Andrej Miklosik[1], **Dana Vokounova**[2]

EXAMINATION OF SELECTED VARIABLES RELATIONSHIP TO THE SEARCH RESULTS POSITION*

The article aims at identifying the impact of 2 different factors on the search engine rankings page position of a web page. These included the compatibility of the web page code regarding the W3C standards and the number of total links (external and internal) on the web page. For this purpose an empirical study on 21 faculties of Slovak universities offering economically oriented educational programs was performed. The authors have validated 4 hypotheses which define the relationship between 2 selected SEO factors and the website SEO performance (measure through the SERP position and Google PageRank).

Keywords: *E-marketing, Internet marketing, search engine marketing, search engine optimization, web visibility.*

Андрей Міклошик, Дана Вокоунова

ВПЛИВ ДЕЯКИХ ЗМІННИХ НА РЕЗУЛЬТАТИ ВИДАЧІ В ПОШУКОВИХ ЗАПИТАХ

У статті досліджено вплив двох різних факторів на позиції сайту в результатах пошукових запитів. Вони включають сумісність коду веб-сторінки із стандартами Консорціуму Всесвітньої мережі і загальну кількість посилань (зовнішніх і внутрішніх) на веб-сторінці. Виконано емпіричне дослідження сайтів 21 факультету словацьких вузів з економічно орієнтованими програмами освіти. Перевірено 4 гіпотези, які визначають взаємозалежність між двома вибраними факторами пошукової оптимізації і пошукової ефективності веб-сайту (виміряно за допомогою позиції SERP і Google PageRank).

Ключові слова: *електронний маркетинг, Інтернет-маркетинг, пошуковий маркетинг, пошукова оптимізація, веб-видимість.*

Табл. 7. Літ. 12.

Андрей Миклошик, Дана Вокоунова

ВЛИЯНИЕ НЕКОТОРЫХ ПЕРЕМЕННЫХ НА РЕЗУЛЬТАТЫ ВЫДАЧИ В ПОИСКОВЫХ ЗАПРОСАХ

В статье исследуется влияние двух различных факторов на позиции сайта в результатах поисковых запросов. Они включают совместимость кода веб-страницы со стандартами Консорциума Всемирной сети и общее количество ссылок (внешних и внутренних) на веб-странице. Выполнено эмпирическое исследование сайтов 21 факультета словацких вузов с экономически ориентированными программами образования. Авторы протестировали 4 гипотезы, которые определяют взаимозависимость между двумя выбранными факторами поисковой оптимизации и поисковой эффективности веб-сайта (измерено посредством позиции SERP и Google PageRank).

Ключевые слова: *электронный маркетинг, Интернет-маркетинг, поисковый маркетинг, поисковая оптимизация, веб-видимость.*

---

[1] PhD, Ing., Senior Lecturer, Department of Marketing, Faculty of Commerce, University of Economics in Bratislava, Slovakia.

[2] PhD., Ing., Senior Lecturer, Department of Marketing, Faculty of Commerce, University of Economics in Bratislava, Slovakia.

* The article originated as the output of the project VEGA no. 1/0418/11 "Sustainable marketing and sustainable consumption".

Introduction. Experts on search engine marketing and particularly search engine optimization are arguing on the importance of partial variables or factors on the performance in search engines. In our empirical study we aim at providing exact evidence supporting these hypotheses. The objective of the article is to prove a direct relationship between:

- the compatibility of the web page code regarding the W3C standards;
- the number of total links (external and internal) on a web page;
- the SERP (search engine results page) positions;
- Google PageRank (GPR) of a web page.

To confirm or exclude the existence of mutual connection between these variables, we evaluated the results of primary data collection on the sample of 21 faculties of universities in Slovakia, providing economically oriented education. Based on our previous research and experience in this area we formulated 4 hypotheses:

H1: *There is significant correlation between the web page code compatibility to W3C standards (HTML and CSS) and the positioning of the web page in search engines.*

H2: *There is significant correlation between the total number of back links (internal and external) and the positioning of the web page in search engines.*

H3: *There is a significant correlation between the web page code compatibility to W3C standards (HTML and CSS) and the Google PageRank of the web page.*

H4: *There is significant correlation between the total number of back links (internal and external) and the Google PageRank of the web page.*

1.Crucial factors influencing the SERP positions. Search engine marketing (SEM) is an important tool of the Internet marketing and e-marketing activities. Marketing managers are aware of the importance of positioning in the search results to acquire more traffic from search engines. Search engine optimization (SEO) as a preferred method of gaining sustainable traffic on the web site is being analyzed by the experts continually. Many authors have published their studies in this area offering innovative approaches to SEM tools and implementation (Abou Nabout, Skiera, Stepanchuk & Gerstmeier, 2011; Ho, Lu, Ho & Peng, 2011; Chien, 2011; Kucuk, 2011; Miklosik, 2006). Marketing managers in companies and organizations need to use its tools much more intensively in the upcoming years. Definitely, the most frequently used methods of SEM are search engine optimization (SEO) and pay per click (PPC). In the introduced IBM study more than 62% of CMOs stated they plan to increase the use of search engine optimization in the near future. The study 2010 Response Rate Trend Report performed on the sample of 473 companies revealed that companies prefer to use PPC as a sales gateway. The average cost per click (CPC) was at $3.79 by the 3.81% conversion rate. The real conversion rate on the company's website was slightly below 5% (DMA, 2011).

Within the SEO experts there is a continuing discussion on the importance of each factor for the overall SERP (search engine results page) rankings. As the algorithm of Google as the world number one in the field of full text search is changing dynamically, the factors importance and weight in the final algorithm changes in time. However, there are many basic principles which remain constant. There are many factors and tools for SEO which directly determine the visibility in search engines. The recent research study by "SEOmoz" identified the key factors of search engine ranking and their importance. We list them in Table 1.

Table 1. **Search engine ranking position factors**

Factor	Importance/weight (%)
Page Level Links Metrics	21,45
Domain Level Link Authority Features	21,13
Page Level Keyword Usage	14,93
Domain Level Keyword Usage	10,73
Page Level Social Metrics	7,22
Domain Level Brand Metrics	6,78
Page Level Keyword Agnostic Features	6,74
Page Level Traffic/Query Data	6,26
Domain Level Keyword Agnostic Features	4,92

Data source: SEOmoz (2011).

With the massive development of Internet social networks we are experiencing the change of users' behaviour which is reflected in the search engines algorithms. Google announced that it is currently giving more emphasis on the activity of a company in social networks. This means a significant change in the direction of evaluating the relevance of a web page. It is not only important to be recognized widely (which means having a high number of highly relevant back links) but also to be able to attract and hold the attention of social media users. This fact will be part of our future research in this area, however, for the purpose of this article we have decided to measure the impact and relationship between the traditional SEO factors. We have chosen the number of total links which, as indicated in Table 1, is widely recognized being the number 1 factor influencing the SEO performance and the compatibility of web-site code with W3C standards, which is not included in TOP factors, however is intensively discussed within the SEO community in recent years.

2. Examining the relationship of two selected indicators to the SERP positions:

2.1. Methodology. In January and February 2012 we realized a complex collection and analyses of primary data for this study. We analysed the selected indicators of 21 web pages of faculties with economically oriented educational programs in Slovakia. The research was accomplished in 2 major phases:

- PHASE 1 – data gathering and processing;
- PHASE 2 – data analyses and findings summarization.

PHASE 1. For the research purposes we collected the data in 4 groups:

1. Compatibility of web page coding to the W3C standards (further as W3C).
2. Ranking in search engine results for selected key words (further as SERP).
3. Number of links pointing to the web pages (further as back links).
4. Google PageRank of the title page (further as GPR).

The objective of this phase of data gathering and processing was to determine the value (performance) of each web page for each of the indicators group, on the scale from 0 (no performance) to 100 (best performance). We describe the methods used to acquire and process the data in these 4 indicator groups as following:

1) Compatibility of web page coding to the W3C standards. This indicator consists of 4 subindicators. We measured the values of all of them for the title page of each web site. The first one is number of HTML errors. This number indicates the serious problems in the code which make the site incompatible to the recent W3C xHTML 1.0 Transitional standard. The second indicator is the number of HTML warnings. They refer to a potentially problematic part in the code. To collect the data for these 2 indi-

cators we used the Markup Validation Service from the World Wide Web Consortium (W3C, 2012b). The third indicator was the number of CSS errors. We measured the errors according to the CSS v 2.1 standards. As the fourth indicator we took the number of CSS warnings. To collect the data from these 2 indicators we used the CSS Validation Service from the World Wide Web Consortium (W3C, 2012a).

After collecting the source data we processed them to fit the scale. We rotated the number ranges so the web page with most errors gets 0 points and the solution with 0 errors receives maximum number of points. After that we scaled the number ranges to the 0–100 scale. As we are aware of the different importance of each of these 4 subindicators, we applied weights to calculate the final score in this indicator. The weights reflect the seriousness of potential issues which can be caused by the incompatibilities with HTML or CSS (Table 2).

Table 2. **Subindicators and their weights within the W3C indicator**

Subindicator	Weight
Number of HTML errors	40%
Number of HTML warnings	20%
Number of CSS errors	25%
Number of HTML warnings	15%

Data source: empirical study.

2) Ranking in search engine results for selected keywords. We decided to measure the ranking in the search engine Google solely. The reason lies in its market share, which is currently over 98%. No other full text search engine is being really used at the moment to search for web pages and information. The performance in Google (www.google.sk) was measured through identifying the SERP ranking of each faculty web site for the defined set of keywords. We decided to use 8 keywords which are closely connected to faculties and courses they are offering. The other criteria to select these keywords was based on identifying how potential customers (students) think when trying to find an appropriate faculty for their future studies in the Internet. The list of the applied keywords is shown in Table 3. In the data gathering we used Slovak originals of the keywords, however for understanding and comparability of the study results we are listing their English translations.

Table 3. **Keywords used to determine the SERP position**

Keyword ID	Keywords – English	Keywords – Slovak
A	University	Univerzita
B	College	Vysokaskola
C	University (abbrev.)	VS
D	Faculty	Fakulta
E	Economic Faculty	Ekonomicka fakulta
F	Economy studies	Studium ekonomie
G	Faculty+management	Fakulta+manazment
H	University studies	Univerzitne studium

Data source: own empirical study.

There are 3 functional methods for determining the SERP position available: manual counting in a web browser, using the position identification tool available at Seotools (Seotools, 2012) and using the SERP tool from SEOMoz (SEOmoz, 2012a). We analysed the results and differences using these 3 methods on the sample

of 5 faculties and determined the most appropriate method of manual counting which enables us to reach precise results and search for the results within the positions 1–50.

We checked the position for every keyword and every faculty, 400 positions in total. We processed the values into a table. If the faculty positioned 51+ in SERP, we used the value 51 which is 1 more than the worst position taken into account. This enables us to evaluate average values and determine the overall development.

After collecting the source data we processed them to fit the scale as well. We rotated the number ranges so the web page with best positioning in Google (position number 1) 100 points and the solution with no positioning (51+) receives 0 points. The final score in this indicator was calculated as an arithmetic average of partial 8 values, calculated for each keyword separately.

3) Number of links on web pages. We supposed the number of links pointing to the web page to determine its performance in search engines significantly. We counted the total number of internal links (links pointing to other subpage within the analysed domain) and external links (links pointing to the web site from other domains – external back links). We also gained the number of total linking root domains (unique root domains containing at least one link to the analysed web page). We used the analytic tool Open Site Explorer to obtain the source data (SEOmoz, 2012b). We abstracted from differences between follow and no-follow links. The acquired data were processed to fit to the 0–100 scale.

4) Google PageRank of the title page. We collected the data of GPR for each of the analysed domains. The data were modified from the 0–10 scale to the desired 0–100 scale.

PHASE 2. In the second phase we analysed the final data and calculated their interdependence. We used the statistical tools to calculate the correlation in the acquired values for each of the criteria. We list the final indicator values used to analyse the correlation and the validity of the defined hypotheses.

Table 4. **Final scoring in the 4 main indicator groups after data processing**

Faculty ID	Acronym	W3C	SERP	Back links	GPR
1	OF	81.46	18.33	22.94	42.86
2	NHF	48.93	48.33	61.47	85.71
3	FMV	77.56	41.67	35.64	85.71
4	FHI	97.02	35.00	35.64	71.43
5	FPM	100.00	63.33	38.53	85.71
6	EUKE	88.65	63.33	39.39	85.71
7	EFUMB	59.40	76.67	35.64	85.71
8	FMUK	88.61	100.00	59.63	85.71
9	FEMUNIAG	92.07	96.67	100.00	85.71
10	SELYE	96.44	35.00	18.10	100.00
11	FMKUCM	62.27	0.00	57.98	71.43
12	EFTUKE	96.36	43.33	0.43	100.00
13	FMUNIPO	78.21	46.67	0.98	100.00
14	VSEVTNUNI	95.57	0.00	1.47	57.14
15	FVSUPJS	95.41	0.00	9.51	100.00
16	PANEURO	98.76	68.33	18.04	71.43
17	VSEMVC	46.62	88.33	0.12	85.71
18	VSM	99.78	36.67	74.72	85.71
19	ISMPO	98.80	26.67	9.08	85.71
20	SM	98.77	0.00	44.91	57.14
21	BIVS	77.20	31.67	0.06	71.43

Data source: Own empirical study.

2.2. Results. We analysed the partial results to confirm or reject the 4 major hypotheses we set:

H1: *There is a significant correlation between the web page code compatibility to W3C standards (HTML and CSS) and the positioning of a web page in search engines.*

We supposed that coding of a web site which is according to the standards will not significantly help web page positioning in Google. To process the evidence on this hypothesis we analysed the correlation between 2 data ranges – W3C and SERP. To provide more evidence we took the 5 best web pages with minimum HTML and CSS errors. We calculated how they score in Google on average and compared this to the average score of the rest, less successful faculty web sites. We list the results in Table 5. The best web pages according to W3C reached the average score in SERP 39, on the other hand, the less successful pages according to W3C reached better positions in SERP. This was also confirmed by calculating the correlation coefficient, which was 0.2195.

Table 5. **TOP 5 web pages in W3C vs. the rest, SERP performance comparison**

	W3C	SERP
TOP 5	99.22	39.00
6–21	80.11	45.31

Data source: Own empirical study.

Based on the presented results we accept the hypothesis H1. There is no positive correlation between the web page code compatibility to W3C standards (HTML and CSS) and the positioning of the web page in search engines. Moreover, the correlation is negative; meaning the positive results in W3C may affect the SERP performance negatively. However the negative correlation is low (over CC = -0.3).

H2: *There is a significant correlation between the total number of back links (internal and external) and the positioning of the web page in search engines.*

Based on our knowledge and the opinion of key SEO experts we expected the number of back links being a significant factor of good placement in the search engines results. We analysed the data ranges back links and SERP. From Table 6 it is clear that web pages having most back links perform in Google much better compared to the rest of web pages. We calculated the correlation coefficient, which is 0.3027. This confirms a moderate positive correlation between the number of back links and the search engine rankings.

Table 6. **TOP 5 web pages in back links vs. the rest, SERP performance comparison**

	Back links	SERP
TOP 5	70.76	56.33
6–21	42.63	23.33

Data source: Own empirical study.

We also tested the correlation between the number root linking domains and the SERP. We expected the number of root linking domains to have higher impact on the positioning in search engines than the total number of back links. The correlation coefficient in this case was 0.3044 which is very similar to 0.3027 by the number of total links.

Based on the presented results we ***accept the hypothesis H2***. The correlation between the indicators back links and SERP is positive with medium significance level.

H3: *There is a significant correlation between the web page code compatibility to W3C standards (HTML and CSS) and the Google PageRank of the web page.*

When formulating this hypothesis we expected the efforts in cleaning the web page code and aligning it to standards to have a minimum impact on the Google PageRank of the web page. The Google PageRank is calculated based on the number of important criteria. SEO experts do not include the code W3C compatibility among these criteria.

The correlation coefficient between the data ranges W3C and GPR is -0.0386. The coefficient is negative, however, the number is so low that we can state no correlation between these indicators. They proved to be independent. Thus, ***we accept the hypothesis H3.***

H4: *There is a significant correlation between the total number of back links (internal and external) and the Google PageRank of the web page.*

Finally, according to the generally accepted opinion that back links do influence the search engine rankings positively, we expect they would have a positive impact on the Google PageRank score as well.

The correlation coefficient calculated with the data from back links and GPR data ranges is -0.2195. This low negative correlation means that no significant correlation between these indicators has been proved. Furthermore, we can expect slightly negative impact on Google PageRank when trying to build masses of back links from the same root domains.

We reject the H4 hypothesis. The correlation is negative with low significance.

Because of this interesting result we calculated the correlation coefficient between the scores in number of linking root domains and the Google PageRank as well. In this case a positive correlation can be observed. The correlation coefficient is 0.1058.

Discussion. The results of our empirical study prove that there is a correlation between the total number of back links and the performance in search engines. This positive correlation was proved in the SEOmoz research published in 2011. On the other hand, we have confirmed that back links do not influence the Google PageRank value positively at all. We think these results are interesting to be further compared by other researchers using data in different sectors and countries. Our methodology is universally applicable and we expect further discussions and research studies in this area.

We can summarize the following **main conclusions of our study**:

- Generally SEO experts tend to point out the errors in the web page HTML and CSS. However, our results have proved that these errors do not affect the positions in search results nor the Google PageRank.

- SEO experts think that the number of linking root domains is much more relevant for good SERP results than the number of total back links pointing to a web page. In our study we prove that these factors are important, however, there is no significant difference between their importance.

- Building a back link campaign has to be managed reasonably. We proved that acquiring back links from many different root domains affects not only the SERP

rankings but also the Google PageRank positively. However, gaining many back links from the same domain can affect the Google PageRank negatively.

References:

Abou Nabout, N., Skiera, B., Stepanchuk, T., Gerstmeier, E. (2011). An analysis of the profitability of fee-based compensation plans for search engine marketing. International Journal of Research in Marketing.

Chien, M. (2011). Consumer decision process based on search engine marketing.

DMA (2011). 2010 Response Rate Trend Report Retrieved from http://www.the-dma.org/cgi/dispannouncements?article=1451.

Ho, L.H., Lu, M.H., Ho, H.Y., Peng, T.F. (2011). A study of website optimization strategy and implementation. Advanced Materials Research, 268–270: 829–834.

Kucuk, S.U. (2011). Towards integrated e-marketing value creation process. Journal of Direct Data and Digital Marketing Practice, 12(4): 345–363.

Miklosik, A. (2006). New Challenges In Internet Marketing. Paper presented at the Business Enterpreneurship in the Conditions of Knowledge Economy Theory and Praxis, Bratislava.

SEOmoz (2011). 2011 Search Engine Ranking Factors, 2012-02-02 from http://www.seomoz.org/article/search-ranking-factors#predictions.

SEOmoz (2012a). Keyword Rankings. Retrieved 2012-02-01 from http://pro.seomoz.org/campaigns.

SEOmoz (2012b). Open Site Explorer. Retrieved 2012-01-05 from http://www.opensiteexplorer.org.

Seotools (2012). Position Analyzer. Retrieved 2012-01-25 from http://www.seotools.sk/poscheck.php.

W3C (2012a). CSS Validation Service. Retrieved 2012-02-01 from http://jigsaw.w3.org/css-validator/#validate_by_uri.

W3C (2012b). Markup Validation Service. Retrieved 2011-01-15 from http://validator.w3.org/check?uri=www.fses.uniba.sk&charset=%28detect+automatically%29&doctype=Inline&group=0.

Appendix 1. **Complete list of faculties included in the empirical study**

Faculty ID	Faculty name	Acronym	URL
1	Obchodna fakulta Ekonomicke juniverzity v Bratislave (EU)	OF	obchodnafakulta.sk
2	Narodohospodarska fakulta EU	NHF	nhf.euba.sk
3	Fakulta medzinarodnych vzt'ahov EU	FMV	fmv.euba.sk
4	Fakulta hospodarskej informatiky EU	FHI	fhi.sk
5	Fakulta podnikoveho manazmentu EU	FPM	fpm.euba.sk
6	Podnikovohospodarska fakulta EU v Kosiciach	EUKE	euke.sk
7	Ekonomicka fakulta UMB	EFUMB	ef.umb.sk
8	Fakulta manazmentu UK v Bratislave	FMUK	fm.uniba.sk
9	Fakulta ekonomiky a manazmentu Slovenskej pol'nohosp. univerzity	FEMUNIAG	fem.uniag.sk
10	Ekonomicka fakulta, Univerzita J. Selyeho	SELYE	selyeuni.sk
11	Fakulta masmedialnej komunikacie UCM v Trnave	FMKUCM	fmk.ucm.sk
12	Ekonomicka fakulta Technicke juniverzity v Kosiciach	EFTUKE	ekf.tuke.sk
13	Fakulta manazmentu Presovskejuniverzity	FMUNIPO	unipo.sk/fm
14	Fakulta socialnoekonomickych vzt'ahov Trencianske juniv. A. Dubceka	VSEVTNUNI	tnuni.sk
15	Fakulta verejnejs pravy Univerzity Pavla Jozefa Safarika	FVSUPJS	fvsupjs.sk
16	Fakulta ekonymie a podnikania Bratislavske jvysokej skoly prava	PANEURO	pacneurouni.com
17	Vysoka skola ekonymie a manazmentu verejnej spravy v Bratislave	VSEMVC	vsemvs.sk
18	Vysoka skola manazmentu v Trencine	VSM	vsm.sk
19	Vysoka skola medzinarodne hopodnikania ISM Slovakia v Presove	ISMPO	ismpo.sk
20	Sales Manager Akademie	SM	sales-manager.sk
21	Bankovni institut vysoka skola	BIVS	bivs.sk

Data source: Own empirical study.

Стаття надійшла до редакції 12.10.2012.