UNDERSTANDING GOOGLE ALGORITHMS AND SEO IS ESSENTIAL FOR ONLINE MARKETER

Guide for interns in Golden Moments

Marjut Lievonen

Bachelor's/Master's thesis

November 2013

International Business

Marketing

ABSTRACT

Tampereen ammattikorkeakoulu

Tampere University of Applied Sciences

International Business

Marketing

MARJUT LIEVONEN:

Understanding Google algorithms and SEO is essential for online marketer

Guide for Interns in Golden Moments

Bachelor's thesis 60 pages, appendices 1 page

October 2013

_____

The marketing has changed remarkably after the introduction of internet and Google. More and more businesses rely on online marketing which gives rise to the need of new online marketing professionals. In order to become a professional, the marketers need to understand the environment they are working in; in online marketing case they need to understand the algorithms and SEO as well as their target audience in order to be successful.

This Bachelor thesis aims to give basic understanding of Google algorithms and how they affect SEO. The thesis provides a start-up material for marketers aiming for online business professionals and gives them the basic knowledge through practical information, examples and explanations.

The thesis focuses on Google's algorithm based ranking systems in search and introduces the major changes in the algorithms as well as some recent changes. The thesis focuses also on SEO on practical level, introducing the SEO rules.

The mention of the role of internet users in relation to algorithms and SEO was a necessary part of the thesis as the algorithms and SEO are done for the internet users and leaving out the users would have been neglecting an important aspect in online marketing.

_____

Key words: google, algorithms, seo,

CONTENTS

ABBREVIATIONS AND TERMS

Words used in the thesis

SEO	Search Engine Optimization
CTR	Click Through Rate
PPC	Pay Per Click
search engine	Programs that search documents based on keywords and return the results for the search of the wanted keywords. Google, Bing and Yahoo! are the most known search engines.
Google Panda	Google algorithm that focuses on demolishing low-quality content pages from Google search results.
Google Penguin	Google algorithm that focuses on fighting against spam sites and erasing them from Google search results.
page rank	Rank of the page on Google search, for example the placement on first page, second page etc.
PageRank	Algorithm named after Larry Page.
crawling	The process of robots circling the web and organizing the information on webpages.
spiders	Search engines' crawling robots which crawl the web and index pages.
Content farm	A website that looks to see what popular searches are in a particular category such as news or popular help topics and then produces content that matches specifically those searches. Usually these kinds of sites spend as little time and money to creating content as possible.
white hat marketer	A webmaster or marketer who is trying to get his or her website to become successful in

	Google search by following SEO rules and Google guidelines and by having good content on the website.
black hat marketer	A webmaster or marketer who tries to get his or her page to appear on good rankings in Google through tricking Google algorithms and not focusing on "clean" SEO.
Matt Cutts	A Distinguished Engineer in Google webspam team.
Barry Schwarz	Barry Schwartz is News Editor of Search Engine and owner of RustyBrick, a New York based web consulting firm. He also runs Search Engine Roundtable, a popular search blog on very advanced SEM topics.
Danny Sullivan	Danny Sullivan is a widely known SEO specialist and "search engine guru". He is a Founding Editor of Search Engine Land.
SERP	Search Engine Result Pages, The pages that come as results from search
UGC	User generated content on a webpage. Content that is not written by the website owner but a user in the page.

1 INTRODUCTION

The world of marketing has evolved greatly during the last two decades, after the introduction of internet and especially after the introduction of Google. The online marketing experienced a boost in the early 2000's and from there on the online marketing has evolved rapidly and continues to do so. For the companies to become and stay successful in the world of online marketing they must be constantly updated and aware of the latest changes and what is to come in near future.

The purpose of my thesis is to supply the company, Golden Moments, with the information regarding the basic knowledge of algorithms and how they affect SEO in order to help in the training of new interns. The thesis's goal is to provide a good start-up material for new interns in Golden Moments. I will cover what are algorithms and how these algorithms affect SEO. Are algorithms something that a marketer should understand and why?

In this thesis I focus mostly on Google's algorithm based ranking systems and their effect on SEO. First I will introduce algorithms, what they are and what is their meaning in online marketing. I will introduce the biggest algorithmic changes in Google history as well as give some information on how they have affected the online marketing.

After covering the algorithms I will discuss SEO and how it is affected by the algorithms. I will cover the SEO rules and the reasons behind them. I will also provide some practical examples of how to do SEO well.

In this thesis I will also discuss the role of customers in relation to SEO and algorithms shortly. How does it benefit the SEO to understand the users? What is the relation between internet users and algorithms? I will answer these questions in my thesis.

2   GOLDEN MOMENTS

This thesis is written for the company Golden Moments for request from the company co-owner, Paul Cheatle. The thesis is to supply Golden Moments with start-up material for new interns who do not have prior knowledge of algorithms and SEO. The thesis is written with the aim of supporting the training of new interns in the company.

Golden Moments is a company based in UK, providing gift experiences, such as Ferrari driving experience, parachuting and other. The company was founded by identical twins, Paul and Steve Cheatle in 1998 and has grown to be a pan-European company with activities in more than 10 countries in Europe. The company offers more than 7000 experiences and is striving to offer the best quality for the customers.

Currently, Golden Moments offers internship opportunities for students in international markets. The company's foreign markets are handled by the interns so the internship offers responsibility and independent working environment. Each intern has their own market to take care of, for example Spanish intern takes care of the Spanish market and the Finnish intern focuses on Finnish market. The company's office, which is located in UK, is international as the office is filled with international interns and foreign languages.

The company helps the interns to gain work experience as well as to learn about SEO. Golden Moments utilizes SEO for its marketing and therefore also trains interns to use SEO marketing effectively. Golden Moments provides SEO training individually in meetings where the interns get to know the tricks little by little. As mentioned before, this thesis is to aid and support the SEO training in these meetings and to give the boost for the interns to start investigating the algorithms also on their own.

3   GOOGLE

3.1   Introduction to Google

Google's story begins with two men meeting in Stanford University in 1995. According to Google Company main page (Google Company "Google's mission is to organize the world's information and make it universally accessible and useful") the pair had built a new search engine which was called BackRub by 1996 which used links "to determine the importance of individual web pages". From there the two men started developing their ideas that would eventually grow to be Google. They were Larry Page and Sergey Brin.

The company Google was founded on September 4th in 1998 California, United States of America. The company tells on their company webpage (Google "Google's mission is...") that the name "Google" is a word play from the original word "googol" which is a mathematical term. Larry Page and Sergey Brim chose the word play of the term as their company name since the term hints to infinite and it suited the pair's mission to organize infinite amount of information on the web.

The company has expanded their mission to organize information on the web to include the entire world. Google tells in its Google Company Philosophy article "Ten things we know to be true" that they now have offices in more than 60 countries and that they maintain more than 180 internet domains. In the same article Google tells that more than half of Google's results are served outside of United States. This means, also, that Google provides results to users in different languages. Google search interface is offered in more than 130 languages and the search engine has given the users the possibility to restrict the results shown to their own language only. Google also offers translation tools and many other applications and products in different languages so the user experience would be better.

Google has announced that its mission is "to organize the world's information and make it universally accessible and useful" (Google Company "Google's mission is..."). From the very beginning they wanted to give the search engine user the best experience possible. Google does not limit this to only search but to everything they do; when they

provide new tools they want them to be easy to use and to work well. Google has emphasized that everything they design is meant to serve the user (Google "Ten things...").

3.2 "Do no evil"

We should never forget that Google is a business. The revenues for the business come from the sale of advertising displayed on Google page and other websites. The revenues also come from offering search technology to companies. Google's mission is to give the user the best possible search experience and the company believes this can be done without malicious means. Google believes that the best search experience for the user comes by providing only good content to the user. Therefore Google is very strict on their policy of "not doing evil" with which is meant bad SEO and providing bad quality content on the web pages. Google has provided guidelines for SEO marketers to follow in order to provide pages with good quality to users and Google follows these principles closely. I will discuss the policies below. (Google "Ten things...")

One of these policies revolves around relevancy. They have therefore left their homepage very simple and clear and do constant work on the search so that the pages on the search are not sold placements and thus as relevant as possible. Google has developed algorithms to find the most relevant pages for each search query from the web. I will discuss more about algorithms later on in this thesis.

The same policy on relevancy is applied to advertisements, according to Google. Google does not want the advertisements to be distracting, quite the opposite; Google wants the advertisements it shows to be relevant as well. Google does not allow those ads that are not relevant for the search results to show. Sometimes there might not be any ads at all for certain search queries.

Google has announced on the same article as above, "Ten things we know to be true", that Google doesn't like ads that are distracting to the user. Therefore flashy ads, such as pop-up advertising is not accepted by Google. They have discovered that simple text advertisements that are relevant for the user draw higher CTR than random ads. The advertisements are thus readily targeted for those interested in the topic.

It might pop in people's minds to just buy a spot from Google to be visible on the first page. But it doesn't work that way. Google proudly states that it never manipulates rankings to place some advertisements or search results higher or more visible. Google strives to be democratic and objective so it is not possible to buy better placement. The better placement can only be acquired with good quality web page and good SEO.

The fact that Google doesn't manipulate its ranking doesn't mean it wouldn't help the businesses to get visibility in search results. Google does help businesses with advertising but it does not promise any placements or success. In 2000, Google launched AdWords, which is a service for creating online ad campaigns. The program has been developing during years and now includes also mobile and video ads in addition to the more traditional ads. With this program businesses can plan advertisement campaigns, check keywords and other.

3.3 Google search

Google believes it is best to do one thing extremely well. This one thing they want to focus on is search. They have invested in large research groups that try to solve the problems on search and make the search even better. Google has dedicated itself to give a "seamless experience for millions of people" (Google "Ten things..."). Not only does Google want to offer all the information possible for the user, they also want to do it fast. They are constantly working on making the search and programs even faster in order to please the user.

As Google's mission is to make the search experience as nice as possible for the user, they do not believe in returning just anything from the search query. The results from the search must be relevant and to achieve that Google relies on more than 200 signals and techniques to determine which sites are the best for that specific search query. These signals and techniques include Google's algorithms, which is the focus of this thesis.

The world is changing, so is Google. Google is a company that is insatiable. They want to make the search better and better. As Google says in its company philosophy "Ten

things we know to be true": "We see being great at something as a starting point, not an endpoint. We set ourselves goals we know we can't reach yet, because we know that by stretching to meet them we can get further than we expected."

4 ALGORITHMS

What are algorithms and what do they have to do with online marketing? Google gives a definitive answer as to what algorithms are: "You want the answer, not trillions of webpages. Algorithms are computer programs that look for clues to give you back exactly what you want." (Google InsideSearch "Algorithms") As the web has widened to include trillions of webpages, there are most likely thousands if not even more pages which have helpful information. But, if there are so many pages, how will we ever find the information we want from those? Nobody will spend time browsing through thousands of pages. For this reason, Larry Page and Sergey Brin started developing computer processes and formulas that make the search easier, algorithms.

Google's algorithms have developed greatly and now they rely on more than 200 unique signals that enable the users to find what they are looking for in the web. For a marketer, it is these signals that are important to know and to understand, or at least part of them as Google does not reveal all the signals to public knowledge. It is these signals, upon which the algorithms define whether the page is shown in first results or not. The signals are for example keywords on websites, the region, and freshness of site content as well as the ranking of the page in internet search.

The whole process behind search starts with crawling and indexing pages. For this, Google and other search engines have developed their own robots to crawl the web. These are sometimes also called spiders. Google's crawling robot is called "googlebot". These crawling robots go from one page to another and analyze its contents and determine whether it is relevant site or not.
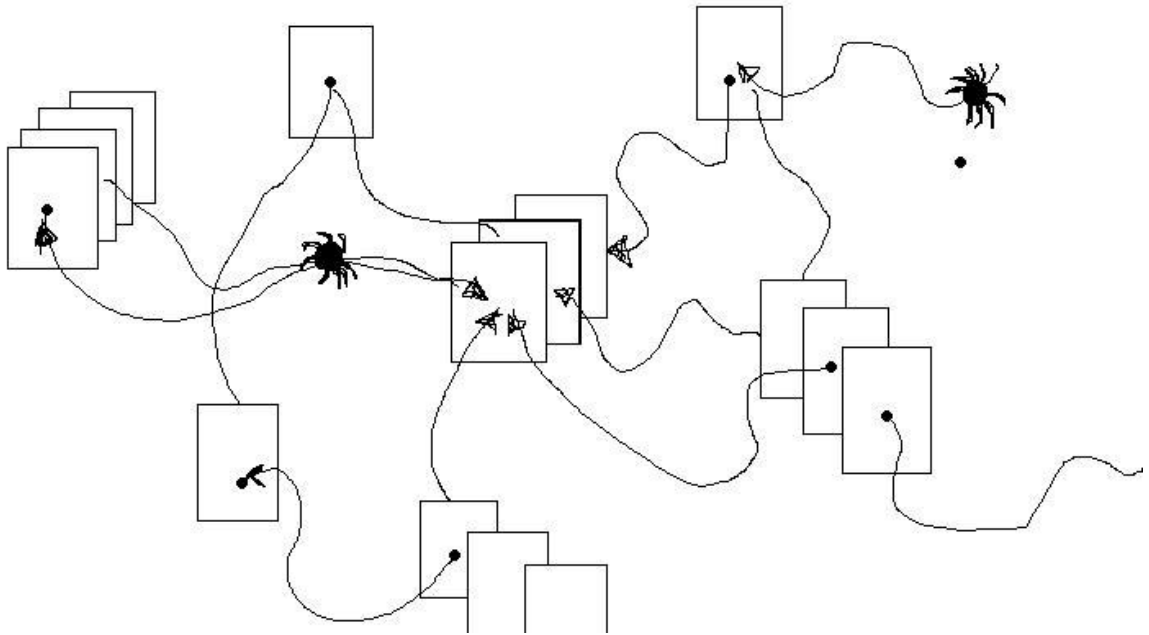
PICTURE 1. Spiders crawling and indexing the web.

The spiders determine the site content based on the before mentioned signals. All pages that spiders crawl will get indexed but the ranking of the page depends on how the algorithms have interpreted the content and how they see the page as a whole. If the page appears to be non-relevant and more like spam site the ranking will go down. On the other hand if the page has original and good quality content with good linking to it the rankings might be higher.

4.1   Short history on algorithms

Google's history started from developing the first search engine algorithm, PageRank in 1997. From there on, Google has developed its algorithms in their mission to enhance the search experiences of the users. Google's goal is to provide the answers the searchers are looking for in the first search results. For this goal, the company continues to make their algorithms better and better all the time. One of the well known SEO experts, Danny Sullivan has said "Google changes its algorithm on a regular basis, but most changes are so subtle that few notice" in his article "Google Forecloses on Content Farms with 'Panda' algorithm update" (2011) in Search Engine Land.

Google has stated in many occasions that their goal is to help people find the best information on the web. For this reason Google keeps developing their algorithms so

they could help websites to concentrate on creating good quality content rather than try and get visits to their webpage through "tricks and gimmicks" as Google calls them. Despite being very open about new updates to algorithms Google does not reveal all details due to "black hat" marketers who try to fool the search results with malicious ways and thus get visits on their sites.

This constant cat-mouse game causes pressure on Google and the constant development. Internet and the way search is conducted will also change with time and so the search engines must keep up. This results to algorithms changing. But, even with the threat of spammers catching up, Google continues being open about the changes since they want the relevant good quality sites to show in search results.

It is thus important that the marketing personnel are constantly up to date with different algorithm changes. Google does inform about great changes and as well as some smaller changes and in order to be able to keep the web pages in good rankings in Google search results, the marketers must be active and follow the changes. To help the marketer to keep up with all the chances, even without having the deep technical understanding, Google provides videos in Youtube with the user name GoogleWebmasterHelp and releases blog posts in Google Webmaster Blog. There are also other useful sites and blogs that discuss the changes. One of these sites is searchengineland.com. This thesis also refers to the webpage from time to time.

4.2   Google algorithms major changes

Google has developed its web searching abilities and one of the most known changes to this web search is PageRank. According to Google, PageRank counts the number and quality of links to a page to create an estimate of the importance of the website. The importance is determined with the assumption that if the website is relevant it has more links from other websites. I will discuss about PageRank more in the following part.

Other great changes in Google history have been Google Panda, an algorithm which focused on getting rid of low-quality websites on the search results and Google Penguin, an algorithm that focuses on reducing spammy websites. I will discuss about Panda and Penguin updates more in detail in the following parts.

4.2.1 PageRank

When the discussion turns to algorithms the first thing that comes to mind, is PageRank. What is it exactly? What does it do, or why should the marketing personnel know about PageRank? What good does understanding PageRank do for marketing person? These are important questions and important to be answered.

PageRank is an algorithm for Google search engine, where websites are ranked on its search results. The higher the PageRank the more important it is considered in search engines. PageRank was developed by Google founders Larry Page and Sergey Brin. According to Google, Page once illustrated the so called perfect search engine as something that "understands exactly what you mean and gives you back exactly what you want" (Google Company "Our products and services"). Google has developed its search engines to correspond to these words and Google's goal is to make information finding as easy as possible for the user. This said, Google has taken on the challenge to make search engines smarter and faster.

Marziah Karch explains PageRank in her online article "What Is PageRank and How Do I Use It?" in About.com Guide that PageRank is patented by Standford and the name for the algorithm comes from Larry Page's surname. PageRank differs from previous search engines in that previously keyword density was the most important. According to Karch the previous search engines were possible to cheat by stuffing keywords to the page. PageRank then again, sees the pages that have many links to it as important. PageRank considers links as "votes"; when a page is linked to another page a vote is given to that linked page.

Karch points out that PageRank is special also in a way that it ranks the pages that have linked the other page. Pages with low PageRank don't give as much value for the vote as the higher ranking pages do. Quantity is not the issue in linking. It is the quality of the pages that matter.

As nice as PageRank sounds, it is not flawless, Karch reminds in her article. So called "Black Hat" webmasters or marketers have found some ways to raise their page ranking

by fooling the PageRank. There are different ways to try to fool PageRank, one of them being "link farming", which means that a certain page has gathered links as its content to attain better ranking. Google has taken measures towards this kind of act which has resulted to pages with too many low quality links have lost their page rank value. If you find your own site being linked from these kinds of sites, it does not necessarily affect your site, but you shouldn't link back to these sites.
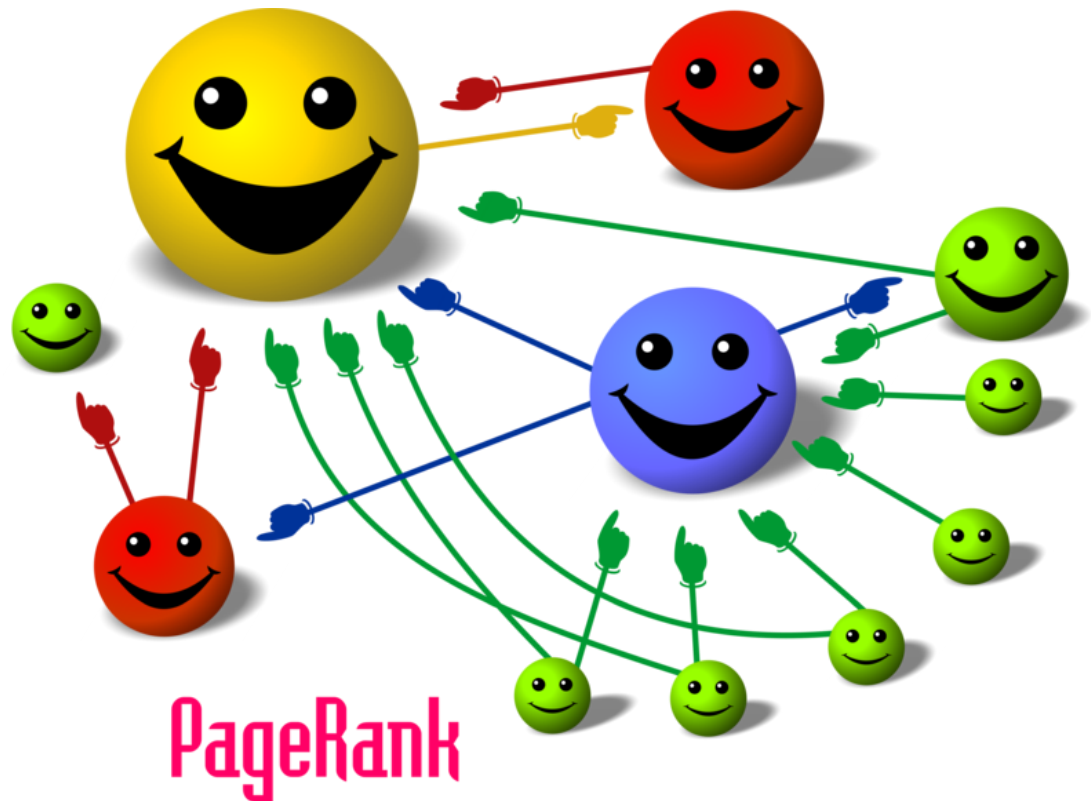
PICTURE 2. Artistic view of PageRank by Felipe Micaroni Lalli. Source: http://en.wikipedia.org/wiki/File:PageRank-hi-res.png

Marziah Karch explains in her article that PageRank is measured by scale which goes from one to ten. The ranking is assigned to individual pages within the website, not the entity of the site. If you want to raise your page ranking you need to get relevant links to your website. One opportunity is to trade links, for example with business partners.

Answer to question of why do marketing personnel need to know about PageRank and understand how it works, is that when marketing personnel know how this algorithm works, they can utilize it. One great way of utilizing this algorithm is link sharing. By sharing links with right partners the page ranking might become better and better.

Equally important to understanding the PageRank algorithm is to understand with whom to share links. If marketing personnel does not understand the PageRank link sharing and the only places which link to your company page are link farming sites, the results will not become better. Quite the opposite, you might even wound up damaging your own web page search results. Therefore, before any action is done it is good to know the algorithms and the opportunities as well as threats that they present.

4.2.2   Google Panda

Now that we have better understanding of PageRank, we can move on to the next big algorithm change in Google history. This other algorithm is called Google Panda, which received a lot of attention when it was launched in 2011 since it affected so many sites. Google Panda was launched in February 2011 and according to Danny Sullivan in his article in Search Engine Land "Google Forecloses On Content Farms With "Panda" algorithm update" (2011) its main target was to detect "content farms" and to block them from showing in Google search results. Sullivan mentions that officially Google does not say that they are targeting content farms with the update but he also tells that Google's Matt Cutts had hinted toward these sites. With "content farm" is meant sites with "shallow or low-quality content" as Google describes them in its Official Blog post "Google search and search engine spam" (2011). Sullivan tells in the same article as above that Google Panda was also meant to act against scraper sites, which as the name implies, scrape content from other sites and tries to get good ranking. This results to the websites having the original content possibly not showing in Google first page and naturally, the webmasters and marketers of the original content are enraged.

Google's Principal Engineer Matt Cutts and Google Fellow Amit Singhal stated in Google blog "Finding more high-quality sites in search" (2011) about the new Panda update: "We can't make a major improvement without affecting rankings for many sites. It has to be that some sites will go up and some will go down. 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. Therefore, it is important for high-quality sites to be rewarded, and that's exactly what this change does."

The new Google algorithm, Panda had an impact on 11.8% of English language webpages according to Cutts and Singhal when it was first launched. By April 2011 the algorithm had been rolled out to all English-language websites globally and by August the same year the change had been expanded to include all the languages with the exceptions of Japanese, Korean and Chinese.

The new Panda algorithm allowed sites with good quality content and original content to rank higher in search engine while the low quality pages suffered losses in rankings. The change was so major it caused a lot of discussion among webmasters. It was evident at that time already, that sites with good quality content survived with better results than those of low quality content. This was exactly what Google had been aiming at. The purpose of the Panda algorithm was to get rid of low-quality sites that provide thin content and that are not as informative and useful for the user.

In total there have been six updates to Panda after its release. All of the "major" updates were conducted throughout the year 2011, making the year very memorable for those struggling with SEO. There have been more updates to Panda even after the first stormy year, and the updates keep coming. Nowadays 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.

What was to be learned from Google Panda? The year 2011 taught the webmasters and marketers a great lesson when it comes to SEO: to produce and maintain good quality content on your website. As Amie Baumwell says in her article "Panda 2.5 does some damage" in Ever Spark Interactive (2011) Panda update's goal was to get rid of content farms. This means that Google wasn't targeting genuine information sharing or business sites. Therefore, she encourages to make sure that the content on the website is "valuable, contributing something, and isn't overly stuffed with keywords". Amie Baumwell also suggests focusing on branding. She reminds that branding across social media platforms and blogs remains to be important.

There are also other SEO specialists who give out tips for those who had suffered from Panda updates. One of them is Vanessa Fox, a Contributing Editor at Search Engine Land. Her previous work has been Google's Webmaster Central. In her article "Google Panda Update 2.4: Panda Goes International, In Most Languages" (2011) she reminds

the webmasters and marketing personnel to look at their pages objectively and think about certain issues I will cover below.

Vanessa Fox asks in her article, whether you need all of the multiple sites on your webpage. Is it necessary to have two almost identical pages for the same topic that answer basically the same thing? As she describes it, "it's one thing to have separate pages on 'best chocolate cake recipe' and 'best pumpkin pie recipe' and quite another to have separate pages on 'best chocolate cake recipe' and 'ideal chocolate cake recipe'". She also points out the importance of originality in the text context of the pages. If the content on your webpage is copied from another source, Google algorithms may rank it lower.

PICTURE 3. Summary of good quality content.

In conclusion, to avoid Google Panda to rank the webpage lower in search results, it is best to make sure the webpage has good quality content. Good quality is defined as having unique and original content while all the information being credible. The information on the page should actually be informational and not just some vain blabbering. So called "thin content" sites that do not really have anything valuable to tell or are only scraping the surface of the subject, are not considered as good content sites by Google. The content on the page therefore should not be shallow but provide information that the readers want to know. In addition to the text content being good, also the page should be user friendly when browsing; meaning the navigation ought to

be easy and there shouldn't be multiple pages which have the same information repeated.

Once these issues have been checked to be all right and the site pleasant to visit, Google Panda shouldn't have anything to complain either. And what is most important, the internet users would find the page nice to visit and might even recommend the page to others by linking to it, which further increases the ranking of the page. Fox reminds in her article that it is not enough to concentrate on making only one or two things great on the webpage and leave the rest only satisfactory. The webpage ought to be good as a whole, not only parts of it if you want to succeed in rankings.

4.2.3　Google Penguin

Penguin is an algorithm from Google that was targeted against webspam and it was launched in April 2012. According to Google's Matt Cutts in the Official Google Webmaster Central Blog post "Another step to reward high-quality sites" (2012), the Penguin will impact about 3% of search queries. This percentage was for English, German, Chinese and Arabic languages whereas for languages which have been detected to have more spam, such as Polish, the impact is about 5%.

Google announced this algorithm's goal was to decrease rankings for sites that were violating Google's quality guidelines. Mainly these violations, as Danny Sullivan has listed in his article "Google Launches 'Penguin Update' Targeting Webspam In Search Results" (2012), are keyword stuffing, link schemes, cloaking and purposeful duplicate content.

What is this keyword stuffing? With keyword stuffing is meant an action where the keywords are being used excessively in order to manipulate the site's ranking in search results. Google thinks that filling the pages with keywords too much is negative experience for the site user (Google Webmaster Tools "Keyword stuffing"). The negative experience for the user comes from the fact that by stuffing keywords to the text as much as possible the text becomes incomprehensive and often repetitive. In other words the text does not offer anything relevant for the reader. The incomprehensiveness and lack of unique and relevant content is considered as a signal to lower the page's ranking. Therefore Google recommends using keywords in context appropriately to

ensure pleasurable reading and browsing experience for the user. Google has set an example of the kind of text one should not produce that would be marked as keyword stuffing:

"We sell custom cigar humidors. Our custom cigar humidors are handmade. If you're thinking of buying a custom cigar humidor, please contact our custom cigar humidor specialists at custom.cigar.humidors@example.com."

Link schemes according to Google Webmaster Tools "Link schemes" are also aiming to change the search result ranking and regarded as violation of Google's webmaster guidelines. Link schemes can vary in its practices, for example buying or selling links is considered link scheming and marketers and webmasters must be increasingly careful with the sites that offer hosting links against a payment. Having been tangled in a link scheming network might affect your webpage badly and the rankings of your site may go lower. Google reminds that creating good content is the best way to get relevant links to your page. Good and unique content will gain popularity in internet and the links will naturally follow.

Therefore, despite it being sometimes very useful and considered good relationship management, sharing links with partners may not always be a good thing to do. When exchanging links it is good to keep in mind that excessive link exchanges are not seen with good eye by Google. Google also includes "large-scale article marketing or guest posting campaigns with keyword-rich anchor text links" as link scheming (Google Webmaster Tools "Link scheming"). Once again Google provides an example of this:

"There are many [wedding rings](#) on the market. If you want to have a [wedding](#), you will have to pick the [best ring](#). You will also need to [buy flowers](#) and a [wedding dress](#)."

Any text advertisements and advertorials that pass PageRank, whether paid or not, are considered to violate Google's guidelines as mentioned in the above Webmaster Tools post. On the other hand, Google reminds that PPC advertising links that do not pass PageRank are completely fine. You can achieve this by adding a rel="nofollow" attribute to the <a> tag or redirecting the links to an intermediate page that has robots.txt file added and thus blocks the search engine robots from crawling it.

Another thing that is considered spamming is cloaking. Google itself admits cloaking is difficult issue but definitely punishable action. With cloaking is meant the action of presenting different content to human users than is presented to search engines. An example that Matt Cutts has presented in Google Webmaster Tools "Cloaking" is a website that presents to search engines it contains cartoons but in reality when the human user comes to the page the content is something entirely else, for example porn. This is definitely a negative experience for user as the search result doesn't match with the searched topic. Google strongly reminds that there is no such thing as acceptable cloaking. The crawlers should be treated the same as human users at all times to give a truthful image of the webpage.

The last one on the spam list that Sullivan has listed in his article " Google Launches 'Penguin Update' Targeting Webspam In Search Results" (2012) is duplicate content. Google has taken actions against duplicate content already on its earlier algorithm update, Panda, but has launched Penguin to tackle the pages with duplicate content as well. The duplicate content is as its name suggests content on the page that matches the content on other pages either completely or is very similar. Google reminds in its Webmaster Tools post "Duplicate content" that mostly the duplicate content is not originally deceptive. Duplicate content may appear in discussion forums which have pages for regular traffic as well as for mobile devices. Also printer-only versions of webpages are acceptable duplicate content.

There is, unfortunately also malicious duplicate content on web pages which tries to manipulate the search engine rankings towards better. This may result to the user having the same content displayed repeatedly making the user experience becoming negative. Google sees this as a violation against their mission and reacts to prevent this from happening.

Google gives out some tips on how to address the duplicate content properly so that also algorithms understand what is going on and do not mark the pages as spam. Firstly, Google "does not recommend blocking crawler access to duplicate content on your webpage". Google explains that if the search engines can't crawl the duplicate content pages they also can't identify the URLs are pointing to the same content and are treated as separate pages. Therefore, Google suggests using rel="canonical" link element, the

URL parameter handling tool or 301 redirects to mark the pages as duplicate content and allow the crawlers to come and go.

Google Penguin algorithm has been updated several times since its first launch. By March 2013 the Penguin had been updated three times already and the Penguin 2.0 update was in fact, the fourth launch. Matt Cutts refers to the previous updates as "data refresh" in his blog post "Penguin 2.0 rolled out today" (2013), and not actual update unlike this fourth one so it has been officially named as Penguin 2.0.

Matt Cutts announced Penguin 2.0 will impact around 2.3% of English-US queries but the result varies by languages. Again the languages with more webspam will be more affected. The affected sites might in actuality be even more now, since Google launched another update on October 4th 2013, Penguin 2.1. Matt Cutts said in his tweet on 4th October 2013 that the update will affect around 1% of searches to "noticeable degree".

As hard as Google tries to fight webspam, the Penguin update hasn't been perfect. Danny Sullivan argues in his article "Google launches 'Penguin Update' Targeting Webspam In Search Results" (2012) that the techniques Google is using to detect webspam aren't new. Google has been fighting webspam for long time and as Sullivan puts it: "what's really happening is that Google is rolling out better ways that it hopes to detect such abuses."

To conclude the Google Penguin algorithm, I find it important for the online marketers to know about the changes it has brought with it and how the search is being developed constantly. What malicious marketing ways may have worked previously are most likely useless already now or in the future due to algorithm changes. Penguin has impacted many sites fatally since it has targeted spam sites that do cloaking, keyword stuffing and other spam methods deemed Google quality guidelines infringements.

4.2.4   Google Hummingbird

Google has launched yet another algorithm update very recently, while this thesis was still in writing-process. One of the first ones to submit information on this algorithm change was Danny Sullivan from Search Engine Land. Sullivan tells in his article "FAQ: All About The New Google 'Hummingbird' Algorithm" (2013) that Google has developed a new search algorithm called Hummingbird. Google started using

Hummingbird in September but the announcement of the change came a month later, in October 2013. The name originates from being "precice and fast", just like hummingbirds are. Hummingbird is considered one of the biggest changes in Google history.

Hummingbird is reportedly not an actual update as much as it seems to be an entirely "new engine". Sullivan explains that Hummingbird is a brand new engine that uses some parts of the old engine systems such as Panda and Penguin. The appearance of the Hummingbird, despite being "a new engine" does not outlaw "old engines" such as PageRank. According to Sullivan Hummingbird utilizes PageRank as well as other factors to its advantage. This algorithm does not affect SEO. Hummingbird only allows Google to process the signals in a new way, not affect the results like Panda or Penguin did.

The new thing that comes along with Hummingbird is "Conversational Search". With this search the user may search information in a conversational way with the search engine. An example of this is to search for "The lead actress in movie King Kong". The search engine will give the answer both in written form as well as spoken. If you are interested to know more, for example the age of the actress you may go on asking "What is her age?". Unlike previously, conversational search remembers the previous search and understands that with "her" is meant the actress of the previous search. This enables a conversation with the search engine, which in part enhances the user experience.

Sullivan reports that Hummingbird pays more attention to each word in the query, meaning that the query as a whole is taken into consideration in search. The objective is to get the matching pages rank higher, rather than pages where only few words appear.

4.2.5   ESA, rumor or reality?

As Google keeps developing and making changes to its algorithms it is natural that the SEO practitioners are nervous and anxious about what will come. This is also perfect platform for rumors to be born. One of these rumors is the rumor about a completely

new algorithm change that hasn't been launched in big scale before. This new rumored algorithm is called ESA, which is abbreviation from Equivalency Score Algorithm.

According to Humaira Khan in her web article "SEO – Massive Changes Expected in 2013" says that ESA is expected to be one of the major changes in year 2013. She states that the ESA will be working the same as Google Adwords Quality Score. Hitesh Patel also claims in his article "Google Equivalency Score Algorithm" (2012) that ESA would enable the pages with equal page rank to be displayed randomly. This would mean that the results might differ each time the user is searching something with the same keywords. Also Steve Fitzpatrick describes the ESA similarly in his article from November 2012, "Will the Google 'Equivalency Score' Algorithm Change Kill Your Business?"

If this algorithm indeed was to come it is rather odd that well respected SEO analysts such as Danny Sullivan, Eric Enge or Barry Schwartz have not written a single article about it. They haven't noted this "new algorithm" at all. Also from Google there has been no notification of such algorithm. Therefore I personally do not think this kind of algorithm is to be launched in near future.

The articles assumed that the change will come in 2013 but as of now, fall 2013, it has not yet happened and also Google has not given any signs for this kind of change. To further explain my suspicions towards these articles and ESA change, the authors of the articles have not based their assumptions on any real evidence other than Google Adwords policy. Google Adwords team and Google search engine team are not cooperating or changing information with each other as a company policy for fair development, so there is no relevant background for this algorithm for the moment. The authors whose opinion is that the ESA algorithm is coming soon are not basing their assumptions on any relevant or trustworthy information source.

I believe this algorithm is merely a rumor at this point but as the algorithms do change constantly it is best to keep following the news and updates. I included this algorithm rumor in my thesis since I believe it is important for SEO marketer to be up to date even about rumors and learn to differentiate the rumors from reality.

4.3. Why marketers need to know about algorithms?

Now that I have covered the biggest and the most known algorithmic changes, it is time to think why exactly did I cover all these? What was the actual use of reading? I firmly believe that when the marketing personnel knows the algorithms and understands them better the utilization of these algorithms is possible, thus resulting in better positioning in search engines. Understanding algorithms is the key element to effectively work in online marketing platform.

Understanding the algorithms and the changes in them provide a great background for SEO understanding. When the marketer understands the algorithms behind SEO, it is easier to follow the SEO rules if you know why the rules are as they are. In my opinion, understanding algorithms and being up to date about the changes in them provides more insight to SEO and helps being up to date to what is happening in online marketing and business world.

In the following chapter I will discuss more how algorithms affect the SEO.

5   Algorithms impact on SEO

Now that I have covered the algorithms it is time to think how this is useful when it comes to SEO. How does understanding the algorithms contribute to the SEO practicing? How are algorithms and SEO related?

Well, understanding the algorithms does contribute a lot to SEO. For online marketer, knowing the little tricks for the best possible results is often based on understanding the algorithms behind search. For SEO practitioners it is also important to understand crawling and indexing which in turn helps them to decide what to do to reach their goals, according to Eric Enge, and Stephan Spencer, Rand Fishkin and Jessie Stricchiola in "Art of SEO: Mastering Search Engine Optimization (2009, 30). Also Matt Cutts is of opinion that the marketers ought to understand the way the search work in a YouTube video "What are the top 3-5 SEO areas where webmasters make the most mistakes?" (2013). Barry Schwartz wrote a short article based on Matt Cutts's video of the top SEO mistakes web marketers do. In the article "The Top Five SEO Mistakes According To Google's Matt Cutts" Schwartz has written down the mistakes Matt Cutts has listed. One of these common mistakes is "Not using webmaster resources and learning about how Google works and what SEO is about".

I feel that this serves as great reason why online marketers should acquire at least a basic understanding of the algorithms since they are the "behind the scenes" factor of Google search. Also, without algorithms there wouldn't be effective SEO since the spamming sites might get more attention than the real informative sites. Algorithms fight spam continuously so the "white hat" SEO practitioners could concentrate on creating great content for the users.

5.1.   Creating content

It is important to remember that the algorithms concentrate determining that the content on the page is unique, as Eric Enge and others mention in their book "Art of SEO: Mastering search engine optimization" (2009, 33). This is not the only thing the algorithms are programmed to do though, as I have discussed earlier. Algorithms work to bring out the most relevant pages to match the search query. Algorithms do this by

analyzing and checking the content on the page and make judgments upon the results of analyzing.

As Eric Enge et al. emphasize in their book "Art of SEO: Mastering Search Engine Optimization" the "content is king" (2009, 86). Then, what is this content and how to create good content? To know what to create and how is also based on understanding the algorithms and what they seek. Since algorithms rank pages based on their relevance to search queries the content is the most important key issue. Therefore, as Enge and co. call it, the "theme" of the page "plays a huge role in ranking" (2009, 33). This means that good content is consistent. If the theme of the page is in harmony the possibilities for the page to rank higher are better.

The content consists of the following attributes.

5.1.1   Linking

Links are one important aspect of the content of the page. They shouldn't however, be the main issue for the page. Matt Cutts has listed link building in his top SEO mistakes as Barry Schwartz has written them down in the article "The Top Five SEO Mistakes According To Google's Matt Cutts" (2013). In the article Schwartz told Matt Cutts to have given the following advice: "Don't think about link building, think about compelling content and marketing".

Therefore, you should concentrate more on the other aspects of content instead of links. The links will follow if your other content is good, since people will link to it. If your page only contains links and no other content, the algorithms will most likely regard your page as link farm, which I explained in algorithms part of the thesis.

5.1.2   URLs

URLs are a topic that is discussed a lot. Some say they are important part of SEO, some say there is no relation to SEO anymore. Matt Cutts has admitted in several GoogleWebmasterHelp videos in YouTube that URLs might affect the rankings, but he

also reminds that the ranking is based on over 200 signals, not URLs or any other single signal entirely.

In his video posted in YouTube "Does the position of keywords in the URL affect ranking?" (2009) Matt Cutts mentions that when he publishes an article he uses the first three to five words of the title of the article as the URL. I believe this is a good practice if Matt Cutts himself uses it and also, it is logical to include the title of the article of the page to the URL so it is easier to find even if it doesn't contribute to the SEO.

Also, as Matt Cutts has mentioned in his videos, the users might be more willing to click an URL that is clear and has connection to the topic rather than incomprehensive URL. A nice clean URL gives the user more distinctive image of what the page is about than an URL which has no relation to the page.

5.1.3   Language tag

As companies get more international, they might want to offer the pages in several languages. The companies want to offer the services to people in certain region and in certain language. In its Webmaster Tools post "Multi-regional and multilingual sites" Google advices the companies to keep the content for each language on separate URLs, and to think about cross-linking the different language versions of the pages. With cross-linking the user can get from one language page to another with single click. Google reminds about user-friendliness and suggests avoiding automatic redirection based on the user's perceived language. These redirections might prevent both the users and search engines from seeing all the versions of the webpage.

The more aware Google is about the targeted country or language of the site the better it can utilize this information in search results. As stated by Google itself in its Webmaster blog post "Multi-regional and multilingual sites" they generally use the following basics to determine the targeted country of a website:

ccTLDs which comes from words "country-code top-level domain names". This means the ending of the domain that is specified for each country. For example for Sweden the

domain is .se and for Germany it is .de. The domain is a strong signal for search engines as well as human users to which country the site is targeted at.

Geotargeting settings with is the tool in Google Webmaster Tools to tell Google that the site is targeted at specific country. This tool is used only when the domain name is generic and not tied to any country. For example if the domain is .net and you want to target Belgian audience, geotargeting tool is useful. On the other hand, if the page is targeted for more than one country there is no reason to use this tool. Google also uses the server location to determine the target country. Often the server is physically near to the users and thus a signal for the intended audience. Then again, there are sites that use distributed content delivery networks or are hosted in another country, so Google does not treat this as definitive signal.

Naturally, Google does use other signals as well, such as local addresses or phone numbers, the local language and currency, links from local sites as well as the use of Google Places tool. What Google does not use in determining a country target are locational meta tags such as geo-position or distribution or HTML attributes for geotargeting.

Google also utilizes language tags in its algorithms to determine the target audience of the page. When it comes to language tag, Google advices to use rel="alternate" hreflang="x" annotations to help Google serve the correct language to users in search queries (Google Webmaster Tools "rel="alternate" hreflang="x""). Google recommends using the annotations when only the template of the page (for example navigation or footer) is translated to another language while the main content is in single language. Forum posts use this kind of method for their sites rather often.

Also it is good to use the language attributes when writing similar content within a single language, for example Spanish. When there are Spanish readers in Mexico and in Spain, it is good to add the attributes to target the audience better. Google also points out that it is good to use the attributes if the page is fully translated to another language.

Google advices to think about the URLs carefully, even when it doesn't actually use the URLs to determine the language of the page, but the content. The URL provides the human users the information of the page content, also the language. Google reminds that if the company has multiple language versions of a URL, all pages must use

rel="alternate" hreflang="x" to indentify different language versions, including itself. An example that Google has provided is as follows: "For example, if your site provides content in French, English, and Spanish, the Spanish version must include a rel="alternate" hreflang="x" link for itself in addition to links to the French and English versions."

Google has prepared the following table for the URL use and reminds that it might be the best to consider using a URL structure that makes geotargeting the pages to different regions easy.

TABLE 1. Google has provided a table for different language URL usages. Source: https://support.google.com/webmasters/answer/182192?hl=en&ref_topic=2370587

URL structure	Example	Pros	Cons
ccTLDs	example.ie	- Clear geotargeting - Server location irrelevant - Easy separation of sites	- Expensive (and may have limited availability) -Requires more infrastructure -Strict ccTLD requirements (sometimes)
Subdomains with gTLDs	de.example.com	-Easy to set up -Can use geotargeting -Allows different server locations - Easy separation of sites	-Users might not recognize geotargeting from the URL alone (is "de" the language or country?)
Subdirectories with gTLDs	example.com/de/	-Easy to set up -Can use Webmaster Tools -Geotargeting -Low maintenance (same host)	-Users might not recognize geotargeting from the URL alone -Single server location -Separation of sites harder
URL parameters	site.com?loc=de	Not recommended	-URL based segmentation difficult -Users might not recognize geotargeting from the URL alone -Geotargeting in Webmaster Tools is not possible

5.1.4   Title tag

Usually the first thing the crawlers as well as users see is the title tag of the page. The page title shows in Google search as the header for the page and is the first landmark for the algorithms to determine if the page is relevant for the search query or not. Eric Enge et al. say that the page title is "one of the most important factors in ranking a given web page" (2009, 36).

Also Matt Cutts has given an advice as according to Barry Schwartz in the before-mentioned article "The Top Five SEO Mistakes According To Google's Matt Cutts". The advice Matt Cutts has given is not to "forget to think about the title and description of your most important pages". This reminder alone should be enough for the webmasters and marketers to acknowledge the importance of the titles. Below I have an example of the search results for "white tiger" and how the titles for the search look like.

PICTURE 4. The search query for "white tiger" to demonstrate the title tag. Title tag is marked with red.

I have marked the title tag with red to distinguish it clearly. As you can see, the highest ranking page for the search query has the exact keywords I have typed. The same thing is with the other results. The algorithms have analyzed these pages and have determined its content to be relevant for the query and therefore these pages have been presented to me. But, as White Tiger is such a large range keyword, the algorithms are not sure if I wanted to know more about the animal or a club of the name "White Tiger" or comics. Therefore there are many results provided for me where I can choose the best option.

It is important to have some sort of extra information on the title tag for the user to know whether the page is indeed about the animal, comics or a club. Also, for branding purposes it is good to have your brand name (company, organization, product) in the title for the algorithms as well as users start associating certain keywords with your brand.

```
1  <!DOCTYPE html>
2  <html lang="en" dir="ltr" class="client-nojs">
3  <head>
4  <meta charset="UTF-8" /><title>White tiger - Wikipedia, the free encyclopedia</title>
5  <meta name="generator" content="MediaWiki 1.22wmf16" />
6  <link rel="alternate" type="application/x-wiki" title="Edit this page" href="/w/index.ph
7  <link rel="edit" title="Edit this page" href="/w/index.php?title=White tiger&amp;action=
8  <link rel="shortcut icon" href="//bits.wikimedia.org/favicon/wikipedia.ico" />
```

PICTURE 5. The title tag as the algorithms see them. The title also has other text besides keywords; in this case it has "Wikipedia, the free encyclopedia".

According to SEO Moz post "Title tag" (http://moz.com/learn/seo/title-tag), the "optimal format" for the title tag is the following:

Primary Keyword - Secondary Keyword | Brand Name
or
Brand Name | Primary Keyword and Secondary Keyword

The title tag is therefore divided into two parts, the other part including the keywords and the other, the brand name. In the example above the page title was indeed in this form, "White tiger – Wikipedia, the free encyclopedia".

SEO Moz also advices to use less than 70 characters for the title tag as this is the limit Google displays in search results. This seems logical since the title cannot be too long. As we can see from the figure above, the two last search results seem to have too long titles so the rest of the titles are cut and what are left are three dots. This may not be pleasant for the user since now something relevant of the page info remains unseen and at some point this may affect the search results negatively.

On the other hand, the page title shouldn't be too short either. Too little information on the page title leaves a vague image for the page. Even if the page ranks in Google first page the user may not come to the page but chooses the competitor's if the competitor has more informative and catchy page title.

5.1.5 Meta description

Enge et al. say in their book "Art of SEO: Mastering Search Engine Optimization" (2009, 38) that Meta tag does not make any difference to search engine rankings but where it does make difference is when the page shows in search results. A good, informative meta description is more likely to earn more clicks to the page than poor and vague description. The meta description serves as an advertisement or summary of the page for the user and therefore the more accurate the description the better for the user and for the marketer. Below is an example of meta description.

PICTURE 6. Meta description illustration.

For this example I used "Bengal tiger" in search query. As you can see, there are many results for the keywords used and all that is left is that I choose the page I want to read. Now this is where the meta description counts the most. I have marked the meta description part with red to distinguish from the other tags.

The first page is Wikipedia site about the tigers. In the meta description is explained about the tiger and it is very clear that this page is about the animal, Bengal tiger and not, let's say, a rapper of the same name. The description is long and Google has cut part of it away and indicates it with three dots at the end. The other result below Wikipedia is National Geographic page and the description tells what you can expect on this page: pictures, videos, photos etc. This description is short yet informative and there is even space left!

It seems both are good descriptions since they tell the user exactly what the page is about. Now it is only up to the user to choose which page to view.

```
<!-- Page metadata -->
<meta name="description" content="Learn all you wanted to know about Bengal tigers with pictures
<meta name="keywords" content="bengal tiger profile, facts, information, photos, pictures, sound

<meta name="author" content="National Geographic Society" />
```

PICTURE 7. Meta description in html form as how the algorithms see it.

It is important that the meta description reflects the page content. If the meta description is different from what the page is actually about, Google might mark the page as trying to practice cloaking, which was explained before in algorithms. If Google judges the page as practicing cloaking the rankings might go lower or Google might eliminate the page from the search query altogether.

5.1.6   Text on the webpage

Text on the webpage is the most important for the user as they are usually looking for information on the web. The algorithms also "read" the text that the marketer produces and evaluates the quality of the text on the page. Below is the text part of the page as the robots see it. Once they detect the "article text" attribute the algorithms concentrate on analyzing the text part that is provided. The algorithms go through the pages and "mark" the keywords and contents from the page text and determine the relevance of the page.

```
<div class="article_text">
    <p>Tigers are the largest members of the cat family and are renowned for their power
</div><!-- .article_text-->
<script type="text/javascript">timerCount('article_body');</script>
</div><!-- .article_body-->
```

PICTURE 8. Algorithms only see this view of the page text.

One issue that is analyzed is the uniqueness of the text. The algorithms check the uniqueness and if they detect that the content on your page is copied from somewhere else, your page will most likely go down in rankings. On the other hand, if the text is uniquely produced and does not have the signs for it to have been copied elsewhere the ranking may rise. Therefore the marketers ought to concentrate to create unique content

that is relevant for the user and that has value. Copied content gives the user a negative experience since he or she may already have read the same text on some other page and this is something that Google wants to avoid. Also, with copied content there is the issue of possible copyright infringement. As a proud white hat SEO practitioner you must never copy the text from another page.

The algorithms also check that the text is of good quality. This is good news for the marketer since now the skill for creating great marketing text comes to use. If you want to enhance your chances of succeeding in SEO, you need your basic marketing skills. Google has in many occasions told marketers to create great text that is pleasant to read and the kind of text that searchers want to read. In most cases the text that people want to read, and maybe bookmark, has good grammar and structure. Often the lack of spelling mistakes is much appreciated, so if possible, no spelling mistakes at all in marketing texts is job well done.

In the text that you write, remember to use the keywords but use them wisely! Do not fill the text with the keywords too much, so that the text would become incomprehensive. There has been some discussion in internet about how often the keywords should repeat in the text but worrying about keyword density has become old-fashioned approach to SEO. Matt Cutts reminds in his video in YouTube "What is the ideal keyword density of a page?" (2011) that the SEO practitioners shouldn't worry about issues like keyword density and follow some artificial rules about them, but rather create good quality content where the keywords appear naturally.

Matt Cutts emphasizes that the first time mentioning the keyword is good, you may mention it twice or three times, but these two latter repeats do not bring in the same value as the first mention of the keyword. As discussed earlier in algorithms, keyword stuffing is considered a spam so if there are too many keywords in the same text, there is a serious threat of Google's algorithms interpreting your page as spammy and not relevant. This will bombard your good SEO intentions.

When inputting content to the page, the webmaster or marketer must remember, that not everything is visible to the search bots like they are to human eye. There are things that search robots do not "see". Enge et al. have listed these in their book "The Art of SEO – Mastering Search Engine Optimization" (2010, 40). The first things in the list are the

images.   Also Harold Davis mentions in his book "Google –Kuinka varmistat verkkonäkyvyytesi" (2007, 54) that the images mean nothing for the crawling robots. Robots can detect there is a picture, but they cannot "see" what is inside it. For example the search engines cannot see whether there is an image of a fruit or text. Enge et al. mention the same thing but add that robots can determine pixel color and even determine if the images have pornographic content based on how much flesh tone there is in the image.

Therefore it is crucial to remember to name the images and use the alt= attribute where to input the name of the picture. Use naming such as "red apple" if the picture is about red apple, "green apple" for green one etc. It is also a good idea to include a text below the picture to explain something about the image so the crawling robots get the idea of the image and the page.
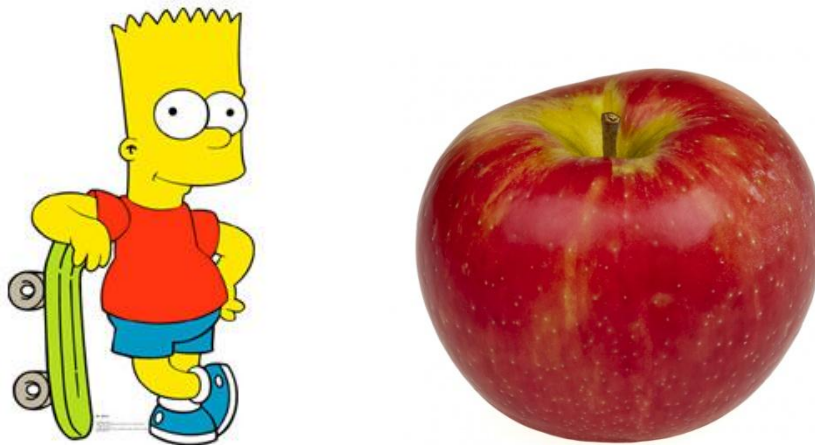
PICTURE 9. The search engines cannot see whether there is an apple in the picture or Bart Simpson. They can only detect the color red, for instance.

Another thing that Enge et al. have listed is Flash files. This is outdated information though, since Google has announced in their Webmaster Tools blog post "Flash and other rich media files" that googlebot "can index almost any text a user can see as they interact with any Flash SWF file on your site, and can use that text to generate a snippet or match query terms in Google searches". Also, Google has stated that you do not need to do any special action for Flash files.

Enge et al. mention also that audio and video files are difficult for search engines to read. Google has stated in the same blog post as above that googlebot cannot crawl the video file contents and reminds that it is important to provide information about the videos on the webpage. There are also other things included to what search engines cannot read or "see", such as some programs and scripts but I will skip them from this thesis.

5.1.7 Keywords

When it comes to choosing keywords, the task is in no way simple and understanding algorithms and users are extremely important. What are then wrong or right keywords? Mike Moran and Bill Hunt say in their book "Search Engine Marketing Inc.: Driving Search Traffic to Your Company's Web Site" (2009, 268) that the keywords you would use in search are often different from those of your customers since you know the keywords to find your site with, but the customers do not. The searchers may not know the exact words for what they are looking for and therefore might type something that might sound silly to some, or they might type something that could have two possible meanings.

Moran and Hunt give an example of this kind of search behavior. When the searcher inputs "home Improvement" to search query, does he mean Tim Allen's TV show or is there intention to remodel the house? According to Moran and Hunt, the search engine's researchers say that making sense of the searcher's query is their biggest frustration. As difficult as finding the right keywords is, it is not impossible. Moran and Hunt remind that it is possible to improve the search marketing, just think about the "need behind the query" (2009, 84).

Matt Cutts has mentioned "not including right words on the page" as one the SEO mistakes as Barry Schwartz has reported in his article "The Top Five SEO Mistakes According To Google's Matt Cutts". According to the article the example given by Matt Cutts is the following: "don't just write, 'Mt. Everest Height' but write, 'How high is Mt. Everest?' because that is how people search". This summarizes to the very old fact that the marketer must know his or her target audience, how they would use the language, what are the "trending" words etc.

But, how to choose the right keywords that attract traffic? One advice that Moran and Hunt give in their book is that the keywords you choose should not be "too hot" or "too cold". The advice Moran and Hunt give is to find the keywords that are "just right" (2009, 271-273).

With "too hot" keywords they mean that the keywords are too broad for you to use or too popular to give you any relevant traffic. Also, some keywords are already too fixated on some certain brand. If we think about Golden Moments, despite the company selling gift experiences, it might not be a very good idea to use "gift experiences" as the only keywords to attract traffic. Since Golden Moments is relatively small business in the field of gift experiences and there are many other similar companies the keywords "gift experiences" might be too popular term. This is especially since most people associate "gift experiences" with Smartbox.

Also, the kinds of keywords that have multiple meanings are considered "too hot" by Moran and Hunt (2009, 273). They give an example of "security" meaning different things to different audiences; also the term is rather broad. They suggest adding some qualifying words to the term so it becomes more specific. This way the chances for the right target audience is better. They also mention acronyms are not the best possible keywords, especially the kind of acronyms which might have several meanings depending on field. In addition to the acronym it might be wise to add the "longer" version as keywords as well to attract the right kind of traffic.

Also, it is wise not to focus on "too cold" keywords either. "Too cold" in this case mean keywords which are not popular and might not have much competition. These kinds of keywords are the kind that people do not use actively. It is simply not wise to use old-fashioned keywords that will not bring in visitors.

Instead of trying to use too popular keywords or those that are either old-fashioned or otherwise not used by people, it is better to try to think of what your site really offers. In Moran and Hunt's words: "sticking to keywords that truly reflect your site will make it easier for you to attract search rankings and will pay off in more conversions than if you do anything else" (2009, 276).

There are helping platforms to find the keywords, such as Google Adwords, KEI (Keyword Effectiveness Index) and other, but the best advice is to stick to the keywords that reflect the page and that describe the products the best. It might be a good idea to brainstorm with your colleagues whenever you have a chance and think about new keywords, do research on the suggested and fitting ones and then choose the best keywords for your page.

Also, the good news is that the searchers on the web have become more accustomed to the search and are therefore using more detailed search queries. This means that even if the keywords do not match perfectly, if there are some words that match to the content on your page the chances for the visitor landing on your page is slightly higher. This does not rule out proper effort done for good SEO though.

5.2. Google Penalties

Google develops their algorithms all the time to answer the demand for more relevant results, faster. As Google has mentioned in their Inside Search post "Fighting Spam", millions of spam sites are being created every day. And Google continues fighting the spam with their algorithms and by manual review.

As Google's mission is to provide the user the best experience possible, sites which use techniques such as cloaking or keyword stuffing or other sorts of manipulative methods of getting better ranking is bad, for all parties. By doing this the relevant websites get lost in the mass, or in Google's words, they "get buried", which is always sorrowful for the site owners who have done things right but their sites end up not easy to find.

For the reason of keeping the search pleasant and to be able to give out relevant results Google has developed Google Webmaster tools for the marketers to know what good SEO is. For the violation of these guidelines Google has created penalties for the site owners. I find it important to mention that the rankings going lower due to algorithms are not the same as Google penalties. Google penalties are manual work by the spam team members of Google and are due to serious guidelines violations.

Penalties are not something the webmasters can just laugh at or ignore. Especially if the site has been created for the purpose of business, the penalizations and the threats must be taken seriously. Once the page has been penalized the progress to fix the situation is slow and may even be impossible to fix, depending on for what reason the site has been penalized. Algorithmic penalizations are more difficult to fix and might take long time to recover.

The manual penalizations might be easier to fix but it doesn't make the process necessarily any faster or make the recovery happen at all. Google notifies the violation of the guidelines in which case the webmaster should immediately think about their actions. If you indeed are guilty of spamming, cloaking or any other violations, take Kristine Scharzinger's advice ("Pure Spam: What Are Google Penalties & What to Do to Recover") and admit having being caught and move on. Fix the mistakes and hope for the best. Though, according to her the chances of recovering from major violations are rather slim.

If on the other hand you have not been doing anything wrong but you have been either hacked or the violation is due to user generated content (UGC) Schazinger advices to turn to SEO specialist who has specialized in penalizations. Also, Google has their own reconsideration request form which the webmasters can fill and send to Google for another review of their page. According to Google all the requests are processed and the webpage owners will "know how it's going" (Google Inside Search "Fighting Spam").

Despite the penalizations being serious issue, the manual penalizations usually have "expiration date" as Matt Cutts calls them in GoogleWebmasterHelp video "When are penalties lifted?" (2011). The expiry dates differ greatly depending on how grave the penalty has been. Google has also mentioned this in their webmaster blog post "Fighting Spam". They have stated that manual actions do not last forever and once the website owner has fixed the violations to non-violating content, they can request reconsideration from Google.

Rachel Searles and Brian White from Google Search Quality team give tips on how to make the reconsideration request for Google in the video "Tips on requesting reconsideration" (2009) in Youtube. First of all, they advice the webmasters to admit the mistakes they have done and fix those mistakes. Secondly the webmasters should

fix the mistakes in a way that there is no need to worry for the future violations. Searles and White tell that when sending the request only telling Google that "I am using the Google guidelines now" is not enough information to process the request properly. When doing the reconsideration request, the webmaster must tell what the penalty was for and how they have fixed it, in as much detail as possible. Searles and White say that one message to the reconsideration forum is enough, spamming is not going to work.

I believe it is important for marketers to know about the penalizations beforehand since fixing the situation after the penalty has been issued takes time and damages the site reputation no matter how quickly the issues are fixed. Also I believe that understanding the things that result to penalties (cloaking, spamming, hacking etc.) will provide the marketers better morale to do SEO properly and follow the algorithm changes and be up to date in general of what is going on in online business environment.

6 UNDERSTANDING USER BEHAVIOR IN INTERNET

Originally, search engines started developing algorithms to make the search easier and for the users to find the information they needed without browsing through numerous pages before finding what they were looking for. The developers of algorithms find it extremely important to understand the user behavior in order to further enhance the user experience in search. According to Enge et al. in the book "Art of SEO: Mastering Search Engine Optimization" (2009) the search engines invest considerable amount of resources into understanding how people search.

Therefore, also marketers who use these algorithms and SEO for their benefit also shouldn't forget the importance of understanding the users. As a marketer we should never focus too much on only algorithms or other technical stuff. The reason why we are working on SEO is always to attract customers. We are not working to attract just any kind of customers, but the targeted customers who would be interested in the products. Therefore, as important as it is to understand the algorithms to practice successful SEO, equally important is to understand user/searcher behavior. In the end, algorithms work in order to make the search experience better for the user. Also SEO revolves around user's needs and behavioral issues, for example successful keywords come from the words that the customers type when searching for something.

Basic marketing skills are needed in online marketing as well. The SEO always comes down to planning and segmenting, just like any other marketing form. There is a difference, though, as Moran and Hunt see it in their book "Search Engine Marketing Inc.: Driving Search Traffic to Your Company's Web Site" (2009, 271), one of the biggest "advantages" as they say is that search marketing is rather a "pull" strategy than the traditional "push" media such as print ads or TV ads. The customers are in more active role in online marketing.

I agree to their view since in internet the customers in most part know what they want and are looking for information regarding that. The customers are already searching for information and most likely also the places where to buy. With right keywords the customers will find their way to the online business webpages and the "pull" is in action. There is necessarily no need for the company to push the advertisements of its

products when the SEO is done well. The customers are "pulled" in to the webpage with good SEO and the true interest from the customer's part.

It is important to notice that the customers will not suddenly appear to the webpage with just any keywords and with just any kind of SEO. The old marketing knowledge comes to use in here regarding segmenting and targeting. According to Enge et al. in the book "Art of SEO; Mastering Search Engine Optimization" understanding the psychology of the target audience is one of the most important issues when it comes to building an online marketing strategy (2009, 2). It is certainly a different story to attract youngsters to the webpage than retired people. First of all, the way the search is conducted and the keywords these two groups use might differ from each other greatly. Below I will explain the different kinds of searches to give some image of how people search in general.

Enge et al. have divided the searches that users conduct into three categories (2009, 6). The first category is navigational query, with which is meant the search of a certain webpage. Enge et al. have described this as the kind of search where the user is looking for certain company or brand but may not be sure about the URL. So the user types the name or what he or she remembers of the name to search query.

PICTURE 10. Navigational query as illustrated by Enge et al. in "Art of SEO; Mastering Search Engine Optimization" (2009, 6).

This kind of search is often of low value, but only if the search is about other companies or brands and not yours. On the other hand, if the user wants to come directly to your page, the value of this navigational search is high. Therefore it is important to build brand value also online and make the site easily found and pleasant for the user to search.

The second search category that the users conduct is transactional query. Transactional query happens when the user is searching the web in order to buy something. This query is rather wide and involves also things like creating Gmail account, paying bills or finding a restaurant from the area where to dine. (Enge et al., 2009)

Figure 1-4. Transactional query

PICTURE 11. Transactional query illustrated by Enge et al. (2009, 7).

The value of this search is very high since the purpose for the query is to buy something, or prepare for buying. The money transaction may not happen immediately, or even online, but this query may push the user forward to buying your product. On the other hand, if the product, the page or something else doesn't please the user, he or she may turn away from your product and search for competitor's sites.

The third category is informational query and as its name suggests, is used when looking for information on something. The range for this query is extremely wide as it includes everything from checking the local weather to details on cameras or the names of actors. The goal of this search is primarily information and the query is often non-transactional even if the user was looking at a certain product.
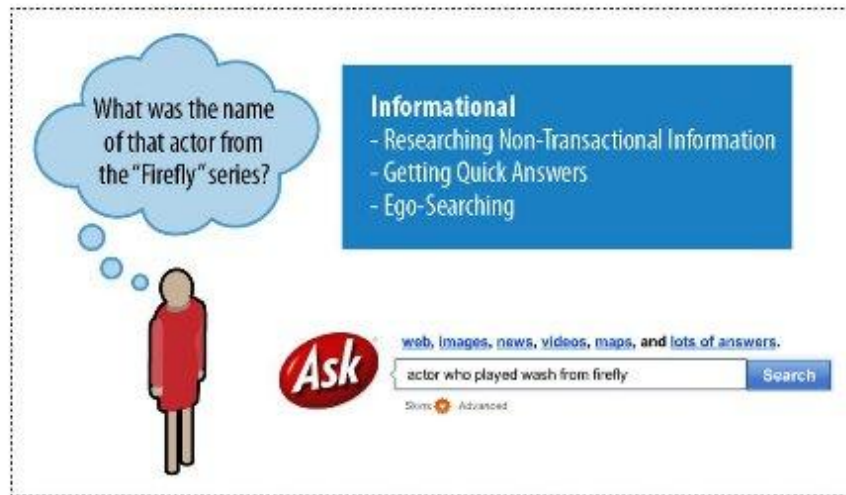
Figure 1-3. Informational query

PICTURE 12. Eric Enge et al. have illustrated the informational query in their book "Art of SEO; Mastering Search Engine Optimization" (2009, 7).

The value for this search is medium even though it is undeniable that the value may rise higher if the content on, for example product page, is so good that the customer may want to purchase it despite not originally meaning to buy it at that moment. The opportunities this search offers raise the value of it. For example the possibility of journalists getting interested in the product or your brand serves as great value if the attention is positive. Also the other users, if they find the information on your site valuable, might link to your site and thus attract attention to it. As we learned from the algorithms part, the more valuable natural back links to your site the better the ranking.

According to Enge et al. a research has been made about search queries by Pennsylvania State University and the Queensland University of Technology and the results show that over 80% of the search queries were informational. This leaves little more than 10% for navigational and transactional queries (2009, 8). Therefore I believe it is worth the effort to focus on making the pages as pleasant to browse and as informational as possible. There is always a chance that the informational query changes into transactional along time.

The problem with attracting traffic is often choosing the right keywords. The way the users type in the keywords is often the main headache for the marketers. Enge et al. has mentioned that the data the users type in to search engines is very limited, only two to

three words. The limited data in search queries results to too general query so the intent of the query remains unclear for the engines. This is due to users, unlike the marketers; do not have the keen knowledge of how the search engines work. (2009, 8)

I find important to mention why it is better to be higher up in the search results in Google first page than lower. Below is a graph of where the internet users look first when they see the Google results page. The areas marked red are where the people look at the most and longest.
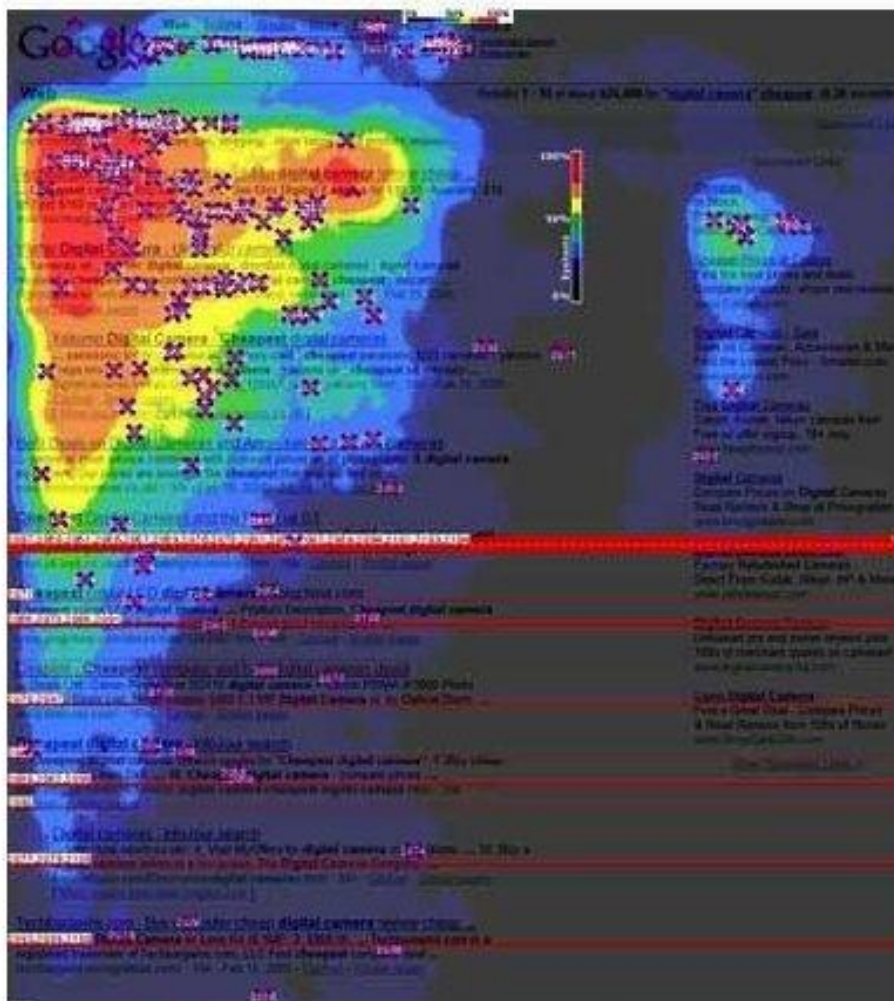

Figure 1-8. Enquiro eye-tracking results

PICTURE 13. The research made by Enquiro, Eyetools and Didit as displayed in Enge et al. book "Art of SEO: Mastering Search Engine Optimization." (2009, 15)

The research made by research firm Enquiro, Eyetools and Didit show that the users spend the longest time looking at the top-left area of the search results page. This means

that the first and second results get the most attention and the lower results are focused on remarkably less. Therefore, only getting to the first page of Google is not necessarily enough, but the companies should strive to get the first position to get more visits to their page.

In conclusion to this part, it is important not to forget the basics of marketing even when the platform is transferred online. The basics of understanding the customers and segmenting and targeting are no less needed than in traditional marketing. Targeting is increasingly important since with targeting it might be easier to choose the right and accurate keywords with which to attract the right kind of traffic. With careful planning and targeting it is also easier to please the targeted audience since the page offers what the customers want to see and the page layout is most likely also designed to fit the audience preferences.

7   CONCLUSION

In this thesis I have introduced the algorithms and given basic knowledge regarding them. I have explained that the algorithms are the "behind the scenes" of the search and therefore it is important to understand the algorithms and how they work. As I have explained, the algorithms work behind the search in order to organize all the information available on the web and to be able to give accurate results to the search query conducted. Algorithms do not only organize the information, they also rank the pages for the purpose of providing relevant results to the searcher. This means that the algorithms check the pages for the content and either deems it relevant, and the ranking goes higher or non-relevant and the ranking goes lower. Naturally, as marketers we want the rankings of our page go higher so we need to be aware of the factors that make the pages either rise or fall in the eyes of Google.

Algorithms are, in other words, the base for the SEO. The rules for SEO come from the facts of what algorithms judge as good or bad qualities of the webpage. It is easier for the marketer to follow the SEO rules and understand how good SEO is done when the algorithms are already understood and backing up the knowledge for the online environment. Also, despite there being rumors of SEO changing and how certain SEO tricks will enhance the rankings, with good understanding of the algorithms the marketer will have better insight to what is in actuality a rumor and what things should be taken more seriously.

So, the algorithms and SEO are quite understandably connected as these two complement each other. What do users have to do with either of these then? First of all, the algorithms were created for the benefit of the internet users. Algorithms exist since the search engines wanted to make the search easier and more pleasant for the user. The whole idea behind algorithms is to work for the user to find what he or she wants to find easily and conveniently.

SEO was born for the users to find a certain page that relates to the search query conducted. SEO is strongly connected to users as well since it is the users we really want to attract to our webpages. As mentioned before, the users do not just happen to find our page, they find the page since it has been displayed as relevant result for certain keyword. Therefore the marketers need to have the SEO right in order for the

algorithms to think that the content is relevant for the search conducted with certain keywords and for the webpage to rank well in Google search. Keywords are thus the strongest link between the SEO and the users. The marketers need to know their target audience and users in order to choose the right keywords that would attract the users to their webpage. If the marketers do not want to trust in luck only, they need to understand how the target audiences behave in internet and what kinds of things they search for. Also, knowing with which keywords the users search is important to get them to find that company's webpage and not competitors'.
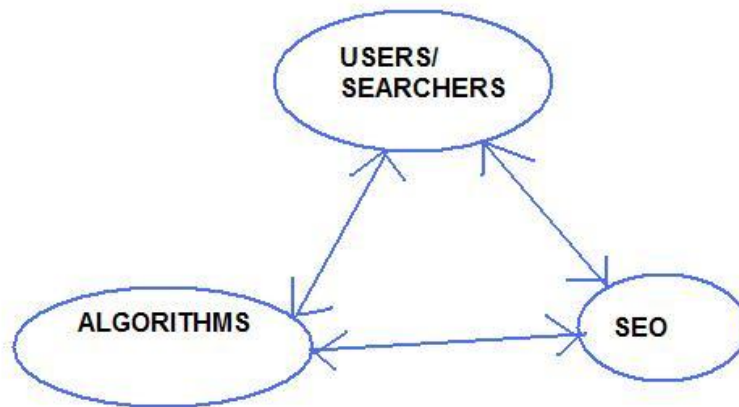
PICTURE 14. The algorithms, SEO and users are connected.

In my opinion understanding the triangle of users, algorithms and SEO is extremely important for an online business. After all, without the internet users there would be no need for the algorithms. Without algorithms there would be no need for the SEO. Without SEO and algorithms the users most likely wouldn't utilize the internet search. These three ends of a triangle are the base for working and ever developing internet and search culture.

As final conclusion to this thesis, it is not enough that the marketers try to throw keywords into their webpages and hope for the best, but they really need to work to understand the environment where they work. The understanding is good to start from the understanding of the target audience and to carefully plan the marketing strategy

like it is planned in traditional marketing as well. After that is it good to start learning the basics of algorithms which are the behind the scenes factor for search and for many features in internet. From there, when the algorithms have been understood, it is natural way to start learning the SEO and utilize the knowledge from understanding the target audience and algorithms to create great content that will rank high in Google search.

REFERENCES

Amie Baumwell. *Panda 2.5 Does Some Damage.* [Online] Posted 3rd October 2011. Available from: http://www.eversparkinteractive.com/panda-25-damage/ (Accessed 17th September 2013)

Barry Schwartz. Search Engine Land. *The Top Five SEO Mistakes According To Google's Matt Cutts*. [Online] Posted 29th April 2013. Available from: http://searchengineland.com/the-top-five-seo-mistakes-according-to-googles-matt-cutts-157574 (Accessed 2nd September 2013)

Danny Sullivan. Search Engine Land. *FAQ: All About The New Google "Hummingbird" Algorithm*. [Online] Posted 26th September 2013. Available from: http://searchengineland.com/google-hummingbird-172816 (Accessed 9th October 2013)

Danny Sullivan. Search Engine Land. *Google Forecloses on Content Farms With "Panda" algorithm Update.* [Online] Posted: 24th Feb 2011. Available from: http://searchengineland.com/google-forecloses-on-content-farms-with-farmer-algorithm-update-66071 (Accessed 7th July 2013)

Danny Sullivan. Search Engine Land. *Google Launches "Penguin Update" Targeting Webspam In Search Results*. [Online] Posted 24th April 2012. Available from: http://searchengineland.com/google-launches-update-targeting-webspam-in-search-results-119295 (Accessed 3rd September 2013)

Danny Sullivan. Search Engine Land. *Penguin 5, With The Penguin 2.1 Spam-Filtering Algorithm, Is Now Live.* [Online] Posted 4th October 2013. Available from: http://searchengineland.com/penguin-2-1-and-5-live-173632 (Accessed 9th October 2013)

Enge E., Spencer S., Fishkin R., Stricciola J.S. *The Art of SEO; Mastering Search Engine Optimization*, United States of America: O'Reilly Media Inc. 2010.

Google. Google Company. *Google's mission is to organize the world's information and make it universally accessible and useful*. [Weblog] Available from: http://www.google.com/about/company/ (Accessed 3rd June 2013)

Google. Google Company. *Our products and services*. [Weblog] Available from: http://www.google.com/about/company/products/ (Accessed 3rd June 2013)

Google. Google Company. *Ten things we know to be true*. [Weblog] Available from: http://www.google.com/about/company/philosophy/ (Accessed 3rd June 2013)

Google. Google Webmaster Tools. *Flash and other rich media files*. [Weblog] Available from: https://support.google.com/webmasters/answer/72746?hl=en (Accessed 10th September 2013)

Google. Inside Search. *Fighting Spam*. [Weblog] Available from: http://www.google.com/insidesearch/howsearchworks/fighting-spam.html (Accessed 2nd October 2013)

Google. Webmaster Tools. *Cloaking*. [Online] Available from: https://support.google.com/webmasters/answer/66355 (Accessed 24th July 2013)

Google. Webmaster Tools. *Duplicate content*. [Online] Available from: https://support.google.com/webmasters/answer/66359 (Accessed 24th July 2013)

Google. Webmaster Tools. *Keyword stuffing*. [Online] Available from: https://support.google.com/webmasters/answer/66358 (Accessed 24th July 2013)

Google. Webmaster Tools. *Link Schemes*. [Online] Available from: https://support.google.com/webmasters/answer/66356 (Accessed 24th July 2013)

Google. Webmaster Tools. *Multi-regional and multilingual sites.* [Weblog] Available from:
https://support.google.com/webmasters/answer/182192?hl=en&ref_topic=2370587
(Accessed 9th September 2013)

Google. Webmaster Tools. *rel="alternate" hreflang="x"*. Weblog. [Online] Available from: https://support.google.com/webmasters/answer/189077?hl=en (Accessed 9th September 2013)

GoogleWebmasterHelp. *Does the position of keywords in the URL affect ranking?* [Video] 2009. Available from: http://www.youtube.com/watch?v=gRzMhlFZz9I (Accessed 11th September 2013)

GoogleWebmasterHelp. *Tips on requesting reconsideration* [Video] 2009. Available from: http://www.youtube.com/watch?v=G959-2RlPyA (Accessed 2nd October 2013)

GoogleWebmasterHelp. *What is the ideal keyword density of a page?* [Video] 2011. Available from: http://www.youtube.com/watch?v=Rk4qgQdp2UA (Accessed 15th September 2013)

GoogleWebmasterHelp. *When are penalties lifted?* [Video] 2011. Available from: http://www.youtube.com/watch?v=ES01L4xjSXE (Accessed 2nd October 2013)

Harold Davis. *Google –Kuinka varmistat verkkonäkyvyytesi.* Trans. Marko Niemi. Helsinki. 2007.

Hitesh Patel. *Google Equivalency Score Algorithm.* [Online] Posted 28th December 2012. Available from: http://www.seo.co.uk/google-equivalency-score-algorithm (Accessed 18th September 2013)

Humaira Khan. *SEO – Massive Changes Expected in 2013.* [Online] No posting date available. Available from: http://smashinghub.com/seo-massive-changes-expected-in-2013.htm (Accessed 20th September 2013)

Kristine Schazinger. SearchEngineWatch.com. *Pure Spam: What Are Google Penalties & What to Do to Recover* [Online] Posted 21st June 2013. Available from: http://searchenginewatch.com/article/2276498/Pure-Spam-What-Are-Google-Penalties-What-to-Do-to-Recover (Accessed 2nd October 2013)

Marziah Karch. About.com Guide. *What Is PageRank and How Do I Use It?* [Online]
No posting date available. Available from:
http://google.about.com/od/searchengineoptimization/a/pagerankexplain.htm (Accessed
6th July 2013)

Matt Cutts & Amid Singhal. Google Official Blog. Finding more high-quality sites in
search. [Weblog] Posted 24th February 2011. Available from:
http://googleblog.blogspot.fi/2011/02/finding-more-high-quality-sites-in.html
(Accessed 20th July 2013)

Matt Cutts, @mattcutts, Twitter. Tweeted 4th October 2013. Available from:
https://twitter.com/mattcutts/status/386231794883780609 (Accessed 5th October 2013)

Matt Cutts. Google Official Blog. *Google search and search engine spam*. [Weblog]
Posted 21st January 2011. Available from: http://googleblog.blogspot.fi/2011/01/google-
search-and-search-engine-spam.html (Accessed 20th July 2013)

Matt Cutts. Official Google Webmaster Central Blog. *Another step to reward high-
quality sites*. [Weblog] Posted 24th April 2012. Available from:
http://googlewebmastercentral.blogspot.co.uk/2012/04/another-step-to-reward-high-
quality.html (Accessed 23rd July 2013)

Matt Cutts. *Penguin 2.0 rolled out today*. [Weblog] Posted 22nd May 2013. Available
from: http://www.mattcutts.com/blog/penguin-2-0-rolled-out-today/ (Accessed 15th July
2013)

Moran M., Hunt B. *Search Engine Marketing, Inc. Driving Search Traffic to Your
Company's Website*. 2nd Edition. United States of America: Pearson plc, IBM Press.
2009.

Picture 2. PageRank by Micaroni Lalli. Available from:
http://en.wikipedia.org/wiki/File:PageRank-hi-res.png (Accessed 21st July 2013)

Steve Fitzpatrick. *Will the Google 'Equivalency Score' Algorithm Change Kill Your
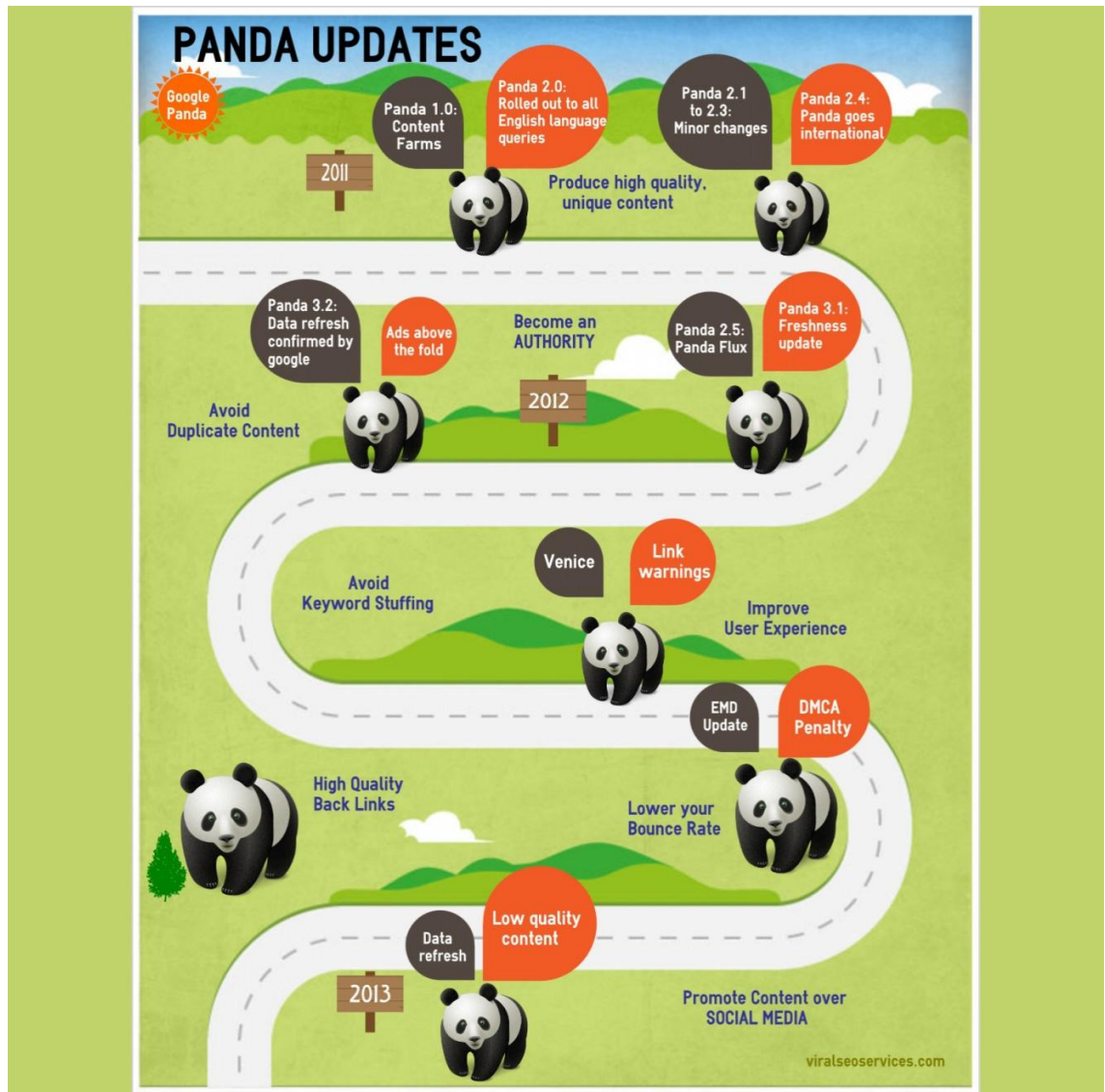Business?* [Online] No posting date available. Available from:

http://digitaldomination.com/google-algorithm-equivalency-score/   (Accessed   5th
September 2013)

APPENDICES

Appendix 1. Panda Updates

The visualization of Google Panda updates (image source http://visual.ly/google-panda-algorithm-updates)