Saimaa University of Applied Sciences
Business Administration, Lappeenranta
Degree Programme in International Business Management

Simo Tahvanainen

SEARCH ENGINE OPTIMIZATION, THE IMPACT ON CONSUMER SALES IN READY-MIXED CONCRETE

CASE RUDUS LTD

Master's Thesis 2012

The purpose of this study was to find out if there is a correlation between search engine visibility improvement and consumer sales in the ready- mixed concrete business. The theory part handles the importance of the Internet and e-marketing activities in general, and is also intended to be a guideline for e-marketing personnel in Rudus Ltd. Because of the case company's needs, search engine marketing is handled more extensively than the research question requires.

The research part was done by collecting data from various sources and by comparing the results to find out if there is a correlation between improvements and sales. The study was a constructive study where the case construction was Rudus' Website and the search engine marketing process.

The study was completed in the autumn of 2011. The theory part was written between September and November of 2011. Rudus' Website and consumer sales survey was carried out between April 1st 2011 and November 31st 2011. This period was compared to the situation between April 1st 2010 and November 31st 2010. There were no other significant changes in consumer sales practices between 2010 and 2011. This is an important factor when judging the relevance of the correlation. Also, Rudus' e-marketing organization and practices were handled during the process.

Keywords: Search Engine Optimization, SEO, Search Engine Marketing, SEM, Concrete

ABBREVIATIONS

CPC	Cost per Click
CSS	Cascading Style Sheets
DNS	Domain Name System
HTML	HyperText Mark-up Language
HTTP	HyperText Transfer Protocol
IAB	Internet Advertising Bureau
KWA	Keyword Advertising
PDF	Portable Document Format
PPC	Pay-Per-Click
RMC	Ready-Mixed Concrete
SEM	Search Engine Marketing
SEO	Search Engine Optimization
URL	Unified Resource Locator
XHTML	Extensible HyperText Mark-up Language

TERMS

Algorithm	Is an effective method expressed as a finite list of well-defined instructions for calculating a function. Algorithms are used for calculation, data processing, and automated reasoning. In simple words an algorithm is a step-by-step procedure for calculations.
Body copy	Body text is the term for the text forming the main content of a book, magazine, Web page or other printed matter. HTML element named <body> that serves to delimit the body matter from the front matter (or in HTML, the <head>) that contains metadata such as the page title. The typesetting of the Web page is carried out by document body elements within this.
Bounce Rate	Bounce rate represents the percentage of visitors who enter the site and "bounce" (leave the site) rather than continue viewing other pages within the same site.
Bowtie-theory	Is a theory that explains sites linking to each other.
Conversion rate	The ratio of visitors who convert casual content views or Website visits into desired actions based on subtle or direct requests from marketers, advertisers, and content creators. If the prospect has visited a marketer's Website, examples of conversion actions might include making an online purchase or submitting a form to request additional information.
Cookie	Is used for an origin Website to send state information to a user's browser and for the browser to return the state information to the origin site. The state information can be

	used for authentication, identification of a user session, user's preferences, shopping cart contents, or anything else that can be accomplished through storing text data on the user's computer.
Cost Per Click	Also called as Pay per Click. An Internet advertising model used to direct traffic to Websites, where advertisers pay to the publisher (typically a Website owner) when the ad is clicked. With search engines, advertisers typically bid on keyword phrases relevant to their target market. Content sites commonly charge a fixed price per click rather than use a bidding system. PPC "display" advertisements are shown on Websites with related content that have agreed to show ads.
Description	The description attribute provides a concise explanation of a Web page's content
Directory	A directory on the World Wide Web. It specializes in linking to other Websites and categorizing those links.
Flash-players	Is software for viewing multimedia, Rich Internet Applications and streaming video and audio, on a computer Web browser or on supported mobile devices.
HTML	Is written in the form of HTML elements consisting of tags, enclosed in angle brackets (like <html>), within the Web page content. HTML tags most commonly come in pairs like and , although some tags, known as empty elements, are unpaired, for example <img>. The first tag in a pair is the start tag, the second tag is the end tag (they are also called opening tags and closing tags). In between these tags Web designers can add text, tags, comments and other types of text-based content.
HTML mark-up	The predominant mark-up language for Web pages. HTML elements are the basic building-blocks of Web pages.
Iframes	Iframes allow a visual HTML Browser window to be split into segments, each of which can show a different document. This can lower bandwidth use, as repeating parts of a layout can be used in one frame, while variable content is displayed in another.
JavaScrip code	JavaScript is a prototype-based scripting language that is dynamic, weakly typed and has first-class functions. It is a multi-paradigm language, supporting object-oriented, imperative, and functional programming styles.
Keywords	One of the most commonly used meta elements.

Metadata	Data about the containers of data.
Page	Also called as Web page. Is a document or information resource that is suitable for the World Wide Web and can be accessed through a Web browser and displayed on a monitor or mobile device. This information is usually in HTML or XHTML format, and may provide navigation to other Web pages via hypertext links.
PageRank	PageRank is a link analysis algorithm, named after Larry Page and used by the Google Internet search engine, that assigns a numerical weighting to each element of a hyperlinked set of documents, such as the World Wide Web, with the purpose of "measuring" its relative importance within the set. The algorithm may be applied to any collection of entities with reciprocal quotations and references.
Page View	A page view is a request to load a single HTML file ("page") of an Internet site. Page views may be counted as part of web analytics.
Site	Also called Website, is a collection of related Web pages containing images, videos or other digital assets.
Text browser	A text-based Web browser is a Web browser that renders only the text of Web pages, and ignores graphic content. Usually, they render pages faster than Web browsers that include graphics because graphics need more bandwidth than text.
Title tag	Title element represents the title of the document. In the HTML syntax, most elements are written with a start tag and an end tag, with the content in between. Tags are composed of the name of the element, surrounded by angle brackets. An end tag also has a slash after the opening angle bracket, to distinguish it from the start tag.

CONTENTS

1   INTRODUCTION

Today, almost any service or product purchase process starts with a Web
search. In fact, 88 % of people start looking for a product or service online
(Tulos.fi, 2011). In Finland, the search tool is almost invariably the Google
search engine. In 2010, 98% of searches were done with Google (Mäkelä,
2011). Rudus Ltd has noticed the significance of search engine growth, but at
the same time has discovered that the Rudus Ltd Website's (www.rudus.fi)
visibility on Google is not good enough. This is a problematic issue since Rudus
Ltd is Finland's market leader in aggregates, crushing and ready-mix concrete,
yet the visibility of its Websites is not on same level.

The purpose of this study is to investigate Rudus Ltd's Website's search engine
optimization and search engine marketing impact on Rudus Ltd's consumer
sales in ready-mixed concrete. Rudus Ltd's markets are mainly in Finland for
Finnish customers, so the search engine being dealt with is limited to Google,
which has the dominant position in Finland. The theoretical element will be
dealing with search engine marketing, search engine optimization, related Web
analytics and other tools for online businesses.

After the theoretical element, a closer look will be taken at search engine
marketing and optimization from Rudus Ltd's point of view, following with Rudus
Ltd's ready-mix concrete business review and a closer look at facts that can be
measured in sales comparison. Finally a comparison is done between Rudus
Ltd's ready-mix concrete consumer sales during a selected period of time to
search engine optimization and number of potential customers entering through
Google to Rudus' Website during the same period. Those figures are then
compared to historical data of sales and visitors before optimization. The
purpose of this thesis is to find out the impact on sales when the company's
presence in search engines is optimized.

The author of the thesis is working in Rudus Ltd as a Regional Director in the
Northern Finland Unit and is involved in a wide range of development projects

concerning Rudus Ltd. One of his main duties is to be part of the Rudus ready-mix concrete management team. Due to this role he has a good preconception on Rudus' current situation in consumer sales and challenges being faced with the company's visibility in Google. The subject is interesting and very current. The results of the study might also be used in other fields of business. Search engine usage is increasing all the time and there are several companies facing these same challenges in their reach ability.

Background

Rudus Ltd is a market dominant company on aggregate and ready-mixed concrete in Finland, but after several brand name changes it has lost the brand advantage in consumer sales. Rudus Ltd has noticed that its reach ability through Google is very poor when consumers are searching only through products, without using the brand name. The biggest problem is that Rudus doesn't have any product linkage in their brand name when compared with its competitors (Ruskon Betoni, Luja Betoni…etc.). This is why the domain doesn't have a direct link to product searches. Consumer sales in the Rudus ready-mix concrete –unit (RMC) means customers that buy concrete for their own needs, mostly house/cottage builders and farmers.

Rudus Ltd has started in April 2011 a project to maximize their presence in Google when customers are searching for products from their portfolio. This study would support that work and would examine thoroughly the impact on consumer sales when the company's presence is optimized.

Internet marketing is a fresh, constantly growing and evolving form of marketing. Internet marketing means banner advertising, search engine optimization, search engine marketing (SEM), blog marketing and social media usage in commercial communication (Mäkelä, 2011). Search engine optimization (SEO) in turn is a form of e-marketing that has evolved around the Internet search engines. SEO aims to raise the Website to a high and prominent position in search engine results if possible.

SEO is one form of SEM, meaning measures to the site that improve site placement in search results as much as possible. The importance of SEO in marketing is growing fast simultaneously with consumer e-marketing habits. It can be called truly directed marketing which doesn't include any so–called "useless" contacts like in the more traditional medias.

Search engines are a very versatile channel in the marketing sense. Consumers are using search engines to compare alternatives, clarify the problems faced in purchase decisions and find the right place to shop. Studies show that 50 % of search engine searches are related to purchasing (Seppänen, 2011). It offers several options for the marketer to pull customers towards their products and services when allocating their portfolio correctly for the search engines and search engine users.

More and more people are today looking for product and service information via the Internet. Hence almost every company has their own Web pages. Companies have made huge investments for Web pages to fulfil the latest requirements and to be as user-friendly as possible. Companies are justifying their Web presence by marketing related reasons. The aim is to sell the company's products and services and to make them knowledgeable for the potential customers. Surprisingly, contrary to conventional marketing rules, the e-marketing success is still rarely measured. Were the target segments reached? What was the result of the campaign? SEO is giving answers to these questions. Site's SEO-process improves site visibility and it strengthens the link between company offerings and customer needs. SEO is a continuous marketing process. The right indicators need to be set in order to measure success and to continue the development process. Web analytics programs are tools used to gather information and statistics about the number of visitors and their activities on the Website. This informs marketers if the customers are finding the pages and what is the rate of return compared to other marketing activities. (Paananen, 2006.)

2   OBJECTIVES AND SCOPE

This study examines what the search engine marketing is about, how well does it affect the number of Website visitors and does the change correlate with the change in consumer sales. The study is limited to deal only with search engine marketing and search engine optimization in the field of Internet marketing. In the Rudus' survey part there is also a part from paid directories. (Chart 7.) Other Internet marketing forms would have been for example banner advertising and E–mail marketing. Other forms are also interesting and it would be useful for Rudus to study the benefit of those in some other context. In this study those are not studied because the topic would have been too broad to handle successfully.

This work was focused on improving the sites ranking in the search engine that has the most users and commercial value among the target customers of the Rudus Ltd's sites(www.rudus.fi). In other words, Google as Google's market share in Finland is overwhelmingly dominant. Google practically owns the markets and other companies combined have less than 2 percent of the market share. Search engine markets in Finland are divided as follows: 4/2010 – 4/2011.

- Google,   98,11 %
- Bing,   1,45 %
- Yahoo!,   0,25 %
- Ask Jeeves,   0,06 %
- AltaVista,   0,04 %
- Muut,   0,09 %.
    (Mäkelä, 2011)

This work handles only Rudus RMC's consumer sales in Finland, so the search engine is limited to Google. In Finland, SEO is already commonly called "Google optimization". SEO has an impact on all search engines even though it would primarily have been targeted to Google. Web pages will reach good positions in other search engines as well. For example the Microsoft owned

search engine, Bing, gives value to many of the same things as Google. (Mäkelä, 2011)

Internet site visibility can also be raised with several other optimization / marketing methods, such as increasing the company's visibility in social media like Facebook, Twitter, Google + or YouTube. Google also appreciates blogs or use of Google's map service. These themes will not be handled in this study.

This research will be made in co-operation with Rudus and it is made by examining the results of search engine usage and user behaviour in Rudus home pages. Rudus home page usage is examined before the optimizing work and after the work is in progress. The consumer sales during the same period are summed up and the results are compared.

Research method in this thesis is constructive research, i.e. applied research methodologies, and it is structured as a study type work where the subscriber company, Rudus, has an important role. Constructive research includes for example device ,system ,principle, etc… designing and evaluation(Seppänen, 2004). In Rudus case, the construction under investigation is their Website and the process of search engine marketing. The thesis includes also a correlation comparison between Website visitor amounts and consumer sales.

Search engine marketing is still young as a discipline. Itt has been around since 1997 , hence there is not much published literature available. Source material includes some published literature, but is focused on Web publications, as well as Master's and Bachelor's degree thesis studies.

3   THE IMPORTANCE OF THE INTERNET

Today, almost every field of business has a presence in the Internet through the company's own Web pages and different sorts of advertising forms. The amount of information on the Internet has grown exponentially. Number of sites in the Internet in 1996 -2010 is presented in chart 1 (Netcraft, 2010).
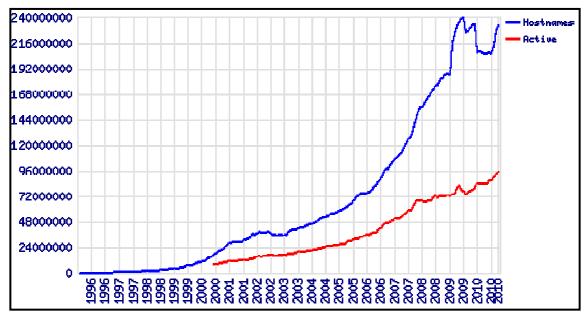

Chart 1. Number of hostnames in the Internet in 1996 -2010 (Netcraft, 2010)

It appears from Chart 1. that the abandoned Web pages number has remained about half of the active sites number since recording was started in 2000. At the same time, as the amount of information on the Internet has increased, so has the use of Internet. Chart 2 presents the prevalence of Internet usage in percents of different age groups in Finland in 2009. All of the information on the Internet has been brought to consumers by separate search engines. Finns make daily millions of searches with Google's search engine and over half of the people use search engines always when using the Internet (Paananen, 2006). The Internet's superior role as an information channel has been a foregone conclusion for a long time, but in recent years the digital marketing, made in the Internet, has gained shares when compared to traditional marketing channels.
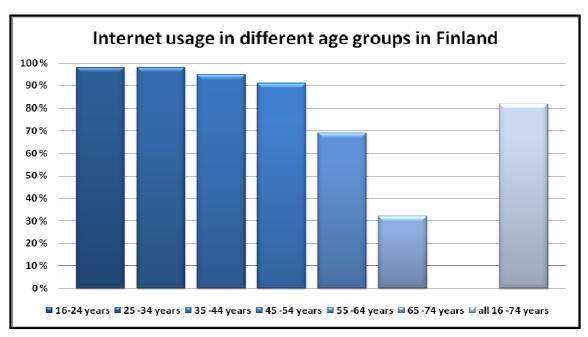
Chart 2. Internet usage in different age groups in Finland (Tilastokeskus, 2010)

In chart 3. is presented comparison research about the number of leads from marketing in different marketing channels.

"On Demand Marketing: As compared to interruptive nature of advertisement on TV and other medias, where ad is shown to even those people who don't want to see, Internet provides On-Demand Information. People search on net for what they want and when they want it. There is low cost and high return on investments. It is one hundred times cheaper than conventional marketing tools like TV or newspaper advertisements. Just compare the cost of printing a full newspaper Ad for one Day vs. the cost of e-mailing to 25,000 e-mail addresses. Internet is clearly the winner. (Wisdomtalks, 2010.)

	Online Marketing	TV	Radio	Newspaper & Magazine
Amount Invested($)	10,000	50,000	10,000	20,000
Leads(Avg.)	700 - 800	500 - 600	200 - 300	300 - 400

Chart 3. Comparative research of leads from advertisement in different Medias. (Wisdomtalks, 2010.)

In the United Kingdom, the Internet Advertising Bureau (IAB) is reporting that online advertising has overtaken television. According to Reuters, advertising

spent on the Internet grew 4.6 % to a total of 23.5 % of the total market(Wisdomtalks, 2010).
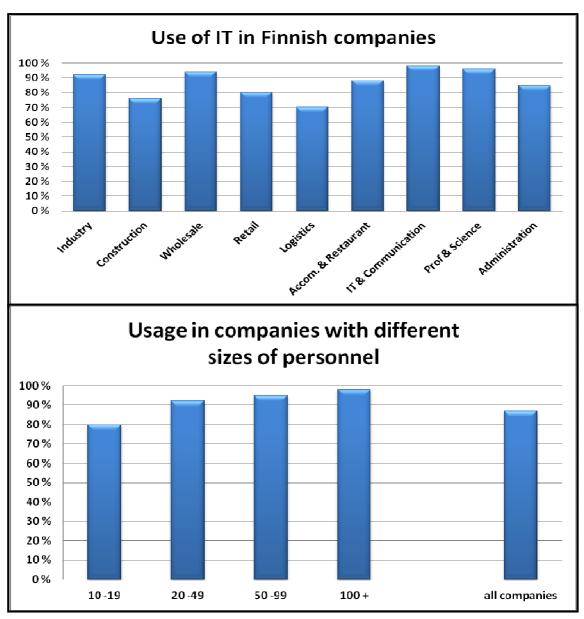

Chart 4. Use of Information Technology in Finnish companies in 2010 (Tilastokeskus, 2010)

Finnish companies have also discovered the importance of Internet visibility for business. The number of homepages has been growing steadily and the Finnish Statistics Office report that 87 percent (Chart 4) of all companies in Finland have their own Website.(Tilastokeskus, 2010.)

According to Finnish Statistics Office, 87 % of all companies employing at least 10 persons had their own Web pages in spring 2010 (Chart 4). The prevalence

of Web pages is increasing the bigger the company is. Of larger companies (+100 employees), 98 % have Web pages but in smaller companies (10 -19 employees) Web pages can only be found from 81 % of the companies.

When examined by the industry it is noticeable that Web pages were most frequently used in information and communications sector(Chart 4). The next highest industries were the professional and scientific activities, wholesale and industrial companies with more than 90 per cent of companies having Web pages. Most sparsely Web pages were used in logistics (71 %) and construction (76 %) industries. The online shopping possibility in their homepages was in 15 % of companies (Chart 5). Value of online shopping dropped from the year 2009, but the share maintained the same 6.1 %. Most of online shopping is business to business shopping, and in 2010 81 % was in this category. (Tilastokeskus, 2010.)

Home pages were commonly used for product marketing. In companies with less than 10 employees, 87 % used them for this purpose (Chart 5). Product catalogues or price lists were found from 38 % of all Web pages and anonline shopping possibility was offered in 17 % of the companies sites.
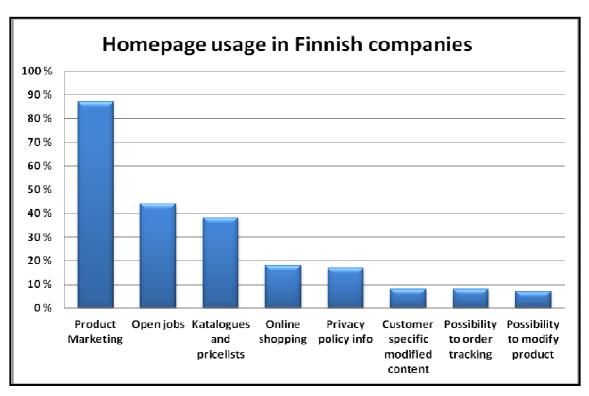

Chart 5. Homepage use in Finnish companies in 2010 (Tilastokeskus, 2010)

Marketing investments used in digital marketing have grown steadily over the last few years. Average company's digital marketing budget will be almost 30 % of the total marketing budget in 2012, says Aalto University's research (Chart 6).
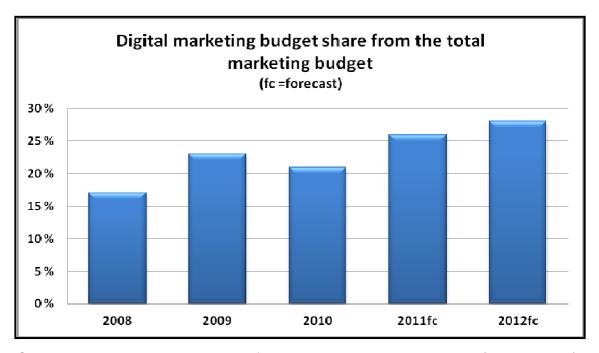

Chart 6. Digital marketing budget from the total marketing budget (DIVIA, 2011)

The importance of search engines

Search engines are Websites that allow users to search for information by using different types of searching keywords. Keywords help the search engine to create a list of sites that contain information related to searched words. There is a high capacity demanding and complex system behind this simple action. Search engines can be divided broadly into three essential elements; data collection, database, and user queries processing.

There are thousands of different search engines in the world. The most well known of these is Google, even though it is not as popular everywhere as it is in Finland. The most popular search engine in Russia is Yandex, and Baidu is used in China. Google in considered to be fast, accurate and informative as well as the most comprehensive search engine on the Internet. Google's advantage in Finland is also its ability to apply Finnish Web pages exceptionally well.

Google offers also separate images, videos, discussion groups, map and scientific knowledge specialized searches. (POHTO, 2009.)

Approximately 400 million Web searches are performed every day and that means that every single second there are more than 4500 searches (Enge et al. 2009). This requires automated and super efficient systems from leading search engine companies.

Search engines are based on an automated system that indexes and organizes separate Websites in the Internet. Search result produced by search engine is a list of the suitable results matching the keywords. A single Web page's rank is based on the detailed search terms' frequency, proximity, phrases, etc. matching the searched words. The technology used and the quality of links are affecting the result alongside content-related factors.

Today, 88 % of the people start product or service searches from the Internet. It is worth noting that the power in services and products marketing has shifted from companies to consumers. "It's not what you say for your company anymore. It's what they say about you." (Seppänen, 2011.)

4 INTERNET MARKETING

Internet marketing refers to all the marketing tasks made in the Internet. Originally the term included only banner advertising, but based on Mäkelä it also consists of e-mail advertising, search advertising, search engine optimization, the use of electronic directories, blog marketing and social media usage as a message channel.

The roots of Internet marketing are in the early 90s, when the activity was mainly limited to text-based product lists and product presentations. Later development went through graphical ads to customer service, sales and stakeholder management (Paananen, 2006). The Internet has become a viable place to trade or do business. It has taken its place among traditional media marketing channels.

Internet marketing and its sections are presented in Chart 7. Search engine optimization (SEO) and Keyword advertising (KWA) are therefore part of the search engine marketing (SEM). The ultimate goal is the same in both of measures, to bring visibility in search engines. In SEO it is done by raising the classification and bringing the site up in so-called organic search results. In KWA the visibility is applied through paid and sponsored links. Usually paid results appear separately from the real search results and the ranking is based on offers given to separate search words/ phrases. Both activities are supporting each other for example by including a keyword-analysis. SEO is also linked to paid directories. Paid directories are link collections, into which companies can add themselves to get more visibility for their contact information in the Internet. These directories are often treated seriously and they are well valued in the business world. Therefore the rating of those is also high and links from directories increase Web page's rating considerably. (Paananen, 2006.)

Chart 7.Internet Marketing (Paananen, 2006)

There has been significant growth in social media, e-newsletters and community usage in recent years. The most popular forms of e-marketing in 2011 were company's own Website, search engine marketing, direct e-mail, electronic newsletters and online advertising. The share of usage increased from the previous years in all marketing types. Results are presented in Chart 8. However, POHTO's study shows a clear distribution to large and small companies. Smaller companies do not take advantage of the Internet even though previously it has been assumed digital channels to be an excellent marketing channel for smaller and more agile companies. (POHTO, 2009, Reponen, 2009.)
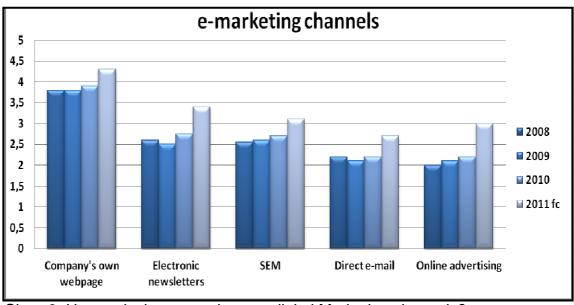

Chart 8. How actively companies use digital Marketing channels?

0= Not at all, 5= All the time, (DIVIA, 2011)

Online advertising has developed by leaps and bounds since the first banner ads were displayed. In the last 15 years online advertising has developed faster than any other marketing form, mostly because of Internet's and its users' progress. Today there are several new media types in use in addition to those that have already been mentioned, for example multimedia display advertising and streaming advertising (Google, 2011). Simultaneously, when marketing opportunities have grown, the companies have organized their digital marketing utilization. Divia's research participants (310 Finnish marketing professionals) related that in 2011, already 41 % of their digital marketing investments were budgeted (Chart 9). (Divia, 2011.)
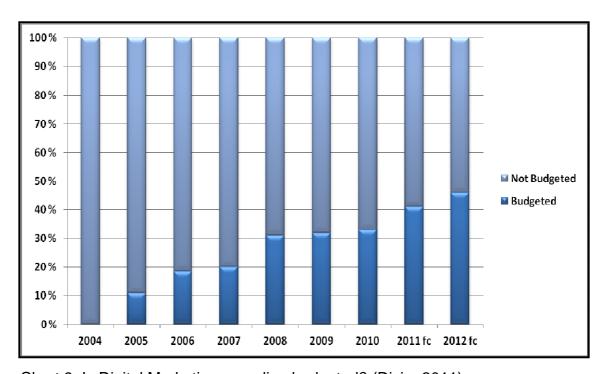

Chart 9. Is Digital Marketing spending budgeted? (Divia, 2011)

4.1 Search Engine Marketing (SEM)

Search engine marketing is divided into search engine optimization and paid keyword advertising(Chart 7). As mentioned in chapter 4, the visibility in KEA is applied through paid and sponsored links. Image 1 shows that paid results appear separately from the organic search results. The pricing in keyword advertising is based on the number of clicks the ad gets (Reponen, 2009, Seppänen 2011). SEO influences on search engines' organic results visibility

and paid KWA gives visibility in Google's advertisement sections. Organic results are also shown in Image 1.

In Finland the digital marketing industry is still quite small and activities of qualified expertise are not yet widely available. On a general level, it can be said that currently the financial stakes are higher in KWA even though the traffic is much higher on the organic side of the result (Image 4) (Liuhto, 2010). SEM is usually started with KWA, because it is faster and easier than SEO.
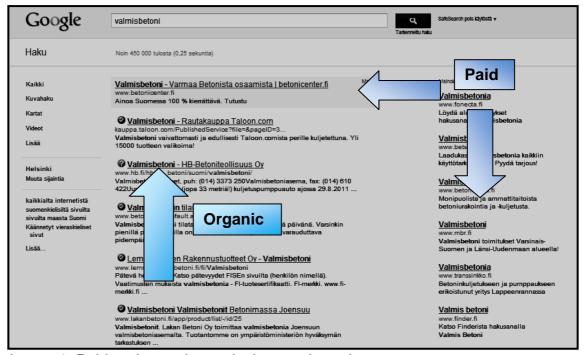

Image 1. Paid and organic results in search engine.

There are easy tools for advertisers to guide keyword advertising themselves. For example in Google's AdWords, advertisers select keywords that describe their company or products. When users search in Google by using keywords, word-related ads are displayed among the search results. Ads in Google are priced by an auction basis which takes place every time when the user types a search query. Advertisers pay only when users click their ad and the system ensures that cost is the minimum necessary to maintain their position. Advertisers can also track the results of their campaigns immediately. The Google Analytics tool provides for advertisers compound measurements of how users find their way to advertiser sites, what they do in the site, do they shop or became a member, and where do they proceed once they leave the site. This

information enables the marketers to experiment and continuously improve their campaigns. They can try different keywords, ad-text, follow the value of keywords and test the efficiency of different front-page layouts. When advertisers are familiar with customer behaviour and trends, it is possible to optimize the path from search to sales. The goal is to optimize the path so that advertisers reach their customers and meet their needs, to find new customers and to improve the return on their investment. (Google, 2011.)

4.2 Keyword Advertising (KWA)

Keyword Advertising means practically paid visibility alongside with organic results (Chart 7). It is also called Pay Per Click advertising (PPC). The Price is based on auction, for example 0,30€/click and customer pays to Google only when the ad is clicked. This type of marketing is very well targeted and cost effective reaching only the interested customers. Google ranks ads by quality points. That means the machine compares the search phrase and the equivalence of the site and presents the results matching the search most relevantly.   To prevent abuses and for competitive reasons, Google is not giving their indexing publicly, but the assumed structure is handled more detailed in the following chapters.

This is how Google is presenting its services to advertisers:

The AdWords system works best for everybody; advertisers, users, publishers and Google too when the ads we display match our users' needs as closely as possible. We call this idea 'relevance'." "We measure relevance in a simple way: Typically, the higher an ad's Quality Score, the more relevant it is for the keywords to which it is tied. When your ads are highly relevant, they tend to earn more clicks, move higher in Ad rank and bring you the most success.(Google, 2011.)

4.3 Marketing practices

The keyword advertising process begins with key words and phrases definition. Key words and definitions are intended to help searchers to find the site. The determination should be done carefully and it is good to remember that if the keywords are used in advertising campaigns, those words should also be found

from the campaign site. Keywords definition process is also made in search engine optimization process and it is dealt with in more detail in chapter 5.4.1. Next step is to determine the content in sponsored link, to plan the measurement practices and to set the daily budget for advertisements. The value of keyword ad is defined based on supply and demand. The more there are auction participants for a word/phrase, the higher the price is. The price is presented as the highest price the advertiser is willing to pay for a click.

The rank of an ad consists of auction price and quality points. In other words the higher the score in the target site the less is needed to pay from ad's visibility. On the other hand, if your site doesn't have good quality points the ad doesn't show no matter how much you pay.

Here are few general conclusions:

- The price of one click is usually about 0.10-0.90 €, depending on the industry in which the advertising is carried out.

- The best visibility related to investment is to get big clicking traffic with a small click price and high daily budget.

- When the clicking traffic is small, it is important to have a click price that wins the auction.

- If the ad is not clicked five times in the last thousand times when it is presented, it is completely eliminated from the circulation. (Seppänen, 2011.)

5 SEARCH ENGINE OPTIMIZATION

Search engine optimization means modifying a Website to a format so that the search engines classify the site as well as possible. Ranking is a value into which the search engine bases its result display order. The goal is to get your site to be located at the top of the organic search results in all site-relevant keywords. In practice, SEO is trying to make Website search engines friendly by editing the HTML mark-up and text content to match better the user searches. The aim is to modify the site structure and content so that search engines appreciate your site more than other sites with same keywords and search phrases. All search engines have their own algorithms into which the engine bases its results. (Reponen, 2009.)

Why is search engine optimization done?

- Effectiveness of SEO is easy to measure. It is inexpensive and its effects are long-term because the SEO makes the entire Website systematically search engine friendly.

- In customers' eyes, you exist only if you can be found from Google. Google's list is a sort of "ranking list" in viewers' eyes.

- You are there for the customer always at the right time

- SEO pursues new customers, increased sales, and maximization of contacts.

- The aim can also be dissemination of information, creating a stronger brand and revenue of advertising growth (in cases when site gets revenues from advertisers!). (Seppänen, 2011.)

5.1 General information about the SEO

SEO has a crucial role in Internet Marketing. SEO is next important after Website functionality and usability. It has a direct link to KWA and its cost effectiveness. Effective SEO means search engine friendly Website technology, quality of content and accuracy of linked pages. Of course, there also needs to be sufficient high quality traffic.

The aspects mentioned above are all part of Google's indexing factors. Factors determine search results and there are more than 200 of those. One of the most important site quality measuring factors is Google PageRank. PageRank is Google's patented algorithm, which analyzes more than 500 million parameters out of 2 billion mathematical terms.. In practice PageRank means sites importance in search engine's view. PageRank grows from every link that leads to the site, especially if the linking site has high PageRank. It is also affected by the content, structure of the site and site's internal linking. PageRank's figure is between 0-10 and it gives a good picture of the Website's quality. Finnish Websites' PageRanks are normally 1-4, but can also be higher like Kela's sites, 8. (Mäkelä, 2008.)

Image 2. PageRank, The bigger the ball the better the rank,
(Wikipedia, 2011)

This value shouldn't however be stared at aimlessly, because Google updates it only a few times a year and it is just one of the about 200 factors that defines the search result. In Chart 10 is SEOMOZ's researches estimation of the sections affecting to Google ranking.
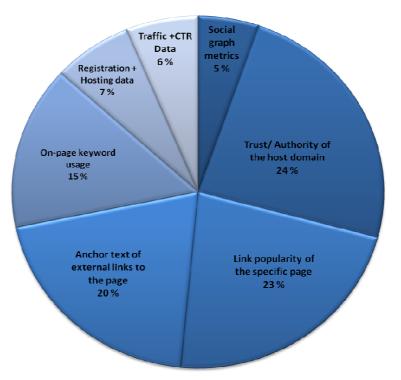
Chart 10. Component's of Google's ranking algorithm Seomoz, 2009)

Here are results of leading SEO professionals' determination about most important ranking factors:

- Keyword use in a title tag
- Anchor text of inbound link
- Global link authority of the site
- Age of the site
- Link popularity within the site's internal link structure
- Topical relevance of inbound links
- Link popularity of the site in topical community
- Keyword use in topical community
- Keyword use in body text
- Global link popularity of the sites that link to the site

 (Enge et.al. 2009.)

SEO is often based more on empirical than theoretical research because the search engines update their algorithm several times during the week and do not tell about changes or algorithms' direct structure to the public. The change and update density of the Google algorithms has been estimated to be about 10

times a week (Reponen, 2009). Although the algorithm changes happen at a frequent pace , the basis of SEO work is still the same, and by going through all basic phases of SEO, search engines will respect the site better in their classification. No matter how the algorithm changed during the last week.

In Finland search engine optimization can also be perceived as Google optimization due Google's dominant market position. However, SEO has a positive impact on all search engines even though it would have primarily been targeted just for Google. For example the Microsoft owned search engine, Bing, appreciates many of the same things as Google. (Mäkelä, 2011.)

History of SEO

SEO is considered to be invented in 1997, when search engine specialists Danny Sullivan and Bruce Clay started to produce public reports on the subject. The first publications dealt with search engine algorithms and how different search engines organized their search results. Marketers and Website administrators were inspired to think how they might increase their ranking in results after reading those reports.

In 1998, Google was founded and caused a change in search engine marketing field. Google's founders were American Stanford University students Larry Page and Sergey Brin. They got an idea of a search engine, which is still Google's mission "organizing the world's information and making it universally accessible and useful." Google's early years popularity secret was in its unique algorithm. Page and Brin had experience of academic research environments and they believed that to be the key to search engines as well. They thought that the best solution would be to index the entire network to database and to analyze who is linking to whom. The idea came from the practice of the academic world, where scientists will increase their own researches' confidence by quoting parts from earlier researches. Google's algorithm is still based on Page's and Brin's basic idea. A Website into which is linked most links from other high definition pages is a relevant page itself. This algorithm is called PageRank which was handled

more precisely in last chapter. Today, the algorithm Google uses is much more complicated than the early years' version. (Reponen, 2009.)

5.2 General information about search engine

Search engines are dependent on the information retrieval science concept that was developed in the 1950s. When developing first search systems scientists discovered two elements, relevance and popularity, enabling a large part of data retrieval functionality. The relevance of document increases simultaneously when searchers chosen keywords amount increases in documents text content. The more a specific term or phrase is mentioned in a document, the more important it is. The significance grows if the searched key term is an important part of the structure of the document like the main or sub title. Document's popularity on the other hand reflects the relative importance. The popularity is measured by the amount of citations from the original document. Today search engines have applied the concept of relevance and popularity to the new format, to document-analysis and link-analysis. In document-analysis the site is considered through the information it contains and term relevancy is interpreted through their locations. Most important locations from search engines point of view are the title, meta-data, tags and body text. Link-analysis is measuring the amount of primarily links from external sites, but also the relevance of the linking sites. Other factors are age of the linking site and the linking pages reliability. (Reponen, 2009.)

Such search engines consist of three parts, the search robot, site database and results displaying application. Search engines perform several processes before results are presented to the user. Search robot's task is to browse through the pages in the Internet and add them to a database or index base. Search robots behave very much like a standard Web browser, but the robot sees the page only like a text browser. Search robot asks for Web pages from the Web server, downloads the page and saves it to the directory. In the third part, the results displaying application looks for corresponding sites from the index and displays the results in order of classification. From SEO point of view the most interesting parts are robots and result displayer. (Paananen, 2006.)

Google's search robot, Googlebot, follows Internets hyperlinks when moving to another site. Instead of page appearance, the Googlebot examines only the structure of the site, content, and external factors such as linking. New sites administrator can add network address to Google's service in order to make sure that page can be found from the site directory. Another possibility to get a page to Google's directory is to wait that Googlebot finds its way to the page through external link. Most of the pages in directory have ended there when Googlebot has found them through external links. Googlebot will save the copy of the sites content to the directory. When the directory includes a copy of all Web pages Googlebot has found, it is instantaneous to present the results after search. When user performs a search the query process looks through the pages that include searched words and organizes search results to hierarchy. (Reponen, 2009.)

Google

Google's mission is to organize the world's information and make it universally accessible and useful. Beginning in 1996, Stanford University graduate students Larry Page and Sergey Brin built a search engine called "BackRub" that used links to determine the importance of individual Web pages. By 1998 they had formalized their work, creating the company you know today as Google. Since then, Google has grown by leaps and bounds. From offering search in a single language it offers today dozens of products and services, including various forms of advertising and Web applications for all kinds of tasks, in scores of languages. Starting from two computer science students in a university dorm room, they now have thousands of employees and offices around the world. (Google, 2011.) Google was named after word "googol", which means mathematical term for a number that starts with 1 and followed by 100 times 0. The name reflects the enormous amount of existing information and Google's mission to organize all that information.

Daily, about 72 % of Finnish Internet users use Google search engine. Another fact that tells about Google's popularity in Finland is that 77 % of Google users

are using Google searches before important purchasing decisions. (Reponen, 2009.)

5.3 Consumer behaviour in the Internet

According to Paananen 2006, Internet is in crucial role in people's consumption decision- making process, and more than 90 percent of companies use Internet when making business to business purchasing decisions. Internet and search engines have therefore became an important factor in the purchase decision-making process. It is also an important factor on consumer behaviour.

Comparing to traditional advertising, instructions, and salespersons memory and professional skills, the Internet provides unlimited data around the clock. What is exceptional in the Internet is its interactivity, which needs to be taken into account when making research from the consumer behaviour point of view. Search situations can be roughly divided into three main categories: Navigation searches (10 %), information search (80 %) and trade-oriented searches (10 %). (Tulos.fi, 2011.)

Navigation searches means searches which aim to find a particular site. Such a site can be for example a company's site or a certain site where the searcher has already visited. SEM doesn't have much influence on these searchers because they know exactly what they want. Search engine marketers however need to take this group into account and to make sure that their sites can be discovered from the main search engines. Even though there is not much potential in navigation searches, they need to be taken account due to their large share.

Information seekers are searching for information about a particular topic. The difference to navigation is the fact that information seekers believe there is information about the topic, but they don't know where. Searchers often start with very general terms and then refine the terms to cut the irrelevant results. Information searches are most important searches from SEM point of view. In this phase the customer has normally identified the need and is looking for a

solution to fulfil it. By providing clear information-rich sites with proper keywords and optimization, these searchers can be contacted.

In the last group there are searchers who are no longer searching for information, but want to buy the product, download the file, join the mailing list or do something else comparable. From the SEM point of view they are tough customers who know what they want and use search phrases with exact product information. The problem areas are in descriptions and search engine results. By creating good descriptions for the product-lists and by bringing special offers and other sales increasing aspects to search engines will help to get better results. (Paananen, 2006.)

Internet browser software vendors attempt to combine the address bar and the search box. This development will run the users to a new habit. This is a sufficient action to write the brand or product name to the address bar. Then Google will show the desired page in their result page. This helps the users when they don't have to remember entire addresses. On the other hand this development makes the search engine position even more powerful. (Reponen, 2009.)

Only a minor group of searchers look for results from other pages than the first one (Chart 11). Only the three first positions get everyone's attention. So, SEO is actually visibility management. (Eisenberg, 2008.)

An eye tracking specialized Research Company, Thinkeyetracking, has done research about Google users eye tracking habits in 2005 and in 2008. The results are presented in Image 3 and they show that in 2005 persons tracked the whole first page before they decided to click the link. Within the next three years, either Google had become much better at showing what the viewer wanted to see or using Google had become a habit since the user found that results are relevant and right answers were found from the first results. The study participants were asked what would they do if the required information can't be found from the first page. Out of the respondents 87 % told they would modify the search term. In Chitika's research from 2010, this figure seems to be

already 98 % (Chart 11). This is clear evidence that only first page matters, and actually it is crucial to be in the top three. (Eisenberg, 2008.)
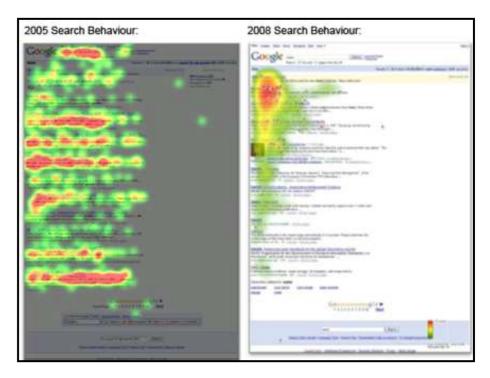

Image 3. Eye tracking comparison from consumers behaviour in 2005 and 2008
   (Eisenberg, 2008)

In chart 11 is presented Chitika Network's research about traffic from different search result positions. For traffic that goes through Google, the 4th position gets only 7,7 % of entries. That is almost 5 times less than traffic in position 1.
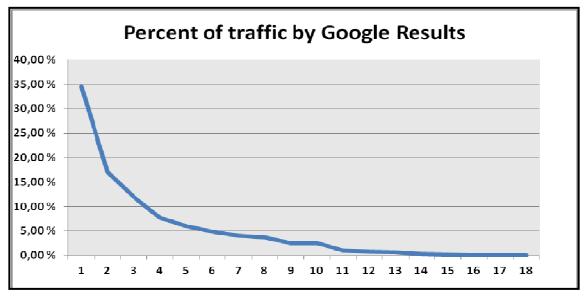

Chart 11. Percent of traffic by Google result (Based on a sample of 8,253,240 impressions across the Chitika network)(Ruby, 2010)

Consumer attitudes towards advertisement on search engines have also been researched. Average numbers of advertisement performance are as follows: 90 % of the clicks are made to organic results and 10 % to paid ads (Image 4). Also, an eye tracking observation map reveals the low proportion of the search results (Image 5).
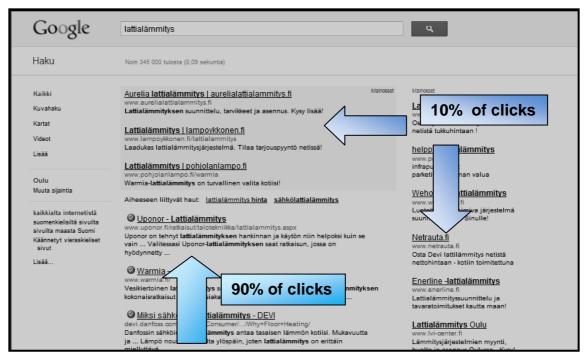
Image 4. Ad clicking popularity vs. organic search results (Seppänen, 2011)
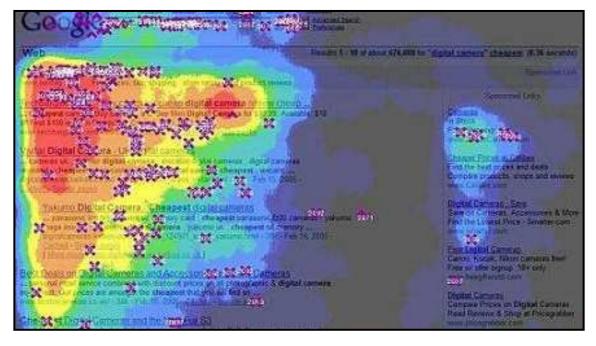
Image 5. Eye scanning, ad vs. organic results (Waraas, 2008)

The first positions in search results are collecting the most clicks and for that reason the cost for those positions is higher. The price determination for paid results was handled in Chapter 4.3.

Today, Google presents also other search related media among organic text format results. This media includes news, images, maps, blogs, videos, etc. Clicking percents for these are higher than paid advertising clicking percents. In Image 6 is presented that 31 % of the viewers are clicking images with 17 % for videos and 36 % for other expanded search results. Google owns YouTube and therefore it is assumed that it favours the search results that match with YouTube videos. (Seppänen, 2011.)
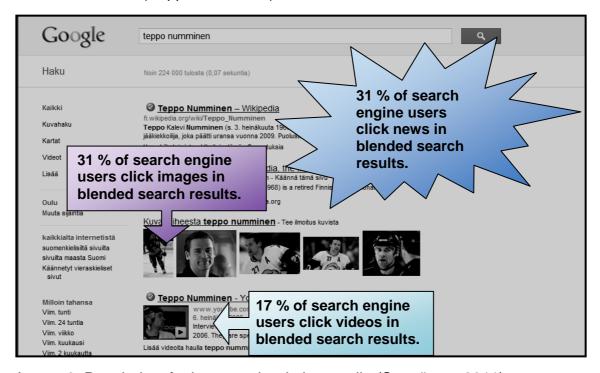
Image 6. Popularity of other search relating media (Seppänen, 2011)

It is important to monitor visitor behaviour on your site. Search engine users leave valuable information about themselves when searching or visiting Websites. To choose the visitor tracking system and to set its objectives is the prerequisite to reach results with Internet marketing. Intended Web analytics tools are available for a variety of purposes, but mainly in Finland the ones needed are the free Google services or some more precise Web analytic tools like Snoobi. Google's Web master tools are used in definition and development of site. Tools are primarily used in order to affect on search queries placement

and search robot rights. Tools show which individual pages Google has indexed, which search terms have brought users to pages and with which keywords Google is ranking the page. Most advanced visitor tracking software have the ability to detect the visitors IP-address and by analyzing that is able to find out decision makers from the visitor's organization. Snoobi is a Finnish visitor tracking software with Web analytic tools and by using it the company can measure the real advantage of its Website usage and other Internet activities. Visitor tracking services allow marketers to choose the right promotional tools as well as to analyze and optimize existing Web pages. Rudus' pages visitor tracking parts in this work are done with Snoobi.

In practice, JavaScrip code is downloaded to the followed Web page. It connects analytic software and downloads tracking information about visitors. An increase in visitors doesn't necessarily mean a SEO process. The goal should always be to increase the sales and attract only potential customers to the site. With analytic software it is possible to track visit duration, entry pathways and use of search keywords among visitor numbers. When tracking e-commercial activities it is possible to find the point where users most likely suspend the purchasing process. This helps to improve sites so that the purchasing process would become finalized more often. One interesting statistic from analytic software is a bounce rate. The Bounce rate tells the number of visitors leaving the page immediately after entering the site. Normally this tells that the visitor notices immediately that this is not the page they were looking for. This is a very important figure and it tells immediately if the site is not optimized with relevant search terms. (Reponen, 2009.)

It is important to remember that it is not enough to just have a visitor tracking system. The obtained information needs to be interpreted correctly and actions should be taken on the site based on visitor behaviour. This is very important especially in sites whose revenue model is based on advertisement or e-commerce.

5.4 Optimization process

The SEO optimization process, based on Paananen, is presented in Chart 12. In practice this means the process by which the site is edited so that the classification of the page is as high as possible and it can be found with desired keywords.

The process starts from a baseline survey where the current situation and improvement demanding areas will be clarified. Keyword analysis is made after baseline survey. The goal is to find out the words that match with most used products or services. After analysis comes the site optimization part which is divided into technical and content parts. In the technical part, possible visibility problems due to server settings are fixed. Also other nonstandard methods of implementation and other issues that interfere with search robots like flash-players, videos and photos are fixed. In the content part, sites are modified to match search engines and user needs concerning the appearance, keyword frequency, how well keywords can be found from the URL, page title, headings and other important HTML-elements and server response times (Image 10). Mainly in this part the focus is on text content modification to match keywords. When the site itself is in order it is time to perform link analysis. In link analysis the site is linked to other industry-related directories. Directories can either be free or paid directories and there can also be public directories. After linking, the site is ready for indexing. It needs to be added for search engines indexing list and tracking set up. Because of the rapidly changing environment in the Internet the optimization process is an iterative process, and the work needs to be started again from the beginning to get the site even more optimized. (Paananen, 2006.)

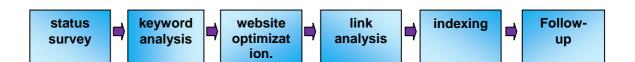| status survey	→	keyword analysis	→	website optimization.	→	link analysis	→	indexing	→	Follow-up

Chart 12. Optimization process (Paananen, 2006)

According to Tulos Helsinki, the importance of optimization work is divided as follows: Background work meaning keyword-definition, competitors sites scouting, own sites current state analysis and Google tools use is ca. 50 % (Image 7). Body text and linking is ca. 25 % and the last quarter goes for Meta description and headings.
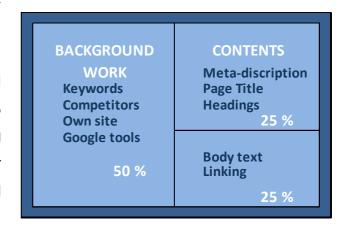

Image 7. Optimization process inputs distribution (Tulos Helsinki, 2011)

There are several ways to start the process, but one good way to start is by checking that the robots are visiting the site and finding all the pages. The inspection can be performed by searching for search robot traces from the server log files or by searching pages directly with search engines. If the robot has not visited the page it can be reported to search engines and the robot will inspect the site in few days. Most sites will be indexed automatically and if your site is not indexed it is worth checking so into the cause. There can be obstacles for robots like non-standard methods in implementation or something in the content that the robot can't read.

When the site is indexed for the first time the actual optimization process can start.

Keyword analysis

According to Ben Norman, " The most important part of optimizing your Website is the selection and implementation of your keywords. If you don't get this part of the optimization process right then everything else you do will be in vain. Target the right keywords and your Website will be a success; target the wrong keywords and you will miss out the valuable traffic you are trying to canvas". (Norman, 2009.)

Optimization work should start with keyword analysis. Analysis should find out the keywords in company's business. The keywords are the ones which would be important to rank high in search results and which customers are actually using when performing searches. When analyzing, it is good to remember the facts from Chart 14; only one word is rarely enough, and 2-4 is better. It is also good to notice that misspelled searches may be good opportunities to guide the searcher to your own site cheaply. The source for analysis can be for example company's product groups and service descriptions services that keep statistics of keywords in the Internet. Generally the company's personnel are the most qualified group to list keywords, but it also is important to get proposals from potential customer as well. External persons often can see the subject from a broader perspective, therefore their keywords are somewhat different from the company's own.

There are several methods to find out the right keywords. Those can be a result of traditional brainstorming session or produced by premium search tools like WordTracker that tells what terms are most prevalent in your business field. Keyword analyzing is often an underestimated process and it is often gone through too quickly. People think that we know what words searchers are using or that it is enough if the site can be found with the product name. Normally in these cases the customer already knows what he is looking for. (Paananen, 2006.)

Keywords can be roughly divided into head- and long tail –terms. For example, "truck", is a head-term and "used yellow concrete truck" is a long-tail term (Image 8). Short, one-word terms collect most of the traffic through search engines, but the competition is harder and the customer who arrives at the site with a short term doesn't necessarily have the same purchasing potential as the one using longer and more precise terms. Searchers who used the search word "truck" would most likely see lot of relevant results in the result page, but if he has a clear target to purchase a used concrete truck, he would most likely use more precise search terms. When the right words are found, those are added to content, link- texts, and headlines.

As Chart 13 shows, the long term searches were increasing their share strongly in 2007-2009 but it seems that the development has slowed down or stopped in 2010-2011. Nevertheless, it is not enough to optimize the main terms only. The site's text content should support the increasing number of search words, because potential customers can discover the sites with surprising combinations of keywords.

"People will find you through the oddest combination of keywords. Most of these words seem nonsensical till you understand that people looking for your products and services do not always know the correct terminology or industry jargon to accurately describe what they need". (Kate Russell, 2010.)

There is much less competition in long-tail searches than short ones (Image 8). That makes long terms cheaper and the customers entering the site with long terms have much higher purchasing potential than customers entering with the short ones. Long terms should be considered especially when product keywords are thought. Image 8 shows how the short and long term affects the competition situation and costs.
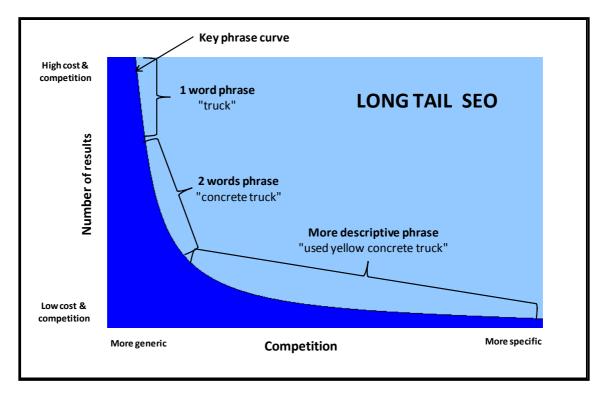
Image 8. Long tail SEO.
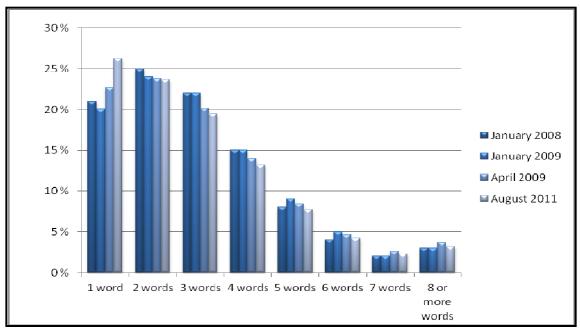
Chart 13. U.S. keyword breakdown by percentage of clicks. Data is based on all search engines on a four week rolling period from the Experian Hitwise sample of 10 million US Internet users. (Experian Hitwise, 2011)

Conversion rate tells the ratio of visitors reaching the goal that is set for the site. It can be calculated as shown in formula 1.0. The goal can be for example, e-commerce purchase action or contact. Chart 14 tells what is the average conversion rate by keyword phrase length. The Chart shows that the best conversion is obtained with three and four word phrases.

$$\text{Conversion rate} = \frac{\text{Number of Goal Achievements}}{\text{Visits}}$$

(1.0)
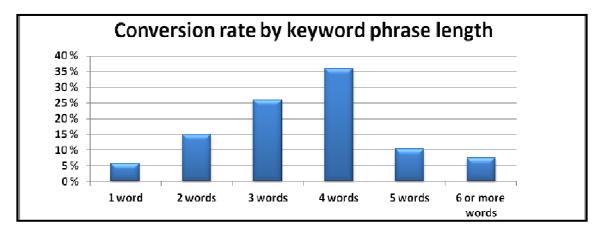
Chart 14. Conversion Rate by keyword phrase length (Baggott, 2010)

41

It can be challenging to define keywords for long tail-searches. However, successful SEO makes the entire Website systematically search engine friendly, and even the unexpected search phrases might still give good positions. It is all up to relevant information matching the keywords. Successful long-tail keywords direct the user to the right site. It is good to remember the potential behind long tails. They are cheaper than short ones (Image 8) and the conversion rate is much better (Chart 14), but still they are not that much less used among searchers (Chart 13).

5.5 Structural changes

In practice, structural changes means changes to issues that the crawler robot reads, but a normal user would not take into account. Changes are made to the platform of the Website, headlines, meta data and site structures. The above-mentioned facts have to be in order and correspond to the content in order to get a good indexing result. The robot interprets the site according to above-mentioned issues. Even if the content would match with the search, the robot would interpret it incorrectly and leave the page away from search results. This is exactly the problem in Rudus' case, and is why this project was started in the first place.

The structure of a Website is described with the HTML mark-up language and search engines respect keywords in some parts of the site structure more than in others. A Website should always match the standard. Starting point can be that it needs at least to follow the XHTML 1.0 specification, which enables platform independence for the used Web browser. The network address (domain address) should include a keyword or several keywords if it is possible. Also the age and suffix of Web address should be taken into account. The longer the domain has existed, the more search engines appreciate it. Search engines and users prefer a local suffix. In Finland that is the ".fi" suffix. (Reponen, 2009.)

Network addresses, as well as file names should contain keywords. For example the imagined concrete e-commerce site's "valupörssi's" address http://www.valuporssi/sivu1.html does not offer any significant keywords for the search engine. Moreover searchers won't get any tips about the site's content from the word "sivu1". In general, the use of hyphens has been found to be the way to separate words from each other in addresses. Based on Reponen's study in 2009, this is a sample address http://www.valuporssi.fi/lattiabetonit/. Many content management systems will shape addresses to search engine-friendly mode. For example to a shape like: http://www.valuporssi.fi/tuotteet/betonit/rudus/, when the keywords are "betonit" and "rudus". On the other hand, the word "tuotteet" could be left out completely, even if it would be one of the sub-pages in site's structure. Often the address is not search engine optimized if there is no content management system in use. A poorly implemented address is for example the following address: http://www.valuporssi.fi/index.php?sivu=tuotteet&alasivu=betonit&yritys=rudus, The example address is too long, uncertain, and includes unnecessary parameters from search engine and user points of view.

Search engine friendly address is :

www.valuporssi.fi/betonit/rudus/

| **one-peace domain** | **shallow directory structure with relevant words** | **keyword in page- name** |

Non search engine friendly could for example be:
 Xyz3.valuporssi.fi/a/?q=31z&mg=vv7z&sd=8020=cC8
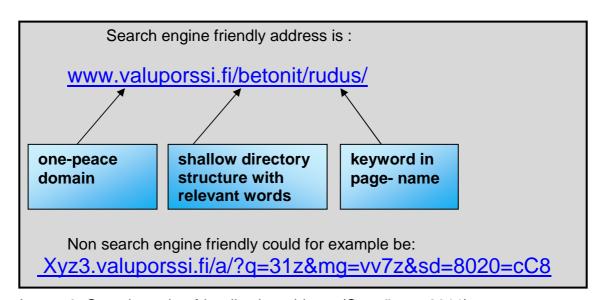
Image 9. Search engine friendly site address (Seppänen, 2011)

5.6   Text content

Text content means site's content that visitors can read and see. After well made technical optimization, the old content is easy to edit to optimized form.

Content optimization refers to site's text content production or editing so that it contains keywords and phrases that came up in keyword analysis.

In search engine optimization the text content is the most important factor. Content includes the headings, title tag, body copy, image tags and links. Web page should contain most important keywords as naturally as possible. Content should primarily offer information that is interesting for the visitor. In recent years the blogs have become common way to generate lots of fresh content to site regularly. (Reponen, 2009.)

According to Paananen, studies show that people do not see information in Internet as relevant as in some other sources. Text is read more by eyeing. Reading of printed text is also 30 % faster than reading from the screen. By adding keywords to the content, the text often becomes more compact and interesting issues are easier to find.

5.6.1 Head Lines

The differences in search results format are mainly in lengths of title-element and meta-description (Image 10). Title shouldn't be more than 69 characters, because search engines do not show any longer titles. Meta name "description" on the other hand can be 156 characters long. Meta keywords (that are not displayed in search results) significance in optimization sense doesn't make any difference anymore. Nevertheless it is advisable to enter a couple of important keywords also there.

The single most important variable in optimization is the title of document. It shows in top of browser and in search engine result pages (Images 10 and 11). The title is described with the HTML document element <title>, which is located in the <head> section(Image 10). Title should describe the contents of the page and it should contain keywords. (Reponen, 2009, Mäkelä, 2011.)

Description tag is a brief description on the site and it shows as a search result description below the result link (Image 10). It is important to consider carefully

the content of the description text. Attractive content gets users to click it more often and even if it doesn't raise search position in optimization sense, the good description might attract more clicks. (Seppänen, 2011.)
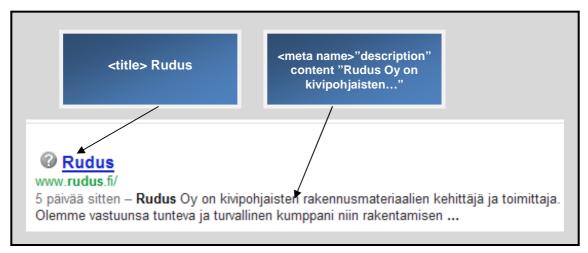

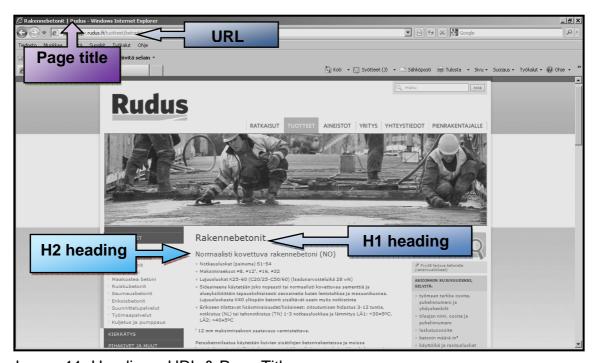Image 10. Title tag & description

Image 11. Headings, URL & Page Title

Page should always contain top head-element, <H1>, and if necessary sub headings <H2> - <H6>. Headings affect search results when the searched word is presented in heading. Top headings should coincide with the structure content of the page and the irrelevant words shouldn't be used in headings. From search engine point of view it is most evident to present only one <H1> heading. It is also logical that this heading is the top heading for the page's text

content. The page can contain several subheadings from same level. In Image 11 is presented top heading <H1> "Rakennebetonit" and sub heading <H2> "Normaalisti kovettuva rakennebetoni (NO).

5.6.2 Body copy

When using HTML, the body copy is written inside the paragraph element (p-element). Text content should be written grammatically correct and it should be natural and addressed for the readers. Body text should contain ca. 200-300 words/page. The key phrase should be repeated at least 2-3 times. The keyword should be reasonably presented throughout the text and if possible with different inflections. However, the text should stay natural and keywords shouldn't be too over-emphasized. Keywords presented especially in the beginning of the text are useful (Seppänen, 2011). Keywords can be emphasized or in bold print in order to point them out for users. Search engines are also taking this into account (Reponen, 2009). It is almost impossible to be ranked to a high position in search results if the key phrase is not mentioned once in the body of the text. In longer texts it is important to use also synonyms to keywords as all users won't necessarily use the most obvious or logical search words. Every page should have unique text content in order to avoid "double content" penalty from search engines.

The most important thing in the body is to check that execution follows standards and that elements for HTML formatting are in use. Several sites' technical implementers use CSS (Cascading Style Sheets) classes to format body copy, headings and necessary highlights. When using CSS the search engine respected HTML-heading elements are not exploited and it is not possible to reach for the best possible rating. This can be fixed by using HTML-heading elements. Attention should also be paid to other page elements. Many of them, such as image and link elements include heading and image definitions helping the usability. If the Image is also a link in the site, navigation must use the alt-attribute. Otherwise, the search robot cannot find it. (Paananen, 2006, Seppänen, 2011.)

5.6.3 Links

In addition to site itself, search engines use also external factors when ranking sites. One of the most important factors is links from external sites. In link analysis the links to site are identified, new potential linkers named and site is added to general and industry-specific directories. For external links there are both free and paid Internet directory services available. External links can also be acquired for example from the company's partners. If it is possible the link text should be a keyword rather than the company name. Different stakeholders are valuable link sources. Only a few companies will link their customers, subcontractors, etc. Websites, but normally when requesting the link has been arranged. It is possible to proceed with the industry information sites and discussion forums with the same way. Links from the pages with lot of users are very important factor for the site's visibility. It is also common to trade or buy links from high PageRank pages. Generally speaking, the better it is the more there are incoming links. On the other hand, it needs to be noted that links shouldn't come from wrong sites from the search engine point of view. These sites are link farms and sites that are focused only for a link change. In addition to external links there are also outgoing links and links to the sites own subpages. (Paananen, 2006, Reponen, 2009.)

Leonhard Euler developed the graph theory in the late 1700s. In computer science, a graph is an abstract data structure that is meant to implement the graph and hypergraph concepts from mathematics. In graph theory, things are described as nodes and routes between them creating a net as shown in image 12. (Hayes, 2000.)
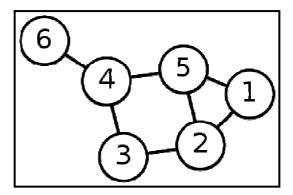
Image 12. Graph theory (Wikipedia, 2011)

Net presentation's advantage is that the network itself can be removed from the content and view only the popularity. Several natural phenomenon, social relations and Internet adhere to this theory. Internet is based on the so-called scale-free network theory, where network is starting on one node and growing without limits. In the 1900s, researches pointed out that it is the characteristic for natural networks, like the Internet, to create links unevenly between nodes. For example: many know who is the President of Finland, but only a few know a concrete truck driver from Tornio. Links exponential distribution strengthens when new nodes are created. When a baby learns to know new people, it is more than likely that the baby probably learns to know the Finnish president rather than a concrete truck driver from Tornio.

Therefore, the amount of links is growing exponentially in popular sites, and unpopular ones won't necessarily grow at all. In the end of last century Altavista, Compaq and IBM extended the theory to the Internet and developed so-called Bow-tie theory which divides Internet into the following four different page types:

- Core sites: Key sites on the Internet. Sites that are strongly externally linked and they are also linking to external sites.

- Origination sites: Sites that are linking to core sites, but core sites are not linking back to them. Normally pages with poor quality or new sites that haven't yet risen to be a core site.

- Termination sites: High quality sites that contain lots of information and services. For example the core sites directories are often linking to these sites.

- Disconnected sites: Pages and sites that are not linking nor linked to other pages. There are pages that are in purpose outside others, and sites where Internet marketing has not been done.

Above mentioned sites linking to each other are presented in Image 13. Original sites are link collections and individuals' sites which are not linked from other sites, but they provide links to other sites. Core sites are forming the centre and they are both linking and to be linked with other sites. Termination sites are large company sites and sites containing lots of information. They don't necessary need to link to the other sites to be popular. (Sherman, 2000, Paananen, 2006.)
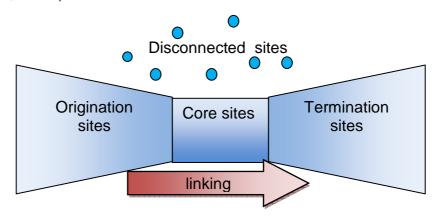

Image 13. Internet pages divided into 4 pieces based on bow-tie theory

Some of the sites are disconnected sites that are not linked to other sites. From SEO point of view it is important to realize the basics from network theory, because the amount of links is almost directly proportional to visitor numbers. When collecting the links it is important to remember that a single link from a core site is more important than several from disconnected sites.

Over time, search engine companies have realized that if ranking is based on page contents only, the system is vulnerable for abuses and search results can be manipulated. Therefore search engines are indexing sites today also through the external links and their quality. Search engines might also penalize sites if they notice irregularities. For example in 2006 Google removed some Finnish companies from its index because they had used the Finnish Webfinder.fi – service that had tried to manipulate search results. (Hakukoneoptimointi.info, 2006.)

Google was presented for the first time in publication "The Anatomy of a Large-Scale HypertextualWeb Search Engine". The significance of that publication

was in the way the external links number was utilized. The criteria for the classification were its independence from manipulation and its ability to point out the most useful sites for majority. Similar methods had been used when measuring publication dignity by counting the quotes to the publication. The current classification methods are considerably more complex than the first versions of Google and it has become very difficult to manipulate the current results. In order to maintain their independence and efficiency, the search engines don't tell their ranking algorithms to the public. Some of the most important factors affecting these algorithms have been resolved through experience. Several search engines are able to combine the content to subject, and to classify the link through content relevantly. Also the links from sites effect ranking results. If the site has lots of links to high PageRank sites, it can be respected as a high rank directory site respected among the users. The best practice have proven to be honest actions plus good quality sites. That is a key for incoming external links and moving towards the core sites group. (Paananen, 2006.)

Internet directories or link directories are directories in the World Wide Web. They are specialized in linking to other websites and categorizing those links. Some say that they are precursors of the search engines, but actually they are not search engines and do not display lists of web pages based on keywords. The directory lists sites by category and subcategory. Although search engines are handling many more sites, directories have maintained their share due to high-quality links and easy usability. It is notable that directory users are divided, just like search engine users. Some of the directories have focused only on specific industries or countries while others are general directories. There are both commercial and free directories. Free and ethically acting directories are useful especially for new sites, even if they won't bring much new traffic to sites. Directories are highly valued in search engine indexes and links from them increases also the target site's ranking. (Wikipedia, 2011, Paananen, 2006.)

It is difficult to determine how much ranking increases from linking, but indicative results can be found by clarifying the following aspects:

- Are the links direct? If the links go through a commercial server or have been carried out with script language, the classification is not growing.

- Is the directory's site handling your topics and are you able to decide the ad text?

- How many links are there in the site? The more there are links the less the classification rises in PageRank.

- Who is linking to the directory and what is the classification of the directory.

When starting the linking process or choosing partners for SEO it is good to remember that experience is especially important when collecting links. After several SEO processes, the optimizer has accumulated an extensive list of various potential linkers. He probably knows the public directories and knows which paid services are worth the money.
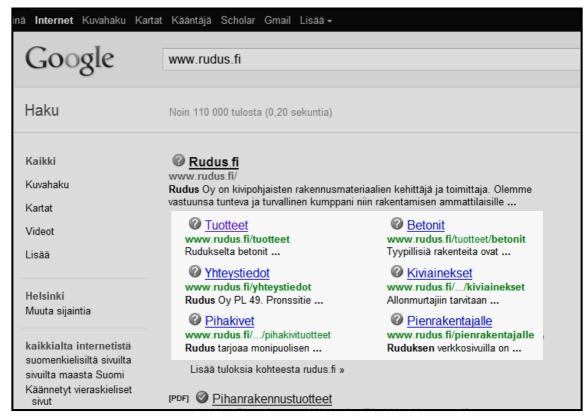
Image 14. Site links in www.rudus.fi

The idea behind internal links is to assist site visitors when moving in site and to keep them longer in sites. Keywords should be taken into consideration when Internal link anchor texts are chosen. A good way to start is to put all links in a separate section in the footer, for example at the bottom of a page.

Google analyzes the link structure of the site and displays the titles as a site links list, as presented in Image 14. Site links are displayed with search results only if they are expected to be useful for users. Google displays up to 12 site links per site. The links are not displayed if the site structure doesn't allow indexing algorithms to find good internal links or they don't match relevantly to the search. The site's Webmaster can also remove site links with Webmaster tools. The value of link is affected by the anchor text, text content surrounding the link and the topic of the linking site. Link, officially called hyperlink, consists of the address and text that describes the link, also called anchor text:<a href ="http://www.rudus.fi">Rudus/>. Search engines are using the anchor text to determine the content of the target site. For example, if thousands of linkers would link to Rudus' Websites as following " Good Service, search engines would value the address http://www.rudus.fi also to good positions when searchers are searching with "good service", even though the site itself wouldn't include that phrase once. (Reponen, 2009.)

Google measures the value of links by its PageRank algorithm which has also been discussed in paragraph 5.1. Basic idea of the algorithm is to describe using a numeric value on a scale from 0-10 the vote result the site is getting from other sites. The average value for sites is between 3 and 5. PageRank value increases in proportion with the significance of the linking pages. The biggest PageRank value is normally in the site's front page where most of external links are linking. Increases and decreases of the PageRank values are updated once a month and the values can be monitored with the Google Toolbar-application. For example, Rudus' front pages' PageRank is 6 (Google, 2011, Reponen, 2009.)

PageRank value inheritance can be prevented by giving the command "nofollow" for the anchor element attribute that describes the link. For example:

Rudus. In this case, the search robots are following this link, but the PageRank value is not inherited from the site. Studies have shown it is likely that Google's robot Googlebot follows nofollow link, but does not save it to the index, unless the destination page has already been indexed. (Reponen, 2009.)

5.7 Indexing

In SEO language Indexing means the site's addition to search engine indexes, or site notification for search engines. Even if the site is optimized well, the whole work is in vain if the site cannot be found from search engine indexes.

When the new site is opened to the Internet it might take a long while before the site can be found from search engines. However, the indexing process can be accelerated by informing the search engines about the site. Before informing, it is wise to check if the site can already be found from search engines, for example by searching with the site name. Depending on the implementation, however, it is possible the site cannot be found with the company name nor product names even if it would have been indexed. If the site is not found the indexing can be checked with the "site"-function in search engines. By searching with the search phrase site: www:rudus:fi, the search engine is searching information only from the address www.rudus.fi. If the site cannot be found with this function it is likely that the site is not indexed. In these cases it is wise to add sites with search engines indexing services that can be found from search engine home sites. Google's indexing requests can be sent to: http://www.google.com/addurl.html.

Adding services do not automatically index the site, but they add sites to search robot's browsing queues. This is the practice how robots can get to sites with no external linking. Afterwards, the site is set to queue the indexing process and can be monitored with server log info. If the search robot visits the site, it can be found from the user-agent field.

Search robots are programs that go through the Internet's graphic structure page by page. Their target is to go through as large part of the Internet pages

as possible and to enter copies of those to a database engine that generates an optimized local copy of the page to match the searches. The optimal target is that the robot gets instructions to find the first page and after that it follows the internal and external links and moves forward until it has downloaded the whole site. Problems might occur in page code interpretation, timeliness of information, handling of the server load and malicious intents against robots. Search robots are playing an important role when measuring the effectiveness of the search engines. Accurate information of search engine activities is hard to find due to strict confidentiality policies of the companies. Simplified structure of the Robots activities is presented in Image 15. This model is based on Paananen's presented model in his Master's thesis "Hakukoneoptimointi Internetmarkkinoinnin tukena, 2006". Basically the robot has a list of URL:s that it needs to read through. It goes through the addresses in the list by downloading and interpreting them. The content is sent for search engine databases and the new links are added to a URL-list.

In the first part a list of Web addresses in URL-mode is maintained (Unified Resource Locator). The engine lists the addresses where the search robot moves next. URL list is also responsible to keep track of pages that have already been visited, when they were visited last time, and do they need to be revisited.

It is practically impossible to index the entire Internet due its size and fragmentation. This is why the right indexing strategy has a huge impact on search results quality and reliability.

Read URL-list

Contact address from the URL-list

Download page

Interpret the content

Add pages links to the URL-list
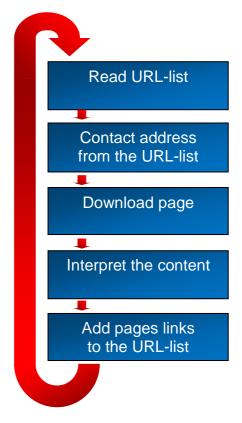
Image 15. Simplified Robot's functions. (Paananen, 2006)

In phase 2, the robot picks a page from URL-list and then downloads it. Aspects that need to be taken into consideration are: robots exclusion protocol compliance, load peaks avoidance and the optimization of DNS (Domain name system). In the mid 90's robots weren't that welcome to all servers. They caused lot of extra load, disrupted the votes and were repeatedly searching for the same files. Currently robots are more intelligent, but still some pages are not welcoming them. The remedy for this problem is the robots exclusion protocol, which has been carried out with a robots.txt.file. (Paananen, 2006.)

Another important part of the search engines from the SEO point of view is the processing of the indexed information. The process of searches can be divided into three main phases (Image 16). In the first phase the search engine analyses used keywords and searches for the corresponding pages. When the pages are found they will be ranked and organized. In the last phase the results are displayed in ranked order.
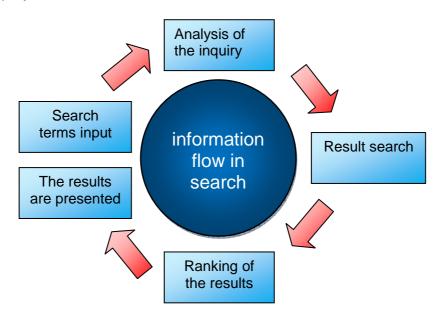
Image 16. Information flow in search process (Paananen, 2006)

The used search bar is called an inquiry at this point. For example, if the user writes to Google search bar "order and delivery", the line "order and delivery" is used in the inquiry. When analyzing the search it will be broken down to individual words or search terms. After that the search engine looks for

information containing searched phrases. Hence, search engine would most probably look for information with phrases "order delivery" and "delivery order".

After the search the search engine performs the ranking. To ensure the reliability of the results and making the manipulating more difficult, the search engine companies are not giving out any precise information about this phase. Some of the affecting variables are presented in chapter 5.1. (Paananen, 2006.)

If the site has existed for a longer time, it is likely that most common search engines have already indexed the site. If not, the reason may be that the search robot cannot move in or find the site content. It is also possible that the search engine has for some reason banned the site. If the site is banned, there is, or has been in the past, some unethical optimization techniques in use. Site won't be banned easily and it requires frequent and ongoing abuse. The use of unethical methods leads to punishment before the complete removal from the index (See Chapter.5.6.3 "Links"). It is also good to follow site indexing as part of the SEO-process. If possible irregularities occur in the site, the information of those comes fast enough to make changes without penalties or banning. There are several methods that are classified as abuses and those are developing as the search engines develop. The most common ones are cloaking, hidden content, hidden links and spamming to blogs and quest books.

In the longer term, the best results are achieved by implementing the site's customer friendliness and honesty. Search engines also give instructions that keep sites away from banned lists by guiding optimization techniques that are not against the search engine code of practice. Google's guidelines are available at:

http://www.google.com/support/webmasters/bin/topic.py?topic=28800

If the page cannot be found from the index, the most common reason is that robot's cannot move freely in the page. When new sites are created, there will be no problems if SEO technical rules are followed. In the older pages problems might occur for several reasons. For example:

- There might be robot visits preventing ".robots.txt" still active or the search engine has given up indexing the site if it has been prevented in the past
- "Robot traps" or links implementation with JavaScript code
- PopUp windows, frames etc.
- Long parameter queues in links caused by publication systems
- Non valid code
- HTTP Cookies enforcing

It is also good to ensure that access from the main page to every single sub-page is simple by following the links. Page maps improve both usability and indexing.

5.8 Follow-Up

To ensure the development there needs to be set metrics for follow up. There are useful visitor monitoring software to track site usage. These tools are presented in paragraph 5.3., "Consumer behaviour".

There needs to be continuous Website visitor and target analysis to reach the objectives. Website visitors and objectives make no difference to other company key figure measuring protocols. Like company's key figures measuring develops the business, Website analysis can help developing these activities.

Website analytic tools were presented on page 35. Tools are easy to use and they give valuable information to develop Websites. The follow up and development should support the conversion of the site. For example in Rudus' sites the conversion i.e. target is to get potential customers to proceed to the "contact" section. There are also other important key performance indicators that should be examined. Those could be the bounce rate, time spent on site, entry methods, share of consumer visitors, etc

6 EVALUATION

This thesis work does not include the actual research part.

The research question in this study was to find out is there a correlation between consumer sales and search engine optimization. The answer to this question was found. Even though the answer is exact, it should be perceived as an illustrative guideline because the reasoning process includes many assumptions.

Due to the case company's needs it was essential to include into the survey also the site's technical evaluation and e-marketing processes functioning in Rudus. SEO and SEM parts are well handled in both theory and research part. But there are still lot of e-marketing activities that need more precise examination.

The most difficult part of the study was to limit it somehow. On the other hand it would have been easier to limit the study strictly to SEO, but eventually that would have not given enough answers for Rudus to reorganize their e-marketing.

7 SUMMARY

In the theory part of this study, the importance of Internet and e-commercial activities were pointed out with statistics. Search engine marketing and search engine optimization were presented and also consumer behaviour in search engines was clarified. All these topics were handled with supporting statistics. Theory part is a useful tool for everyone participating in the SEO process in organizations. The technical part of optimization is not handled in a detailed manner, because this is a business administration related master's thesis.

The research part was made for case company Rudus Ltd. Research was made by collecting the needed results from various sources. There was a survey period in Rudus' Website and consumer sales between April 1st 2011 and December 31st 2011. This period was compared to situation between April 1st 2010 and December 31st 2011. There were no other significant changes in consumer sales practices between 2010 and 2011. This is one important factor when judging the relevance of the correlation.

This study was instructive and interesting for its author.

CHARTS

IMAGES

FORMULAS

REFERENCES

Baggott Chris, 2010, Debating the Value (and Meaning) of "Great Content" for SEO http://www.seomoz.org/blog/debating-the-value-of-great-content(Accessed on 29th October 2011)

DIVIA, Digital Marketing round table 2011, Aalto University http://www.aaltoee.fi/images/stories/Liitteet/digibarometri_2011.pdf (Accessed on 27th October 2011)

Eisenberg, B. September 5th, 2008, Why rank #1 in Google, http://www.grokdotcom.com/2008/09/05/why-rank-1-in-google/(Accessed on 28th October 2011)

Enge, E. Spencer,S. Fishkin, R. Stricchiola, J. C. 2009.
The Art Of SEO, mastering search engine optimization, O'Reilly

Experian Hitwise, 2011
http://www.hitwise.com/us/about-us/press-center/press-releases/google-share-of-searches-66-percent-in-sept-2011/ (Accessed on 5th December 2011)

Hakukoneoptimointi.info, 2006,
http://www.hakukoneoptimointi.info/2006/04/19/webfinder-pelista-pois/

Hayes, B. 2000, American Scientist.
https://www.americanscientist.org/issues/issue.aspx?id=3279&y=0&no=&content=true&page=3&css=print (Accessed on 20th January 2011)

Liuhto, S. 2010, Googlen hakukoneoptimointi ja sen vaikutukset yrityksen sivustoliikenteeseen. Laurea- ammattikorkeakoulu.

Mäkelä, M. 2011, Verkkosivujen hakukoneoptimointi tehokkaan Internetmarkkinoinnin keinona. Seinäjoen ammattikorkeakoulu.

Netcraft, 2010, October 2010 Web Server Survey,
http://news.netcraft.com/archives/2010/10/12/october-2010-web-server-survey.html (Accessed on 30th October 2011)

Norman, B. 2009, Get to #1 on Google,
In easy steps limited.

Paananen, J. 2006, Hakukoneoptimointi internetmarkkinoinnin tukena, Lappeenrannan teknillinen yliopisto

POHTO, Parhaat hakukoneet ja hakukoneluettelot.
http://www.oulu.fi/pohti/hakuindeksi.html (Accessed on 18th October 2011)

Reponen, J. 2009, Digitaalinen markkinointi: löydettävyyden merkitys ja kehittäminen. Laurea, Leppävaara.

Ruby, D. 2010 The Value of Google Result Positioning
http://insights.chitika.com/2010/the-value-of-google-result-positioning/(Accessed
on 28th October 2011)

Russell,K.2010, A Search Box by Any Other Name …
http://blogs.nd.edu/web-musings/a-search-box-by-any-other-name/(Accessed
on 5th December 2011)

Seomoz, 2009 Perfecting Keyword Targeting & On-Page Optimization
http://www.seomoz.org/blog/perfecting-keyword-targeting-on-page-
optimization(Accessed on 5th December 2011)

Seppänen, K. 2011 Genisys Ltd, Hakukoneet Markkinoinnissa,
Genisys.

Seppänen, V. 2004, Konstruktiivinen Tutkimus.
media.tol.oulu.fi/video/jtmk/konstruktiivinen_tutkimus.ppt

Sherman, C. 2000, Information Today, Inc.
New Web Map Reveals Previously Unseen 'Bow Tie' Organizational Structure
http://newsbreaks.infotoday.com/nbreader.asp?ArticleID=17813#bowtie
(Accessed on 20th January 2011)

Tietotekniikan käyttö yrityksissä 2010, Tilastokeskus.
Google,2011 http://www.google.com/about/corporate/company/
(Accessed on 25th October 2011)

Tilastokeskus, Internetin käytön yleiset muutokset.
http://www.stat.fi/til/sutivi/2010/sutivi_2010_2010-10-26_kat_001_fi.html
(Accessed on 10th January 2011)

Tulos.fi, 2011,
http://www.tulos.fi/hakukoneoptimointi/ (Accessed on 1st October 2011)

Waraas Jon, 2008, Heatmap Of Google Results
http://www.jonwaraas.com/heatmap-of-google-results/(Accessed on 28th
October 2011)

Wikipedia, Graph Theory
http://en.wikipedia.org/wiki/Graph_theory (Accessed on 10th December 2011)

Wikipedia, PageRank
http://fi.wikipedia.org/wiki/PageRank (Accessed on 30th October 2011)

Wisdomtalks December 10, 2009 ,Internet Marketing Vs. Conventional
Marketing (TV, Newspaper, etc)
http://wisdomtalks.com/internet-marketing-vs-conventional-marketing-tv-
newspaper-etc/ (Accessed on 18th October 2011)