# Search Engine Optimization (SEO): Website analysis of selected faculties in Croatia

#### Iva Gregurec, Petra Grd

Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
{iva.gregurec, petra.grd}@foi.hr

Abstract. Search engine optimization is an Internet marketing strategy, and as such, very important in building websites. Websites need to be optimized in order to atract users and to sell products and services. This paper will give an overview of selected SEO elements, and analyse websites of Croatian faculties related to computer science and engineering.

**Keywords.** SEO, website, informatics, optimization, search engines, SEM

## 1 Introduction

Nowadays it is almost impossible to function without the Internet, both for the users advertisers. Interactive media, particularly Internet, allow an easy flow of information, which enables users to participate in, and modify the form and content of the information they receive in real time. Traditional media are a one-way media, companies are not aware of the exact response to it. On the other hand, the new media allow users to perform a variety of functions (receive information, make inquiries, respond to questions and make purchases) and for companies it is easier to know the response to their campaign [10]. The Internet as the key to major technological changes in the media image of the 21st century have enabled a communication revolution: the ability to send and retrieve information everywhere has changed the way we work and live today [16, 472].

Companies need to communicate with consumers every day; they need to know what their wishes, needs, and preferences are. They also need to provide the best product/services for individual customer. And with the help of multi-channel marketing they may use different marketing channels to reach a customer. One way of communication with customers is through corporate web pages. And if company aims to take part on the first page or first row when people enter the keyword in the search engines they must use search engine marketing (SEM) techniques because of lower costs [26, 488] and one of them is search

engine optimization (SEO). SEM is an element of internet marketing promotional mix besides viral marketing, social networking and virtual communities and traditional elements of promotional (advertising, sales promotion, personal selling, direct marketing, public relations and publicity) [8, 349].

Many advertisers hire agencies to run their search engine marketing campaigns and include them in their marketing plans. In these plans, the advertiser pays the agency a fee for each conversion (i.e., acquired customer) but the agency pay all of the search engine marketing costs [15, 68].

#### 2 Internet

Internet is a main place to find information about any area and it is known as a wide community that includes millions of users and becouse of that more and more people and organizations now publish on the Internet [27, 692]. And that is the reason why internet is named as one of today's most important communication tools, but also as a new selling channel [26, 487]. So because of that internet has become a wide pool of information (,,internet information is characterized by the instability of its contents, decentralization of locations, multiplicity of forms, diversity of user groups, and dynamics of the environment" [27, 691]) where users dive to try and find relevant pages while webmasters fight for the visibility of their sites [12, 7]. The amount of information on the internet continues its' tremendous growth, and the size and number of sources appear to be growing at an exponential rate [27, 691]. Although Internet includes hundred thousands of web pages, it is better to use search engines, such as Google, Live, Yahoo!, MSN Search, etc., in order to reach wanted data or information [26, 487]. "In this two-sided universe, search engines are a third party, acting as intermediaries, connecting (hopefully) users to the web pages containing the information they search for [12, 7]."

# 3 Search engines

The world of the internet was transformed with the development of search engines in the mid-1990s [27, 665]. These tools provided access to the numerous resources on the Web and their primary function is displaying relevant links for a user's search [12, 7]. "Web sites, which index and class other web sites according to their keywords, explanations and contents and make it easier and faster to reach obtained site-search results, are called as search engines [26, 487]." It is too important for search engines to give response for lots of people, especially marketers. Search engines have established themselves as critical gatekeepers of information are "free and easy to use, and they tend to project an image of objectivity [12, 7]." After all, they do not promote their own content, but they are not as neutral as they might seem: their popularity makes them attractive to advertisers and they are even becoming the leading advertising medium through their sponsored lists of results [12, 7]. As the internet has matured, search engines have occupied an increasingly powerful position in the online world, by both channelling the attention of millions of users, and generating revenue for web sites through contextual advertising programmes, such as Google's AdSense" [11, 21].

Search engine is some kind of software, which collects data (web site URL, some keywords or keyword groups that define the content of the web site, the code structure that forms the web page and also links provided on the web site) about web sites which are indexed and stored on a database [26, 488]. Search engines may return many millions of documents for each user query (enterd keyword), but the user only looks at a selected few [11, 21] so it is very imortant for companies to be on the first page, to have good SEO.

"There are various factors which can contribute to visibility of a webpage in a search engine results list, for example, webpage metadata structure, webpage content, hyperlink cited status, search query expansion, and other possible factors [27, 666]." When we enter keyword or keyphrase search engine may use hundreds of factors in ranking websites, including the following considerations [25, 39]:

- Age of the site and its content, links and reputation of linking sites
- Regularity with which content is updated
- Uniqueness of content
- Title and meta tags
- Citations and research sources
- Depth of documents in the site
- Use of sub-domains, use of keywords in subdomains and volume of content on subdomains, and negative scoring for such activity

- Other affiliations of linking site with the linked site
- Quality of HTML coding, presence of coding errors

# 4 Search engine marketing (SEM)

Search engine marketing (SEM) is defined as , placing messages on a search engine encouraging clickthrough to a web site when the user types a specific keyword phrase [7, 34]. "Today SEM is arguably the most important digital marketing channel for customer acquisition. The fact is that all of us turn to a search engine when we are seeking a new product, service, some entertainment or just information [7, 282]. "The key to successful SEM is achieving keyphrase relevance since this is what search engines strive for – to match the combination of keywords typed into the search box to the most relevant destination content page [8, 378]." "SEM is great for targeting audience at the moment of intent. It can help create a level playing field where small companies can be listed alongside well-known brands to increase their awareness and drive response [7. 34]." The two main disciplines of SEM are Search engine optimization (SEO) - the process of trying to rank highly a given web page or domain for specific keywords [11, 22] and Pay-Per-Click (PPC). Third type that is mentioned in the literature is trusted feed including paid-for inclusion [7, 34][20, 9].

# 5 Search engine optimization (SEO)

There are many definitions of SEO, but they all say the same thing, SEO is "a structured approach used to increase the position of a company or its products in search engine natural or organic results listings for selected keywords or phrase [8, 376]." "Search engine optimization (SEO) is a set of techniques aimed at improving the ranking of a website in search engine listings, thereby making it more likely that the end user will choose to visit the site [25, 39]." "SEO gathers the techniques leading to a better indexing of one's site by one or several targeted search engines [12, 7]." It is also defined as "the process of identifying factors in a webpage which would impact search engine accessibility to it and fine-tuning the many elements of a website so it can achieve the highest possible visibility when a search engine responds to a relevant query [27, 666]." Its' role is to achive the highest position or ranking on search engines, such as Google, Yahoo! and MSN Search and so on, after a specific combination of keywords or key phrase that is typed in. "The final goal is for the site to be better ranked by one or several targeted search engines and therefore appearing higher in their results lists for specified requests [12, 6]." The position or ranking is dependent on an algorithm used by each search engine to match relevant site page content with the key phrase entered [7, 283]. Many companies utilize search engine optimization (SEO) tools designed to enhance the process and this can help produce sustainable results.

"Search engines compile an index of words on web sites by sending spiders or robots to crawl around sites that are registered with that search engine. The search engine algorithm weights the index according to different parameters and then stores the index as part of a database on a web server. This index is searched when potential customers type in keywords [8, 376]."

The key to a good optimization is to choose the keywords or key phrases on which the site should be optimized [12, 7]. "The most elegantly designed site with the most useful information on the web is of no value to those who can not find it [25, 39]."

Sellers can improve their listings on the search engines by modifying their site codes to make them more relevant and therefore more search-engine compatible. Most on-line sellers, however, do not invest in search engine optimization (SEO) to get higher search-results rankings for their listings, but instead prefer paid placements [20, 9]. They explain that SEO is more expensive than paid placements. On the other hand buyers use search engines to look for information and they tend to trust and follow links displayed in the editorial section of the search-results page.

Search engine optimization (SEO) can be divided into two categories: organic and paid optimization. Organic SEO refers to achieving good search rankings for a website without paying for it. And paid SEO or search engine advertising (SEA) may be best known in the form of Google Adwords and Overture.

#### **6** SEO elements

Search engine optimization has many elements (over 200 hundred), some of the elements analysed in this paper are:

- Page load time,
- Clean URL
- Redirecting (duplicate content)
- Meta description
- Sitemap
- Robots.txt file
- 404 error
- Title tag
- H1 tag

#### 6.1 Page load time

Page load time can be defined as a number of seconds the page needs to load all its elements in a

browser. The influence of page load time on search engine ranking is not as great as other elements, but is significant. Google PageRank algorithm also considers page load time in its calculations. On the other hand, slow page load time (some say over two seconds) creates difficulties for users to effectively browse a website.

#### 6.2 Clean URL

Search engines are much more at ease with URLs that don't contain long query strings. URL like http://www.example.com/4/basic.html can be indexed much easier, whereas its dynamic form, http://www.example.com/cgibin/gen.pl?id=4, can confuse search engines and cause them to miss possibly important information contained in the URL, and thus preventing you from getting the expected ranking [3].

With clean URLs, the search engines can distinguish folder names and can establish real links to keywords. Many of the SEO professionals agree that dynamic URLs are not very good for web spiders, while static URLs have greater visibility for them.

#### **6.3 Duplicate content (redirecting)**

Duplicate content in SEO exists for a long time and search engines have been getting smarter in figuring out the best page of duplicate content to display. It is less to worry about today, than before, but knowing this issue exist, advice is given on how this issue should be fixed. Some say simply block your duplicate content pages, some say redirect them [2].

#### 6.4 Meta description

Meta descriptions are HTML attributes that provide explanation of the contents of web pages. They are commonly used by search engines on search result pages to display preview snippets for a given page [4]. For example:

<meta name="description" content="Fakultet
organizacije i informatike, Sveučilište u Zagrebu" />

Length of the meta description content should not be over 155 characters [9].

Before, information in a Meta description could help a page rank for the words that were contained within it. Today, neither Google, Bing, nor Yahoo! use it as a ranking signal [4]. In terms of rankings, metadescription could be omitted. But, there are 3 important ways that Meta descriptions are being used today that make them an important part of your SEO and overall online marketing strategy[4]:

 They can be used as the description (or part of the description) of the page if it shows up in the search results.

- They are often used as part of the descriptive information for pages when Google shows "extended sitelinks" for the site.
- They are often used as the default description in social media marketing links such as Facebook and Google+.

#### 6.5 Sitemap

Sitemaps are map of a web site. On one page the structure of a site is shown, its sections, the links between them, etc. Sitemaps make navigating web sites easier. They are an important way of communicating with search engines [23]. There are two types of sitemaps. One type is for human visitors, and the other for spiders. Using sitemaps has many benefits, not only easier navigation and better visibility by search engines. Sitemaps offer the opportunity to inform search engines immediately about any changes on your site [23].

Sitemaps can be used on websites of any size however if a website has more than 16 pages, implementing a sitemap is essential from a search engine optimisation point of view as it makes it easier for Google and other search engines to crawl the site and identify what every page of the site is about [19].

#### 6.6 Robots.txt

Robots.txt files inform search engine spiders how to interact with indexing web site content. By default search engines are greedy. They want to index as much high quality information as they can, and they will assume that they can crawl everything unless told otherwise [21]. If robots.txt file doesn't exist, server logs will return 404 errors whenever a spider tries to access robots.txt file [21].

There are 3 primary reasons for using a robots.txt file on your website [14]:

- Information that doesn't want to be made public through search - In situations where there is content on a website which shouldn't be accessed via searches, the robots.txt will prevent search engines from including it in their index.
- Duplicate Content Often similar content is presented on a website under various URLs. Duplicate content can incur penalties by search engines which is bad from a SEO point of view. The robots.txt file can help control which version of the content the search engines include in their index.
- Manage bandwidth usage Some website's have limited bandwidth allowances (based on hosting packages).
   As robots use up bandwidth when indexing a site, in some instances – people might want to stop some user

agents from indexing elements of the site to conserve bandwidth usage.

#### 6.7 404 error

404 error occurs when server can't find the page requested. One common misconception is that it is a SEO best practice to simply 301 redirect pages that return a 404 status code to the homepage of the given domain [5].

Some of the reasons why 404 page might be shown are [1]:

- When a URL is typed incorrectly.
- There is a broken link on a web page and the user clicks on it.
- Another website that has a link to page on your website that does not exist.
- A page has been removed from a website or renamed and somewhere there is a link to that page and the user has clicked on the link.

### 6.8 Title tag

Title tag is one of the most important elements in search engine optimization. It describes content of a web page. Title tag appears in three key places [6]:

- Browser Title Tags show up at the top of a browser and in applicable tabs.
- Search Result Pages Title tags also show up in search engine results.
- External Websites External websites (especially social media sites) will use the title of a web page as its link anchor text.

Maximum number of characters displayed in search results is 70. All characters above 70 will not be displayed.

#### 6.8 H1 tag

Header tags (one of them is H1) are the second most important on page SEO factor. Sites optimized with quality H1 significantly better ranking. Search engine crawlers look for H1 description (along with other important SEO elements), to determine the relevance of a page. Omitting this tag isn't good SEO practice [17].

# 7 Search engine ranking

Results of a good SEO is good search engine ranking of a website. In this paper websites are analysed according to two search engine ranking algorithms: Google PageRank and Alexa rank.

#### 7.1 Google PageRank

PageRank is an algorithm that calculates a web metric which shows how reputable a particular page is according to Google [24]. PageRank takes into consideration quality and the quantity of the incoming links, with other parameters such as the number of outgoing links per page, the position/visibility of the links and more [24].

PageRank values are on a scale from 0 to 10.

#### 7.2 Alexa rank

Alexa rank sorts the websites according to their position of getting traffic. The websites that get more traffic get ranked better than others. Sometimes it happens that the search engines cares the website with a good Alexa ranking most and sort the search results according to the rankings respectively [22].

# 8 Web sites analysis

Topic of this chapter is analysis of websites according to elements mentioned earlier. Six websites were selected. These websites are web sites of Croatian faculties which teach computer science. In every subchapter overview of the analysis will be given, and negative values explained. Negative values are classified in two categories:

- Fix immediately negative values which influence ranking significantly
- Fix over time negative values which have less influence on ranking

Pear analytics tool for website analysis was used [18].

# **8.1** Faculty of organization and informatics (FOI)

FOI has one Fix immediately element. There is no XML sitemap in the root or in robots.txt file. As was said earlier, sitemaps are a way of communication with search engines.

Also, there are two Fix over time elements:

- Robots.txt file does not exist
- Page load time should be below 2 seconds

Table 1. FOI analysis

SEO element	Value
Alexa rank	43 962
Google PageRank	7
Page load time	8,294220
Clean URL	yes

Duplicate content (redirecting)	yes
Meta description	yes
Sitemap	no
robots.txt	no
404 error	ok
Title tag	yes
H1 tag	yes
Fix immediately	1
Fix over time	2
TOTAL FIX	3
These are ok	8

# **8.2** Faculty of Electrical Engineering and Computing (FER)

FER has two Fix immediately elements:

- There is no XML sitemap
- Error in duplicate content non-www version of the website is not redirected to the www version

Also, there are two Fix over time elements:

- There is no H1 tag
- Robots.txt file does not exist

Table 2. FER analysis

14010 2.1 211 41141 / 515	
SEO element	value
Alexa rank	43 962
Google PageRank	7
Page load time	1,680061
Clean URL	yes
Duplicate content (redirecting)	no
Meta description	yes
Sitemap	no
robots.txt	no
404 error	ok
Title tag	yes
H1 tag	no
Fix immediately	2
Fix over time	2
TOTAL FIX	4
These are ok	7

#### 8.3 Faculty of science (PMF)

PMF has two Fix immediately elements:

- There is no XML sitemap
- Error in duplicate content non-www version of the website is not redirected to the www version

Also, there are three Fix over time elements:

- There is no H1 tag
- Robots.txt file does not exist
- Page load time should be below 2 seconds

Table 3. PMF analysis

SEO element	Value
Alexa rank	43 962
Google PageRank	7
Page load time	2,265887
Clean URL	yes
Duplicate content (redirecting)	no
Meta description	yes
Sitemap	no
robots.txt	no
404 error	ok
Title tag	yes
H1 tag	no
Fix immediately	2
Fix over time	3
TOTAL FIX	5
These are ok	6

### 8.4 Polytechnic of Zagreb (TVZ)

TVZ has one Fix immediately element: Error in duplicate content – non-www version of the website is not redirected to the www version.

Also, there are three Fix over time elements:

- Google PageRank is below 7
- Page load time should be below 2 seconds
- Alexa Rank is not in top 100 000

Table 4. TVZ analysis

SEO element	TVZ
Alexa rank	390217,00
Google PageRank	6
Page load time	5,500435
Clean URL	yes
Duplicate content (redirecting)	no
Meta description	yes
Sitemap	yes
robots.txt	yes
404 error	ok
Title tag	yes
H1 tag	yes
Fix immediately	1
Fix over time	3

TOTAL FIX	4
These are ok	7

# 8.5 Faculty of electrical engineering Osijek (ETFOS)

ETFOS has the most negative elements. There are two Fix immediately elements:

- There is no XML sitemap
- Error in duplicate content non-www version of the website is not redirected to the www version

And there are six Fix over time elements:

- There is no H1 tag
- There is no meta description
- Robots.txt file does not exist
- Google PageRank is below 7
- Page load time should be below 2 seconds
- Alexa Rank is not in top 100 000

Table 5. ETFOS analysis

Element/Faculty web site	ETFOS
Alexa rank	391177,00
Google PageRank	6
Page load time	6,782722
Clean URL	yes
Duplicate content (redirecting)	no
Meta description	no
Sitemap	no
robots.txt	no
404 error	ok
Title tag	yes
H1 tag	no
Fix immediately	2
Fix over time	6
TOTAL FIX	8
These are ok	3

#### 8.6 University of applied sciences (VERN)

VERN does not have any Fix immediately elements, but it has four Fix over time elements:

- There is no H1 tag
- There is no meta description
- Page load time should be below 2 seconds
- Alexa Rank is not in top 100 000

Table 6. VERN analysis

Element/Faculty web site	VERN
Alexa rank	738707,00
Google PageRank	7

Page load time	2,007111
Clean URL	yes
Duplicate content (redirecting)	yes
Meta description	no
Sitemap	yes
robots.txt	yes
404 error	ok
Title tag	yes
H1 tag	no
Fix immediately	0
Fix over time	4
TOTAL FIX	4
These are ok	7

#### 9 Results

Previously mentioned results are ranked according to number of positive elements of every website. If this number is equal, number of Fix immediately elements is taken into consideration (Table 7, Fig 1.).

Table 7. Ranking of analysed faculties

Rank	Faculty
1.	Faculty of organization and informatics
2.	University of applied sciences
3.	Polytechnic of Zagreb
4.	Faculty of electrical engineering and
	computing
5.	Faculty of science
6.	Faculty of electrical engineering Osijek

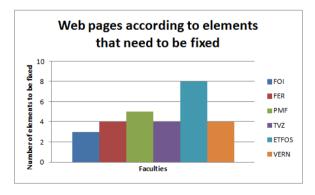


Figure 1. Websites according to elements that need to be fixed

As it was mentioned earlier, the result of a good SEO is good search engine ranking. In Alexa Rank the lower is better (Fig.2). In PageRank higher is better (Fig.3).

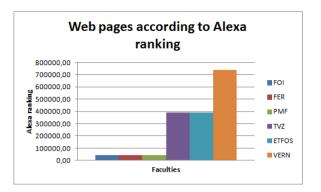


Figure 2. Websites according to Alexa ranking

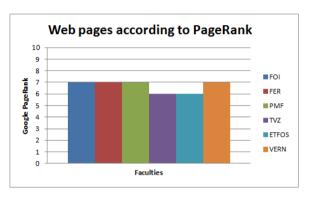


Figure 3. Websites according to Google PageRank

Another important element is page load time. And only FER has a positive Page load time element (Fig.4).

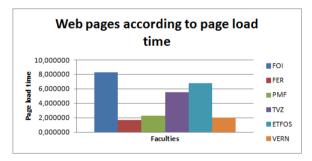


Figure 4. Websites according to page load time

#### 10 Conclusion

In order to attract users, websites need to be optimized. Not only for human users, but also for search engines. This paper gave an overview of a few most common elements of search engine optimization. Web sites of six Croatian faculties were analysed, and results have been given. It can be concluded that most of the six websites analysed pay attention to SEO as an important tool for attracting users.

### References

- [1] Agile SEO, Optimizing a 404 page for SEO, http://agileseo.net/optimizing-404-error-pages-for-seo-158, downloaded: April 25th 2012.
- [2] Arriola B, Duplicate content: Block, redirect or cannonical, http://www.seomoz.org/blog/duplicate-content-block-redirect-or-canonical, downloaded: April 25th 2012.
- [3] Avangate, How to make your URLs SEO friendly, http://www.avangate.com/company/resources/article/url-rewriting.htm, downloaded: April 25th 2012.
- [4] Best practices for search engine optimization, Meta description tag, http://www.seomoz.org/learn-seo/meta-description, downloaded: April 25th 2012.
- [5] Best practices for search engine optimization, HTTP Status Codes, http://www.seomoz.org/learn-seo/http-status-codes, downloaded: April 25th 2012.
- [6] Best practices for search engine optimization, Title tag, http://www.seomoz.org/learn-seo/title-tag, downloaded: April 25th 2012.
- [7] Chaffey, D.; Smith, PR.; eMarketing excellence Planing and optimizing your digital marketing, 3rd edition, Butterworth-Heinemann, Oxford, 2008, page 34
- [8] Cheffey, D.; Ellis-Chadwick, F.; Mayer, R.; Johnston, K.; Internet marketing Strategy, Implementation and Practice, Prentice Hall, Essex, 2006, page 349
- [9] deGeyter S. Everything you need to know about meta description and keyword tags, http://www.searchengineguide.com/stoney-degeyter/seo-101-part-3-everything-you-need-to-kn.php, downloaded: April 25th 2012.
- [10] Della Penna, M.: Exclusive PMN Research: Brands Struggle to Reach & Engage Younger Gen Y's Within Social Networks, 4th March 2009, page 20-21 Available at: [http://thepmn.org/BlogRetrieve.aspx? BlogID=1874&PostID=54641, [Accessed 22nd April 2010]
- [11] Evans, M.P.; Analysing Google rankings through search engine optimization data, Internet Research, Vol. 17, No. 1, 2007, pp. 21-37, page 21
- [12] Gandour, A.; Regolini, A.; Web site search engine optimization: a case study of Fragfornet, Library Hi Tech News, 2011, pp. 6-13, page 7
- [13] Jensen, M.B.; Online marketing communication potential Priorities in Danish firms and advertising agencies, European Journal of Marketing, Vol. 42, No.3/4, 2008, pp. 502-525, page 504
- [14] LeMerle R. The definitive guide to robots.txt for SEO, http://blog.ineedhits.com/tips-advice/the-definitive-guide-to-robots-txt-for-seo-20398484.html, downloaded: April 25th 2012.
- [15] Nabout, N.A.; Skiera, B.; Stepanchuk, T.; Gerstmeier, E.; An analysis of the profitability of feebased compensation plans for search engine marketing, International Journal of Research in Marketing 29, 2012, pp. 68-80, page 68

- [16] Neumann M., O'Murchau I., Breslin J., Decker S., Hogan D., MacDonaill C.: Semantic social network portal for collaborative online communities, Journal of European Industrial Training, Vol. 29, No. 6, 2005, pp. 472-487, page 472
- [17] Notes and Points on Search Engine Optimization, H1 Tag for SEO is it a myth?, http://seonotes.wordpress.com/2008/04/15/h1-tag-for-seo-is-it-a-myth/, downloaded: April 25th 2012.
- [18] Pearanalytics, Free SEO analysis, http://www.pearanalytics.com/, downloaded: April 25th 2012.
- [19] Roi.com.au, Sitemap for SEO, http://www.roi.com.au/seo-tips/the-importance-of-having-a-site-map-3/#.T5pcf9lN\_To, downloaded: April 25th 2012.
- [20] Sen, R.; Optimal Search Engine Marketing Strategy, International Journal of Electronic Commerce, Vol. 10, No. 1, 2005, pp. 9-25, page 9-10 [21] SEO Book, Robots.txt tutorial, http://tools.seobook.com/robots-txt/, downloaded: April 25th 2012.
- [22] TechShouters, Significance of Alexa Rank for Your Website You Must Know, http://www.techshouters.com/significance-of-alexa-rank-for-your-website-you-must-know/, downloaded: April 25th 2012.
- [23] Webconfs, Importance of sitemaps, http://www.webconfs.com/importance-of-sitemaps-article-17.php, downloaded: April 25th 2012.
- [24] WebSEO analytics blog, Is Google PageRank still important in Seach Engine Optimization?, http://www.webseoanalytics.com/blog/is-google-pagerank-still-important-in-seach-engine-optimization/, downloaded: April 25th 2012.
- [25] XXX, Journal of Visual Communication in Medicine, March 2006, Vol. 29, No. 1, pp. 39-40, page 39
- [26] Yalçin, N.; Köse, U., What is search engine optimization: SEO?, Procedia Social and Behavioral Sciences 9, 2010, pp. 487-493, page 488
- [27] Zgang, J.; Dimitroff, A.; The impact of webpage content characteristics on webpage visibility in search engine results (Part II), Information Processing and Management 41, 2005, pp. 691-715, Page 692