to find a webpage that related to their search query, keyword is necessary to be used. In other words, do not have keyword means do not have a chance for Google to list this webpage on Google search result page to compete the rank with other webpages.

Thirdly, there is no keyword in the title of the home page. By scanning the source code of the case website's home page, it can be seen that the home page uses 'TechnoUpdates.Org' as the title of the page, which means users only able to see this website by typing the title name as search query. Based on Chapter 5.2.2 in this thesis, keywords are recommended to put in the title tag and meta description tag to allow users to understand what the website is

about. There is meta description with the keywords for the home page; however, the keywords are too broad, which hard to fight for a good rank in Google's search result.

Fourthly, there is no <h1> and <h2> tag that are being used in home page. As discussed in Chapter 5.2.2 in this thesis, <h1> and <h2> tags catch more attention from Googlebot. Therefore, keywords should be used with <h1> and <h2> tags. However, <h1> tag is only suggested to be used once in each webpage.

Finally, there is no sitemap included in the source code. As sitemap is used for helping Google to find all contents that are hosted on the website. Therefore, including sitemap is important, because it is needed for Google to find and perceive content better, in order to improve future crawling by Googlebot.

Putting aside all the negative aspects of the case website from SEO perspective, there are some useful features. Firstly, the website has included a feature that allows users to find a website either with 'www' or without 'www', which delivers user-friendly experience.

Secondly, the contents provided on the case website are up-to-date, as latest article is on 8th of Dec, 2014. Moreover, the website offers also older articles for visitors' entertainment.

Thirdly, the percentage of duplicate content is about 1%, although the unique contents are general level information. The small percentage of duplicate content can avoid the penalization from Google Panda and Google Penguin algorithms.

Finally, faster page loading time delivers user-friendly experience. Based on the result of Siteliner, the average page load time is only 275ms, which the average for all other website is 1,079ms (Siteliner 2014). The case website's average page load time is better than 82% of all other websites (Siteliner 2014).

6 CONCLUSIONS

There is an increasing number of websites concentrating on SEO in order to fight for the first place in the search result page. The Black Hat SEO utilizes loopholes of algorithms, and therefore Google keeps changing the algorithms with the aims to fight the Black Hat SEO, in order to deliver a healthy web ecosystem to users. Consequently, using the White Hat SEO and considering of providing user-friendly experience will make the website possess a high ranking in search result.

In this thesis research, the basic knowledge of algorithms has been introduced. It is important to understand how algorithms impact on the website in order to avoid Google's penalization. The primary goal of Google is to provide a relevant result to users as fast as possible, and protect users from having a bad search experience. The changes in all Google algorithms are used to match Google's primary goal. In the other words, the purpose of the algorithms work is to organize all the information available and to be able to give the best relevant results to the search query conducted. Understanding the changes in algorithms will help webmasters to distinguish what Google allows and what Google rejects. In addition, factors in creating a website with the White Hat SEO have been discussed in this thesis research. Moreover, the suggestions of fixing the penalized website also are explored in this thesis research.

In order to get a high rank on Google search result, webmasters should make sure the website is able to be crawled by Googlebot, and have lucrative content to gain repeat visitors. Furthermore, Google encourages webmasters to use the White Hat SEO to optimize a website, in order to deliver users a friendly search experience. The list below is a summarization of the common White Hat SEO techniques, which helps webmasters in using the right way to do SEO and create a high-quality website.

- Creating original, engaging and valuable content instead of using duplicate content;

- Using natural and clear keywords or phrases to stand out of the text in content and without losing readability;
- Creating page title and leading text to clearly reflect the contents of the webpage;
- Setting up a structured data markup to help search engine identify the visible content of the webpage correctly;
- Making valuable comments on relevant blogs without solely expecting to get a backlink;
- Writing high-quality guest blogs for authority blogs; and
- Making validated code ensure the website can be displayed correctly in all main browsers.

In conclusion, to get a long-term high ranking in Google, webmasters should avoid using the Black Hat SEO. However, webmasters should use the White Hat SEO to make a high-quality website available for users. Increasing the search volume through enhanced marketing or creating the excellence content could lead to a larger percentage of visitors to the website. For example, valuable content on the blog could drive more subscriptions. Moreover, enjoyable user experience could result in referrals and repeat visitors. In addition, rather than focusing solely on ranking, webmasters should pay additional attention to users. Google's algorithms keep changing. However, Google's goal will always be the same, and therefore webmasters should perpetually modify and improve the website based on Google's objective, in order to improve customer service and to receive a high page ranking.

REFERENCES

- Alex & Cutts, Matt 2013. Unnatural Links to site. Google Webmasters. Referenced 15 January, 2014.
<https://www.youtube.com/watch?v=6UFHCw4Q_FQ>
- Backlink Watch 2014. Referenced 11 December, 2014.
<<http://www.backlinkwatch.com>>
- Brin, Sergey & Page, Lawrence 2014. The Anatomy of a Large-Scale Hypertextual Web Search Engine. Referenced 10 November, 2014.
<<http://infolab.stanford.edu/~backrub/google.html>>
- Chopra, Narenden 2012. 5 Reasons why Mac Users should buy upcoming Leap Motion. Referenced 11 December, 2014.
<<http://technoupdates.org/reasons-mac-user-buy-leap-motion> with keywords>
- Chopra, Narenden 2014. Ways that Technology and Marketing Are Symbiotic. Referenced 11 December, 2014.
<<http://technoupdates.org/ways-technology-marketing-symbiotic>>
- Croft, W. Bruce & Metzler, Donald & Strohman, Trevor 2009. Search Engines: Information Retrieval in Practice. Referenced 20 January, 2014.
<http://www.amazon.com/Search-Engines-Information-Retrieval-Practice/dp/0136072240/ref=sr_1_1?s=books&ie=UTF8&qid=1417250879&sr=1-1&keywords=search+engines+information+retrieval+in+practice>
- Cutts, Matt 2009. Matt Cutts: Gadgets, Google, and SEO. PageRank sculpting. Referenced 20 July, 2014.
<<https://www.mattcutts.com/blog/pagerank-sculpting/>>
- Cutts, Matt 2010. How Search Works. Google Webmasters. Referenced 20 July, 2014.
<<https://www.youtube.com/watch?v=BNHR6IQJGZs>>
- Cutts, Matt 2011a. Cloaking. Google Webmasters. Referenced 20 July, 2014.
<<https://www.youtube.com/watch?v=QHtnfOgp65Q>>
- Cutts, Matt 2011b. What is the ideal keyword density of a page. Google Webmasters. Referenced 15 January, 2014.

<<https://www.youtube.com/watch?v=Rk4qgQdp2UA>>

Cutts, Matt 2012. What is Google's view on guest blogging for links? Referenced 20 July, 2014.

<<https://www.youtube.com/watch?v=IMxC3wQZOyc>>

Cutts, Matt 2013a. Should I use rel="nofollow" on internal links to a login page? Google Webmasters. Referenced 15 January, 2014.

<https://www.youtube.com/watch?v=86GHCVRReJs&list=UUWf2ZINsCGDS89VBF_awnvA>

Cutts, Matt 2013b. If I write about another article, where should I link to the original source? Google Webmasters. Referenced 15 January, 2014.

<https://www.youtube.com/watch?v=ZLLylwUIXcc&list=UUWf2ZINsCGDS89VBF_awnvA>

Cutts, Matt 2013c. Why isn't my site's PageRank changing. Google Webmasters. Referenced 15 January, 2014.

<https://www.youtube.com/watch?v=M7glS_ehpGY&list=UUWf2ZINsCGDS89VBF_awnvA>

Cutts, Matt 2013d. How does Google handle duplicate content? Google Webmasters. Referenced 15 January, 2014.

<<https://www.youtube.com/watch?v=mQZY7EmjbMA>>

Cutts, Matt 2014a. Is there a version of Google that excludes backlink as a ranking factor? Google Webmasters. Referenced 19 October, 2014.

<https://www.youtube.com/watch?v=NCY30WhI2og&list=UUWf2ZINsCGDS89VBF_awnvA>

Cutts, Matt 2014b. What is a "paid link"? Google Webmasters. Referenced 19 October, 2014.

<<https://www.youtube.com/watch?v=zuplbMyMfBI>>

Cutts, Matt 2014c. The importance of content in the body of a page. Google Webmasters. Referenced 19 October, 2014.

<<https://www.youtube.com/watch?v=8DNzUWgHtKc&index=25&list=PL-the5o2x8TO16kLg179RMD12E8i72U58>>

Cutts, Matt 2014d. Will backlinks lose their importance in ranking. Google Webmasters. Referenced 19 October, 2014.

<https://www.youtube.com/watch?v=iC5FDzUh0P4&list=UUWf2ZINsCGDS89VBF_awNvA>

Cutts, Matt 2014e. Matt Cutts: Gadgets, Google, and SEO: About Me. Referenced 19 October, 2014.

<<https://www.mattcutts.com/blog/about-me/>>

Cutts, Matt 2014f. Matt Cutts: Gadgets, Google, and SEO: The decay and fall of guest blogging for SEO. Referenced 19 October, 2014.

<<https://www.mattcutts.com/blog/?s=guest+blogging>>

Douvas, Alexi 2012. Creating great content that performs well in Google search results. Google Webmasters. Referenced 19 October, 2014.

<<https://www.youtube.com/watch?v=y8s6Y4mx9Vw>>

Enge, Eric & Spencer, Stephan & Stricchiola, Jessie & Fishki, Rand 2012. The Art of SEO, Second Edition. Referenced 19 October, 2013.

Google 2010. Welcome to Google's Search Engine Optimization Starter Guide. Referenced 19 October, 2013.

<<http://static.googleusercontent.com/media/www.google.fi/en/fi/webmasters/docs/search-engine-optimization-starter-guide.pdf>>

Google 2011a. Google search and search engine spam. Referenced 19 October, 2014.

<<http://googleblog.blogspot.fi/2011/01/google-search-and-search-engine-spam.html>>

Google 2011b. Finding more high-quality sites in search. Referenced 19 October, 2014.

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

Google 2012a. Another step to reward high-quality sites. Referenced 19 October, 2014.

<<http://googlewebmastercentral.blogspot.fi/2012/04/another-step-to-reward-high-quality.html>>

Google 2012b. Page layout algorithm improvement. Referenced 19 October, 2014.

<<http://googlewebmastercentral.blogspot.fi/2012/01/page-layout-algorithm-improvement.html>>

- Google 2014a. Google facts. Referenced 22 April, 2014.
<<http://www.google.com/about/company/>>
- Google 2014b. Algorithms. Referenced 22 April, 2014.
<<http://www.google.com/insidesearch/howsearchworks/algorithms.html>>
- Google 2014c. Irrelevant Keywords. Referenced 22 April, 2014.
<<https://support.google.com/webmasters/answer/66358?hl=en>>
- Google 2014d. Hidden text and links. Referenced 22 April, 2014.
<<https://support.google.com/webmasters/answer/66353?hl=en>>
- Google 2014e. Doorway Pages. Referenced 22 April, 2014.
<<https://support.google.com/webmasters/answer/2721311?hl=en>>
- Google 2014f. Affiliate program. Referenced 20 April, 2014.
<<https://support.google.com/webmasters/answer/76465?hl=en>>
- Google 2014g. Scraped Content. Referenced 20 April, 2014.
<<https://support.google.com/webmasters/answer/2721312?hl=en>>
- Google 2014h. Webmaster Guideline. Referenced 20 April, 2014.
<<https://support.google.com/webmasters/answer/35769?hl=en>>
- Google 2014i. Preventing comment spam. Referenced 20 April, 2014.
<<http://googleblog.blogspot.fi/2005/01/preventing-comment-spam.html>>
- Google 2014j. Link Schemes. Referenced 20 April, 2014.
<<https://support.google.com/webmasters/answer/66356?hl=en>>
- Google 2014k. Our products and services. What we do for you. Referenced 20 April, 2014.
<<http://www.google.com/about/company/products/>>
- Google 2014l. What are backlinks and how do I use them? Referenced 20 April, 2014.
<<https://support.google.com/blogger/answer/42533?hl=en>>
- Google AdSense 2012. AdSense Tips: Consider Your Content When Adding New Units. Referenced 19 October, 2014.
<<https://www.youtube.com/watch?v=G8lpZFJTpWk>>
- Grappone, Jennifer & Couzin, Gradiva 2011. Search Engine Optimization: An Hour a Day, 3rd edition, Forword by Danny Sullivan, Editor-in-Chief of SearchEngineLand.com. Referenced 12 January, 2014.

- Hill, Tim 2014. What does Matt Cutts really say about keyword density. Referenced 21 November, 2014.
<<https://www.youtube.com/watch?v=T4yJ3JgloAw>>
- Huang, Betty 2012. How malicious parties can spam your site. Google Webmasters. Huang Betty is from Google's Search Quality Team. Referenced 21 November, 2013.
<<https://www.youtube.com/watch?v=kewJSFjKfPw>>
- Karch, Marziah 2014. What Is PageRank and How Do I Use It. Referenced 20 July, 2014.
<<http://google.about.com/od/searchengineoptimization/a/pagerankexplain.htm>>
- Kim, Larry 2014. Organic Keyword: What is an Organic Keyword? Referenced 16 Jan, 2014.
<<http://www.wordstream.com/organic-keyword>>
- Krishna, Radha 2013. DoFollow Blog Commenting Sites List. Referenced 11 December, 2014.
<<http://www.9to5blogger.com/dofollow-blog-commenting-sites-list/400/>>
- Lee, Jennifer 2012. Using Sitemaps to help Google find content hosted on your site. Google Webmasters. Referenced 16 Jan, 2014.
<https://www.youtube.com/watch?v=TxahHF_MD2Q>
- Ledford, Jerri L. 2009. Search Engine Optimization Bible, 2nd Edition. Referenced 10 April, 2013.
<<http://www.centarzaedukaciju.com/katalog/IT%20tutorials/SEO%20Bible%202011.pdf>>
- Lievonen, Marjut 2013. Understanding Google algorithms and SEO is essential for online marketer. Referenced 10 November, 2014.
<http://publications.theseus.fi/bitstream/handle/10024/67859/Lievonen_Marjut.pdf?sequence=1>
- Mcdaniel Jr., Carl & Gates, Roger 2010. Marketing Research Essentials With Spss. 7th Edition. Hoboken, NJ: John Wiley & Sons INC. Referenced 16 April, 2013.
- Moz 2014. Google Algorithm Change History. Panda 4.1 (#27) –September 23, 2014. Referenced 11 December, 2014.

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

Page Rank Checker 2014. Referenced 11 December, 2014.

<http://www.prchecker.info/check_page_rank.php>

Sandy & Cutts, Matt 2013. Unnatural links from your site. Referenced 16 Jan, 2014.

<https://www.youtube.com/watch?v=3mdMXdQ_VyU>

Siteliner 2014. Referenced 11 December, 2014.

<http://www.siteliner.com/technoupdates.org?siteliner=site-dashboard&siteliner-sort=scan_time&siteliner-from=1&siteliner-message=frequency>

Sullivan, Danny 2007. What Is Google PageRank? A Guide For Searchers & Webmasters. Referenced 15 Jan, 2014.

<<http://searchengineland.com/what-is-google-pagerank-a-guide-for-searchers-webmasters-11068>>

Sullivan, Danny 2011. Google Forecloses On Content Farms With “Panda” Algorithm Update. Referenced 16 Jan, 2014.

<<http://searchengineland.com/google-forecloses-on-content-farms-with-farmer-algorithm-update-66071>>

Sullivan, Danny 2012a. How Google Measures ‘High Quality’ Content. Entrepreneur. Referenced 16 Jan, 2014.

<<https://www.youtube.com/watch?v=GMOuXbtcdmU>>

Sullivan, Danny 2012b. The Top Heavy Update: Pages With Too Many Ads Above The Fold Now Penalized By Google’s “Page Layout” Algorithm. Referenced 16 Jan, 2014.

<<http://searchengineland.com/too-many-ads-above-the-fold-now-penalized-by-googles-page-layout-algo-108613>>

Sullivan, Danny 2012c. Google Launches “Penguin Update” Targeting Webspam In Search Results. Referenced 16 Jan, 2014.

<<http://searchengineland.com/google-launches-update-targeting-webspam-in-search-results-119295>>

Sullivan, Danny 2013. Google Still World’s Most Popular Search Engine By Far, But Share Of Unique Searchers Dips Slightly. Referenced 16 Jan, 2014.

<<http://searchengineland.com/GOOGLE-WORLDS-MOST-POPULAR-SEARCH-ENGINE-148089>>

Wyszomierki, Michael 2012. Webspam Content Violations. Google Webmasters. Referenced 16 Jan, 2014.

<<https://www.youtube.com/watch?v=yFxNda5Z4eE>>