

Analýza a návrh odporúčaní na zlepšenie nájditeľnosti webových sídiel knižníc v internete

Makulová, Soňa

Filozofická fakulta Univerzity Komenského, Katedra knižničnej a informačnej vedy, Bratislava

Úvod

Cieľom príspevku je na základe analýzy nájditeľnosti webových sídiel knižníc v internete navrhnúť odporúčania na zlepšenie. Vychádza sa z analýzy informačnej architektúry webového sídla a určení pozitívnych a negatívnych faktorov majúcich vplyv na zlepšenie radenia výsledkov vo výstupe. V štúdiu sú uvádzané príklady webových sídiel knižníc zapojených do súťaže TOP WebLib 2006.

Nájditeľnosť webového sídla a jej vzťah k informačnej architektúre a optimalizácii stránok pre vyhľadávače

V štúdiu **Analýza faktorov vplyvajúcich na nájditeľnosť webových sídiel** (Makulová 2006 a) sme zdôraznili a detailne analyzovali faktory, ktoré majú vplyv na nájditeľnosť webových sídiel. Sú to predovšetkým:

- enormné a neustále narastajúce množstvo webových sídiel a digitálnych dokumentov (LYMAN – VARIAN, 2004),
- nízka úroveň informačnej gramotnosti používateľov informácií internetu,
- nedodržiavanie štandardov World Wide Web konzorcia (ďalej W3C) a zásad informačnej architektúry pri tvorbe webového sídla.

Pri dnešnom enormnom náraste webových sídiel už nestačí mať iba stránku v internete, ale je dôležité, aby stránka bola aj nájditeľná a podľa možnosti sa umiestnila na hľadané kľúčové slová na popredných miestach vo vyhľadávačoch. Ako uvádza aj Peter Morville (2002) v článku *The Age of Findability*, nájditeľnosť nie je limitovaná iba obsahom alebo webom. Nájditeľnosť znamená navrhnúť také systémy (vrátane fyzických), ktoré umožnia nájsť ľuďom potrebné informácie. **Nájditeľnosť sa vzťahuje na kvalitu, že sme lokalizovateľní.** Na úrovni položky môžeme hodnotiť do akého stupňa je jednoduché určitý objekt lokalizovať, prípadne objaviť. Na systémovej úrovni môžeme analyzovať ako fyzikálne, prípadne digitálne prostredie podporuje navigáciu a informačný prieskum. Nájditeľnosť nie je limitovaná iba na web, ale vzhľadom na to, že web obsahuje niekoľko miliárd položiek, začína

byť veľkou výzvou. Nájdiťnosť nie je synonymom informačnej architektúry, ale je jej cieľom zároveň s používateľnosťou, kredibilitou a prístupnosťou.

My sa budeme zaoberať predovšetkým nájdiťnosťou webových sídiel. Nájdiťnosť webového sídla je dôležitá, pretože stále viac návštevníkov prichádza na stránky z vyhľadávačov a ako výsledky prieskumov ukazujú, až 90 percent respondentov neprejde pri listovaní vo výsledkoch cez tretiu ponúkanú stranu (iProspect Search Engine User Behavior Study 2006).

Na nájdiťnosť vplýva predovšetkým informačná architektúra webových sídiel a optimalizácia webových sídiel pre prieskumové stroje (SEO – search engine optimisation), čiže vylepšenie pozície vo výsledkoch vyhľadávania fulltextových vyhľadávačov.

Informačnú architektúru chápeme v súlade s Louis Rosenfeldom a Petrom Morvillom nasledovne (2002, s. 4):

- kombinácia organizačnej, navigačnej a popisnej schémy informačného systému,
- štrukturálny dizajn informačného priestoru s cieľom uľahčiť vykonávanie úloh a umožniť intuitívny prístup k obsahu,
- umenie a veda štruktúrovania a klasifikovania webových sídiel a intranetu tak, aby ľudia dokázali nájsť a riadiť informácie,
- vznikajúca disciplína a spoločenská prax s cieľom zaviesť princípy dizajnu a architektúry do digitálneho sveta.

V odbornej literatúre sa stretávame s viacerými prístupmi k definovaniu SEO. Podľa Ramosa a Cotovej je SEO súbor metód použitých k získaniu vyššieho umiestnenia u vyhľadávacích nástrojov (RAMOS - COTA, 2004, s. 11). Thurow v publikácii Search Engine Visibility definuje SEO ako proces dizajnu, písania, kódovania (HTML), programovania a skriptovania webového sídla s cieľom získať vysoký ranking pre dané kľúčové slová v prieskumových strojoch s cieľom pomôcť potenciálnym zákazníkom nájsť webové sídlo (THUROW, 2003, s. 18).

SEO by sme nemali chápať iba v úzkom poňatí, ale SEO chápeme ako vedu a umenie zároveň, ktorej cieľom je získať vyššie umiestnenie pre definované kľúčové slová a frázy. Zdôrazňujeme predovšetkým etické princípy SEO optimalizácie, ktoré zdôrazňujú budovanie kvalitného obsahu a dodržiavanie štandardov W3C a zásad informačnej architektúry. Získanie vyššieho umiestnenia vyžaduje integrovaný prístup, pričom sa využívajú aj automatizované nástroje analyzujúce spôsob formulácie termínov používateľmi internetu (napr. Wordtracker, ETARGET a pod.), správanie používateľov pri vyhľadávaní informácií, analyzujú sa transakčné logy. Mimoriadny dôraz sa kladie

na čistý HTML kód stránky, ktorý prejde validátormi a nakoniec princípy informačného dizajnu pred dizajnom grafickým (Makulová 2006a).

Analýza faktorov SEO optimalizácie

V odbornej literatúre sa uvádza, že pri radení záznamov na výstupe Google berie do úvahy viac ako 100 faktorov. Stretávame sa s rôznorodou klasifikáciou. K zaujímavým prístupom patrí ich členenie na pozitívne a negatívne faktory, ktoré sa nachádzajú na stránke, alebo mimo stránky (Google Ranking Factors - SEO Checklist).

1. Pozitívne SEO faktory na stránke
2. Negatívne SEO faktory na stránke
3. Pozitívne SEO faktory mimo stránky
4. Negatívne SEO faktory mimo stránky

V súčasnosti sú údaje o takmer 118 faktoroch, ktoré vplyvajú na SEO optimalizáciu a tým aj na nájdniteľnosť webových sídiel. Pokúsime sa zhrnúť aspoň tie najdôležitejšie a ich využitie demonštrovať na príkladoch knižníc prihlásených do súťaže **TOP WebLib 2006**.

Pozitívne SEO faktory na stránke (on page faktory)

Kľúčové slovo v URL adrese

Robot vyhľadávacieho nástroja vysoko hodnotí, ak sa v URL adrese nachádza kľúčové slovo.

Pozitívne príklady:

<http://www.starlib.sk/sk/prirastky-novych-knih/>

<http://www.starlib.sk/sk/partneri-kniznice/>

<http://www.kniznicatrnava.sk/publikacie/index.html>

http://www.kniznica-cadca.sk/HTML/vyhodnotenie_cinnosti.htm

Negatívny príklad:

Novohradská knižnica v Lučenci <http://www.nklc.sk/> (URL adresa sa nemení, pretože je použitá technológia rámcov).

Kľúčové slovo v doméne

Robot vyhľadávacieho nástroja hodnotí vysoko pozitívne, ak sa nachádza kľúčové slovo v doméne.

Pozitívne príklady:

Kysucká knižnica, <http://www.kniznica-cadca.sk>

Galantská knižnica, <http://www.galantskakniznica.sk>

Kľúčové slovo v titule

Značka **Title** patrí k najdôležitejším a najvýznamovejším prvkom pre optimalizáciu webového sídla pre všetky prieskumové stroje. Prvok Title sa vkladá do hlavičky HTML dokumentu a vidíme ho ako popis stránky nad prehliadačom. Vyhľadávače berú Title ako kľúčový zdroj o obsahu stránky a preto by mal byť dostatočne výstižný. Informácia vložená v značke <title> sa ukladá aj ako informácia v obľúbených položkách internetového prehliadača.

Preto je veľmi dôležité, aby bol názov dostatočne zmysluplný a výstižný a jasne hovoril, o čom stránka pojednáva. Preto aj **rozdielne stránky webového sídla by mali mať rozdielne názvy**. Veľmi dôležité je umiestniť dôležité slovo na začiatok názvu, ideálne je začínať slovom, ktoré najviac vystihuje obsah stránky. Veľkou chybou, ktorú často nájdeme na slovenskom internete je názov „Vitajte na stránkach XY!“. Podobne v prípade viacjazyčných verzií by nadpis nemal začínať určitým a neurčitým členom (A, An, The, Die, Der, Das a pod.). Názvy dokumentov sa líšia od písania názvov pre papierové médiá predovšetkým v dvoch hlavných rozdieloch (Nielsen, 2002, s. 128):

- Online názvy sa často zobrazujú mimo kontext ako súčasť zoznamov vo vyhľadávaní, v systéme obľúbených záložiek a v iných navigačných pomôckach.
- Aj keď sa názov zobrazí v kontexte, treba brať do úvahy, že monitor neposkytuje také možnosti čítania ako klasické papierové médium.

Preto by sme pri písaní názvov dokumentov mali dodržiavať podľa Nielsena nasledovné zásady (2002, s. 128 - 129):

- Jasne povedať o čom webová stránka pojednáva pomocou zrozumiteľných výrazov.
- Používať jednoduchý jazyk a nie sofistikované názvy.
- Vyhybať sa názvom začínajúcim neurčitým členom, prvé slovo, prípadne skupina slov by mala byť dostatočne informačne bohatá.

- Nezačínať názvy všetkých stránok rovnakým slovom a nepoužívať rovnaké názvy na všetkých stranách webového sídla.

Pozitívne príklady:

Staromestská knižnica <http://www.starlib.sk>

<title>Staromestská knižnica - VÝPOŽIČNÝ PORIADOK</title>

<title>Staromestská knižnica - PRÍRASTKY NOVÝCH KNÍH</title>

Negatívne príklady:

<title>Úvod - Mestská knižnica mesta Piešťany</title> zbytočné slovo úvod,

<TITLE>Novohradská knižnica v Lučenci</TITLE> prvok Title je rovnaký na všetkých stránkach webového sídla a nemení sa ani URL adresa, pretože sa používa technológia rámcov.

<title>SEK EUBA</title> Titul Slovenskej ekonomickej knižnice Ekonomickej univerzity v Bratislave, je rovnaký na všetkých stránkach a obsahuje iba skratku.

Kľúčové slovo v metaprisku description

Ak sa nachádza v hlavičke dokumentu metaprisk description, Google ho využije pri zobrazení výstupného záznamu. Metaprisk description by mal odrážať obsah každej stránky webového sídla.

Pozitívne príklady:

Slovenská knižnica pre nevidiacich v Levoči <http://www.skn.sk>

<meta name="description" lang="sk" content="Slovenská knižnica pre nevidiacich Mateja Hrebendu v Levoči je jediným a jedinečným celonárodným knižnično-informačným centrom pre zrakovo hendikepovaných ľudí Slovenskej republiky i Slovákov žijúcich v zahraničí">

Negatívne príklady:

Kysucká knižnica v Čadci nemá žiadne metaprisky <http://www.kniznica-cadca.sk/>

Knižnica Józsefa Szinyeieho v Komárne nemá žiadne metaprisky <http://www.kjszkom.sk/>

Malokarpatská knižnica v Pezinku nemá žiadne metaprisky <http://www.kniznicapezinok.sk/>

<meta name="DESCRIPTION" content="Mestská knižnica mesta Piešťany" nedostatočne vystihuje činnosť knižnice. Môže mať až 200 znakov.

Slovenská ekonomickej knižnica Ekonomickej univerzity v Bratislave nemá žiadne metaprisky <http://www.sek.euba.sk/>

Kľúčové slovo v metaprívku keywords

Podobne ako v predchádzajúcom prípade, je výhodné ak sa nachádza v metaprívku maximálne 10 kľúčových slov. Dôležité je, aby sa kľúčové slová aj nachádzali v texte, v opačnom prípade sa považujú za spam. Podobne by sa kľúčové slovo nemalo vyskytovať viackrát.

Hustota kľúčových slov v tele textu

V odbornej literatúre sa uvádza, že optimálna hustota kľúčových slov by mala byť 5-20% (všetky kľúčové slová/celkový počet slov). V opačnom prípade sa väčšia hustota kľúčových slov už môže považovať za spam.

Individuálna hustota kľúčových slov

V ideálnom prípade by mala byť 1-6% (každé kľúčové slovo/celkový počet slov).

Kľúčové slová v záhlaviach H1, H2, H3

Prieskumový stroj hodnotí vysoko použitie kľúčových slov v záhlaviach H1 – H6.

K ďalším pozitívnym faktorom na stránke patrí proximita kľúčových slov (čím sú bližšie, tým lepšie), frázy kľúčových slov, prominencia kľúčových slov (najdôležitejšie by mali byť na vrchu stránky, podľa možnosti zvýraznené, alebo písané väčším fontom), kľúčové slová v alternatívnych popiskoch k obrázkom a pod.

K dôležitým faktorom patria aj interné spojenia v rámci navigácie a externé spojenia v rámci navigácie.

Navigácia – interné spojenia

Patria sem kľúčové slová v spojeniach, ktoré vedú na stránku, kľúčové slová v interných prepojeniach webového sídla, validita všetkých interných prepojení (dá sa overiť validátorom <http://www.dead-links.com/>).

Navigácia – externé spojenia

K veľmi dôležitým patria spojenia na externé stránky. Odkazovať by sme mali iba na kvalitné stránky. Text spojenia by mal byť dostatočne výstižný a validitu spojení by sme mali pravidelne kontrolovať. Spojení by nemalo byť viac ako 100.

Ďalšie pozitívne faktory na stránke

Trieda doménového mena

Vyšší ranking získavajú domény org, gov, edu a info, čo súvisí aj s kredibilitou webových sídiel.

Veľkosť súboru

Veľkosť súboru by nemala byť väčšia ako 100K. Preferujú sa súbory < ako 40K.

Spojenia v URL adrese

Na spojenie kľúčových slov v URL adrese je vynikajúce používať pomlčku. Treba dať pozor, aby ich nebolo viac ako 2, v prípade, že je ich viac ako 4, už sa to môže považovať za spam.

Aktualizácia webových sídiel

Často aktualizované webové sídla majú vyšší ranking ako sídla, ktoré sú aktualizované iba zriedkavo, respektíve nie sú aktualizované vôbec.

Aktualizácia – množstvo zmien v obsahu

Udáva sa ako pomer starých stránok k novým stránkam.

K ďalším dôležitým pozitívnym faktorom patrí aktualizácia spojení, dĺžka URL adresy, používanie synonymných termínov, vek webového sídla (čím staršie, tým lepšie), vek stránky versus vek webového sídla (nové stránky na staršom webovom sídla získavajú väčšiu pozornosť).

Negatívne faktory na stránke (on page faktory)

K negatívnym faktorom patria predovšetkým tie, ktoré ovplyvňujú čitateľnosť stránky robotom prieskumového stroja. K najčastejším prekážkam, ktoré bránia správne indexovať webové sídlo sú:

používanie Java Scriptu, technológie flash, tvorba úvodných (tzv. splash pages) vo flash animácii, technológia rámcov, využívanie obrázkov namiesto textov, používanie obrázkových máp, používanie rozbaľovacieho menu (malo by ísť iba o doplnkovú navigáciu), dynamické stránky s viac ako troma parametrami, stránky, ktoré vyžadujú na sprístupnenie vyplnenie formulárov, časté zmeny URL adries, duplicitné URL adresy, validnosť HTML kódu a veľkosť webovej stránky (Smička, 2004, s. 26 - 39).

Technológia rámcov

V našich príspevkoch sme viackrát zdôrazňovali, že W3C neodporúča využívať technológiu

rámčov, pretože prieskumové stroje nedokážu stránku zindexovať, stránka nespĺňa zásady prístupnosti a nedá sa urobiť značka v bookmarkoch (URL je konštattné.) Začiatkom marca technológiu rámcov využívala Novohradská knižnica v Lučenci (<http://www.nklc.sk/>) a Verejná knižnica Jána Bocatia, v Košiciach (<http://www.vkjb.sk/>).

Text je väčšinou prezentovaný iba v grafickej podobe

S podobnými prípadmi sme sa stretávali hlavne v minulosti. Dnes sa už našťastie uprednostňuje informačný dizajn pred dizajnom grafickým.

Pridružené webové sídla (affiliate site)

Išlo predovšetkým o webové sídla združované pod jedného vlastníka s úbohým obsahom. K typickým predstaviteľom patrili cestovné sídla a sídla ponúkajúce predaj kvetov.

Penalizácia za príliš veľkú optimalizáciu

Používa sa termín OOP (over optimisation penalty). K najčastejším patrí veľká hustota kľúčových slov, časté používanie prvku H1 s kľúčovým slovom a opakovanie kľúčových slov v metaprvkoch.

Spojenia na zlých susedov

Hlavne v prvých rokoch, keď sa do úvahy na výpočet relevancie bral iba počet spojení, vznikali tzv. farmy liniek, ktoré na seba odkazovali.

Kradnutie obrázkov alebo textu z iných domén

Predovšetkým Google kladie stále väčší dôraz na dodzrievanie Digital Millennium Copyright Act (<http://www.google.com/dmca.html>) a na svojej stránke uvádza aj podrobný návod, čo treba v prípade porušenia autorských práv spraviť.

Ďalšie negatívne faktory na stránke

K ďalším negatívnym faktorom na stránke patria presmerovanie návštevníkov prostredníctvom refresh metaprvkov, používanie nevhodných slov, prílišné prepájanie obsahu prostredníctvom spojení,

Pozitívne faktory mimo stránky (off page faktory)

PageRank

K pozitívnym faktorom mimo stránky patria predovšetkým všetky odkazy, ktoré smerujú na

stránku. Odkazy môžu byť z ďalších webov, ako aj z webu vlastného. Odkazy sa vo významnej miere podieľajú na PageRanku stránky (PR). PR, je obchodná značka ktorá patrí a je patentovaná spoločnosťou Google. PageRank bol vyvinutý na Stanfordskej univerzite Larry Pagom (odtiaľ pochádza meno Page-Rank) a Sergejom Brinom ako súčasť ich výskumného projektu nového druhu vyhľadávacieho nástroja. Projekt začal v roku 1995 a viedol k úspešnému prototypu Google v roku 1998.

Systém PR je založený na miliardách prichádzajúcich a odchádzajúcich odkazov, ktoré tvoria web. Ak web A odkazuje na web B, dáva mu svoj hlas a tým aj zvyšuje jeho dôležitosť. Čiže je to niečo ako systém citačného indexu. Google neberie do úvahy iba počet liniek, ktoré vedú na stránku, ale tiež aj aký má stránka rating. Ak je stránka autoritou v svojej oblasti, je hodnota spojenia vyššia ako zo stránky s nízkou popularitou. Podobne záleží tiež, na aký počet sídiel daná stránka odkazuje. PageRank nie je to isté ako umiestnenie stránky vo výsledkoch vyhľadávania (SERP – search engine results page). Tento závisí od obsahu stránky a celej rady ďalších faktorov. Page Rank stránky zistíme, ak si nainštalujeme Google Toolbar. Na Toolbare sa objavuje zelený prúžok, ktorý meria alebo boduje stránku od 0 do 10.

Uvádzame PageRank niektorých knižníc:

Galantská knižnica 3/10

Kysucká knižnica v Čadci 4/10

Staromestská knižnica 5/10

Mestská knižnica mesta Piešťany 4/10

Knižnica Juraja Fándlyho v Trnave 4/10

Novohradská knižnica v Lučenci 4/10

Slovenská ekonomická knižnica Ekonomickej univerzity v Bratislave, 5/10

Akademická knižnica Akadémie ozbrojených síl generála Milana Rastislava Štefánika, 3/10

Počet prichádzajúcich spojení na stránku

Meria sa niekedy aj ako popularita spojení, ktoré prichádzajú na stránku. Vysoká popularita stránky sa môže vo zvýšenej miere podieľať aj na zvýšenej návštevnosti a tým aj vo zvýšení rankingu daného sídla. Treba dávať pozor na to, ktoré stránky odkazujú na nás a nevyužívať tzv. farmy liniek, ktoré sa považujú za spam. Počet prichádzajúcich spojení môžeme zistiť aj na adrese <http://www.linkpopularity.com>

Staromestská knižnica

Google 26

MSN 254

Yahoo 1048

Slovenská ekonomická knižnica Ekonomickej univerzity v Bratislave

Google 9

MSN 457

Yahoo 561

Počet prichádzajúcich spojení zo strániek s vyšším PageRankom

Je dobré, ak na stránku vedú spojenia zo stránok s vyšším PageRankom, čiže zo stránok kvalitných. Na Slovensku má jeden z najvyšších PR systém NAJ.SK (7/10), systém YAHOO! má PR 9/10.

Text v spojení kotvy prichádzajúcej na stránku

Do úvahy sa berie aj text v kotve spojení prichádzajúcej na stránku. Text v spojeniach viedol aj často k tzv. google bomb. Termín Google bomb je internetový slangový výraz pre určitý druh útoku pre ovplyvnenie hodnotenia dôležitosti danej stránky vo výsledkoch poskytnutých vyhľadávačom Google, často so žartovnými či politicko-kritickými úmyslami. Podľa štýlu, akým funguje algoritmus vyhľadávača Google, sa stránka radí vyššie, ak iné stránky, ktoré na ňu odkazujú, používajú zhodný kotevný text. Google bomba sa vytvorí, ak veľké množstvo stránok odkazuje takýmto štýlom na inú stránku. Vyhľadávací algoritmus Googlu teda prihliada nielen na obsah stránok, ale aj na to, akým textom na ne odkazujú iné stránky, resp. ako cieľovú stránku popisujú iní ľudia. Tento postup má zvyšovať relevanciu výsledkov, ale možno ho aj zneužiť. V zahraničí sa doteraz najznámejšou obeťou podobného internetového útoku stal americký prezident, ktorého stránka sa istý čas objavovala ako prvá pri vyhľadaní spojení "strašidelné zlyhanie". Webová stránka Microsoftu bola zasa dostupná cez frázu "diabolskejší ako sám Satan". Fico sa tak stal prvou slovenskou obeťou "Google bombingu", keď sa jeho stránka nachádzala na prvom mieste na spojení "nechutný populista" (Bella 2004).

K ďalším pozitívnym faktorom mimo stránky patria vek spojení (podľa Google čím starší, tým lepšie), frekvencia zmien textu kotvy spojení na stránke (podľa Google nie je dobré často meniť text kotvy), popularita referujúcej stránky, počet odkazov z referujúcej stránky (ak vedie menej spojení, je to lepšie, pretože sa rozdelí ranking na menej stránok), pozícia spojení na referujúcej stránke (čím

vyššie tým lepšie), hustota kľúčových slov na referujúcej stránke (týka sa hľadaných spojení), obsah prvku titul referujúcej stránky (ide o spoločný obsah), spojenie zo stránky, ktorá sa považuje za experta v danej tematickej oblasti, rovnaká téma obsahu referujúcej stránky (v prípade rozdielnej témy sa považuje odkaz za spam), spojenie využívajúce technológiu javascriptu a obrázkovej mapy (nevýhodné). K dôležitým pozitívnym faktorom patrí aj vek webového sídla (čím staršie, tým lepšie), stromová štruktúra webového sídla, ak sa v rámci navigácie nachádza aj mapa sídla s vhodnými kľúčovými slovami, využívanie tezaurov, príbuzných termínov a pod.

Tvorba mapy sídla

Mapa sídla predstavuje významný prvok systému navigácie na stránke. Mapa sídla je vo všeobecnosti prehľadne zoradený zoznam všetkých stránok, ktoré na sídle môžeme nájsť. Často nepredstavuje len jednoduchý súpis stránok, ale vyjadruje predovšetkým vzťahy medzi nimi a prináša metainformácie o týchto stránkach. V prípade sofistikovanejších máp sídiel sú jednotlivé stránky sídla zgrupované do tematických kategórií, môžu byť odlišované aj farebne. Ich cieľom je, aby používateľ sa v sídle rýchlo orientoval a pochopil štruktúru sídla.

Dobrá mapa sídla by mala byť predovšetkým prehľadná, kompletná, nie príliš rozsiahla, rýchlo načítateľná a pravidelne aktualizovaná. Mapa má veľký význam pre roboty prieskumových strojov. Niektoré roboty môžu mať napríklad naprogramovanú hĺbku indexovania sídla iba na dve úrovne z hlavnej stránky. Teda stránky od tretej úrovne by neboli vôbec zindexované. Mapa stránky takto môže nasmerovať robota aj na hlbšie umiestnené stránky. Dobré redakčné systémy by mali automaticky generovať mapu sídla po každej aktualizácii.

Pozitívne príklady mapy sídla:

Galantská knižnica <http://www.galantskakniznica.sk/sitemap/>

Zemplínska knižnica Gorazda Zvonického <http://www.zkgz.sk/sitemap.html>

Staromestská knižnica <http://www.starlib.sk/sk/mapa-stranok/>

Knižnica Juraja Fándlyho v Trnave <http://www.kniznicatrnava.sk/sitemap.html>

Slovenská lesnícka a drevárska knižnica TU vo Zvolene
<http://www.tuzvo.sk/showdoc.do?docid=2232>

Na stránkach sa nenachádzala mapa sídla

Kysucká knižnica v Čadci <http://www.kniznica-cadca.sk/index.htm>

Mestská knižnica mesta Piešťany <http://www.kniznica.sk/>

Akademická knižnica AOS <http://aos.valm.sk/kniznica/>

Štátna vedecká knižnica Košice <http://www.svkk.sk/index.php>

Stránke je registrovaná v adresároch internetu

Za vysoko pozitívny faktor sa pokladá, ak je stránka registrovaná v adresároch internetu, predovšetkým v DMOZ adresári, alebo tzv. Open Directory Project (<http://www.dmoz.org>). Adresár DMOZ vznikol ako odpoveď na adresár YAHOO!, kde bolo treba čakať veľmi dlho na zaradenie do adresára. V adresári DMOZ bolo vo februári 2007 vyše 4 miliónov sídiel, o ktoré sa staralo 74 719 redaktorov vo viac ako 590 000 kategóriách. Uvedený adresár poskytuje výsledky vyhľadávania aj Googlu a preto sa stránka uvedená v adresári získava vysoký rating. V ostatnom čase sa objavuje tiež veľa námietok voči adresáru DMOZ za príliš zdĺhavé zaradenia stránok. Záleží to samozrejme od redaktora príslušnej časti adresára.

Zo stránok slovenských knižníc boli v adresári DMOZ začiatkom marca 2007 niektoré knižnice:

V časti Top: World: Slovensky: Zdroje

Staromestská knižnica v Bratislave

Zemplínska knižnica Gorazda Zvonického, Michalovce

Verejná knižnica Mikuláša Kováča v Banskej Bystrici

V časti Top: World: Slovensky: Regionálne: Európa: Slovensko: Bratislavský kraj: Bratislava: Vzdelávanie

Staromestská knižnica

Univerzitná knižnica v Bratislave

K ďalším dôležitým adresárom, kde sa dá zaregistrovať patrí YAHOO!

Zo stránok slovenských knižníc boli začiatkom marca 2007 v adresári knižnice:

V časti Directory > Regional > Countries > Slovakia > Reference>Libraries > Academic Libraries

Technical University of Košice Library

V časti Directory > Regional > Countries > Slovakia > Reference>Libraries > National Libraries

Slovenská národná knižnica

Správanie sa používateľov pri návšteve stránky

K dôležitým faktorom patria aj výsledky analýzy správania sa používateľov pri návšteve webového sídla (vzrastajúci počet návštevníkov, kľúčové slová prostredníctvom ktorých nájdu používatelia stránku, čas, ktorý strávia na stránke a pod.).

Údaje o správaní sa návštevníkov môžeme spoľahlivo zistiť, ak stránku zaregistrujeme v systéme NAI.sk (www.naj.sk) na meranie a analýzu návštevnosti stránok. K dispozícii sú rozsiahle štatistiky zahrňujúce základné informácie o návštevnosti, ako napríklad počet návštev, počet návštevníkov, počet zobrazených stránok, počet aktívnych návštevníkov, informácie o tom, odkiaľ ľudia na stránky prichádzajú a ako sa správajú, aké používajú prehliadače, aké majú technické možnosti (podpora flash, java, cookies a pod.), ako a odkiaľ sa k internetu pripájajú a veľa ďalších ukazovateľov spracovaných v prehľadných tabuľkách a grafoch. Tieto údaje možno ďalej využiť na optimalizáciu webových sídiel pre prehliadače, zlepšenie informačnej architektúry, určenia vyhľadávacích zón a pod. Na nasledujúcom obrázku vidíme, z ktorých vyhľadávačov prichádzajú návštevníci najčastejšie na Portál ku katalógom a zbierkam slovenských knižníc, ktorý je jedným z výstupov projektu KIS3G (www.kis3g.sk).



Obr. 1 Z ktorých vyhľadávačov prichádzajú ľudia najviac na portál projektu KIS3G

Ako vidíme na obrázku, zo 6 185 návštev až 65,92% používateľov prichádza z vyhľadávača Google Slovensko. K ďalším vyhľadávačom s menším zastúpením patria Google.com (8,5%), Zoznam.sk (fulltext 7,18%) a Zoznam.sk (katalóg 7,07%).

K dispozícii sú aj mesačné reporty poskytujúce celý rad užitočných informácií, ako napríklad počítače návštevníkov podľa regiónov v SR, podľa spôsobu pripojenia, podľa poskytovateľov pripojenia, informácie o tom, koľko času strávia na stránke a na webe, koľko stránok navštívia za jednu návštevu, ktoré sú najčastejšie hľadané spojenia a pod.

Registrácia domény

K pozitívnym faktorom ovplyvňujúcim umiestnenie stránky patrí aj čas registrácie domény. Čím je doména registrovaná dlhšie, tým lepšie.

Negatívne SEO faktory mimo stránky (off page faktory)

Mnohé z negatívnych faktorov sa považujú za techniky spamu, ako napríklad kupovanie spojení, umelé zvyšovanie návštevnosti, cloaking, kradnutie obsahu webového sídla (domain hijacking), a pod.. Za negatívny faktor sa berie ak na stránku nevedú žiadne spojenia, nespoľahlivosť servera. Vzhľadom na stále väčší počet neetickej SEO optimalizácie sa využívajú spam reporty. Ide o stránku, prostredníctvom ktorej sa nahlasujú vyhľadávačom neetické techniky. Väčšina vyhľadávačov ich má na svojej stránke, ako napríklad Google <http://www.google.com/contact/spamreport.html> (obr. 2)

Report a Spam Result

Exact query that shows a problem (copy this from the Google search box):

Resulting Google page that shows problem (copy the Google URL):

The specific web page or site that is misbehaving:

Type(s) of problem (check all that apply):

- Hidden text or links
- Misleading or repeated words
- Page does not match Google's description
- Cloaked page
- Deceptive redirects
- Doorway pages
- Duplicate site or pages
- Other (specify)

Additional details:

Submit

Obr. 2 Spam report vo vyhľadávači Google

Techniky týkajúce sa celého webového sídla alebo tzv. on-site techniky

V poslednom čase sa v odbornej literatúre stretávame aj s tzv. on-site technikami, ktoré sa týkajú optimalizácie celého webového sídla. Veľmi pekne ich popisuje Dan Thies (2004). Súvisí to

predovšetkým s čoraz zložitejšou informačnou architektúrou sídiel a zložitejším obsahom. Ide o to, že prieskumové stroje identifikujú v rámci webu určité tematicky vyhranené zhluky a pre každý zhluk určujú najrelevantnejšie webové sídla. Čiže my si môžeme pre konkrétne webové sídlo určiť tematické zhluky, do ktorých patrí. Potom je dôležité, aby sme budovali kvalitný obsah webového sídla so správnym rozsvtením a logickým prepájaním jednotlivých vrstiev. Zatiaľ čo sa v minulosti optimalizovala hlavne úvodná stránka, dnes je trendom optimalizovať pre zvolené kľúčové slová stránky tretej úrovne, na ktoré vedie odkaz zo stránok druhej úrovne, alebo z tzv. roadmap pages. Každá stránka tretej úrovne odkazuje na svoju roadmap stránku druhej úrovne a na titulnú stránku. Stránka tretej úrovne je optimalizovaná pre 5 kľúčových slov. Ak je to dôležité, vytvoríme ešte stránku štvrtej úrovne, pre ktorú bude už typický hĺbkový obsah týkajúci sa kľúčových slov 3 úrovne.

Počet jednotlivých vrstiev závisí od veľkosti a tematického zamerania webového sídla.

Záver

V článku sme analyzovali faktory, ktoré sa vo veľkej miere podieľajú na zlepšení nájditel'nosti a tým logicky aj vo zvýšenej návštevnosti webových sídiel knižníc. Už z množstva faktorov je jasné, že ide o veľmi zložitú problematiku, ktorá sa nedá vyriešiť jednorazovo. Zlepšenie nájditel'nosti, ktoré úzko súvisí so SEO optimalizáciou, ale aj celkovou informačnou architektúrou webového sídla je dlhodobý proces. Odporúča sa dodržiavať smerníc pre webmastrov, ktoré sú publikované na stránkach vyhľadávačov (Webmaster Guidelines 2007).

Napríklad najväčší vyhľadávač Google uvádza smernice týkajúce sa dizajnu, obsahu a techniky. Špeciálnu pozornosť venuje smerniciam kvality, ktoré by sme mohli zhrnúť nasledovne:

- Vytvárajte stránky pre používateľov a nie pre prieskumové stroje, ktoré nezavádzajte. Nesnažte sa im prezentovať in typ stránky (cloaking).
- Vyhýbajte sa trikom, ktoré zvyšujú umiestnenie stránky vo výsledkoch vyhľadávania.
- Vyhýbajte sa pochybným návodom na zvýšenie PageRanku. Môže to vyvolať presne opačný účinok.
- Nevyužívajte neautorizované programy na registráciu stránok a kontrolu rankingu (napr. WebPosition Gold™).
- Vystríhajte sa použitia skrytého textu alebo skrytých spojení.
- Nepoužívajte cloaking alebo redirect.

- Nepoužívajte stránky s nerelevantnými slovami.
- Na stránkach prezentujte obsah, ktorý je pridanou hodnotou pre používateľa a motivuje ho k opätovnej návšteve Vašej stránky.

V ostatnom čase sa veľa hovorí o vplyve personalizácii na SEO optimalizáciu (Bruemmer 2007). V podstate ide o to, že vo výsledkoch vyhľadávania sa používateľovi ponúkajú záznamy, ktoré vyhovujú jeho personalizačnému profilu na základe automatickej analýzy histórie vyhľadávania. Napríklad každý, kto využíva služby Google vyžadujúc Google Account (Gmail, AdSense, Google Analytics a pod.) je automaticky zapísaný do troch produktov: Search History, Personalized Search a Personalized Homepage. Podľa výsledkov prieskumu Choice Stream Personalization Survey realizovanom v roku 2006 až 79% respondentov si praje dostávať personalizovaný obsah. Je to tiež jedno z možných riešení problému informačného preťaženia používateľov.

Bibliografické odkazy

1. BELLA, T. 2004. Google: Fico je "nechutný populista". In POČÍTAČE SME, 19. 5. 2004 [cit. 2007-02-26]. Dostupné na internete: <http://pocitace.sme.sk/clanok.asp?cl=1500313>
2. BRUEMMER, C. FEATURE FOCUS: The Impact of Personalization on SEO. 15 February 2007. [cit. 2007-02-26]. Dostupné na internete: <http://www.bruceclay.com/newsletter/volume26/seopersonalization.html>.
3. Google Ranking Factors - SEO Checklist. [cit. 2007-02-26]. Dostupné na internete: <http://www.vaughns-1-pagers.com/internet/google-ranking-factors.htm#positive-on>
4. iProspect Search Engine User Behavior Study. April 2006. [cit. 2007-02-26]. Dostupné na internete: http://www.iprospect.com/premiumPDFs/WhitePaper_2006_SearchEngineUserBehavior.pdf
5. LYMAN, P., VARIAN, H. 2004. How Much Information 2003? School of Information Management and Systems at the University of California at Berkeley. [cit. 2007-03-03]. Dostupné na internete: <http://www.sims.berkeley.edu:8000/research/projects/how-much-info-2003/>.
6. MAKULOVÁ, S. 2005a. Návrh riešenia problémov pri vyhľadávaní informácií v internete alebo od kvantity ku kvalite. In: Knihovna. 2005. Roč.16, č. 1, s. 23 - 43.
7. MAKULOVÁ, S. 2005b. Hodnotenie webových stránok. Ikaros [online]. 2005, č. 11 [cit. 2007-01-31]. Dostupné na internete: <http://www.ikaros.cz/node/2036>. ISSN 1212-5075.
8. MAKULOVÁ, S. 2005c. Prečo je potrebné dodržiavať štandardy World Wide Web konzorcia. In

IT lib. roč. 9 (2005), č. 3, s. 54-55.. [cit. 2007-03-03]. Dostupné na internete:

<<http://www.cvtisr.sk/itlib/itlib053/makulova.htm>>

9. MAKULOVÁ, S. 2006 a. Analýza faktorov vplyvujúcich na nájdiťnosť webových sídiel. In: Information Use in Information Society. International Conference. Bratislava, Slovakia, October 10-11, 2006. Bratislava : Centrum VTI SR, 2006, s. 109 – 118.
10. MAKULOVÁ, S. 2006 b. Systém navigácie na WWW stránkach. In IT lib. roč. 10 (2006), č. 2, s. 37-41. [cit. 2007-03-03]. Dostupné na internete:
<<http://www.cvtisr.sk/itlib/itlib062/makulova2.htm>>
11. MORVILLE, P. The Age of Findability. In Semantic Studios, April 29, 2002. [cit. 2007-02-25]. Dostupné na internete: <<http://semanticstudios.com/publications/semantics/000007.php>>
12. NIELSEN, J. Web.Design. Průvodce od Jakoba Nielsena. Praha : SoftPress, 2002, 382 s.
13. PAPÍK, R. 2001. Vyhledávání informací II. Uživatelské rozhraní a vlivy oboru “human-computer interaction”. In Národní knihovna, 2001, roč.12, č.2, s. 81-90. [cit. 2007-03-03]. Dostupné na internete: <<http://full.nkp.cz/nkk/NKKR0102/0102081.html>>.
14. RAMOS, A., COTA, S., 2004. Insider’s Guide to SEO. How to Get Your Website to the Top of the Search Engines. California : Jain Publishing Company, 2004. 114 s.
15. ROSENFELD, L., MORVILLE, P. 2002. Information Architecture for the World Wide Web. 2. vyd. Sebastopol : O’Reilly&Associates, 2002, 461 s. ISBN 0-596-00035-9.
16. SMÍČKA, R. 2004. Optimalizace pro vyhledávače – SEO. Jak zvýšit návštěvnost webu. Dubany : Jasmínka, 2004. 120 s.
17. Spamdexing [online]. Wikipedia: The Free Encyclopedia. 21. máj 2006 [cit. 2007-03-03]. Dostupné na internete: <<http://en.wikipedia.org/wiki/Spamdexing>>
18. THIES, Dan. Search Engine Fast Start. A Step By Step Guide For Busy People. 2004. [cit. 2007-02-25]. Dostupné na internete: <<http://www.mattcas.com/resources/dl/SEO%20Fast%20Start.pdf>>
19. THUROW, S. 2003. Search Engine Visibility. Berkeley : New Riders Publishing, 2003. 297 s.
20. TKÁČIKOVÁ, D. 2004. Kvalitní dokument jako základ účinného vyhledávání informací. In Informace na dlani 2004. Inforum 2004 : 10. konference o profesionálních informačních zdrojích 25.-27. května 2004, Vysoká škola ekonomická Praha.[CD-ROM]. Praha : Albertina icome, 2004. ISSN 1214-1429
21. Webmaster Guidelines 2007. [cit. 2007-03-03]. Dostupné na

internete:<<http://www.google.com/support/webmasters/bin/answer.py?answer=35769>>

22. ZELDMAN, J. 2004. Tvorba webu podle standardu XHTML, CSS, DOM, ECMAScript a dalších. Brno : Computer Press, 2004, 410 s.

23. Príspevok bol spracovaný v rámci vedeckého grantového projektu VEGA 1/2481/05 Využívanie informácií pri informačnom správaní vo vzdelávaní a vede.