

Optimalizace pro vyhledávače - SEO

Jak zvýšit návštěvnost webu

Ing. Radim Smička

Předmluva

Jelikož papírová verze knihy je již vyprodaná, rozhodl jsem uvolnit tuto knihu v elektronické podobě. Elektronická podoba však neznamená, že lze části knihy kopírovat bez uvedení zdroje nebo bez povolení autora. Pokud shledáte zde uvedené informace za užitečné, budu rád pokud se o této knize zmíníte svým přátelům, ať již ústně nebo třeba na svém blogu.

Praktické ukázky jsou použity z [inzertního serveru Bazos.cz](http://www.bazos.cz) a z [internetového knihkupectví Jasminka.cz](http://www.jasminka.cz) z roku 2004. Od té doby se Bazos rozrostl o více sekcí a slovenskou verzi. Průměrná návštěvnost podle serveru Toplist činí 14000/den. Slovenská verze má pak návštěvnost 2200/den. Knihkupectví Jasminka v současnosti navštěvuje 1200 návštěvníků/den.

The screenshot shows the homepage of Bazos.cz, a Czech classifieds website. The browser window title is "Bazos.cz - Inzerce zdarma, inzeráty, bazar - Opera". The page features a navigation menu with categories: PC, Elektro, Mobil, Auto, Zvířata, Realty, Sport, Knihy, Práce, Děti, Hudeba, Nábytek, Ostatní, and a "Přidat inzerát" button. Below the menu is a grid of 12 category tiles, each with an icon and a list of sub-categories:

- PC**: Notebooky, PC, digitální fotoaparáty, mobily, grafické karty, processory ...
- Auto**: Škoda, Opel, Renault, Náhradní díly, Havarovaná, Motocykly, Skútry ...
- Zvířata**: Psi, Kočky, Koně, Akvarijní ryby, Terarijní zvířata, Chovatelské potřeby ...
- Knihy**: Beletrie, Časopisy, Skripta, Mapy, Kuchařky, Dětská literatura, Učebnice ...
- Realty**: Prodej Bytů 3+1, Prodej Domů, Pronájem Garáže, Podnájem, spolubydlení ...
- Sport**: Horská kola, Silniční kola, Lyže, Inline brusle, Fotbal, Golf, Vodní sporty, potápění ...
- Mobily**: Alcatel, LG, Motorola, Nokia, Samsung, Siemens, Sony Ericsson ...
- Hudeba**: Bicí nástroje, Strunné nástroje, DVD, CD, MC, LP, Vstupenky, Zkušebny ...
- Práce**: Administrativní, Logistika, Finance a ekonomika, Brigády, Řemeslné práce ...
- Elektro**: Autorádía, Domácí kina, Televizory, DVD přehrávače, Ledničky, Sekáčky ...
- Děti**: Autosedačky, Kočárky, Přiblížovač, Hračky, Kojenecké potřeby ...
- Nábytek**: Jídelní kouty, Křesla a gauče, Kuchyně, Postele, Sedací soupravy ...

At the bottom left, it says "Bazos - Inzertní server. Ref:2003. Nápořádá, O nás, Fórum, Reklama, Podmínky, RSS". At the bottom right, it says "16 KB".

1. ÚVOD	7
2. HISTORIE INTERNETOVÝCH KATALOGŮ A VYHLEDÁVAČŮ	9
2.1. ROZDÍL MEZI KATALOGY A VYHLEDÁVAČI.....	9
2.1.1. <i>Katalogy</i>	9
2.1.2. <i>Fulltextové vyhledávače</i>	10
2.2. HISTORICKÝ VÝVOJ.....	12
2.3. SOUČASNÉ TRŽNÍ PODÍLY VYHLEDÁVAČŮ.....	18
2.4. ČESKÉ VYHLEDÁVAČE A KATALOGY.....	19
3. PŘEDPOKLADY PRO OPTIMALIZACI	25
3.1. JAVASCRIPT.....	25
3.2. FLASH.....	26
3.3. SPLASH PAGE.....	26
3.4. RÁMCE.....	27
3.5. KLIKACÍ OBRÁZKOVÉ MAPY.....	29
3.6. ROZBALOVACÍ MENU.....	30
3.7. PROBLÉM S DYNAMICKÝMI STRÁNKAMI.....	30
3.7.1. <i>Praktický příklad optimalizace dynamických stránek</i>	31
3.7.2. <i>Přepisování URL pro Windows IIS</i>	33
3.8. SESSION ID.....	34
3.9. ČASTÉ ZMĚNY V URL.....	34
3.10. AUTOMATICKÉ PŘESMĚROVÁNÍ.....	36
3.11. DUPLICITNÍ URL.....	36
3.12. VALIDNOST HTML KÓDU.....	37
3.13. VELIKOST STRÁNKY V KB.....	38
4. METODY V SEO	39
4.1. OPTIMALIZACE OBSAHU STRÁNKY - ON PAGE FACTORY.....	39
4.1.1. <i>Klíčová slova (keywords)</i>	39
4.1.2. <i>Výběr klíčových slov</i>	40
4.1.3. <i>Nástroje na vyhledávaná slova</i>	41
4.1.4. <i>Konkurenčnost klíčového slova</i>	45
4.1.5. <i>Stop slova</i>	45
4.1.6. <i>Škodící slova (poison words)</i>	46
4.1.7. <i>Dobrá a špatná návštěvnost (traffic)</i>	46
4.1.8. <i>Konverzní poměr</i>	47
4.1.9. <i>Důležité tagy na jednotlivé stránce</i>	47
4.1.10. <i>Optimální počet klíčových slov v textu stránky</i>	51
4.1.11. <i>Klíčová slova na začátku tagu (stránky)</i>	52

4.1.12.	<i>Klíčová slova v URL</i>	53
4.1.13.	<i>Vlastní obsah</i>	53
4.2.	OFF PAGE FAKTORY	54
4.2.1.	<i>PageRank</i>	54
4.2.2.	<i>Algoritmus PageRanku</i>	55
4.2.3.	<i>Model náhodného surfaře</i>	55
4.2.4.	<i>Charakteristika PageRanku</i>	56
4.2.5.	<i>Toolbar PageRank</i>	58
4.2.6.	<i>Skutečný PageRank</i>	58
4.2.7.	<i>Tématicky a předmětově orientovaný PageRank</i>	59
4.2.8.	<i>Model inteligentního surfaře</i>	60
4.2.9.	<i>Topic-Sensitive PageRank</i>	61
4.2.10.	<i>SEO a tématický PageRank</i>	63
4.2.11.	<i>BadRank a PageRank 0</i>	63
4.2.12.	<i>Rozdělení interního PageRanku</i>	64
4.2.13.	<i>Rankingy ostatních vyhledávačů</i>	69
4.2.14.	<i>Mapa webu</i>	70
4.2.15.	<i>Jak získávat zpětné odkazy</i>	70
4.3.	ANCHOR TEXT (LINK TEXT)	72
4.3.1.	<i>Okolí odkazu</i>	74
4.3.2.	<i>Obrázkové versus textové odkazy</i>	74
4.3.3.	<i>Google bomba</i>	75
4.4.	OSTATNÍ OFF-PAGE FAKTORY	77
4.4.1.	<i>Klikavost na vyhledávané slovo</i>	77
4.4.2.	<i>robots.txt</i>	78
4.4.3.	<i>ErrorDocument 404</i>	80
4.5.	JAK ANALYZOVAT VÝSLEDKY OPTIMALIZACE	82
5.	ZAKÁZANÉ METODY V SEO	84
5.1.	CO LZE POVAŽOVAT ZA SPAM	84
5.1.1.	<i>Skrytý text nebo odkazy</i>	84
5.1.2.	<i>Malý text</i>	84
5.1.3.	<i>Klamná a často opakovaná slova</i>	85
5.1.4.	<i>Cloaking</i>	86
5.1.5.	<i>Odkazové farmy</i>	86
5.1.6.	<i>Klamné přesměrování</i>	86
5.1.7.	<i>Doorway pages</i>	87
5.1.8.	<i>Duplicitní web nebo stránky</i>	87
5.1.9.	<i>Tapetování katalogů duplicitními doménami</i>	88
5.1.10.	<i>Jiné klamné taktiky vedoucí k zmanipulování vyhledávačů</i>	88
5.1.11.	<i>Cybersquatting</i>	89

5.2.	SPAM REPORT	90
5.3.	NA CO SI DÁVAT POZOR	90
5.3.1.	Zaregistrování stránek do tisíců vyhledávačů	91
5.3.2.	Trvalé pozice ve vyhledávačích	91
5.3.3.	Garantované umístění	91
5.3.4.	Analýzy pozice ve vyhledávačích	91
5.3.5.	Lepší pozice bez zásahu do obsahu	92
5.3.6.	Reference	92
5.3.7.	Výměna zakázaného odkazu	92
6.	PRAKTICKÁ UKÁZKA POSTUPU SEO	93
6.1.	OPTIMALIZACE SERVERU BAZOS.CZ	94
6.2.	OPTIMALIZACE SERVERU JASMINKA.CZ	100
7.	ZÁVĚR	105
	SEZNAM POUŽITÉ LITERATURY	107
	SLOVNÍČEK	111
	PŘÍLOHY	121

1. Úvod

Internetové vyhledávače se v poslední době stávají stále častějším způsobem hledání informací. Téměř každý, kdo pracuje s internetem, nějaký vyhledávač použil. Tím jak zviditelnit webovou stránku ve vyhledávačích se zabývá Search Engine Marketing (SEM). Skládá se ze dvou základních částí, placené a neplacené. Rozdíl mezi nimi samozřejmě není nutně v celkových finančních nákladech, ale v tom že v placené části SEM se platí za zviditelnění přímo vyhledávači, kdežto v neplacené se platí pouze člověku (firmě), který optimalizaci provádí a ne vyhledávači. Neplacená část se nazývá optimalizace pro vyhledávače (SEO - Search Engine Optimization), někdy také pouze přirozená optimalizace.

Způsobem jak zlepšit svou pozici ve vyhledávačích se lidé zabývali již od vzniku prvního vyhledávače, respektive od vzniku vyhledávače, který začal řadit výsledky vyhledávání podle jiných kritérií, než podle abecedy nebo datumu. V České republice však byla tato možnost zvýšení návštěvnosti v minulosti dosti opomíjena, nebo alespoň se o ní tolik nemluvilo. Částečně to bylo způsobeno tím, že vyhledávače se u nás začaly více rozšiřovat daleko později než v zahraničí, např. ve Spojených státech. Používání vyhledávačů u nás neustále roste, avšak i v současnosti ještě řada lidí, kteří běžně pracují s internetem, prochází při vyhledávání nějaký katalog a neznají možnosti vyhledávače.

Jelikož SEO je v ČR velkou neznámou (do této doby neexistovala žádná česká literatura, která by se daným tématem zabývala), je cílem této knihy popsat tuto oblast v češtině a dokázat, že je to účinná možnost jak zvýšit návštěvnost webových stránek. Důkaz účinnosti SEO bych rád provedl v poslední kapitole na svém vlastním inzertním serveru Bazos.cz a na serveru internetového knihkupectví Jasminka.cz. Pomocí serveru Jasminka.cz bych chtěl také

dokázat, že je to nejenom možnost jak zvýšit návštěvnost, ale také jak zvýšit obrat internetového obchodu.

V této knize se nezabývám placenou částí Search Engine Marketingu (mění se podle aktuální nabídky vyhledávačů) ani srovnáním ceny, kterou stojí přirozená optimalizace, a ceny za inzerci ve vyhledávačích. Základem ceny za optimalizaci by byl čas zaměstnance, který má na starosti SEO (nebo cena externí firmy) a základem pro placenou inzerci je ceník vyhledávačů. Tato kniha je určena pro všechny, které tato oblast zajímá, zvláště pak pro tvůrce webových stránek.

2. Historie internetových katalogů a vyhledávačů

2.1. Rozdíl mezi katalogy a vyhledávači

Tato kniha se věnuje hlavně vyhledávačům. Rád bych ale nejdříve definoval co to vlastně fulltextové vyhledávače jsou a jak se liší od katalogů.

2.1.1. Katalogy

Katalog je web, který obsahuje celou řadu odkazů na jiné webové stránky. Odkazy jsou zpravidla řazeny do tématických oblastí. Mezi nejznámější české katalogy patří Seznam, Centrum a Atlas, mezi zahraniční Yahoo! a ODP.

Záznam do katalogu se provádí registrací do příslušné sekce (ta která se věnuje danému tématu). Po zkontrolování editorem příslušné sekce je záznam přidán do databáze katalogu. V katalogu se hledá buď procházením jednotlivých sekcí anebo pomocí jednoduchého dotazu do vyhledávacího políčka. Pokud se hledá pomocí dotazu, katalog porovnává zadaný dotaz s URL, nadpisy a popisky stránek, které jsou v katalogu registrovány a také s kategorií do níž je odkaz zařazen. Pokud je stránka například v kategorii Inzertní servery je pravděpodobné, že ji katalog při vyhledávání slova inzerce přidá plusové body. Vyhledávání v katalogu nebere v potaz obsah stránky.

Možností jak ovlivnit pozici stránky při vyhledávání v katalogu je tedy hlavně klíčové slovo v nadpisu stránky (popisek má nižší i když nezanedbatelnou váhu). Výběru klíčových slov se věnuji v kapitole č. 4.1.2. Výběr klíčových slov. Problémem je, že editoři dosti často ponechávají v nadpisu pouze název firmy nebo URL. Výhodu v katalogu tak v drtivé většině mají webové stránky, které mají obecný název domény (knihkupectvi.cz, bazar.cz ...).

V popisku stránky je vhodné mít co nejvíce klíčových slov, které mají s daným tématem něco společného. Je nutné se tedy vyhnout slovům typu: internetové stránky, nejrychlejší, nejlepší atd. Editoři je z názvu a popisku odstraňují, protože nepřinášejí žádnou užitečnou informaci. Klíčová slova a fráze v popisku se zvažují stejně důkladně jako v nadpisu. Katalogy v drtivé většině totiž neumí skloňovat ani hledat podobná slova. Co není v přesném tvaru jak to uživatel hledá, katalog nenajde. Samozřejmě popisek musí být složen z vět a ne jen z prostého výčtu klíčových slov. Takovou stránku by editor sekce neměl do katalogu zařadit.

V jednotlivých sekcích se stránky poslední dobou řadí převážně podle hodnocení fulltextového vyhledávače (placené odkazy mají ovšem přednost). Například Seznam řadí podle S-Ranku a Atlas řadí záznamy podle JyxoRanku. Také Centrum používá k řazení odkazů ranking ze svého fulltextového vyhledávače Morfeo.

V katalogu může mít web v jedné kategorii vždy pouze jeden zápis (nebo by alespoň měl). Registrování podstránek, které mají unikátní obsah odlišný od již zaregistrované webové stránky je v jiné sekci většinou povoleno. Co všechno může být v nadpisu, v popisku a do kolika sekcí je možné stránku zařadit mají katalogy napsané v obecných pravidlech registrace. Zlepšování pozice v katalozích se nazývá optimalizace pro katalogy.

2.1.2. Fulltextové vyhledávače

Vyhledávač je software, který hledá a stahuje dokumenty (webové stránky, dokumenty textových editorů, PDF atd.), které indexuje (zpracovává a ukládá) do své databáze. V této databázi pak umožňuje návštěvníkům svých webových stránek vyhledávat pomocí jednoduchých i složitých dotazů. Mezi nejznámější zahraniční vyhledávače patří Google, mezi české pak Jyxo a Morfeo. Protože

rtivá většina současných vyhledávačů prohledává celý dokument, jedná se o vyhledávače fulltextové.

Každý fulltextový vyhledávač se skládá ze dvou částí. První část je robot (také se nazývá crawler, spider, bot nebo pavouk) a druhá část je webové rozhraní. Robot má na starosti procházení webu, stahování souborů a jejich indexaci. Robota by šlo dále dělit ještě na getter, který stahuje soubory a na indexer, který má za úkol soubory zpracovávat a ukládat do databáze. Robot se po webu pohybuje úplně samostatně, nelze mu nařít jak často má na stránku chodit (vyhledávači lze samozřejmě zaplatit za pravidelnou indexaci), ale pouze to, které stránky nesmí indexovat, viz. kapitola č. 4.3.3 robots.txt. Pamatuje si kde byl a v jeho algoritmu je naprogramováno kdy se na danou stránku má opět vrátit.

Do vyhledávačů lze registrovat stránky, ale většinou to postrádá smysl. Robot vyhledávače je schopen stránku najít sám. Registrace bez toho aby na stránku vedl nějaký odkaz nepomůže. Výjimku tvoří české stránky, které mají národní doménu jinou než .cz (například .com nebo .net). Takové stránky české vyhledávače (Jyxo¹, Morfeo) automaticky indexují na základě registrace do některého z katalogů (například pokud je www.example.com registrován v Atlasu, tak ho Jyxo považuje za českou stránku a provede indexaci). Pokud stránka není registrovaná v katalogích je možnost ji zaregistrovat přímo do vyhledávače².

Při vyhledávání se na webové stránce vyhledávače pokládá dotaz, který vyhledávač porovnává se svou databází. Jelikož nikdo, kromě programátorů, nezná přesný algoritmus podle něhož se řadí výsledky vyhledávání, lze se jen domnívat, že je to v současné době u moderních vyhledávačů kombinace toho

¹ Jyxo hledá od května 2004 i české stránky, které nemají doménu .cz a nejsou v Atlasu.

² <http://jyxo.cz/d/submit>

kolikrát a v jaké html značce se hledaný výraz na stránce nachází a popularity stránky (více v kapitole 4.2. Offpage faktory). Popularitě se říká ranking stránky, Google používá PageRank, Seznam S-Rank. Zlepšování pozice ve vyhledávačích se nazývá optimalizace pro vyhledávače (SEO.).

2.2. Historický vývoj

Web se začal rozšiřovat v roce 1993, ale již před tím existoval vyhledávač, který uměl prohledávat ftp servery. Vytvořil ho v roce 1990 Alan Emtage z univerzity McGill v Montrealu a jmenoval se Archie. Původní název tohoto vyhledávače byl "archives", ale Unix vyžadoval krátké názvy.

Pokud si lidé v této době chtěli vyměňovat nějaká data na počítači, tak to většinou dělali pomocí File Transfer Protocolu (FTP). Pokud jste chtěli sdílet data museli jste mít FTP server, když jste chtěli něco stáhnout stačil FTP klient. V podstatě velmi efektivní výměna informací, ale pouze v malé skupině lidí. S přibývajícím počtem serverů se uživatelé mírně řečeno ztráceli. V tom dokonale pomohl právě Archie, který vyhledával soubory, jenž se shodovaly s dotazem uživatele. V podstatě to byla databáze názvů souborů. Popularita Archie se zvyšovala, a tak po jeho vzoru Nevadská univerzita o rok později představila vyhledávač Veronica, který sloužil pro hledání textových informací v Gopheru.

V roce 1993 začíná rozvoj WWW (World Wide Web) a na svět se dostává první browser Mosaic. Tuto informaci asi slyšel každý z nás, ale neméně důležité je, že se také objevil první robot, který procházel web. Naprogramoval ho Matthew Gray a jmenoval se World Wide Web Wanderer. Tento robot měl původně jen počítat aktivní stránky na internetu, a pomocí toho zjišťovat rychlost růstu webu, Matthew Gray ho však později vylepšil tak, aby indexoval URL (ukládal je do databáze - Wandex). Wanderer byl ze začátku spíše problémem, protože někdy stahoval stovky stránek z jednoho webu za krátkou

chvíli, a tak způsoboval přetížení (což se může stát i současným nově naprogramovaným robotům).

Ke konci roku 1993 Martijn Koster vytvořil vyhledávač Aliweb podobný Archie, který indexoval webové stránky. Aliweb umožňoval uživatelům, aby do něj stránky přihlašovali a na tuto žádost je zaindexoval. Nevýhodou bylo, že nebyl propojen s robotem, který by stránky sám hledal. Do konce roku 1993 se objevily tři vyhledávače se svými vlastními roboty - JumpStation, World Wide Web Worm a Repository-Based Software Engineering. První dva indexovaly hlavičku dokumentu a URL. Jejich problém však byl, že dokumenty zobrazovaly tak, jak je našly. Oproti tomu RBSE implementoval ranking systém.

World Wide Web se začíná rychle rozvíjet. V roce 1994 vytvořili Jerry Yang a David Filo ze Standford University Yahoo!. Ze začátku to byl seznam jejich oblíbených odkazů, kterým se snažili dát určitý řád, a tak vytvořili do současné doby nejnámější katalog. 20. dubna 1994 vydal Brian Pinkerton z University of Washington WebCrawler. Což byl první crawler, který indexoval kompletní stránku, a ne jen hlavičky a URL. Později byl WebCrawler koupěn Excite. V roce 1995 přicházejí na scénu další a další vyhledávače: Metacrawler, Excite, Altavista, Inktomi/Hotbot, AskJeeves a GoTo. Vedoucí postavení si však udržovalo vcelku s přehledem Yahoo!, což je ale v podstatě katalog. Z vyhledávačů byla populární hlavně AltaVista. Jak se ale objevovaly stále další a další internetové stránky, byl stále větší problém něco najít (klesala relevance výsledků vyhledávání).

Problém se vzrůstající mírou vyhledaných nerelevantních dokumentů vyřešili ke konci roku 1998 Larry Page a Sergey Brin s vyhledávačem Google. Základní myšlenka byla, že výsledky vyhledávání se řadily podle toho, kolik odkazů vede na dané stránky. Každý odkaz je v podstatě hlas, který tvůrce stránky dává jiné stránce. Kromě obsahu stránky, začalo být důležité, kolik

dobrych odkazů na stránku vede. Jedná se v podstatě o citační analýzu aplikovanou na internet.

V současné době se hodně mluví o potřebě personalizovaného vyhledávání, a o potřebě lokálního vyhledávání. Personalizované vyhledávání je založeno na definovaných oblastech, které uživatele zajímají. Preference může určit přímo uživatel zaškrtnutím jednotlivých kategorií, nebo se dají také získat z oblíbených položek v prohlížeči a z toolbarů (zjišťuje se, jaké stránky uživatel navštěvuje). Například Google má personalizované vyhledávání v současné době v beta testování¹. Stačí zaškrtnout oblast zájmu a Google se snaží vybrat stránky, které mají s daným oborem něco společného (viz. kapitola 4.2.7 Tématicky orientovaný PageRank).

Obrázek 1: Lokální vyhledávání v Googlu

The screenshot shows a Google search interface in a Mozilla Firefox browser window. The search query is 'pizza near New York, NY'. The results are displayed as a map of Manhattan with numbered markers (1-10) and a list of 10 pizza restaurants. The first two results are:

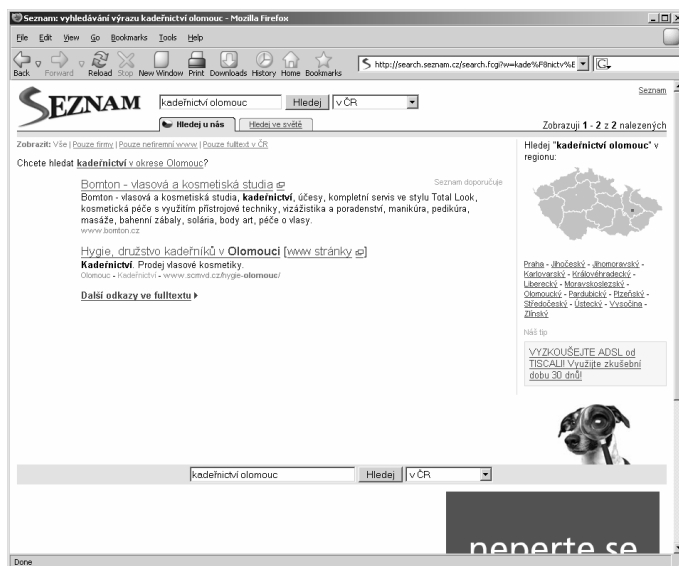
Name	Address	Related Web Pages
Tasty Pizza & Deli (212) 227-1643	149 Church St New York, NY 10007 0.1 mi NW - Directions	New York Restaurants - Italian ... Hartsdale: 1063 Pizza Corp 1053 2nd Avenue - New York, NY. ... zami.com - and more related pages »
Albachiara Pizzeria (212) 267.com	10 Reade St New York, NY 10007	Restaurant Pizza , New York, New York, NY, Superpages com can help you find Restaurant Pizza business listings in ...

¹ <http://labs.google.com/personalized>

Lokalizované vyhledávání je alespoň podle mě daleko užitečnější. Dalo by se tvrdit, že fungovalo již od vzniku vyhledávání, jen v současné době je vylepšené. Pokud do vyhledávače zadám pizza Praha, očekávám, že mi to najde pizzerie v Praze. Pokud do lokálního Googlu¹ v současné době zadám pizza New York vypíše se mi seznam pizzerií, které mají na svých stránkách adresu v New Yorku, a navíc se mi zobrazí mapa kde přesně je najdu.

Lokální vyhledávání nabízí také například Seznam, v tomto případě se však na rozdíl od Googlu vyhledávají pouze firmy registrované v katalogu.

Obrázek 2: Lokální vyhledávání na Seznam.cz



Samozřejmě ostatní vyhledávače nevyimizely, jen nejsou tak populární jako Google, který drží v současné době největší podíl na trhu. I ostatní vyhledávače časem začaly používat vzorec na hodnocení stránek podobný

¹ <http://local.google.com/lochp>

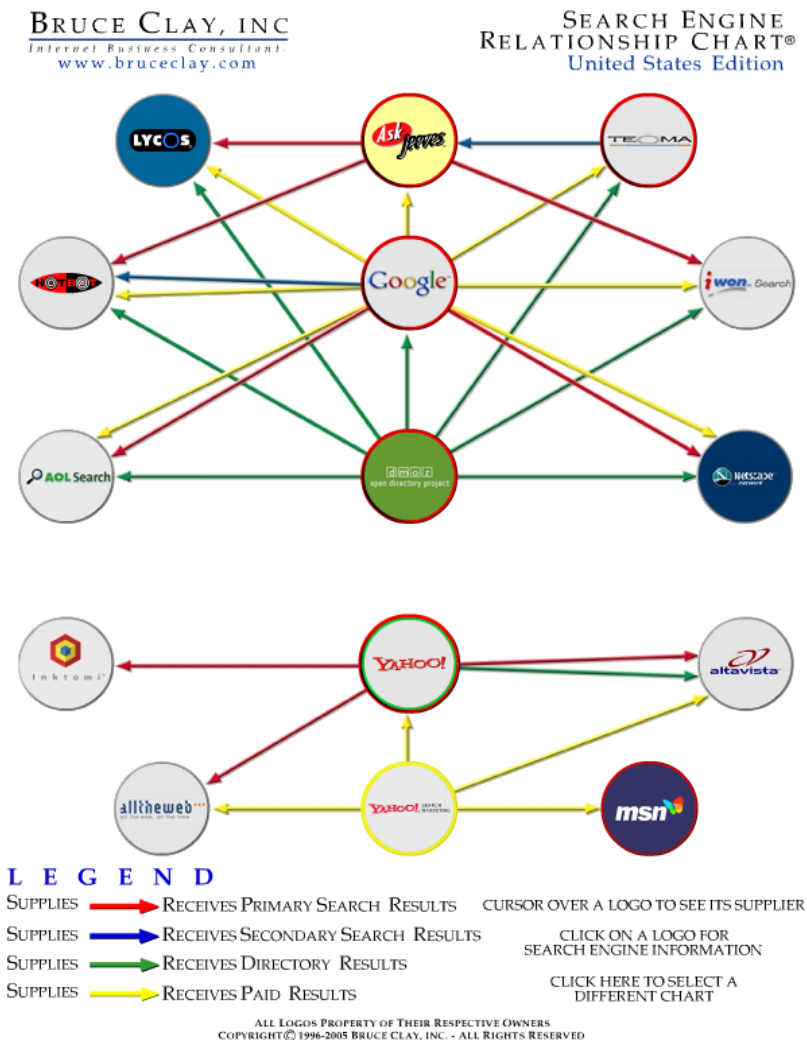
tomu jako použil Google, samozřejmě s trochu jinými vahami. V současné době se hovoří o tom, že Google nemůže kralovat věčně, ale prozatím není vyhledávač, který by byl populárnější. Dohnat Google se snaží hlavně Yahoo!, které koupilo celou řadu vyhledávačů (Altavista, Inktomi, Alltheweb), aby z nich vyladilo jednu technologii, kterou začalo používat na svých stránkách na začátku roku 2004.

Také Microsoft si chce vydobýt své místo v oblasti vyhledávačů. Na začátku roku 2005 opustil vyhledávací technologii Yahoo a implementoval svou vlastní. V nové verzi Windows pak pravděpodobně integruje vyhledávací funkci přímo do prohlížeče Internet Explorer. Podobnou integraci již dnes nabízí prohlížeče Opera a Mozilla (FireFox), které však dávají na výběr mezi jednotlivými vyhledávači.

Celá řada portálů ve skutečnosti nejsou vyhledávače, i když se tak na první pohled tváří. Pouze používají cizí technologii a přebírají výsledky, které pak kombinují s katalogem (nejčastěji s ODP¹ - Open Directory Project). Současnou situaci na trhu nejlépe vykresluje obrázek č. 3 od Bruce Clay. Jedná se v postatě o outsourcing vyhledávacích a katalogových služeb.

¹ <http://www.dmoz.org/>

Obrázek 3: Propojení katalogů a vyhledávačů ve Spojených státech

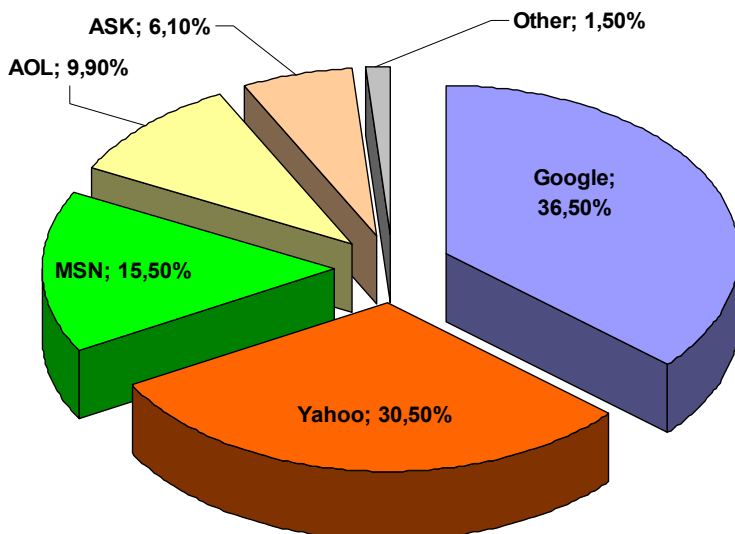


Zdroj: <http://www.bruceclay.com/searchenginereationshipchart.htm>

2.3. Současné tržní podíly vyhledávačů

Pro přehled tržních podílů jednotlivých vyhledávačů jsem použil výsledky měření comScore¹ Media Metrix Search Engine Ratings [22]. Tento výzkum zahrnuje pouze anglicky mluvící uživatele a je za období červen 2005.

Graf 1: Podíl vyhledávačů v anglicky mluvících zemích

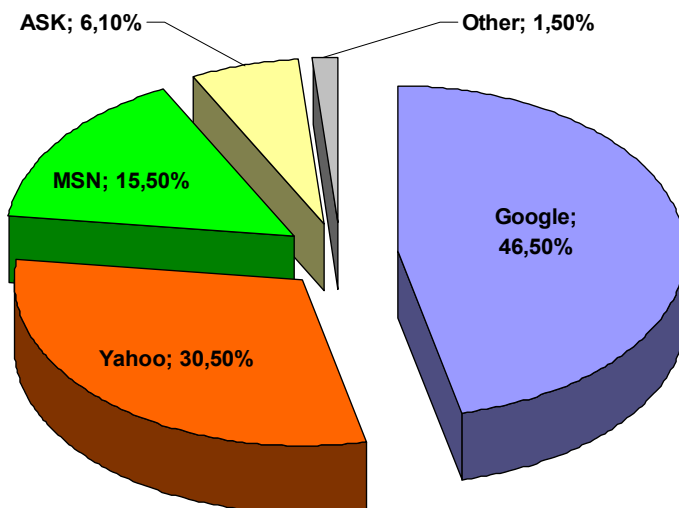


Pokud se tento výzkum převede na podíly technologií, získá Google téměř polovinu trhu (viz graf č. 2). Současní největší hráči na trhu tedy jsou: Google, Yahoo!, Microsoft a ASK. O AOL, které používá technologii Google, se spekuluje, že bude koupeno některým hráčem z velké trojky. Podíl

¹ <http://www.comscore.com/>

vyhledávačů je samozřejmě obtížně měřitelný a existují různé přístupy ¹. O tom, že ale globálně vede Google spory nejsou.

Graf 2: Podíl vyhledávacích technologií v anglicky mluvících zemích



2.4. České vyhledávače a katalogy

České vyhledávače se vyvíjely se zpožděním za těmi světovými, takže se z historie zmíním krátce o Seznamu, Centru a Jyxu a hlavně rozeberu současnou situaci na českém trhu katalogů a vyhledávačů. V roce 1997 naprogramoval Ivo Lukačevič a Štěpán Škrob (spolužáci na ČVUT) fulltext Kompas který používal Seznam jako vyhledávací technologii. V určité podobě fungoval Kompas až do roku 2001, kdy Seznam začal používat Google. Později zakoupil technologii Jyxo, kterou na začátku roku 2005 nahradil svou vlastní - Seznam Fulltextem. Seznam, tak již primárně nevyhledává v katalogu, ale ve Fulltextu.

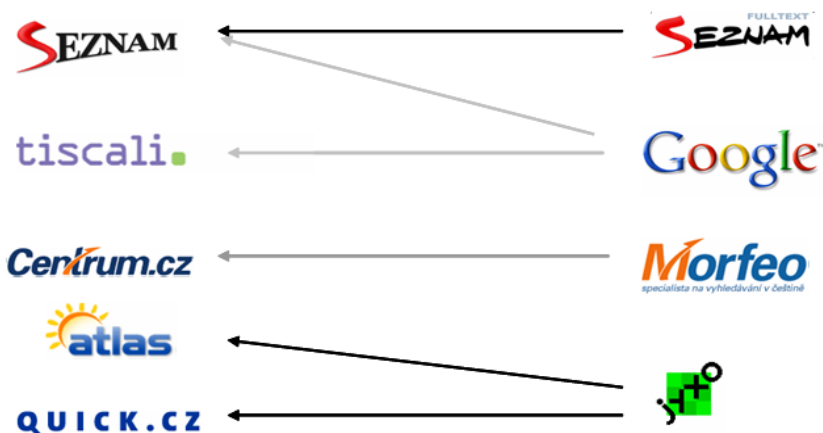
¹ <http://searchenginewatch.com/reports/article.php/2156451>

Centrum začalo fulltext vyvíjet v roce 1999 společně s Martinem Marešem z MFF UK. V té době se používal pouze na serveru Centrum.cz. V srpnu 2002 spustilo Centrum Webfast, který už měl samostatnou doménu. Začátkem roku 2003 Centrum udělalo ve fulltextu zásadní změny (specializace na češtinu - ohýbání slov, synonyma, překlepy), a proto byla změněna značka na Morfeo. V téže roce se z důvodu zkvalitnění výsledků fulltext propojil s katalogem Centra. Na podzim 2003 začalo Morfeo používat pro řazení výsledků také zpětné odkazy (obdoba Google PageRanku).

Další hráč na českém poli je Jyxo, které vzniklo na jaře 2002 (založení firmy). Z velkých katalogů ho od začátku roku 2003 používá Atlas a Quick a od února 2004 do března 2005 pak Seznam. Stejně jako Morfeo zvládá ohýbání slov, synonyma, překlepy a používá JyxoRank (obdoba Google PageRanku).

Jaké vyhledávače používají české katalogy ukazuje Obrázek č.2. Pro upřesnění Seznam v současnosti používá pro český web vlastní technologii Seznam Fulltext a pro zahraničí Google. Tiscali používá technologii Google. Atlas a Quick pak technologii Jyxo.

Obrázek 4: Používání vyhledávačů českými katalogy



Pro přehled podílu vyhledávačů v referreru (URL odkud přišli návštěvníci webu) jsem čerpal data ze serveru Toplist¹ (viz. Tabulka č. 1), který zahrnuje dostatečné množství údajů (téměř 70 tisíc webů), dále pak z Choseho² stránky, která agreguje data z Toplistu a tyto data jsem ještě z kombinoval s údaji ze svého serveru Bazos.cz (pouze pro dopočítání podílu ostatních vyhledávačů). Z údajů vyplývá, že největší význam z hlediska odkazujících serverů má v ČR Seznam a Google. Seznam má podobnou pozici jako v anglicky mluvících zemích Yahoo!, ze kterého i ideově vychází. I když to není možná na první pohled patrné, jsou vyhledávače přímou konkurencí katalogům. Google poslední dobou stále ukrajuje větší část, a to jak Seznamu tak i Yahoo!.

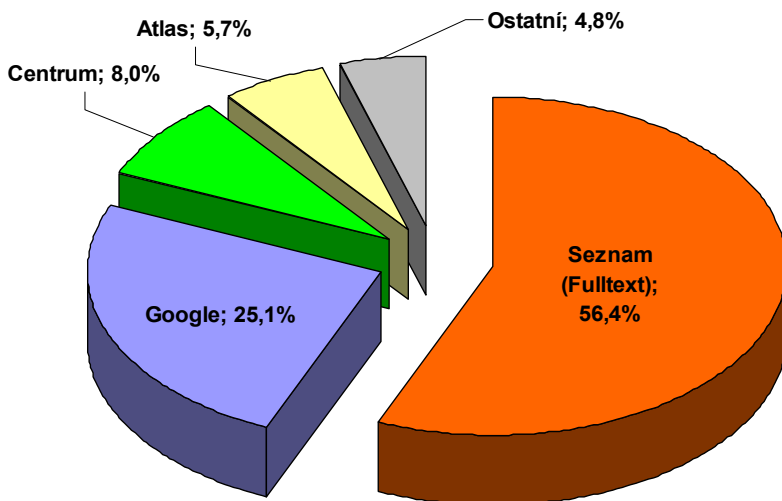
¹ <http://www.toplist.cz/global.html>

² <http://vyhledavace.chose.cz/>

Tabulka 1: Podíly vyhledávačů podle serveru Toplist.cz, říjen 2005

Název serveru	Podíl v procentech
1. Seznam Fulltext	56,4
2. Google	25,1
3. Centrum	8,0
4. Atlas	5,7
5. Ostatní	4,8

Zdroj: <http://www.toplist.cz/global.html>

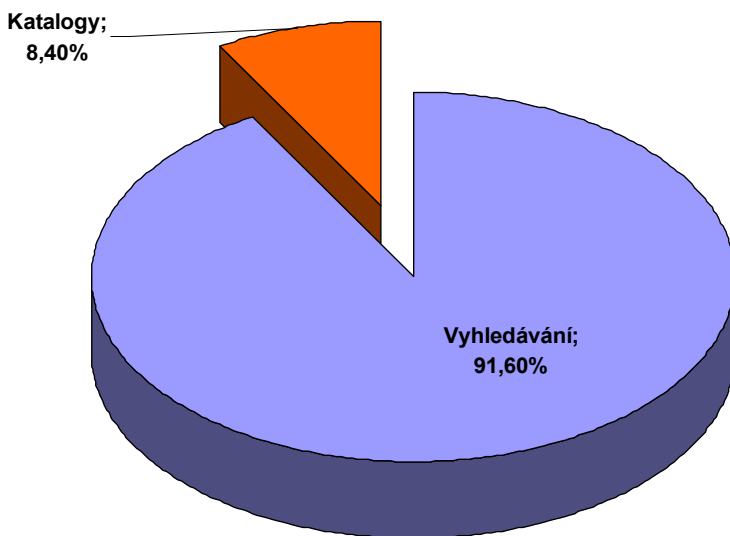
Graf 3: Podíl refererrů v ČR

Pro ověření zda uživatelé v ČR používají také katalogy, jsem použil data ze serveru Navrcholu¹ za květen 2005. Výsledek již v dnešní době nemůže nikoho překvapit. 91,6 procent lidí hledá požadované informace pomocí vyhledávání a jen 8,4 procenta pak pomocí procházení katalogových sekcí. To

¹ <http://www.navrcholu.cz/>

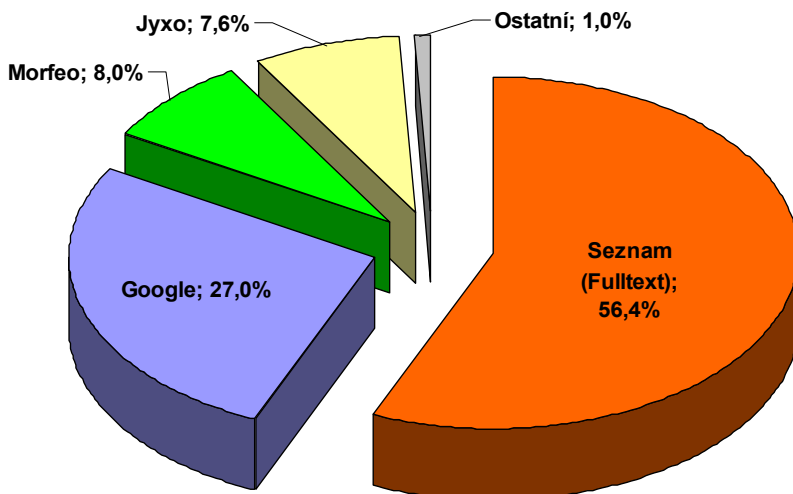
ovšem neznamená, že není důležité se do katalogů registrovat. Stále platí, že základní registrace do katalogů u nás nic nestojí, kromě pár minut času. Navíc se registrací do správných sekcí katalogů, získávají relevantní zpětné odkazy. Doby, ale kdy byly katalogy hlavní zdroje návštěvnosti jsou již ale dávno pryč. Hodně tomu přispělo hlavně přepnutí vyhledávání na fulltext namísto prohledávání katalogu.

Graf 4: Podíl vyhledávání a procházení katalogu



Pokud v ČR sečteme používané vyhledávací technologie má Seznam stále dominantní podíl (viz. Graf č. 5). Googlu bylo připočteno Tiscali. Jyxo se pak skládá z Atlasu, Quicku, Zooho a samotného Jyxa. Ostatní zahrnuje hlavně MSN a Yahoo. Z výsledných dat vyplývá, že v současnosti se u nás vyplatí optimalizovat pro Seznam Fulltext, Google, Morfeo a Jyxo.

Graf 5: Podíl vyhledávacích technologií v ČR



Pro SEO má podíl na trhu vyhledávačů pouze informativní charakter. Navíc zde prezentovaná data nejsou v žádném případě přesná. Srovnávat s tím svoje výsledky je sice možné ale jejich odlišnost ještě nemusí nic znamenat. Srovnání přístupů z vyhledávačů by bylo nutné porovnávat se servery, které mají stejné zaměření. To znamená porovnávat internetový obchod se souhrnnou sekci Internetové obchody na Toplistu, nebo ještě lépe s podobně zaměřeným internetovým obchodem. Technicky zaměřený server bude mít pravděpodobně nejvíce návštěvníků z Googlu, naopak u magazínu pro ženy by se dal očekávat větší podíl Seznamu atd.

3. Předpoklady pro optimalizaci

Vyhledávače vidí stránky přibližně stejně jako textový prohlížeč Lynx¹, takže v případě použití některých technologií (Javascript, Flash, Rámce atd.) nedokáží zaindexovat celý web. Ne všechny vyhledávače mají stejné problémy, některé dokáží indexovat dynamické URL, ale ve Flashi nedokáží rozpoznat odkazy, jiné zase naopak. Každý vyhledávač je prostě jiný. Pro kontrolu, jestli je zaindexován přibližně celý web, se porovnává počet zaindexovaných stránek (zjišťuje se příkazem `site:example.com example`) s celkovým počtem stránek daného webu.

3.1. Javascript

Javascript je programovací jazyk, který se používá na webových stránkách. Vyvinula ho firma Netscape a jeho hlavní účel je rozhybat web na straně uživatele (klienta). Mezi programátory, ani mezi uživateli není příliš oblíben, a to z důvodu, že v každém prohlížeči reaguje trochu jinak. Osobně se ho snažím používat co nejméně, přece jen ladit ho, aby fungoval pro všechny prohlížeče je někdy nadlidský úkol. Vyhledávače Javascript také nemají v oblibě, tedy přesně řečeno ho velmi často neindexují. Pokud je tedy například navigace webu provedena pomocí Javascriptu, hrozí nebezpečí, že vyhledávače nedokáží odkazy najít (jsou schované za Javascriptem). A tak se také může stát, že web vypadá skvěle, funguje skvěle, ale zaindexována je pouze úvodní stránka. Podstatné informace ale mohou být na celém webu.

Samozřejmě jsou situace, kdy je Javascript vhodné použít (například pro kontrolu polí odesílaného formuláře). Když už je ale nutné Javascript použít, je vždy lepší ho schovat do souboru:

```
<Script src="java.js"></script>
```

¹ <http://lynx.browser.org/>

Dělá se to za účelem přehlednosti kódu (přehlednost při další editaci), a pro větší čitelnost pro vyhledávače. Myslím tím, že pokud by bylo Javascriptu na stránce opravdu hodně, mohl by se vyhledávač zmást a stránku vůbec nezaindexovat. Mezi další výhodu "uklizení" Javascriptu do externího souboru patří, že pokud se stejný script používá vícekrát, stahuje ho internetový prohlížeč jen jednou.

3.2. Flash

Flash je aplikace vyvinutá firmou Macromedia. Jedná se o samostatnou technologii, která se často vkládá do webových stránek. Flash tedy není webová stránka, ale něco na způsob obrázku na webové stránce. Tato technologie je vhodná hlavně pro prezentaci, reklamu a vše, kde je primárním cílem zaujmout.

Pro optimalizaci pro vyhledávače je Flash nepoužitelný, vyhledávače ho totiž většinou neumí indexovat, což je jeho základní nevýhoda. Vyhledávací robot si v tomto případě lze představit jako prohlížeč, který neumí zobrazit Flash. Samozřejmě s indexací jsou již první pokusy, ale zatím je to spíše ve fázi vývoje, například Google dokáže Flash částečně číst. Asi největší chybou designu webové stránky je použití Flash pro navigaci. Robot vyhledávače tak zaregistruje první stranu, ale ostatní již nevidí, a tak odejde. Když už je tedy na stránce Flash, je nutné ho doplnit textovými informacemi, popřípadě udělat neflashovou verzi.

3.3. Splash page

Splash page je vstupní stránka, která obsahuje jen veliký obrázek nebo flashovou animaci. Po ukončení animace dost často přichází přesměrování. Toto úvodní přivítání je úplně zbytečné pro uživatele. Přece jen kdo je ochoten čekat 30 sekund na to, než proběhne intro stránky ve Flashi. A pokud to již vytrpí, po druhé to již nebude chtít vidět. Pro uživatele lze nastavit Cookies, to

znamená, že podruhé se již Splash page nezobrazí, ale vyhledávače v drtivé většině Cookies nepoužívají, a tak pokaždé musí stáhnout Splash page.

Obrázek 5: Ukázka Splash page



Některé vyhledávače takové stránky jednoduše označí za spam a odmítnou je indexovat. Neobsahují totiž dostatek kvalitního obsahu. Pokud již na stránce Splash page je, měla by u Flashe obsahovat minimálně možnost přeskočit intro a na spodu stránky by měly být odkazy na hlavní části webu. Jelikož u PageRanku existuje faktor útlumu, který se ve vyhledávacích pohybuje kolem 15 procent, znamená to, že skutečná hlavní stránka na kterou vede odkaz ze Splash page může dostat přibližně o 15 procent méně PageRanku.

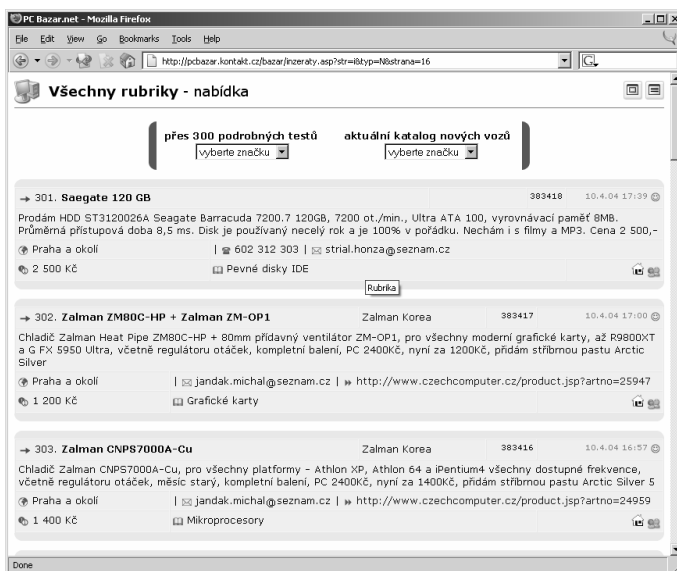
3.4. Rámce

Rámce (frames) jsou dodnes používaný způsob tvorby stránek, vedle tabulek a CSS je to základní možnost jak stránku rozvrhnout. Základní soubor rámců

definuje, v které části se nachází menu (u nás nejčastěji vlevo a nahoře), a v které části se načte vlastní obsah stránky. Po kliknutí na položku v menu se obsah zobrazuje v příslušné části. Rámce se začaly používat hlavně z důvodu, že kód navigačního menu byl na jednom místě a také se pomocí nich dá ušetřit na množství dat přenesených k uživateli.

Hlavní problém rámců je s vyhledávacími roboty. Některé¹ totiž rámce úplně odmítají indexovat a zaindexují pouze část noframes. Pokud již vyhledávače zaindexují obsah jednotlivých stránek v rámcích, je problém při zobrazení uživateli. Nabídnou mu totiž odkaz na konkrétní stránku a ne na rámce.

Obrázek 6: Stránka, které chybí po vyhledání jakékoliv menu



Uživateli se pak většinou zobrazí stránka bez navigačního menu, což značně snižuje pravděpodobnost, že takový návštěvník na stránce zůstane déle (nemá

¹ Velké vyhledávače jako Google nebo Yahoo! dokáží v rámcích rozpoznat odkazy.

na co kliknout a ne každý umí odmazat část URL). Problémy s rámcí nemají jen vyhledávače, méně pokročilí uživatelé mohou mít problém nastavit správný rámec, který chtějí vytisknout (místo obsahu se jim vytiskne menu). Pokud chce někdo odkázat na konkrétní stránku webu, tak to v podstatě nejde, nebo jde, ale pak nelze vidět příslušné menu.

Ačkoliv se zdá, že rámce jsou v současné době CSS již nepoužitelné, není tomu tak. V některých případech jsou rámce velice výhodné, jedná se například o webové chaty. Vždy tedy záleží na konkrétním projektu, ale musí se mít na paměti, že použití rámců může znamenat nemožnost správné indexace vyhledávači.

Pokud jsou již rámce nutnost, optimalizuje se alespoň obsah značky `noframes`. Ten se pak ukazuje ve vyhledávačích. Bohužel lze najít celou řadu stránek, ve kterých je použito něco podobného:

```
<noframes>
    <body>
    Váš prohlížeč není podporován.
    </body>
</noframes>
```

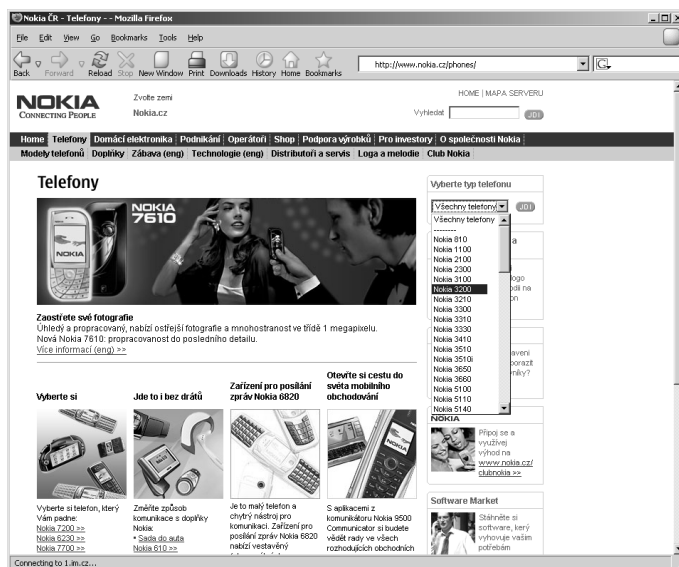
3.5. Klikací obrázkové mapy

Klikací obrázková mapa je obrázek rozčleněný v HTML na menší části. Každá taková část může odkazovat na jinou stránku. Používá se proto, aby se obrázek nemusel v grafickém programu pracně rozdělovat. Tyto mapy však například Jyxo nedokáže projít a stránky, na které nevede jiný odkaz, zůstanou nezaindexovány. I když by si s nimi měla většina vyhledávačů poradit (Google, Morfeo) je vhodné, aby na stránky, na které se odkazuje v obrázkové mapě, vedl alespoň jeden normální (textový) odkaz.

3.6. Rozbalovací menu

Rozbalovací menu (*select a option*) se někdy používají v navigačním systému webu kvůli úspoře místa. Při použití 50 i více abecedně seřazených položek v rozbalovacím menu se uživatel stále orientuje v navigačním systému. Jenže vyhledávací robot není uživatel, nemá tudíž myš, aby si na takové rozbalovací menu kliknul. Existují samozřejmě vyhledávací roboti, kteří přes takové menu projdou, ale je jich zatím menšina.

Obrázek 7: Ukázka použití rozbalovacího menu



3.7. Problém s dynamickými stránkami

Dynamické stránky jsou webové stránky s URL ve kterém jsou proměnné za otazníkem, tedy ve tvaru www.example.com?id=xyz&id2=abc&detail=dfg. Taková adresa má nevýhodu pro uživatele když si ji chce zapamatovat nebo jen opsat na papír. Je v ní celá řada znaků a proměnných ve kterých se dá udělat snadno chyba.

Webový vyhledávač je jako obyčejný uživatel. Je pro něj jednodušší zapamatovat statickou stránku s URL www.example.com/xyz/abc/dfg.html než www.example.com?id=xyz&id2=abc&detail=dfg. Přitom obě dvě adresy mají stejné proměnné a mohou odkazovat na stejný soubor. První se však tváří jako statická stránka bez proměnných, kdežto druhá je na první pohled dynamická.

Dalo by se jistě namítnout, že všechny významné vyhledávače dokáží indexovat dynamické stránky, což je pravda, ale je zde určitý druh penalizace. Například Google indexuje pouze maximálně tři proměnné (dříve dvě). Pokud se jako název proměnné v Googlu do března 2004 vyskytovala proměnná *id* za otazníkem, tak takovou stránku nikdy neindexoval. To že Google nebo Jyxo indexují dynamické stránky ještě tedy neznamená, že neexistují menší vyhledávače, které to vůbec nezvládnou.

3.7.1. Praktický příklad optimalizace dynamických stránek

Na příkladu internetového knihkupectví Jasminka.cz bych chtěl demonstrovat problém dynamických stránek. Web Jasminka.cz jsem programoval někdy před čtyřmi lety, je programován pomocí php a mysql. Na webu má každá kniha svou stránku. URL bylo původně ve tvaru <http://www.jasminka.cz/detail.php?isbn=12345678>. Po pravdě řečeno mě ani nenapadlo, že by vyhledávače neindexovaly všechny stránky. Až když jsem se začal zabývat SEO tak jsem si zkusil, kolik Google indexuje z tohoto webu stránek. Do Googlu stačí dát `site:jasminka.cz jasminka1` (stránky z webu jasminka.cz obsahující slovo Jasminka).

Našel jsem pouze 30 stran, takže zaindexován byl pouze zlomek a žádná stránka s detailem o knize. Problém nebyl s tím na kolikáté pozici web je, ale

¹ <http://www.google.com/search?q=site:jasminka.cz+jasminka&hl=cs&lr=&ie=UTF-8&oe=UTF-8&start=20&sa=N>

s tím že nebyl celý zaindexován. Rozhodl jsem se tedy změnit URL na tvar <http://www.jasminka.cz/kniha/j-k-rowlingova-harry-potter-1-4-komplet.html>. Problematika jak se to dělá pomocí mod_rewrite je popsána v dokumentaci Apache a vysvětlení k dokumentaci je v literatuře č. [8].

Mod_rewrite se ovládá pomocí souboru .htaccess, který musí být umístěn v příslušném adresáři. V mém případě jde o kořenový adresář, kde jsem umístil .htaccess s následujícím obsahem.

```
RewriteEngine on (aktivuje Mod_rewrite pokud
není zapnutý globálně)
RewriteRule ^kniha/(.*)\.html$ detail.php
[L,QSA]
```

Druhý řádek znamená, že každý dotaz, který směřuje na adresář kniha/cokoliv.html si zavolá soubor detail.php, který daný dotaz zpracuje a podsuně stránku prohlížeči. Uživatel pak vidí virtuální adresu. Skutečný soubor s takovým jménem neexistuje. L v hranaté závorce znamená, že pravidla nehodlám řetězit, QSA pak říká, že případné předávané proměnné (tedy vše za otazníkem v adrese stránky) mají být do cílové adresy vloženy. V souboru detail.php se pak musí získat hodnota proměnné z URL, což už je poměrně jednoduché.

```
$url=strip_tags($REQUEST_URI);
$url_array=explode("/", $url);
array_shift($url_array); //první je prázdné
if(empty($url_array)){ }
else{
    if ($isbn==""){
        $isbn=str_replace(".html", "",
$url_array[1]);}
}
```


Při použití `mod_rewrite` nastávají dva celkem dost zásadní problémy. Je celkem jednoduché skript zacyklit a shodit Apache, takže je bezpodmínečně nutné to odzkoušet lokálně nebo to dát zkontrolovat někomu, kdo to umí. Osobně všechny skripty testuji lokálně a server jsem si shodil asi třikrát (pouze na jedné chybě). Druhý problém je s Cookies, aby se také přeposlaly, musí se použít skript podobný tomuhle:

```
rewriteEngine on
rewriteCond %{HTTP_COOKIE} uin
rewriteRule /* http://%{HTTP_HOST}/main.html
[R,L]
```

Po změně adres u některých vyhledávačů trvá dost dlouho než si změny všimnou. Některým i půl roku, ale postupem času stále víc ve svých výsledcích vyhledávání zobrazují nové adresy. Google, po použití `mod_rewrite` na `Jasminka.cz` zaindexoval 99 procent všech stránek.

3.7.2. Přepisování URL pro Windows IIS

Pro přesměrování ve Window IIS primárně slouží ISAPI¹. To je ovšem placené a ne všechny hostings umožňují jeho použití. Samozřejmě existují i jiná řešení, ale jelikož v ASP nejsem tak zběhlý a nejsem si jistý, že bych ho dokázal vymyslet sám, našel jsem si na serveru Interval [12] jak se přesměrování pro Windows vlastně dělá.

Server Interval.cz publikuje celou řadu článků. Každý článek je uložen v databázi, má své číselné ID, textové ID-ASCII, jméno a nějaký připojený text. Normálně se článek vyvolává přes zobrazovací skript, kterému se pomocí nějakého parametru v URL předává požadované ID článku, například <http://interval.cz/clanek.asp?article=1621>. Pokud by však Interval chtěl použít

¹ <http://www.isapirewrite.com/>

URL ve tvaru <http://interval.cz/neviditelny-web-a-jine-pohadky.htm>, tak taková stránka na serveru samozřejmě není, takže dojde k vyvolání chyby 404. Pro ten případ by musel mít server přednastavenou ASP stránku, kterou má předat klientovi. V této stránce by číhala rutina, která odchytlí jméno původně požadované stránky, převede je na vhodný řetězec a zkusí v databázi najít příslušný článek, který bude mít shodné ID-ASCII. Pokud takový najde, použije Server.Transfer pro získání článku z běžné zobrazovací stránky, pokud nikoli, předá chybovou hlášku nebo třeba provede nějakou jinou akci. Podrobněji to rozebírat nebudu, efekt je obdobný jako v případě použití mod_rewrite na Apachi.

3.8. Session ID

Session ID se používají pro identifikaci uživatele. Pokud uživatel nepoužívá cookies předává se mu URL ve tvaru `example.com?id=23kjl23jhjh234j23323j`. Jelikož robot vyhledávače si cookies neukládá, dostává při použití session ID při každé návštěvě jinou URL. Session ID tedy teoreticky vytváří nekonečné množství URL se stejným obsahem. Vyhledávače stránky používající session ID dost často odmítají indexovat. Při jejich použití je tedy nutné rozpoznat robot vyhledávače a dát mu stránku bez parametru.

3.9. Časté změny v URL

URL adresy by měly být co nejvíc stabilní. Je nutné je navrhnout tak, aby se nemusely dlouhou dobu (pokud možno nikdy) předělovat. Jestliže dnes někdo odkáže na článek na určitém webu, měla by tato URL adresa fungovat i za tři roky. To je hlavně z důvodu uživatelů ne kvůli SEO. Pokud se již z nějakého důvodu na webu předělovají URL, například kvůli převodu dynamických stránek pomocí mod_rewrite na statické, je potřeba původní URL přeměrovat na novou. V žádném případě nesmí původní URL přestat fungovat (kód 404).

Než se změny projeví ve vyhledávačích, může to trvat i půl roku, a nefunkční odkazy z jiných stránek se nemusí spravit nikdy. Menší zlo je ponechat funkční obě dvě verze, ale to pak vede k duplicitnímu obsahu (viz kapitola č. 3. 11.). Správné přesměrování je potřeba provést pomocí http hlavičky s kódem 301 - Moved Permanently. Po přesměrování pomocí kódu 301 by měly vyhledávače zaindexovat novou adresu a ranking, který měla původní stránka, dát stránce nové. V PHP se to dělá následujícím způsobem:

```
<?
header("HTTP/1.1 301 Moved Permanently");
header("Location: http://www.example.com/");
header("Connection: close");
?>
```

V ASP pak

```
<%
Response.Status = "301 Moved Permanently"
Response.addheader "Location",
"http://www.yourdomain.com/"
Response.End
%>
```

Při přesměrování se nesmí vynechávat první řádek, tedy kód 301 Moved Permanently, jinak by server nahlásil místo toho 302 Found (moved temporarily), ač to na první pohled vypadá stejně, pro robot to znamená, že původní stránku našel. Zkontrolovat správnost přesměrování lze například pomocí aplikace Ethereal network protocol analyzer¹, ta je dobrá i pro překontrolování posílaných Cookies.

¹ <http://www.ethereal.com/>

Pokud není možné řešení pomocí http hlaviček na straně serveru, lze použít v krajním případě meta tagu:

```
<meta http-equiv="refresh" content="10;  
url=nove-url"> (mělo by být delší než 1 sekunda)
```

Samozřejmě, že takové přesměrování nemusí vždy fungovat, takže je nutné na takovou stránku umístit i odkaz na novou URL adresu.

3.10. Automatické přesměrování

Přesměrování (30x Moved nebo META Refresh) by se mělo používat jen při změně URL adresy. V běžném provozu je vždy lépe se mu vyhnout. Některé vyhledávače¹ mají s přesměrováním problémy, a to může způsobit nezaindexování stránky. V žádném případě není vhodné přesměřovat úvodní stránku na jinou.

3.11. Duplicitní URL

Pro vyhledávače jsou http adresy

- www.example.com/
- www.example.com/index.html
- <http://example.com>
- example.com/index.html

čtyři různé stránky. Z toho důvodu je nutné důsledně dodržovat při odkazování na stránku jednu verzi, jinak dochází zbytečně k rozdělování PageRanku na duplicitní stránky. Stejně tak www.example.com/neco je jiná stránka než

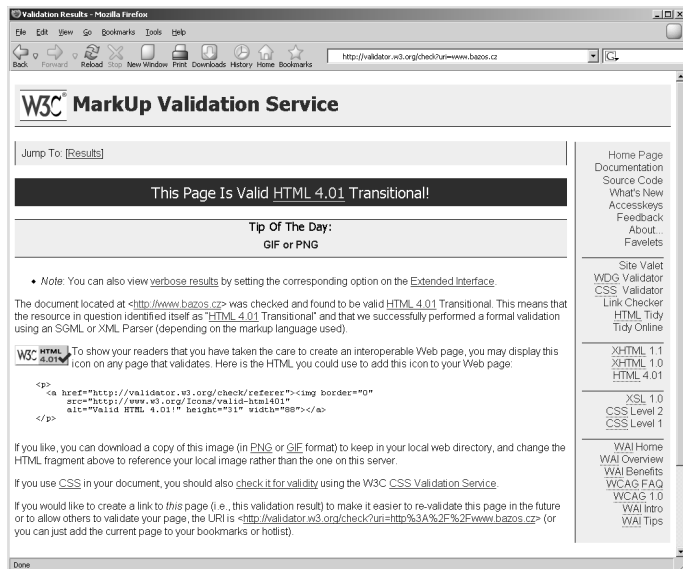
¹ Například Yahoo! ještě na začátku roku 2004 při použití přesměrování zobrazovalo starou URL místo té, na kterou vedlo přesměrování. Novou URL často úplně vyřadilo ze své databáze.

www.example.com/neco/. Vyhledávače¹ vždy zobrazí pouze jednu verzi stránky, která je duplicitní. Měla by to být ta, na kterou se nejčastěji odkazuje. Ostatní jsou ve výsledcích vynechány.

3.12. Validnost HTML kódu

Současné webové prohlížeče dokáží zobrazit HTML kód i s hodně chybami. Když chybí ukončení značky, snaží se ji vhodně doplnit. Robot vyhledávače by to měl zvládnout také, ale může se stát, že při neukončené značce může zaměnit text za HTML značku. Aby se zabránilo takovéto chybě, používají se pro kontrolu validátory (bohatě postačí třeba ten od W3C²). Ačkoliv by se zdálo, že to každý dodržuje, najít web, který by tímto testem prošel není zrovna jednoduché.

Obrázek 8: W3C HTML Validator



¹ Některé vyhledávače si dokáží [example.com](http://www.example.com) a www.example.com spojit dohromady.

² <http://validator.w3.org/>

3.13. Velikost stránky v KB

Velikost stránky není důležitá jen pro uživatele, kteří používají pro připojení na internet modem, ale i pro roboty vyhledávačů. Pokud velikost přesáhne určitou hodnotu, mohou stránku přestat stahovat a zaindexovat pouze stáhnutou část. Velikost lze snížit například používáním CSS stylů, umístěním Javascriptu do externího souboru, nahrazením obrázků texty, atd. Jedno krásné pravidlo říká: “Podařilo se vám dostat velikost stránky na 50 KB? Velmi dobře, dostaňte to na polovinu a bude to v pohodě.“ Nejde ani tak o přesná čísla (Google například stahuje 101 KB kódu bez obrázků), ale pouze o to, že snahou o zmenšení velikosti stránky ušetříte uživatelům čas, a to jistě ocení. Pro testování velikosti stránky je dobrý například Web Page Analyzer¹.

Obrázek 9: Měření velikosti stránky pomocí Web Page Analyzer

The screenshot displays a web browser window showing a 'Web Page Speed Report' for the URL www.bazos.cz. The report is generated by WebsiteOptimization.com. It includes a navigation menu, a newsletter sign-up form, and an announcement. The main content area is titled 'Web Page Speed Report' and provides the following details:

- URL:** www.bazos.cz
- Title:** Bazos.cz - Inzerce, PC bazar
- Date:** Report run on Mon Apr 12 14:43:23 EDT 2004

Diagnosis

Global Statistics

Total HTTP Requests:	6
Total Size:	27685 bytes

Object Size Totals

Object type	Size (bytes)
HTML:	19204
Images:	8481
Javascript:	0
CSS:	0
Multimedia:	0
Other:	0

On the right side of the report, there is an 'Announcement' section stating that the analyzer is available at WebPageAnalyzer.com and includes a 'Speed Up Your Site' graphic.

¹ <http://www.websiteoptimization.com/services/analyze/>

4. Metody v SEO

Metody v SEO lze dělit na on page faktory a off page faktory. V této kapitole jednotlivé metody popisují a v jejím závěru se zabývám možností měření výsledků optimalizace.

4.1. Optimalizace obsahu stránky - On page faktory

On page faktory je cokoli co se vyskytuje na jedné unikátní stránce. To znamená nadpisy, hlavičky, text stránky, interní i externí odkazy na stránce, atd. Podle důležitosti je to pravděpodobně sestupně obsah tagu Title, meta description, H1, H2...H6, , atd. Záleží samozřejmě jak jsou jednotlivá slova na stránce často (hustota), kde se vyskytují (v jakém tagu) a jak daleko jsou od začátku tagu (stránky).

Naprostý základ v optimalizaci je, že každá stránka na optimalizovaném webu musí být unikátní. Vždy je důležité mít na paměti, že vyhledávače hodnotí jednotlivé stránky zvlášť. Takže je nutné soustředit se nejenom na úvodní stranu, to by bylo téměř zbytečné, ale na všechny stránky webu. Například u webu s počtem 500 stránek optimalizujeme ne desítky klíčových slov ale minimálně stovky.

4.1.1. Klíčová slova (keywords)

Výběr klíčových slov je nejdůležitější část SEO. Pokud totiž na stránce není určité klíčové slovo, tak ji pod tímto slovem nemůže vyhledávač ani najít (výjimku tvoří Anchor text). Přesně tak se můj kamarád divil proč nemůžu najít jeho stránky na kterých je Pizzeria v Olomouci. Jenže problém je, že jsem zadával do vyhledávače (konkrétně do Googlu) Pizzeria Olomouc. A slovo Olomouc na té stránce prostě nebylo. Za týden jsem se ho na přesnou adresu zeptal a podíval jsem se z jakého důvodu jsem ji nemohl najít. Celý problém

byl v tom, že Google ještě neumí české skloňování a na té stránce se slovo Olomouc vyskytovalo jen ve větě rozvoz po Olomouci je zdarma. Navíc to bylo jen na jednom místě. Celá problematika klíčových slov se zabývá výběrem relevantních vyhledávaných klíčových slov a jejich následnou optimalizací.

Pokud se web dělá od začátku, stačí vybrat vhodná klíčová slova, zjistit jejich vyhledávanost, a pak je dobře rozmístit po celém webu. Pokud se jedná o firmu, která prodává pět druhů výrobků, je dobré na každý tento druh výrobku mít vlastní stránku, než to celé dávat dohromady na jednu. Každou stránku lze dobře zoptimalizovat maximálně třeba na 5 až 6 klíčových slov, čím konkurenčnější oblast, tím je toto číslo menší (1-2). Příkladem může být firma prodávající židle, stolky a pohovky. Pokud se udělá statický web, který se bude skládat pouze ze dvou stránek a to z kontaktu a z druhé stránky, kde bude popis nabídky dohromady, tak to přece jen není ono. Je lepší to rozdělit a udělat pro židle samostatnou sekci, pro stolky a pohovky také. Pak každou jednotlivou stránku optimalizovat pro klíčová slova dané oblasti.

4.1.2. Výběr klíčových slov

Je dobré si pro začátek promyslet jak by uživatel stránku hledal, tedy to co zadá do vyhledávače. Pro zjištění více nápadů je vhodné zapojit rodinu, přátele a popřípadě i obchodní partnery, čím více různých uživatelů tím lépe. Každý zadává do vyhledávače trochu něco jiného, jinak hledá informatik a jinak by hledala moje mamka. Při výběru klíčových slov se nesmí zapomínat na skloňování a množná čísla, ne všechny vyhledávače je umí odvodit. Základní pravidlo u klíčových slov je, že se musí hledat relevantní slova k danému webu. Pokud je uživatel na daný web nalákán s tím, že si myslí, že koupí rybičky a ve skutečnosti se octne na webu prodávající stolky nic si nekoupí a naštve se. Příště již na odkaz na daný web neklikne.

Při výběru se musí volit specifická slova dané oblasti. Server prodávající autorádia není téměř možné a ani účelné zoptimalizovat na samostatně stojící slovo auto. Uživatel, který hledá auto, pravděpodobně nemá zájem o informaci o autorádiích. Čím specifičtější a vyhledávanější slovo dané oblasti tím lépe.

Reverzní inženýrství

Někdy je jednodušší podívat se na zoptimalizovanou stránku, která je podobně zaměřená jako ta optimalizovaná. Pak se stačí poučit. Nevidím na tom nic špatného, učit se z dobrých příkladů. Jak poznat klíčová slova? Stačí projít Title, description a obsah tagů H1 až H6. Pro pokročilejší analýzu je dobré použít nástroje jako je například Keyword density analyzer.¹

4.1.3. Nástroje na vyhledávaná slova

Samozřejmě člověk není slovník, takže lze použít nástroje které dokáží navrhnout ještě další slova. Podle mě je v současnosti nejlepší nástroj od Googlu, který byl primárně navržen, aby pomohl inzerentům dobře propagovat stránky v placeném systému Adwords. Tento nástroj je zdarma a dokáže hledat i česká slova, které se v Googlu vyhledávají. Jeho nevýhodou je, že přímo neukazuje kolikrát bylo dané slovo v Googlu hledáno, ale jen odhad počtu potencionálních kliků.

Google AdWords Keyword Suggestions²

Pro ukázkou jsem zvolil slovo letenky. Po zadání mi Adwords navrhl následující varianty

letenky

letenky usa

¹ <http://www.ranks.nl/tools/spider.html>

² <https://adwords.google.com/select/tools.html>

levné letenky	letenky levne
levne letenky	letenky londýn
nejlevnější letenky	letenky levné
gts letenky	letenky do usa
letenky cz	letenky levně
nejlevnejsi letenky	letenky last minute
letenky on line	prodej letenky
letenky gts	letenky online

Obrázek 10: Odhad vyhledávanosti slov v Adwords

Google Adwords: Nastavte maximální cenu za proklik - Mozilla Firefox

1. Maximální cena za proklik = Ešátka, kterou byste nanejvýše zaplatili za kliknutí [?]
 2. Vyšší maximální cena za proklik a vyšší míra prokliků = vyšší pozice a více kliknutí [?]
 3. AdWords diskontér automaticky snižuje vaši průměrnou cenu za proklik tak, aby byla pouze o jeden cent vyšší než minimální cena nutná k tomu, aby vaše reklama zůstala zobrazena nad následující níže umístěnou reklamou. Nemusíte tak nadále sledovat a kontrolovat své ceny [?]

Zvolte měnu a maximální cenu za proklik
 USD \$

Klíčové slovo	Prokliky / Den	Průměrná cena za proklik	Cena / Průměrná pozice	Průměrná pozice	
letenky	7,9	\$0,31	\$2,41	2,0	návrh klíčových slov / smazat
letenky cz	< 0,1	\$0,05	\$0,00	1,0	návrh klíčových slov / smazat
letenky do usa	< 0,1	\$0,26	\$0,02	2,0	návrh klíčových slov / smazat
letenky last minute	< 0,1	\$0,05	\$0,00	1,0	návrh klíčových slov / smazat
letenky levne	0,8	\$0,23	\$0,19	1,6	návrh klíčových slov / smazat
letenky levně	< 0,1	\$0,05	\$0,00	1,0	návrh klíčových slov / smazat
letenky levně	< 0,1	\$0,05	\$0,00	1,0	návrh klíčových slov / smazat
letenky londýn	0,2	\$0,22	\$0,04	2,0	návrh klíčových slov / smazat
letenky on-line	< 0,1	\$0,05	\$0,00	1,0	návrh klíčových slov / smazat
letenky online	< 0,1	\$0,05	\$0,00	1,0	návrh klíčových slov / smazat
letenky usa	< 0,1	\$0,05	\$0,01	1,0	návrh klíčových slov / smazat
levne letenky	< 0,1	\$0,05	\$0,00	1,0	návrh klíčových slov / smazat
levně letenky	1,3	\$0,36	\$0,47	1,6	návrh klíčových slov / smazat
nejlevnejsi letenky	0,1	\$0,37	\$0,04	1,8	návrh klíčových slov / smazat
nejlevnější letenky	0,3	\$0,36	\$0,09	2,0	návrh klíčových slov / smazat
Souhrn	10,6	\$0,31	\$3,24	1,9	

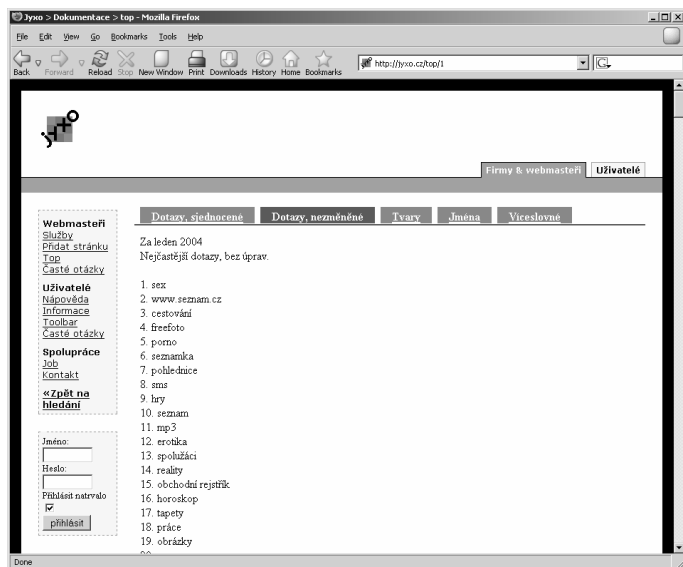
Změnit klíčová slova

Většina slov je na první pohled zřejmých, ale může se zde vyskytovat přeci jen nějaké užitečné. Dále je potřeba vědět vyhledávanost daného slova. Tento údaj

se však nedá úplně přesně zjistit. Vyhledávače ho totiž většinou volně neposkytují. Určitá možnost je použít Google Adwords, které je primárně určeno na reklamu na Googlu. Adwords bohužel poskytuje pouhý odhad počtu prokliků z reklamy v tomto systému. Vypočítat z něho vyhledávanost není jednoduché nebo to vůbec nejde. Takže se dá pouze určit odhad které slovo je nejvyhledávanější, což ale většinou stačí.

Vyhledávanost většinou roste s kvalitou výsledků vyhledávání. Pokud přirozené vyhledávání dává nerelevantní výsledky, lidé na ně málo klikají a příště dané slovo již nehledají (klesá vyhledávanost slova). To samé platí u placené reklamy pokud je kvalitní vzhledem k dotazu víc se na ni kliká.

Obrázek 11: Nejčastější dotazy ve vyhledávači Jyxo¹



¹ <http://jyxo.cz/top/1>

Overture Search Term Suggestion Tool¹

Obdobný nástroj jako má Google, ale není téměř vůbec vhodný pro české výrazy.

Wordtracker²

Dokáže navrhnout podobná slova a navíc zjistit jejich potencionální návštěvnost. Ukazuje také jak těžké bude zoptimalizovat stránku pro dané slovo (konkurenčnost výrazu). Nevýhodou služby je, že zdarma ukazuje jen část výsledků, zbytek je placený. Bohužel ani Wordtracker nedává vždy úplně přesné výsledky.

Pro české výrazy může být lepší najít synonyma pomocí Wordu nebo pomocí českých vyhledávačů Morfeo a Jyxo. Pak jen zkontrolovat jejich potenciál pomocí Adwords. Volba klíčových slov je určitě kritické místo úspěšnosti webové stránky, takže se vyplatí strávit s tím trochu více času a zjistit které slovo optimalizovat. Někdy je daná oblast natolik konkurenční, že nelze dané slovo do první desítky téměř dostat. V tomto případě může být pohodlnější a někdy také levnější (zvláště u českých slov, v zahraničí kopíruje reklama cenu za SEO) si dané slovo koupit.

eTarget³

eTarget je slovenský reklamní systém, který je používán i v Česku. V současnosti se reklama na vyhledávaná slova zobrazuje na Zoohoo.cz, Quick.cz, Volny.cz, Tiscali.cz, Redbox.cz IDnes.cz atd. eTarget dokáže navrhnout ke zvolenému slovu další české alternativy a také zjistit jejich denní návštěvnost. V této knize se nezabývám Pay-per-click (PPC), takže systémem

¹ <http://inventory.overture.com/d/searchinventory/suggestion/>

² <http://www.wordtracker.com/>

³ <http://www.etransfer.cz>

eTargetu a zahraničních konkurentů se zabývám jen jako možností, jak jednoduše zjistit vyhledávanost klíčového slova.

4.1.4. Konkurenčnost klíčového slova

Jak těžké bude zoptimalizovat klíčové slovo udává konkurenčnost výrazu. Pro měření se dost často uvádí počet nalezených dokumentů, to je ale dost nepřesný údaj. Klíčové slovo na které je v Googlu nalezeno 50 000 dokumentů může být daleko hůře optimalizovatelné než slovo na které Google najde půl milionu dokumentů. Jediná přesná možnost je prohlédnout si první desítky nalezených stránek, zda byly zoptimalizovány na dané slovo. Pokud nebyly, a třeba nemají klíčové slovo ani v nadpisu stránky, je toto slovo málo konkurenční i přes to, že Google na něj našel velké množství dokumentů.

Pokud je ale první stovka stránek na dané slovo zoptimalizovaná je konkurenčnost tohoto slova velmi vysoká, bez ohledu na celkový počet nalezených stránek.

4.1.5. Stop slova

Stop slova jsou běžná slova, které nenesou žádnou nebo téměř žádnou informaci. Jedná se většinou o spojky, předložky atd. V češtině se jedná například o a, i, nebo, když, v, na, pod..., v angličtině je to například and, or, but, the, a, an, about Vyhledávače dost často těmto slovům přiřkládají nižší váhu, nebo je úplně ignorují. Dělají to kvůli zrychlení hledání. Je tedy téměř zbytečné dávat stop slova do Title, description nebo keywords.

O tom, zda dané slovo je vyhledávačem filtrováno je nejjednodušší se přesvědčit ve vyhledávači. Například Jyxo vypíše: "Slovo 'a' nenese příliš informací, zčásti jej ignoruji". Google reaguje obdobně: "the" is a very common word and was not included in your search.

4.1.6. Škodící slova (poison words)

Jedná se o slova, která snižují hodnocení příslušné stránky na určitý dotaz. Pro uživatele, který nechce obchod, ale stránku s recenzí produktu je škodícím slovem koupit, cena, prodej... (v angličtině buy, sell, price...). Samozřejmě, že škodící slova mohou být i v recenzi produktu, ale tam bývají většinou pouze v textu. Z toho důvodu některá škodící slova na určité téma vadí pouze v titulku nebo v URL. Obchodní slova škodí pouze pokud uživatel nechce obchod, otázkou je jak to vyhledávače mohou zjistit. Buď uživatel musí vybrat nechci obchod (třeba pomocí *-obchod*) a nebo to nějak vyhledávač zjistí přímo z dotazů. Pokud například uživatel zadá do vyhledávače *název produktu recenze*, asi těžko bude chtít něco koupit.

Zvláštní kategorií jsou škodící slova spojená se sexem. Obecně se předpokládá, že pokud uživatel něco hledá nechce stránky se sexem, takže na ně jsou dosti často uvaleny speciální filtry, které stránku úplně vyřazují nebo minimálně snižují hodnocení. Samozřejmě filtr se neuplatní v případě, že uživatel hledá přímo tato škodící slova.

Další škodící slova mohou být například katalog, fórum, adresář, kontakt, odkazy atd. (samozřejmě v zahraničních vyhledávačích anglické ekvivalenty). Tato škodící slova se často zdůvodňují tím, že uživatel nehledá odkaz, nebo katalog, ale přímo určitou stránku (stránky s odkazy bývají zneužívány odkazovými farmami). Fóra také byla používána k umělému zvyšování PageRanku. Žádný oficiální seznam škodících slov samozřejmě není.

4.1.7. Dobrá a špatná návštěvnost (traffic)

Ne všechny traffic na stránce je dobrý. Proto není vhodné optimalizovat na slova, které nejsou vůči danému webu relevantní. Myslím tím, slovo **A** má při dobré optimalizaci (první místo ve vyhledávačích) potenciálně 1000 kliků za

den, ale nemá téměř nic společného s obsahem stránky. Naproti tomu slovo **B** má potenciálně třeba jen 10 kliků za den a z toho 2 budou skuteční zákazníci.

4.1.8. Konverzní poměr

Počet provedených akcí (v případě obchodu počet objedávek)/počet návštěvníků je konverzní poměr, čím větší procento tím lépe. Příkladem může být webový obchod, pokud se zoptimalizuje na něco co neprodává, dosáhne sice vysoké návštěvnosti, ale žádných nových objednávek. Výsledkem může být jen růst nákladů na hosting. Optimalizace pro klíčová slova, která nemají nic společného se skutečným obsahem stránky je spam, více viz. v kapitole č. 5 - Zakázané techniky optimalizace.

Konverzní poměr se dá tedy zvyšovat pomocí optimalizace na specifická slova, která nejlépe vystihují danou stránku. Pokud uživatel hledá název produktu a k dotazu přidá cena, je předpoklad, že pokud obchod zoptimalizoval své produkty na tuto kombinaci bude konverzní poměr dost vysoký.

Při optimalizaci stránek se nesmí ovšem zapomínat na druhou možnost jak zlepšovat konverzní poměr. Tou je celkové zlepšování stránky. Čím lépe a přehledněji vypadající a jednodušeji fungující stránka, tím pravděpodobněji je konverzní poměr vyšší. Pokud se například zákazník před vhozením položky do košíku musí registrovat, určitě to není jednoduše fungující obchod.

4.1.9. Důležité tagy na jednotlivé stránce

Titulek

```
<head>  
<title>Titulek stránky</title>
```

```
</head>
```

Jednoznačně nejdůležitější tag na stránce je Title. Je v podstatě jediný, který má velký význam ve všech vyhledávačích. Každá stránka na konkrétním webu (i vnořená ve frame) by měla mít jiný Title, který nejlépe popisuje obsah stránky a obsahuje vyhledávané klíčové slovo. Ač by se zdálo, že je to nejjednodušší pravidlo, velké množství webmasterů to nebere na vědomí. Pro toto tvrzení stačí do jakéhokoli vyhledávače zadat "Untitled document" nebo třeba "Domovská stránka" a podívat se na počet takových stránek.

Webmaster takové stránky si jistě nemyslí, že někdo bude hledat Domovskou stránku, pouze nedomyslel, že vůbec existují fulltextové vyhledávače nebo prostě neměl čas se takovými věcmi zabývat (malý rozpočet). Jako Title je tedy nutné například používat název firmy a klíčové slovo. Na místo například Jasminka.cz je dobré použít Jasminka.cz - internetové knihkupectví, nebo něco podobného. Dalo by se namítnout proč nepoužít pouze klíčová slova a vynechat název serveru (z důvodu zvýšení důležitosti klíčového slova, což je nutnost u některých extrémně zoptimalizovaných výrazů). Stránky se však nesmí dělat pouze pro vyhledávače, ale pro lidi. Tvrdím že je dobré budovat značku a tak je nutné aby tato značka byla v atributu Title a navíc na prvním místě. Doporučená délka Title je do 70 znaků (přibližný počet znaků, které se zobrazují ve výsledcích vyhledávání).

Meta description

```
<meta name="description" content="Popis stránky">
```

Meta tag description některé vyhledávače zobrazují u popisku stránky ve výsledcích vyhledávání, takže se vyplatí zde napsat něco smysluplného. Description používá většina vyhledávačů. Stejně jako u titulku je důležité aby

u každé stránky bylo description, které ji nejlépe vystihuje, tedy pro každou stránku odlišné. Doporučená délka je do 250 znaků.

Keywords

```
<meta name="keywords" content="klíčová slova">
```

Meta tag keywords již tak jednoznačný není, většina vyhledávačů (možná i Google) ho nepoužívá. To ovšem neznamená, že je na škodu ho vyplnit klíčovými slovy. Opět každá stránka potřebuje vlastní keywords, stejně jako Title a description.

Pokud se klíčové slovo vyskytuje pouze v meta tagu description nebo keywords a není již ve vlastním textu (v body) pak v drtivé většině vyhledávač danou stránku na toto slovo nenajde. Je tedy zbytečné vkládat do těchto meta tagů něco co není v textu. To platí i o překlepech nebo o psaní slov bez háčeků a čárek, pokud to není jinde než v meta tazích je to ztráta času.

Nadpisy - H1 .. H6

```
<H1>Nadpis stránky</H1>
```

Pokud je něco v nadpisu, mělo by to mít logicky větší váhu. Platí, čím je delší tag H1, tím má klíčové slovo v něm menší význam. Váha je samozřejmě největší u H1, není nutné hledat využití pro H4 a níže, jejich váha je logicky nižší. H1 se smí na stránce opakovat pouze jednou, ostatní nadpisy vícekrát.

Tučný text a kurzíva

```
<strong>tučný text</strong><em>kurzíva</em>
```

Doporučuje se mít na stránce alespoň jednou klíčové slovo tučně a jednou kurzívou. Velký význam to ale pravděpodobně nemá, sledují to jen některé

vyhledávače¹. Hustota tučného textu a kurzívy na stránce by měla být taková, aby byla co největší přehlednost textu (ani hodně ani málo).

Popisky u obrázků

```

```

U každého obrázku by měl být vyplněný atribut alt, který se používá k zastoupení obsahu obrázku. Existuje celá řada uživatelů (včetně mě), kteří mají na modemu obrázky vypnuté. Robot vyhledávače si lze představit jako uživatele, který má vypnuté obrázky. Atribut alt mu pak říká co na obrázku je. Volitelný je atribut title, který se ukazuje když se na chvíli zastaví myši nad obrázkem. Je trošku méně významný než alt a měl by obsahovat obecné shrnutí obsahu obrázku.

U obrázků o velikosti 1x, které slouží pouze ke grafickým účelům je nutné nechat atribut alt prázdný. Použití klíčového slova u takového obrázku by mohl vyhledávač shledat jako spam.

Zbytečné tagy na stránce

Mezi zbytečné tagy na stránce patří například

```
<meta name="revisit-after" content="14 days">
<meta name="robots" content="index, follow">
<!-- poznámka -->
```

Žádný vyhledávací robot si nenechá diktovat, kdy má stránku navštívit (pokud se mu nezaplatí). Frekvence návštěv je v drtivé většině dána popularitou webu,

¹ Například Google nepřikládá tučnému textu téměř žádnou váhu, naopak u Yahoo! to nějakou váhu má, ale také ne příliš velkou.

tedy počtem kvalitních odkazů, které na stránku vedou a aktuálnosti obsahu (jak často se něco na stránce mění). Také nemá cenu říkat robotu, že stránku může zaindexovat a že má následovat odkazy na stránce. Vyhledávač totiž dopředu počítá s tím, že zaindexovat může vše, co není zakázané. Poznámka se nezobrazuje uživateli, takže vyhledávače v ní nehledají žádný užitečný obsah.

4.1.10. Optimální počet klíčových slov v textu stránky

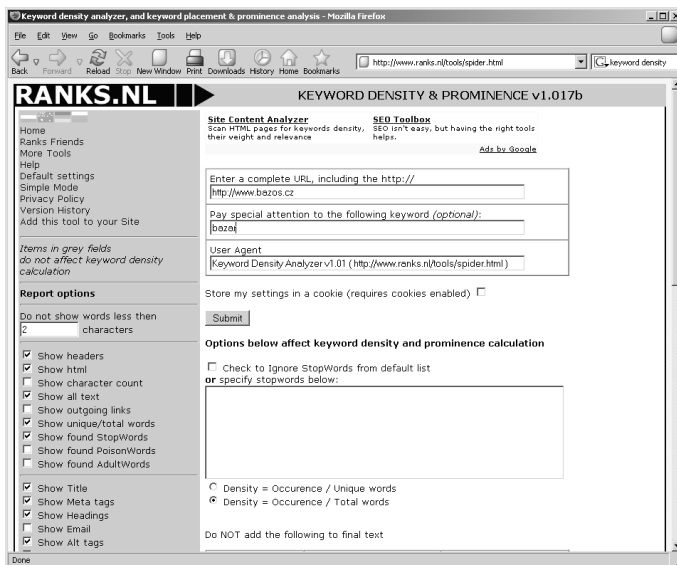
Vyhledávače neurčují důležitost klíčového slova na stránce podle jeho četnosti, ale podle hustoty výskytu klíčového slova. Hustota slova na stránce tedy vyjadřuje podíl jeho četnosti (počtu výskytů) ku celkovému počtu slov celého textu. Pro klíčové slovo se za optimální většinou považuje hustota mezi 2 až 7 %. Pokud hustota převyší určitou hodnotu (pro každý vyhledávač jinou) může to mít záporný efekt. Hovoří se o takzvané přeoptimalizaci neboli OOP (Over Optimized Penalty).

$$\frac{\text{Počet výskytů klíčového slova}}{\text{Celkový počet slov}} = \text{Hustota slova}$$

Hustota klíčových slov bývá často přeceňována, zdaleka však nepatří mezi nejdůležitější kritéria. Pro počítání hustoty slov existuje na webu celá řada nástrojů, mezi ty lepší patří například Keyword density analyzer.¹ Optimální počet slov na stránce lze zjistit tak, že začneme na malé hustotě (2 procenta) a postupně přidáváme klíčová slova (zvyšujeme hustotu). Po přidání jednoho slova je nutné počkat až se změna projeví v testovaném vyhledávači (například v Googlu). Pokud se stránka posunuje nahoru, lze hustotu dále zvyšovat. Bohužel existuje celá řada dalších vlivů (změny na ostatních stránkách), které není možné úplně odstínit. Navíc tento postup může trvat neúměrně dlouhou dobu, za kterou může být ideální hodnota vyhledávačem změněna.

¹ <http://www.ranks.nl/tools/spider.html>

Obrázek 12: Keyword density analyzer



4.1.11. Klíčová slova na začátku tagu (stránky)

Čím blíže je slovo od začátku tagu (stránky), tím má větší důležitost. Vzniká tak ale problém, zda například v tagu Title použít Internetové knihkupectví Jasminka a nebo Jasminka - Internetové knihkupectví. Z hlediska optimalizace je lepší použít první variantu, z hlediska budování značky je jednoznačně lepší druhá. Začátek tagu Title se totiž ukazuje například v taskbaru, kde je uživateli neustále na očích.

Vždy je nutné mít na paměti, že webové stránky se vytváří primárně pro lidi a ne pro vyhledávače. Optimalizace nikdy nesmí vypadat jako by tam nepatřila. Umístit klíčová slova na začátku stránky (tag body) lze celkem jednoduše udělat pomocí CSS stylů.

4.1.12. Klíčová slova v URL

Většina vyhledávačů přikládá URL hodně velký význam. Je tedy vždy užitečné mít klíčové slovo v URL. Větší význam mívá doménové jméno než zbytek URL. Osobně si však nemyslím, že je dobré mít jako jméno domény obecný vyhledávaný výraz. Lze se o tom jednoduše přesvědčit, je lepší název Google a nebo Searchengine? Searchengine bude jistě vyhledávanější slovo, ale značka je značka. Přece jen, má snad někdo název webové stránky vyhledávač? Co se týká zbytku URL za lomítkem, tak zde žádné dilema není. Vždy je tam užitečné mít klíčová slova, které se oddělují znaky - (mínus) a / (lomítko). Podtržítko (_) slouží jako spojovací znak. Vyhledávač čte prodejnu_odevu jako prodejnuodevu (Činí tak například Google. Ve výsledcích sice označí **prodejnuodevu**, ale je to pouze označení).



Nic se ale nemá přehánět, čím kratší URL je tím lépe. Takže v žádném případě nedoporučuji URL typu www.prodej-realit-brno-praha-inzerce-zdarma.cz.

4.1.13. Vlastní obsah

Obsah je na webu to nejdůležitější. Většinou platí, že čím kvalitnější obsah tím méně optimalizace je potřeba. Protože na stránku s kvalitním obsahem častěji lidé odkazují. Obsah se píše vždy pro uživatele a ne pro vyhledávače, měl by tak být co nejvíce přirozený.

Čím více kvalitního a aktuálního obsahu na webu je, tím více uživatelů z vyhledávačů může získat. Web, který má 10 stránek s kvalitním obsahem může získat například 30 kliků z vyhledávačů denně. Web, který má 1000 stránek s kvalitním obsahem může získat 300 a více kliků za den (čísla jsou

uvedená pouze pro názornost, počet kliků z vyhledávačů závisí na mnoha faktorech, například také na tématu webu).

4.2. Off page faktory

Off Page faktory je vše co není přímo na optimalizované stránce, hlavně tedy všechny odkazy, které směřují na danou stránku. Tímto se nemyslí jen odkazy z cizích stránek (webů), ale samozřejmě i z webu vlastního. V podstatě všechny spočítané odkazy podle důležitosti v sobě vyjadřuje ranking stránky. Kromě odkazů řadíme mezi off-page faktory také například anchor text, URL odkazu, title a alt, kterým je věnovaná kapitola č. 4.3.

4.2.1. PageRank

PageRank je obchodní značka, která patří Googlu. Byl vyvinut jako číselný systém hodnocení relativní důležitosti webových stránek. Vytvořili ho zakladatelé Googlu Larry Page a Sergey Brin na Stanford University v Kalifornii. Koncept který použili byl jejich vlastními slovy "to calculate the uniquely democratic nature of the web by using its vast link structure as an indicator of an individual page's value"[2].

Když se to vezme doslova, tak se celý systém opírá o přichozí a odchozí odkazy miliard webových stránek, které tvoří internet. Pokud web A odkazuje na web B, dává mu vlastně hlas, který zvyšuje jeho důležitost. Je to samozřejmě velice zjednodušená definice. Úplně něco jiného je pozice stránky při vyhledávání určité fráze (Search engine results page - SERP). Až kombinace PageRanku, obsahu stránky a řady dalších faktorů určuje pozici na vyhledávané slovo. Důležité je, že PageRank má každá jednotlivá stránka, a ne celý web. Také hodnota přichozích odkazů je různá, a to podle důležitosti stránky, z které tento odkaz přichází, a počtu odkazů na této stránce.

4.2.2. Algoritmus PageRanku

Původní algoritmus PageRanku jak byl zveřejněn jeho tvůrci.

$$PR(A) = (1-d) + d (PR(T1)/C(T1) + \dots + PR(Tn)/C(Tn))$$

kde

- $PR(A)$ je PageRank stránky A ,
- $PR(Ti)$ je PageRank stránek Ti , které odkazují na A ,
- $C(Ti)$ je počet odchozích odkazů na stránce Ti a
- d je faktor útlumu (damping factor), který je mezi 0 a 1.

V současnosti není aktuální PageRank znám, takže o jeho nastavení či tvaru se vedou jen dohady. Měl by ale vycházet z principů této rovnice.

4.2.3. Model náhodného surfaře

Ve svých publikacích dali Lawrence Page a Sergey Brin [2] také velmi jednoduché vysvětlení algoritmu PageRanku. Uvažovali o PageRanku jako o modelu chování, kde surfař kliká na odkazy náhodně bez ohledu na obsah.

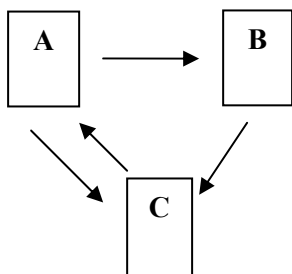
Náhodný surfař navštívuje stránku s určitou pravděpodobností, která je odvozena z PageRanku. Pravděpodobnost, že náhodný surfař klikne na odkaz je výhradně dána počtem odkazů na stránce. To je důvod proč PageRank stránky není kompletně předán dál na stránky na které odkazuje, ale je rozdělen na počet odkazů na stránce.

Pravděpodobnost pro náhodného surfaře, že dosáhne dané stránky, je suma pravděpodobností náhodného surfaře následujícího odkazy na tuto stránku. Tato pravděpodobnost je snížena o faktor útlumu (damping factor d).

Opodstatnění je to, že náhodný surfař nekliká na nekonečné množství stránek, ale časem se unaví a sem tam odskočí náhodně na jinou stránku.

Pravděpodobnost, že náhodný surfař nepřestane klikat na odkazy je dána faktorem útlumu, který je nastaven mezi 0 a 1. Čím vyšší je d , tím větší je šance, že surfař bude dále klikat. Pravděpodobnost, že surfař odskočí na jinou náhodnou stránku, poté co přestane klikat, je pak $(1-d)$. Bez ohledu na počet odkazů vedoucích na stránku pravděpodobnost, že na ni náhodný surfař odskočí je vždy $(1-d)$, takže stránka má vždy minimální PageRank.

4.2.4. Charakteristika PageRanku



Jak PageRank funguje se dá ukázat na příkladu, který jsem našel na německé stránce eFactory [16]. Uvažujeme malý web o třech stránkách A, B a C, kde stránka A odkazuje na B a C, stránka B odkazuje na C a stránka C odkazuje na A. Podle původní verze měl být faktor útlumu na 0,85. Ale aby se to dalo lehce spočítat použijeme 0.5. Po

doplnění dostaneme následující rovnice:

$$PR(A) = 0.5 + 0.5 PR(C)$$

$$PR(B) = 0.5 + 0.5 (PR(A) / 2)$$

$$PR(C) = 0.5 + 0.5 (PR(A) / 2 + PR(B))$$

Tyto rovnice jdou jednoduše vyřešit a zde jsou výsledky:

$$PR(A) = 14/13 = 1.07692308$$

$$PR(B) = 10/13 = 0.76923077$$

$$PR(C) = 15/13 = 1.15384615$$

Suma všech PageRanků je 3, a tak se rovná počtu všech webových stránek. V praxi je to trochu složitější, protože web má miliardy stránek, takže nalézt řešení není tak jednoduché. Právě kvůli velikosti webu nejenom Google používá přibližný odhad PageRanku. Každá stránka dostane na začátku počáteční hodnotu a PageRank ostatních je dopočítán v několika výpočetních cyklech, které jsou založeny na základní rovnici algoritmu PageRanku. Pro náš příklad použijeme počáteční hodnotu 1.

Tabulka 2: Cykly odhadu Google PageRanku

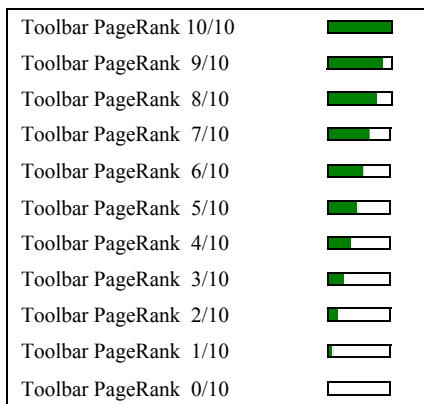
Opakování	PR(A)	PR(B)	PR(C)
0	1	1	1
1	1	0.75	1.125
2	1.0625	0.765625	1.1484375
3	1.07421875	0.76855469	1.15283203
4	1.07641602	0.76910400	1.15365601
5	1.07682800	0.76920700	1.15381050
6	1.07690525	0.76922631	1.15383947
7	1.07691973	0.76922993	1.15384490
8	1.07692245	0.76923061	1.15384592
9	1.07692296	0.76923074	1.15384611
10	1.07692305	0.76923076	1.15384615
11	1.07692307	0.76923077	1.15384615
12	1.07692308	0.76923077	1.15384615

Zdroj: <http://pr.efactory.de/e-pagerank-algorithm.shtml>

Kvalitního odhadu PageRanku lze dosáhnout jen po pár opakováních. Podle Lawrence Page a Sergey Brina je potřeba 100 opakování, abychom dostali dobrý odhad hodnot PageRanků celého webu. To také znamená, že součet

všech PageRanků se stále blíží celkovému počtu webových stránek. Průměrný PageRank je roven 1. Minimální PageRank je dán $(1-d)$. Maximální PageRank můžeme dostat pokud by všechny stránky odkazovaly na jedinou a ta by odkazovala sama na sebe. Hodnota by pak byla rovna $dN+(1-d)$, kde N je celkový počet stránek.

4.2.5. Toolbar PageRank



V roce 2002 vydal Google Toolbar pro Microsoft Internet Explorer (<http://toolbar.google.com/>), na kterém se zobrazuje PageRank. Google Toolbar zobrazuje PageRank ve škále od 0 do 10. Nulový PageRank (PR0) znamená, že na danou stránku vede minimum odkazů (např. nová stránka než ji Google zaindexuje, nový web mívá "not ranked"). V případě, že se jedná o starší stránku, PR0 většinou

znamená, že stránka použila nekalé (zakázané) techniky a je penalizována.

4.2.6. Skutečný PageRank

Skutečný PageRank je trochu jiné číslo, než se zobrazuje v Google Toolbaru. Předpokládá se, že nabývá vyšších hodnot (pro větší rozlišovací schopnost 0 - 1 oproti 0 - 100000). Pravděpodobně je pro převod použita logaritmická funkce, ale to by samozřejmě mohli potvrdit pouze lidé z Googlu, a ti takovou informaci nikdy neuveřejní.

Toolbar PageRank
0/10
1/10

Skutečný PageRank
0.15 - 0.9
0.9 - 5.4

2/10	5.4 - 32.4
3/10	32.4 - 194.4
4/10	194.4 - 1,166.4
5/10	1,166.4 - 6,998.4
6/10	6,998.4 - 41,990.4
7/10	41,990.4 - 251,942.4
8/10	251,942.4 - 1,511,654.4
9/10	1,511,654.4 - 9,069,926.4
10/10	9,069,926.4 - $0.85 \times N + 0.15$

V současné době Google přepočítává PageRank přibližně jednou za měsíc (je to ale nepravidelné, dohromady s PageRankem mění i zpětné odkazy). U ostatních vyhledávačů je to například u Alltheweb jednou týdně a u českého Jyxa se JyxoRank počítá každý den. S postupem času Google pravděpodobně přepočítávání urychlí (je trochu něco jiného spočítat rank po celém světě anebo jen v Česku).

Protože je u PageRanku použita pravděpodobně logaritmická funkce je dosažení zlepšení PR 4 na PR 5 daleko jednodušší než dosáhnout změny PR 7 na PR 8. Ačkoliv se zdá, že význam PageRanku již poslední dobou není tak velký jako na začátku svého uvedení, není to pravda. Klesl jen význam Toolbar PageRanku, skutečný PageRank (zvláště pokud jde o tématický PageRank) svůj význam stále má.

4.2.7. Tématicky a předmětově orientovaný PageRank

Již řadu let se hovoří o tom, že při vyhledávání by se mělo zohlednit téma nebo předmět, jako jedno z kritérií. Mezi hlavní důvody patří eliminace spamových stránek, které mají koupené odkazy. Tím jak implementovat téma do vyhledávání se zabývá celá řada teoretických přístupů. Všechny mají společné to, že stránka není hodnocena jenom podle svého obsahu, ale také podle obsahu dalších stránek. Například obsah celého webu může ovlivňovat ranking

každé jednotlivé stránky. Pokud je to například web knihkupectví a na některé stránce je zmínka o kladivu, tak to kladivo by mělo při použití tématického PageRanku dostat nižší hodnocení než kdyby bylo na webu železářství.

Je také možné, že ranking jednotlivé stránky je odvozen od obsahu stránek, které na ni odkazují. To je jedním z důvodů proč se obecně doporučuje dělat výměny odkazů především se stránkami, které mají podobné zaměření. Jestli opravdu vyhledávače používají téma jako jeden z faktorů, je celkem těžké určit, protože je to většinou obchodní tajemství. Uvedu dva přístupy, které se implementací tématicky zaměřeného PageRanku zabývají. Jedním z nich je model inteligentního surfaře od Matthew Richardsona a Pedra Dominga a druhým z nich je Topic Sensitive PageRank od Tahera Haveliwaly.

4.2.8. Model inteligentního surfaře

V modelu Matthew Richardsona a Pedra Dominga se mění náhodný surfař na surfaře inteligentního. Náhodný surfař klikal na odkazy náhodně, oproti tomu inteligentní surfař sleduje pouze odkazy, které mají něco společného s původním tématem stránky. Potom, co se unaví, tak odskakuje na stránku, která má něco společného s původní. Pro Richardsonova a Domingova "inteligentního surfaře" jsou relevantní pouze stránky, které obsahují slovo, které původně hledal. To ovšem znamená, že pro každé slovo, které se na stránce objeví, je potřeba vypočítat vlastní PageRank založený na odkazech mezi stránkami, které dané slovo obsahují.

Pokud by se PageRank počítal tímto způsobem, přineslo by to řadu problémů. A to hlavně v případech, kdy se vyhledávané slovo nevyskytuje často. Proto, aby se dalo specifické slovo zahrnout do výpočtu, musí být nejenom na dané stránce, ale také na stránkách, které na ni odkazují. To ovšem znamená, že výsledky budou založeny na malém výseku webu a mohou být opomenuté relevantní stránky. Také pokud se použije malý výsek, je vyhledávání

náchylnější na automaticky generované stránky se spamem, a to je jeden z důvodů, proč se vůbec o tématickém PageRanku uvažuje.

Problém je také ve vypočítání takového PageRanku. Pokud bychom uvažovali 100 000 výrazů, byla by potřeba času pro vypočítání tématicky zaměřeného PageRanku oproti originálnímu 100 až 200 větší. Například pokud uvažujeme, že se původní PageRank počítá pět hodin, tak tématický PageRank by se počítal nejméně 3 týdny.

4.2.9. Topic-Sensitive PageRank

Přístup Tahera Haveliwaly [5] vypadá pro aktuální zavedení vyhledávači trochu praktičtější. Stejně jako Richardson a Domingos používá Haveliwala různé PageRanky. Rozdíl je v tom, že PageRank se nepočítá pro každé slovo, ale pro každé téma. Pro každé toto téma odlišuje různé authority pages. Pro svůj experiment si Haveliwala vybral jako autoritativní stránky katalog Open Directory Project (ODP), konkrétně 16 hlavních kategorií. Pokud se tedy počítá PageRank pro téma zdraví, všechny stránky v kategorii zdraví dostanou vyšší hodnotu E v rovnici PageRanku :

$$PR(A) = E(A) (1-d) + d (PR(T1)/C(T1) + \dots + PR(Tn)/C(Tn))$$

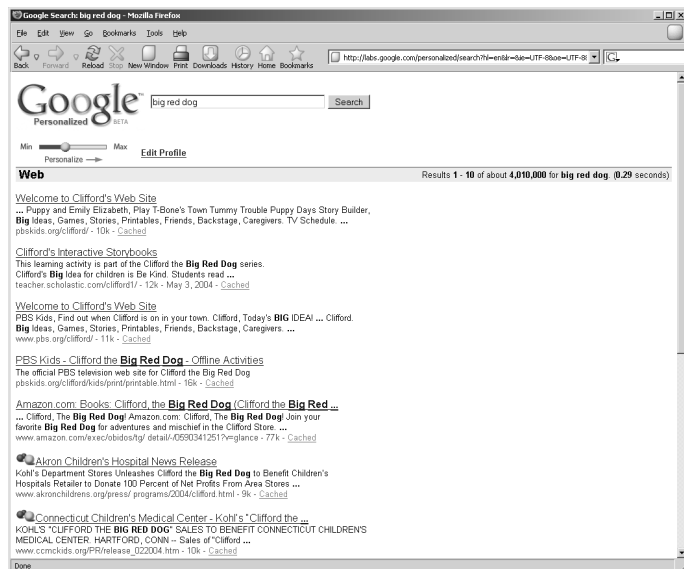
Vyšší PageRank pro zdraví se pak přerozděluje dále. Protože předpokládáme, že stránky mají tendenci odkazovat na stránky s podobným tématem, vyšší PageRank pro zdraví se pak týká všech relevantních stránek. Haveliwala ve své práci zjistil, že výsledky vyhledávání při použití Topic-Sensitive PageRanku jsou vzhledem k dotazu o dost relevantnější, a jeho implementace by prý neměla být příliš náročná. Problémem je ale závislost na katalogu ODP a na jednotlivých editorech (kategorie může obsahovat chybně zařazené stránky, které s tématem nemají nic společného).

Dalším problémem je, že potřebujeme znát uživatelské preference. Pokud nevíme, z jakého tématu chce uživatel výsledky, nemůžeme mu je ani dodat. Hledání by se muselo rozčlenit na jednotlivé kategorie. Pokud hledám internetový obchod, použije se vyšší Topic-sensitive PageRank pro stránky, na které je odkazováno v ODP kategorii nakupování. Mezi možnostmi zmapování uživatelských preferencí jistě patří použití Google Toolbaru, ale ne všichni mají Toolbar nainstalovaný. Metodu s toolbarem by nejlépe mohl implementovat Microsoft ve svém připravovaném vyhledávači. Pokud by do standardního Internet Exploreru přidal svůj Toolbar a pomocí něho vyhodnocoval uživatelské preference. Ale co když uživatel hledá něco, co ještě nehledal?

Přístup Tahera Haveliwalý pravděpodobně používá Google ve svém personalizovaném vyhledávání¹. Jelikož je toto vyhledávání v Beta verzi musí si uživatel vybrat z jaké kategorie chce výsledky. Google tedy zatím nezkoumá údaje z toolbaru, ani pravděpodobně pomocí jiných metod (možné to ale je).

¹ <http://labs.google.com/personalized>

Obrázek 13: Personalizované vyhledávání Google



4.2.10. SEO a tématický PageRank

Je jasné, že dříve či později implementují vyhledávače nějaký způsob tématického PageRanku. Takže pořád platí obecná zásada, že výměna odkazů s podobně zaměřenými stránkami se vyplatí více, než výměna odkazů se stránkou, která má úplně odlišné téma. Pokud by byl implementován Topic-sensitive PageRank, určitě se vyplatí mít zaregistrovanou stránku ve správné kategorii ODP. Pokud by něco podobného použily české vyhledávače, pak by byla nutnost být ve správné kategorii i v českých katalogích (což je ovšem i tak nezbytnost z důvodu, že je stále dost přístupů na stránky z katalogů).

4.2.11. BadRank a PageRank 0

Nulový PageRank neznamena vždy, že stránka je penalizována. Může to také znamenat, že je nová nebo na ni vede minimum odkazů. Při použití neetických

technik, ale mohou vyhledávače samozřejmě stránku penalizovat. Děje se tak buď automaticky nebo stížností na takovou stránku. Podrobnostmi neetických taktik blíže rozebírá zvláštní kapitola č. 5 Zakázané metody v SEO.

Problém penalizace je však spojen i s BadRankem a s takzvaným špatným okolím stránky. Vezměme v úvahu dvě stránky, stránka **A** je penalizovaná, stránka **B** nikoliv. Pokud **A** odkazuje na **B** a **B** neodkazuje zpět, BadRank se nepřenáší. Pokud však stránka **B** odkazuje zpět, dostane od penalizované **A** BadRank. BadRank se tedy přenáší pouze pokud odkazujete na penalizovanou stránku zpět, a to z důvodu, že nelze ovlivnit odkaz z penalizované stránky.

4.2.12. Rozdělení interního PageRanku

PageRank, který je dosažen pomocí zpětných odkazů, je z úvodní stránky přerozdělován do ostatních. Systém interních odkazů lze udělat takovým způsobem, že určité stránce na webu se PageRank zvýší a jiné sníží [21]. Tedy ty stránky, které jsou vhodné lze upřednostnit. Na začátku je nutné si definovat stránky, které chceme preferovat (s vyšším nutným PR) a ty které chceme, aby se jen zaindexovaly, a není nutné aby byly ve vyhledávačích moc viditelné. Pokud je optimalizovaná stránka webový obchod, lze to rozdělit na stránky, které prodávají a na ty ostatní. Samozřejmě ty, které nechceme indexovat vůbec, lze zakázat v robots.txt.

Na stránky u kterých je požadován vyšší PageRank používáme standardní odkazy ve tvaru:

```
<a href="abc.html" title="abc">abc</a>.
```

Místo abc se samozřejmě používá klíčové slovo. Stránka kromě části PageRanku dostává také anchor text a title. Na stránky, které chceme sice

zaindexovat, ale není nutné aby měly vysoký PageRank, se odkazuje pomocí speciálního souboru redirect.php.

```
<a HREF="redirect.php?stranka=dfg.html">dfg</a>
```

Tento soubor nedělá nic jiného, než že přesměruje dotaz na požadovanou stránku. Pokud je v souboru robots.txt zakázané tento soubor indexovat, PageRank se nepřenese. Samozřejmě je nutné, aby na danou stránku vedl alespoň jeden “normální” odkaz, aby byla zaindexována. Pro příklad uvedu web, který má deset stránek. Normální HTML odkazy označím A,B,C,D,E,F,G,H,I,J a přesměrované odkazy pak a,b,c,d,e,f,g,h,i,j.

HTML odkazy	Přesměrované odkazy
A,B,C,D,E,F,G,H,I,J	a,b,c,d,e,f,g,h,i,j

Na začátku předpokládejme, že je k dispozici 100 PR bodů, které byly získány ze zpětných odkazů (odkazy z externích webových stránek). Pokud je potřeba zvýšit PR úvodní stránce, a to z jakéhokoli důvodu, (např. je jediná optimalizovaná, nebo nechceme aby návštěvníci chodili přes jiné), je potřeba udělat navigaci webu například viz. Tabulka č. 3.

Tabulka 3: Možnost přerozdělení PageRanku na úvodní stránku

Stránka A	A	B	C	D	E	F	G	H	I	J
Stránka B	A	b	c	d	e	f	g	h	i	j
Stránka C	A	b	c	d	e	f	g	h	i	j
Stránka D	A	b	c	d	e	f	g	h	i	j
Stránka E	A	b	c	d	e	f	g	h	i	j

Stránka F	A	b	c	d	e	f	g	h	i	j
Stránka G	A	b	c	d	e	f	g	h	i	j
Stránka H	A	b	c	d	e	f	g	h	i	j
Stránka I	A	b	c	d	e	f	g	h	i	j
Stránka J	A	b	c	d	e	f	g	h	i	j

Zdroj: <http://www.seo-guy.com/pagerank-articles/channeling-pagerank.html>

Jak lze vidět, z hlavní stránky se odkazujeme HTML odkazy na všechny ostatní, ale z těch se již odkazujeme pomocí HTML odkazů pouze na hlavní stránku a pomocí přesměrovaných odkazů na ostatní. To vede k hromadění PR na hlavní stránce.

Odkazy pomocí HTML

Hlavní stránka = 127,56 bodů

Všechny ostatní = 49,9 bodů

Na vypočítání jednotlivých bodů je možné použít jednoduchý Google PageRank kalkulátor¹, který jednotlivé váhy dokáže vypočítat.

Odkazy pomocí přesměrování podle tabulky nahoře

Hlavní stránka = 310,98 bodů

Všechny ostatní = 29,52 bodů

Předpokládejme, že stránky A,B,C,D, a E jsou optimalizované a chceme aby měly vyšší PR než ostatní. Takže na ně budeme odkazovat HTML odkazy a na ostatní přesměrovanými odkazy (až na jeden, aby se zaindexovaly).

¹ http://www.webworkshop.net/pagerank_calculator.php3

Tabulka 4: Možnost přerozdělení PageRanku na stránky A, B, C, D a E

Stránka A	A	B	C	D	E	f	g	h	i	j
Stránka B	A	B	C	D	E	f	g	h	i	j
Stránka C	A	B	C	D	E	f	g	h	i	j
Stránka D	A	B	C	D	E	f	g	h	i	j
Stránka E	A	B	C	D	E	F	g	h	i	j
Stránka F	A	B	C	D	E	f	G	h	i	j
Stránka G	A	B	C	D	E	f	g	H	i	j
Stránka H	A	B	C	D	E	f	g	h	I	j
Stránka I	A	B	C	D	E	f	g	h	i	J
Stránka J	A	B	C	D	E	f	g	h	i	j

Zdroj: <http://www.seo-guy.com/pagerank-articles/channeling-pagerank.html>

Hlavní stránky

A = 166,58 bodů

B = 96,48 bodů

C = 96,48 bodů

D = 96,48 bodů

E = 99,98 bodů

Ostatní stránky

f = 17,14 bodů

g = 2,57 bodů

h = 0,51 bodů

i = 0,22 bodů

j = 0,18 bodů

Stejně tak se dá zvýšit PageRank například u stránek A a B.

Tabulka 5: Možnost přerozdělení PageRanku na stránky A a B

Stránka A	A	B	c	d	e	f	g	h	i	j
Stránka B	A	B	C	d	e	f	g	h	i	j
Stránka C	A	B	c	D	e	f	g	h	i	j
Stránka D	A	B	c	d	E	f	g	h	i	j
Stránka E	A	B	c	d	e	F	g	h	i	j
Stránka F	A	B	c	d	e	f	G	h	i	j
Stránka G	A	B	c	d	e	f	g	H	i	j
Stránka H	A	B	c	d	e	f	g	h	I	j
Stránka I	A	B	c	d	e	f	g	h	i	J
Stránka J	A	B	c	d	e	f	g	h	i	j

Zdroj: <http://www.seo-guy.com/pagerank-articles/channeling-pagerank.html>

Hlavní stránky

A = 218,40 bodů

B = 223,90 bodů

Ostatní stránky

c = 95,3 bodů

d = 27,15 bodů

e = 7,84 bodů

f = 2,37 bodů

g = 0,82 bodů

h = 0,38 bodů

i = 0,25 bodů

j = 0,22 bodů

Takhle se dá optimalizovat libovolné množství stránek. Na ty, které chceme aby měly malé PR, odkazujeme pouze jednou pomocí “normálního“ HTML odkazu a PR se koncentruje na těch ostatních. Podobného systému se dá docílit i použitím vhodného navigačního systému bez přesměrovaných odkazů se zakázanou indexací. Pokud je navigační systém hierarchický, hromadí se PR hlavně na stránkách více nahoře v hierarchii, kde jsou většinou více obecná slova, na které je větší konkurence. Čím níže v hierarchii, tím méně konkurenční slovo (specifické pro danou oblast), a z toho důvodu stačí daleko menší PageRank.

Dříve bylo možné tento systém udělat pomocí Javascriptových odkazů, ale stále více vyhledávačů se snaží Javascript, alespoň částečně zaindexovat. Z toho důvodu je účinnější použít přesměrovaný odkaz, který lze zakázat indexovat vyhledávači v souboru robots.txt.

4.2.13. Ranky ostatních vyhledávačů

Většina současných vyhledávačů používá nějakou obdobu PageRanku. Samozřejmě jsou mezi nimi určité odlišnosti, ale pro SEO nemají tyto odlišnosti příliš velký význam. Nic se nemění na tom, že čím více kvalitních odkazů, pokud možno z webů ze stejným tématem, tím lépe. Ne jenom pro lepší pozici ve vyhledávačích, ale také kvůli tomu, že některé katalogy řadí podle rankingu stránky v jednotlivých sekcích. Například Seznam podle S-Ranku, Atlas podle JyxoRanku, Centrum podle Q a Google Directory pak podle PageRanku.

Malou výjimku tvoří WebRank od Yahoo!, což je kombinace odkazů vedoucích na stránku, klikavosti ve výsledcích vyhledávání a návštěvnosti měřené pomocí Yahoo! toolbaru.

4.2.14. Mapa webu

Mapu webu osobně doporučuji pouze jako nouzové řešení, tehdy když web není dobře indexovatelný (používá rámce, Flash, javascripty v menu, viz. kapitola č.3. Předpoklady pro optimalizaci). Pokud uživatel musí procházet mapu serveru, dost často to znamená, že se neorientuje v normálním systému navigace daného webu. Mapa serveru není dobrá technika z důvodu, že je na ní většinou velké množství odkazů. Takže PageRank, který získá, musí rozdrobit na tyto odkazy. Pokud je jich tam 100, v podstatě každý odkaz dostane 1/100 PageRanku. Navíc každý odkaz stejně. Pokud má web server hierarchický systém navigace, stránky nahoře hierarchie dostávají větší PageRank, ty dole, kde většinou klíčová slova nejsou tak významná, pak nižší. Toto vše mapa serveru nedokáže. Navíc například Google na své stránce doporučuje¹ maximálně 100 odkazů na jednu stránku. Takže pokud je na webu více než 100 stránek, musíme mapu rozdělit na více částí (ty pak dostanou ještě nižší PageRank).

4.2.15. Jak získávat zpětné odkazy

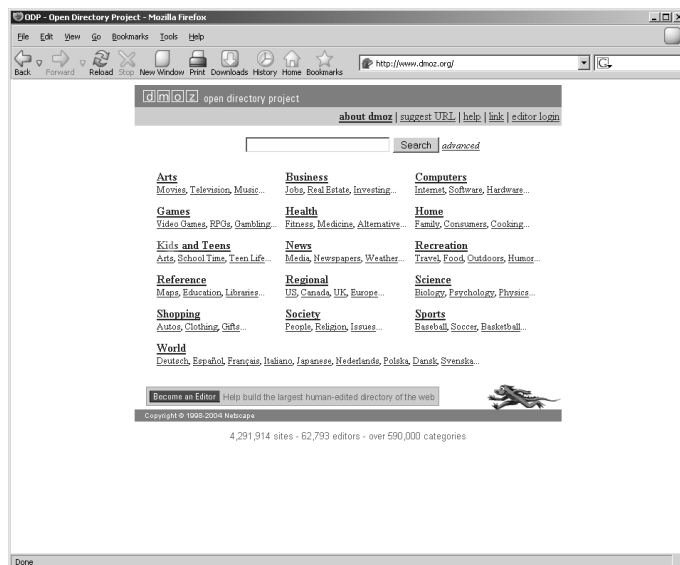
Nejednoduší jak získávat odkazy je jednoznačně registrace do internetových katalogů. Registrace do nich je alespoň u nás většinou zdarma (v zahraničí jsou placené katalogy daleko častější). Mezi nejdůležitější české katalogy, do kterých je nutnost se zaregistrovat, patří Seznam, Centrum, Atlas, Tiscali, Caramba a Quick. Podle mě se však vyplatí zaregistrovat i do těch menších, v žádném případě ne však automaticky pomocí programků jako je Submit Wolf. Jednoznačně se vyplatí ruční práce. Českých malých katalogů je spousta, krátký seznam je například na Katalogy.iglu². Ze zahraničních se

¹ <http://www.google.com/webmasters/guidelines.html>

² <http://www.katalogy.iglu.cz/>

zmíním o DMOZ¹ Open Directory Project, který používá celá řada vyhledávačů (u výsledků zobrazují navíc vyhledávání z katalogu) a katalogů.

Obrázek 14: Důležitý zahraniční katalog ODP



V případě získávání odkazů pro anglické stránky, není pravděpodobně možné se vyhnout registraci do placených katalogů. Mezi neznámější patří Yahoo!, které je pro komerční projekty placené a pro nekomerční sice neplacené, ale pokud není uhrazen poplatek, zařazení trvá neúměrně dlouho (i půl roku).

Registrace do placených katalogů vyžaduje obzvlášť pozornost. Protože tyto katalogy negarantují za poplatek zařazení do dané sekce, ale pouze to, že editor příslušné sekce požadavek zhodnotí a do určité doby rozhodne. Registrace do špatné sekce, neadekvátní popis, duplicitní web nebo jiné prohřešky mohou v tomto případě stát dost peněz.

¹ <http://www.dmoz.org/>

Na internetu nejsou katalogy jediným místem, kde se dají získávat odkazy. Jelikož se již nějakou dobu mluví o tématickém vyhledávání, je užitečné získávat odkazy z podobně zaměřených webů. Při výměně odkazů je bezpodmínečně nutné zkontrolovat PageRank protistrany. Nevyplatí se udělat výměnu s webem, který má o řád nižší PageRank (3 oproti 5). Obecně se doporučuje nevyměňovat odkaz se stránkou s nižším PageRankem než 4, kterého se dá dosáhnout pouhou registrací do katalogů. Je nutné zkontrolovat PageRank stránky, kde bude odkaz umístěn, a ne PageRank hlavní stránky webu. Tento PageRank je pak rozdělen na všechny odkazy, takže pokud je na stránce 50 odkazů, přerozděluje se na každý odkaz 1/50.

Možnost jak zjistit, kde získat odkazy nabízí samotné vyhledávače, a to pomocí zjištění odkazů, které vedou na konkurenční stránky. Do vyhledávače stačí zadat "*link: www.example.com*". Vyplatí se to zadat do více vyhledávačů, protože například Google takhle zobrazuje jen stránky s PageRankem vyšším než 4.

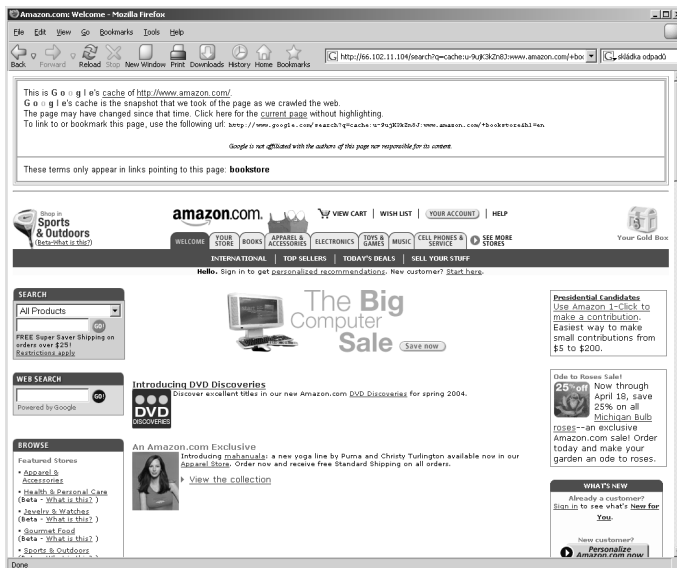
Čím lepší stránka, tím je zpravidla jednodušší získávat zpětné odkazy. Na dobrou stránku totiž každý odkáže mnohem raději než na tu, na které není nic zajímavého.

4.3. Anchor text (link text)

Vyhledávače neberou v úvahu jen faktory, které jsou na stránce, ale i ty co přicházejí z odkazů. Dost často se stává, že vyhledávané slovo na dané stránce vůbec není, i přesto je však daná stránka vyhledána. Je to způsobený právě **anchor textem** (`zde je anchor text`). Pokud se odkaz na stránku přidá například do katalogů s anchor textem `bbbauto`. Získá se jen slovo bbbauto, což asi moc lidí hledat nebude. Pokud se ale jako anchor text použije `Autobazar BBbauto.cz ` lze získat slovo autobazar. Váha anchor textu je v současné době u vyhledávačů celkem

významná, osobně si myslím, že je hned za tagem <title>. Je to prostě nevyužitá příležitost, jak zvýšit váhu určitého slova.

Obrázek 15: Slovo bookstore na Amazonu není



Samozřejmě, že využití anchor textu neplatí jen pro externí odkazy, ale také pro interní (i když pro ně o něco méně - mají menší váhu). Pokud je na stránce seznam produktů, a pro detail je nutné kliknout na odkaz [zde](abc), pak detail produktu získá klíčové slovo zde, což je celkem zbytečné. Pokud jako anchor text je uvedeno například Ferrari 612 Scaglietti, je to jistě vyhledávanější slovo než "zde".

Pro zjištění stránek na které se odkazuje určitým anchor textem se používá v Googlu operátor `allinanchor:text`. Seznam dalších užitečných operátorů lze nalézt přímo na Googlu¹, bohužel `allinanchor` mezi nimi není.

4.3.1. Okolí odkazu

Kromě anchor textu například Google porovnává okolí odkazu a pravděpodobně také téma stránky. Z tohoto důvodu, jsou odkazy na podobně zaměřených stránkách mnohem cennější, než odkaz na stránce, z které vedou odkazy na sex kluby, dovolenou v Chorvatsku a zároveň na prodejnu oděvů.

Odkaz na stránce s hodnotným a stejně zaměřeným obsahem je daleko cennější než odkaz na stránce se seznamem odkazů. Takové odkazy je ovšem mnohem těžší získat. Stránky se seznamem odkazů také většinou mívají nižší PageRank, než stránky s hodnotným obsahem.

4.3.2. Obrázkové versus textové odkazy

Textové odkazy jsou na stránce daleko lepší než odkazy obrázkové. Jednodušeji se dají změnit, pomocí CSS mohou vypadat velice pěkně a jsou téměř ihned stáhnuté i na modemu. Oproti tomu obrázek dlouho trvá, než se stáhne a samotný nenes pro vyhledávače žádnou hodnotu. Při použití obrázků je minimálně nutné použít vlastnosti `alt` a `title`. Uživatel může mít vypnuté obrázky, potom se mu zobrazuje vlastnost `alt`. Pomocí atributu `title` lze dát uživateli informaci o obrázku navíc, například co se stane když na něj klikne. Vyhledávače neví co je v obrázku, takže atributy `title` a `alt` jsou pro ně jediným zdrojem informací. Samozřejmě, že nemají takovou váhu, jako anchor text.

Váhu atributu `alt` jsem testoval v Googlu a v Jyxu. Na stránku [Jasminka.cz](http://www.jasminka.cz) jsem umístil obrázkový odkaz `<img`

¹ <http://www.google.com/help/operators.html>

src=example alt="abcdefg">. Nechal jsem to tak do doby než oba vyhledávače tuto změnu zaregistrují. Google reagoval tak, že s alt tagem počítal (tyto výrazy se vyskytují pouze v odkazech na tuto stránku: abcdefg). Jyxo alt tag úplně ignorovalo. To dokazuje, že je lepší používat textové odkazy, protože některé vyhledávače alt tag v odkazu ignorují.

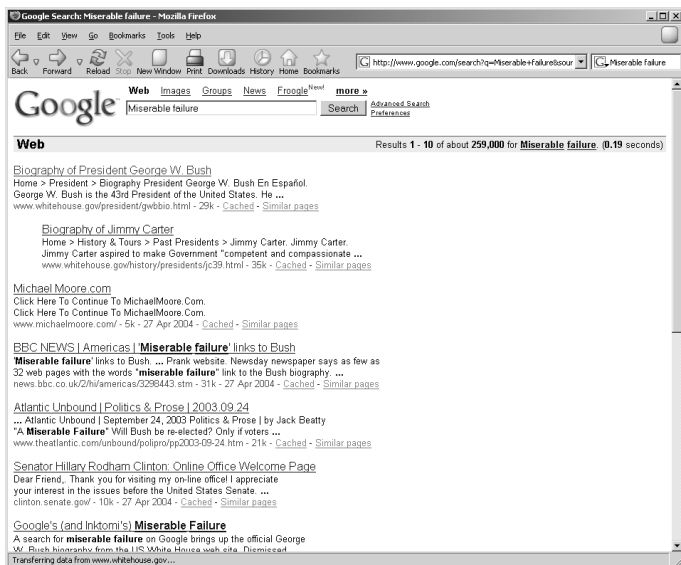
4.3.3. Google bomba

První zmapovanou Google bombu použil Adam Mathes v roce 2001, když nalinkoval stránku svého kamaráda jako "talentless hack". Pokud kdokoliv hledal v Googlu v roce 2001 hackera bez talentu, našel na prvním místě stránku kamaráda Adama Mathese. Celé to zapříčinil anchor text, o kterém se zmiňuji výše. Na postižené stránce se samozřejmě fráze "talentless hack" nevyskytovala.

O tom, že Anchor text je mocná technika se přesvědčil i americký prezident George Bush. Pár lidí (bloggeři) se totiž rozhodlo, že na jeho životopis na adrese bílého domu¹ budou odkazovat slovy "Miserable failure" - osudové selhání. Povedlo se to dokonale, kdokoli dal v prosinci 2003 do vyhledávače Google slovo Miserable failure, objevil se mu na prvním místě Bushův životopis. Což jistě není dobrá vizitka. Na to, aby se o americkém prezidentovi psalo například na BBC² jak uboze selhal, stačilo pouhých 32 odkazů. Ochrana proti tomu ani nemůže existovat. Není možné někomu zakázat odkazovat na cizí stránky takovým textem (i když v Americe je možné vysoudit vše).

¹ <http://www.whitehouse.gov/president/gwbbio.html>

² <http://news.bbc.co.uk/2/hi/americas/3298443.stm>

Obrázek 16: Co ukazuje na dotaz Miserable failure Google

I když se o Google bombě hodně mluví, její význam je minimální. Přece jen, kdo hledá někoho bez talentu, nebo něčí selhání. Je to spíše zábavný způsob ukázky, jak anchor text vlastně funguje.

4.4. Ostatní Off-page faktory

Do této kapitoly jsem zařadil ostatní faktory, které nejsou přímo na optimalizované stránce. Rozebírat budu klikavost na dané slovo ve vyhledávačích, dále pak způsob zakázání indexace souborů pomocí robots.txt a optimalizaci chybové stránky s kódem 404.

4.4.1. Klikavost na vyhledávané slovo

V angličtině se klikavosti říká Click Through Popularity. Některé vyhledávače (v současnosti například Yahoo!) měří kolikrát se kliká na daný odkaz ve výsledcích vyhledávání. Ve výsledcích nejsou přímo odkazy na danou stránku, ale script, který počítá počet přístupů. Poznat se také dá například pokud uživatel na stránku klikne, ale vrátí se zpět tlačítkem Back. Vyhledávač z toho může vyvodit závěr, že dané klíčové slovo neodpovídá tomu, co uživatel přesně hledal.

Snažit se klikat na vlastní výsledky ve vyhledávání nemá smysl, vyhledávače si jistě ukládají i IP adresu a rychle by takovou věc odhalily. Klikavost na určité slovo může být vyhledávačem použito jako velice účinná technika. Pokud by se server www.example.com podařilo optimalizovat na slovo sex, a tento server by byl o něčem jiném, uživatelé by po druhé co vyhledávají slovo sex na www.example.com již neklikli, protože ví, že tento server je úplně o něčem jiném. To se může započíst jako záporné body a server www.example.com klesá na výsledky slova sex.

Pokud se to vezme do detailu je možné se ve výsledcích posouvat nahoru i pomocí lepší klikavosti. Té se dá dosáhnout například lépe zvoleným titulkem, popiskem ale především lepším obsahem, ten zapříčiní, že uživatelé na stránku ve vyhledávači kliknou i za týden.

4.4.2. robots.txt

Zvláštností robots.txt je, že v něm nejsou napsány soubory, které roboti mohou stahovat, ale jsou zde vyjmenovány ty, které jsou zakázány stahovat. Každý robot se, před tím než navštíví server, podívá do robots.txt na to, co by neměl indexovat. Soubor robots.txt musí být umístěn v kořenovém adresáři a píše se malým písmem¹. Většina lidí se spíše snaží roboty nějak nalákat, než jim něco zakazovat. Pokud se tedy každý robot může po webu bez omezení pohybovat, zápis vypadá takhle:

```
User-agent: *  
Disallow:
```

V případě, že chceme některému robotu indexaci úplně zakázat (například Jyxobotu) a ostatním dovolit vše procházet, zápis vypadá následovně:

```
User-agent: *  
Disallow:
```

```
User-agent: Jyxobot  
Disallow: /
```

Zápis *Disallow: /smazat/*, znamená, že nechceme aby roboti stahovali adresář smazat (všechny soubory v něm, například /smazat/abcxxx.html). Zápis *Disallow: /deletei*, pak znamená zákaz indexace souborů začínajících na deletei (například deletei?a=xxx). Jména robotů lze nejjednodušeji získat z logu, stačí se podívat, kdo stahuje soubor robots.txt. Samozřejmě zakázat se dá více adresářů (souborů).

```
User-agent: *  
Disallow: /smazat/
```

¹ <http://www.jasminka.cz/robots.txt>

Disallow: /deletei

Ne všichni roboti se souborem robots.txt řídí a stahují i zakázané soubory a adresáře. Jedná se například o roboty spamérů, kteří se tak snaží jednoduše získat emailové adresy, anebo o chybu robota, který může robots.txt špatně interpretovat. Z tohoto důvodu by neměly být volně přístupné neveřejné informace na webu. Robots.txt lze použít pouze při vlastnictví domény (při možnosti zapisovat do kořenového adresáře). Druhá možnost je doplnit do hlavičky dokumentu "<META NAME="ROBOTS" CONTENT="NOINDEX, NOFOLLOW">", což znamená zákaz indexace stránky a sledování odkazů.

Tabulka 6: Nejznámější roboti procházející český web

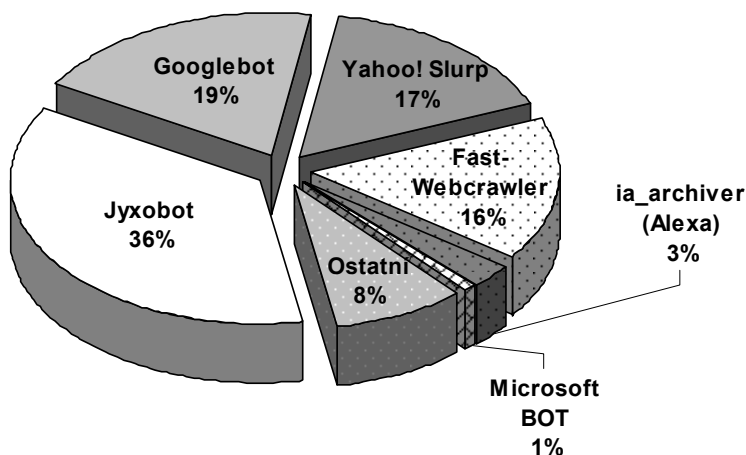
Vyhledávač	Robot
Google	Googlebot
Jyxo (Atlas [24])	Jyxobot
Seznam	Seznambot
Yahoo!	Yahoo! Slurp
Alexa	ia_archiver
AlltheWeb	FAST-WebCrawler
Altavista	Altavista Scooter
Microsoft	Microsoft BOT
Morfeo	Holmes

Google má pro svého robota speciální meta značku, pomocí které ho lze částečně ovládat:

```
<meta name="googlebot"  
content="snippet/nosnippet, archive/noarchive"  
>
```

Výchozí volba, tedy i v případě že tato značka není použita, je snippet a archive. Znamená to, že u stránky je povoleno archivování (při vyhledávání lze v Googlu kliknout na Archiv), a u výsledků vyhledávání se zobrazuje část textu před a za hledaným slovem. Pokud se použije nosnippet, je místo části textu použita meta značka description.

Graf 6: Podíl vyhledávacích robotů podle stáhnutých souborů na serveru Bazos.cz



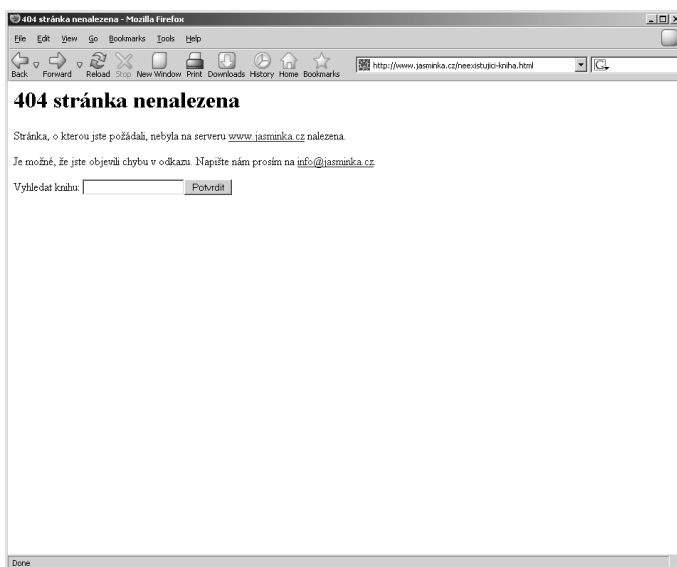
4.4.3. ErrorDocument 404

Pokud není nalezena na serveru požadovaná stránka, objeví se uživateli v prohlížeči dokument s chybovou hláškou 404. Ta by měla být využita pro zorientování uživatele. Měla by na ní být vypsána chyba 404 soubor

nenalezen, odkaz na hlavní stránku serveru, kontakt a popřípadě vyhledávací formulář[19]. Pro servery, které běží na systému Linux se 404 nastavuje pomocí souboru .htaccess, který musí být umístěn v kořenovém adresáři. Obsah .htaccess je pak následující:

ErrorDocument 404 /chyba.html

Obrázek 17: Obsah chybové stránky 404

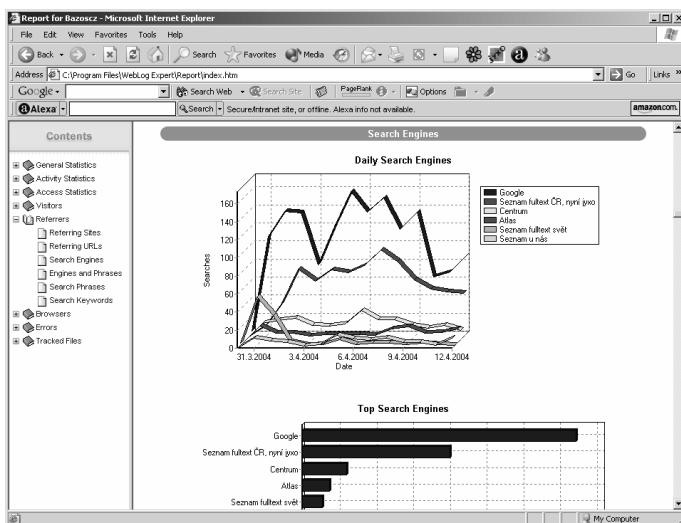


Stejně tak se mohou definovat i jiné chyby (například 403 nebo 500). Nefunkčnost tohoto nastavení může být způsobena zakázáním souboru .htaccess na daném hostingovém serveru. Pokud běží server na Windows IIS je pro nastavení 404 nejjednodušší doklíkat se na dané místo: Ovládací panely - Nástroje pro správu - Internetová informační služba, webový server (vlevo ve stromu IIS). Na něj kliknout pravým tlačítkem, vlastnosti a vlastní chyby.

4.5. Jak analyzovat výsledky optimalizace

Cokoliv co nejde měřit nelze ani řídit. Tato věta platí samozřejmě i pro SEO. Je důležité zjistit jaký účinek na návštěvnost (na počet zákazníků, na konverzní poměr) optimalizace webové stránky přináší. Pro měření návštěvnosti většinou stačí webové nástroje typu Toplist¹ nebo Navrcholu². Dokáží udělat grafy na kterých lze sledovat celkový vývoj návštěvnosti, odkud lidé na stránky přicházejí, kolik času na něm tráví a tak podobně. Podmínkou jejich použití je však umístění kódu na každou stránku webu. Pokud uživatelé přicházejí z vyhledávačů není to v drtivé většině na úvodní stránku.

Obrázek 18: Ukázka výstupu z programu Weblog Expert



Co však webové nástroje nedokáží měřit je počet zákazníků a konverzní poměr. To dokáže pouze důkladná analýza log souboru serveru. Z ní se dá zjistit jak se zákazník na web dostal (zda právě pomocí úspěšné optimalizace),

¹ <http://www.toplist.cz/>

² <http://www.nv.cz/>

zda se takto získaný návštěvník pravidelně vrací, jaký je konverzní poměr jednotlivých klíčových slov zadávaných ve vyhledávačích a podobné věci, které jsou nezbytné pro správné vyhodnocení SEO.

Na analýzu log souboru serveru existuje celá řada nástrojů. Mezi nejznámější patří Awstats¹, Analog², WeblogExpert³ nebo WebTrends⁴. V drtivé většině je potřeba tyto nástroje nakonfigurovat pro české vyhledávače. Osobně na analýzu log souboru používám program Weblog Expert (trial verze je k dispozici zdarma) a kombinuji to s každodenním nahlédnutím přímo do logu souboru. U každého zákazníka se snažím zjistit jak se na web dostal, popřípadě jaké klíčové slovo zadal do vyhledávače (v programu WeblogExpert se dá sledovat soubor potvrzující objednávku). Všechna tyto slova zadávám do vyhledávače a kontroluji na jaké pozici se na ně umísťují mnou optimalizované servery. Pokud je mnou optimalizovaný web ve výsledcích vyhledávání až za konkurencí, nebo až někde na druhé stránce, snažím se dané slovo posunout nahoru.

¹ <http://www.awstats.com/>

² <http://www.analog.cx/>

³ <http://www.weblogexpert.com/>

⁴ <http://www.netiq.com/webtrends/>

5. Zakázané metody v SEO

Některé techniky optimalizace jsou neetické, a z toho důvodu jsou vyhledávači penalizovány.

5.1. Co lze považovat za spam

Většinou jde o to jakým způsobem je daná technika použita. Ne všechny techniky jsou nutně spamem. Obecně se dá říci, že jde o techniky na oklamání vyhledávačů, z kterých neprofituje uživatel.

5.1.1. Skrytý text nebo odkazy

Text, který má stejnou barvu jako barva pozadí se nazývá skrytý text (pro uživatele je neviditelný). Myslet si, že takhle se dá zlepšit pozice na určité slovo, je mylné. Může to platit krátkodobě, ale může být celá řada lidí, kteří nahlásí stránku se skrytým textem pomocí spam reportů. Zbavit se poté penalizace není jednoduché a zabere to čas. Skrytý text v menu, který se objevuje po najetí ukazatele myši na menu není rozhodně spamem. Menu totiž v tomto případě uživatel vidět může.

Za skrytý odkaz se považuje ten, ze kterého není patrné že jde o odkaz. Má tedy stejnou barvu jako obyčejný text, je schován v tečkách na konci vět, v malých obrázcích atd. .

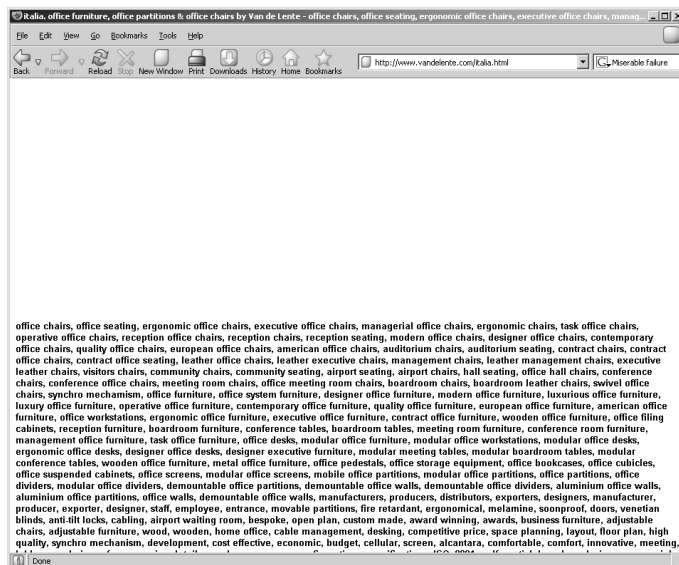
5.1.2. Malý text

Jedná se o text, který je často umisťován na konec stránky. Je natolik malý, že ho uživatel není schopen přečíst. Tato technika se začala rozšiřovat co začaly první pokusy s automatickým rozpoznáním skrytého textu.

5.1.3. Klamná a často opakovaná slova

Za klamná slova jsou považována ta, které nemají nic společného se skutečným obsahem stránky. Příkladem by mohlo být použití slova "sex" na stránce internetového obchodu prodávajícího kancelářské potřeby. Další neetickou možností je časté opakování slov, například jedno slovo použité dvacetkrát za sebou. To je sice také možné považovat za spam, ale navíc to ani nepomáhá dosáhnout lepších výsledků. Každé slovo má totiž na stránce určitou váhu (jak často se opakuje, v textu v nadpisech ...). Pokud to přesáhne například 20 procent, tak nastává přesně opačný efekt, než bylo opakováním zamýšleno.

Obrázek 19: Spam na konci stránky, skrytý nerelevantní text



5.1.4. Cloaking

Cloaking znamená vytváření speciálních stránek pro vyhledávací roboty. Například pokud se Google robotu podsouvá jiná stránka než běžnému uživateli, tak se tomu říká cloaking. Spameři to dělají z důvodu, aby Google dobře indexoval určité slovo. Podstrčená stránka je speciálně optimalizovaná pro určitý vyhledávač. Ten ji pak zaindexuje a možná zobrazí na prvním místě. Teoreticky by se pro každý vyhledávač mohla vytvořit speciální stránka optimalizovaná pro něj. Za cloaking se nepovažuje pokud se uživateli, který používá Operu dává trochu jiné HTML (CSS), než tomu co používá Explorer. V tomto případě jde totiž pouze o různé formátování stejného obsahu.

5.1.5. Odkazové farmy

Odkazové farmy vznikly krátce po tom co se objevil Google. Spameři totiž brzy přišli na to, jakou váhu má PageRank, a tak se ho snažili uměle zvyšovat. Tento způsob se prováděl pomocí prolinkování velkého množství webů dohromady, a tak vznikly první odkazové farmy (link farms). V současné době kdy existuje takzvané špatné okolí stránky s PRO je to velmi nebezpečná technika. Proto je vždy vhodné zvážit s kým se odkaz vyměňuje. Google totiž penalizoval účastníky většiny farem nulovým PageRankem.

5.1.6. Klamné přesměrování

Klamné přesměrování je odkaz, který přesměruje uživatele na jinou stránku než na tu na kterou se domníval, že tento vede. Pokud se to převede do extrému, na stránce je umístěn odkaz, ze kterého si uživatel myslí, že po kliku se dostane ke koupi zubního kartáčku, ale odkaz ho přesměruje na pornografickou stránku.

5.1.7. Doorway pages

Trochu jsem váhal, jak doorway pages popsat, ale když člověk neví, má se zeptat vyhledávače. Pro vyhledání definice jsem použil Jyxo, a to mi na prvním místě vyhodilo nabídku firmy co nabízí Doorway pages za 450 Kč.

Na základě klíčových slov bude vytvořeno 5 html stránek, které jsou optimalizovány pro co nejúčinnější vyhledávání vámi zvolených klíčových slov a frází. Se správnou formulací a výběrem slov a frází vám pomůžeme. Stačí pouze vyplnit tento formulář a registrace dalších 5 stránek do 1000 vyhledávačů, odkazujících na vaše url, proběhne za zcela výjimečnou cenu.

Doorway pages tedy znamená vytvoření více stránek, které nemají jako takové jiný účel, než směřovat na domovskou stránku, a tak zajistit její lepší pozici. Za doorway pages není považováno registrace domény druhého řádu pod více národních domén (například .cz, .com, .eu), což se dělá z důvodu, aby majitel doménového jména za rok nezjistil, že jméno někdo zaregistroval pod .eu jen z důvodu, že ho chce zpětně prodat. Je to čistě ochrana značky (názvu stránky).

5.1.8. Duplicitní web nebo stránky

Duplicitní stránka je ta, která má stejný obsah jako některá další stránka na webu. Není jediný důvod proč by vyhledávače měli zobrazovat duplicitní stránky ve výsledcích vyhledávání více než jednou, ale je fakt, že ne vždy se jim to daří. V případě nahlášení tohoto případu nedochází většinou k penalizaci, ale spíše k překontrolování algoritmu pro kontrolu zobrazení duplicitních stránek.

Duplicitní stránky mohou vznikat i nepozorností. Například nedůsledným používáním stejných URL. Mezi nejznámější příklady patří stránky na kterých

se listuje. Při odkazu na první stránku se používá `www.example.com/category` nebo `www.example.com/category/0`. Při programování se dosti často zapomíná, že to jsou dvě stránky s duplicitním obsahem.

5.1.9. Tapetování katalogů duplicitními doménami

Jedná se o pořízení více domén se stejným nebo jen trochu odlišným obsahem za účelem vícenásobné registrace do jedné sekce katalogů. Tím se následně zvyšuje PageRank. Katalogy by měly na tento spam reagovat vymazáním duplicitních stránek (ne vždy se jim to bohužel daří). Například i na Seznamu, jsou v některých kategoriích duplicitní servery s jinou URL. V tomto případě je nejrychlejší metodou, napsat na technickou podporu žádost o vyřazení duplicitního záznamu. Například Seznam většinou reaguje do dvou dnů vymazáním.

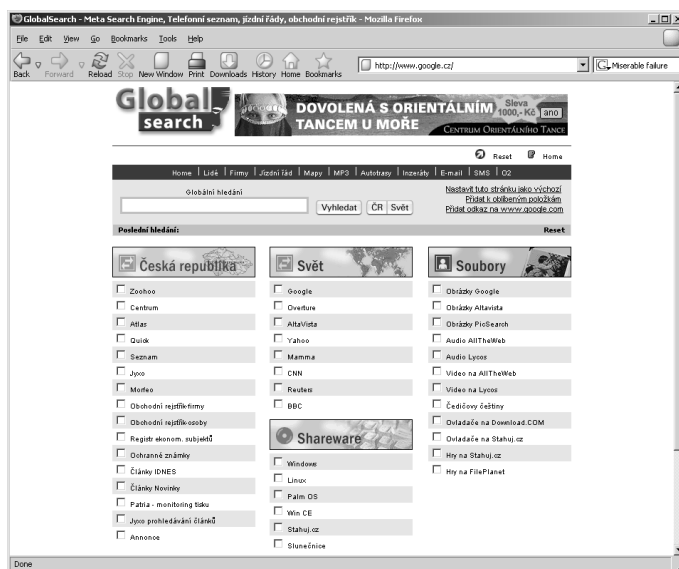
5.1.10. Jiné klamné taktiky vedoucí k zmanipulování vyhledávačů

K dobrým výsledkům vedou dvě cesty. Na jedné straně je možnost použití etického SEO přístupu, a na druhé použití spamu. Obě cesty něco stojí. Náklady na spam by se mohly zdát na první pohled menší, ale není tomu tak. Při použití spam taktiky vždy hrozí propad z předních pozic někam úplně dolů z důvodu, že to někdo odhalí (konkurence nebo přímo vyhledávač). Takže až se na to přijde a stránka bude penalizována, bude nutné pracně spam ze stránky odstraňovat nebo registrovat úplně novou stránku (vyhledávač nemusí penalizaci zrušit ani po odstranění spamu). Představa, že někdo použije spam taktiku například při programování webových stránek pro velké firmy nebo banky jako jsou Škoda auto, Baťa, KB nebo ČS a ty si pak budou muset zaregistrovat novou doménu, je absurdní. Ale i pro malou stránku může spam způsobit propad návštěvnosti a negativní popularitu.

5.1.11. Cybersquatting

Cybersquatting znamená registrace doménového jména za spekulativním účelem. Krásný příklad je registrace domény Google.cz, na které se nenachází vyhledávač Google. Je pouze na Googlu zda danou stránku penalizuje. Pokud tak učiní, může se BadRank rozšířit na všechny stránky v okolí Google.cz, viz kapitola č.4.2.11. BadRank a PageRank 0.

Obrázek 20: Ukázka Cybersquattingu



Rozhodně nedoporučuji vyměňovat odkazy se stránkami, které prokazatelně provádí Cybersquatting. Samozřejmě jsou příklady, kde to není tak prokazatelné, ale v případě Google.cz bych řekl že to nebezpečné být může. Stejně tak by to bylo v různých zkomoleninách tohoto výrazu jako googl, google a tak podobně. Například Yahoo! určitou penalizaci překlepů výrazů Yahoo provádí.

5.2. Spam report

Spam report je stránka pomocí níž se hlásí neetické techniky vyhledávačům. Kdyby neexistoval, pravděpodobně by bylo daleko více lidí, kteří by spam používali jako účinnou marketingovou metodu. Můj osobní názor je, používat spam report tak často, jak to jen jde. Čím dříve je daný člověk (firma) za spam penalizován, tím dříve se poučí, a přestane nekalé taktiky používat. Neberu nahlášení neetických taktik jako špatnou věc. Pokud konkurenční stránka použila spam, a díky němu se dostala před moji, tak neváhám ani minutu a spam report použiji. Ani nepíšu mail majiteli, kde bych ho informoval, že nečiní zrovna nejlépe. Problém je, že on by nebyl naštvaný na sebe, ale na mě.

Většina vyhledávačů má stránku, na které lze spam nahlásit (Seznam¹, Google², Yahoo!³). Větší seznam spam reportů se dá najít například na Search marketing⁴. Co se stane po nahlášení? Například Google penalizuje danou stránku nulovým PageRankem (PR0), což učiní daný web neviditelným. Jelikož Google tvrdí, že se snaží identifikovat spam hlavně automaticky, tak je jednoduchá cesta jak se toho zbavit – odstranit všechny nelegální praktiky použité na stránce. Pro ostatní vyhledávače někdy nezbyvá nic jiného, než "sesmolit" omluvný dopis s přiznáním jejich použití. Samozřejmě, že to bude trvat dost dlouho (i půl roku) a kdoví jestli si ten mail někdo vůbec přečte. Nejúčinnější metoda, jak se zbavit penalizace je samozřejmě vytvoření úplně nového webu, kde již spam nebude použit.

5.3. Na co si dávat pozor

Při optimalizaci pro vyhledávače je nutné si dát pozor na nabídky určitých firem. Jedná se o nabídky typu registrace do tisíců vyhledávačů, zlepšení

¹ <http://fulltext.seznam.cz/url.py/reportScreen>

² <http://www.google.com/contact/spamreport.html>

³ http://add.yahoo.com/fast/help/us/ysearch/cgi_reportsearchspam

⁴ <http://www.search-marketing.info/search-engines/report-spam.htm>

pozice ve vyhledávačích bez zásahu do obsahu webu, trvalé umístění na určitá slova atd. Dostí často takové nabídky znamenají použití spamu a nebo jsou to jen vyhozené peníze.

5.3.1. Zaregistrování stránek do tisíců vyhledávačů

Do vyhledávačů není nutnost se registrovat. Vyhledávače si totiž web najdou sami podle odkazů z jiných stránek. Samotná registrace neublíží, ale ani nepomůže. Některé firmy nabízí také registraci do tisíců katalogů. Platí zásada, že je vždy lepší se do katalogu registrovat ručně. Je nutné zvolit nejlepší sekci (tu kde jsou podobně zaměřené stránky) a to často automatické programy nezvládnou. Vhodných katalogů také nejsou tisíce, jejich počet se pohybuje maximálně v desítkách.

5.3.2. Trvalé pozice ve vyhledávačích

Každá nabídka, která zaručuje trvalou pozici ve vyhledávači je nepravdivá. Jediná možnost jak být trvale na první pozici ve vyhledávači je placená inzerce. SEO je sice dlouhodobá záležitost, ale nikdy ne trvalá. Vždy se může najít konkurent, který bude mít lepší stránky, změní se trochu algoritmus atd.

5.3.3. Garantované umístění

V SEO se nedá nic garantovat na 100 procent, tedy kromě placené inzerce. Takže na nabídky typu garance nebo peníze zpět je vždy nutné se dívat velmi obezřetně.

5.3.4. Analýzy pozice ve vyhledávačích

Nabídka typu, Váš web se nalézá až na 25. pozici v našem katalogu. Jelikož je to pozice až na třetí stránce, nejste dostatečně viditelní. Pro lepší umístění prosím zaplat'te tuto částku xxx . Tato nabídka má nedostatky v tom, že

neuvádí návštěvnost dané sekce, je nevyžádaná a hlavně snaží se mylně přesvědčit majitele o její výhodnosti. Analýzu pozice si každý může udělat sám, stačí se do katalogu podívat nebo zadat do vyhledávače klíčové slovo.

5.3.5. Lepší pozice bez zásahu do obsahu

SEO je o obsahu stránky, o hledání vhodných klíčových slov a o jejich následné optimalizaci. Pokud stránka nebyla v minulosti již někdy kvalitně zoptimalizována jsou zásahy do obsahu v drtivé většině případů nezbytností. Většina takových nabídek znamená, že firma vytvoří doorway pages směřující na optimalizovaný web. Po této “optimalizaci” může být stránka penalizována.

5.3.6. Reference

Reference v případě, že optimalizaci neprovádíme vlastními silami, jsou jistě důležité, ale že firma optimalizovala stránky pro některou z top deseti největších českých firem ještě nutně neznamená, že používá pouze etické SEO. Na místě je spíše otázka jestli firma byla někdy za optimalizaci penalizována a také jakých výsledků dosáhla. Klienti ale dost často nesouhlasí ze zveřejněním.

5.3.7. Výměna zakázaného odkazu

Při výměně odkazu se musí dávat pozor zda protistrana nezakáže na stránce kde je odkaz umístěn následovat odkazy. A to buď zakázáním takové stránky v robots.txt nebo zakázáním pomocí meta znakčky robots, která má přiřazený NOFOLLOW. Přesměrované odkazy se vůbec nevyplatí vyměňovat, u nich není jistota, že nejsou zakázány indexovat. Také se musí dávat pozor zda protistrana odkaz nevymaže.

6. Praktická ukázka postupu SEO

V této kapitole popíšu jak se má při optimalizaci webu postupovat. Jako první ze všeho je potřeba zkontrolovat zda optimalizovaný server používá některé překážky optimalizace (viz. kapitola č.3. Předpoklady pro optimalizaci). Pokud ano, je vždy lepší tyto překážky odstranit, u nového webu je samozřejmě nepoužívat. Kontrolu počtu zaindexovaných stránek ve vyhledávacích lze provést zadáním dotazu "*site:example.com example*" do vyhledávačů. Tento dotaz najde všechny stránky webu *example.com* na kterém se vyskytuje slovo *example*.

Pokud web nemá žádnou překážku optimalizace a je tedy dobře zaindexován je druhým bodem SEO u již existujícího webu časový snímek návštěvnosti, neboli kontrola v minulosti vyhledávaných slov. Je nutné ji provést, aby při optimalizaci nebyla ztracena dříve vyhledávaná slova.

Tuto analýzu je možné provádět pomocí webových nástrojů na měření návštěvnosti (Toplist¹, Navrcholu²) nebo pomocí analýzy log souboru serveru. Webové nástroje sice nejsou špatné, ale chybí jim určité klíčové vlastnosti (mezi hlavní bych zařadil nastavitelnost podle specifických požadavků). Při analýze používaných slov osobně vždy používám log soubor, který zpracovávám programem Weblog Expert³ (na jeden měsíc je na webu k dispozici zdarma na odzkoušení). Je nutné ho však upravit na české podmínky, to znamená změnit konfigurační soubory tak, aby program našel české vyhledávače (náznak konfigurace je v příloze této knihy).

Další část je výběr klíčových slov. To je pravděpodobně nejdůležitější činnost v SEO. Co na stránce není, nemohou vyhledávače ani nalézt (mimo anchor

¹ <http://www.toplist.cz/>

² <http://www.nv.cz/>

³ <http://www.weblogexpert.com/>

text). Vybraná klíčová slova se umisťují do html značek (do Title, description, nadpisů ...). Celou optimalizaci provází snaha o získání co nejvíce kvalitních odkazů (s vysokým rankingem a stejným tématem) na web.

Po optimalizaci je nutné vyčkat na novou indexaci webových stránek. Ta je obvykle prováděna podle rankingu stránky (čím větší PageRank, Jyxorank tím rychleji je stránka zaindexována). Kompletní reindexace může u některých vyhledávačů trvat i měsíce. Hlavní stránky webu by měly být ale většinou reindexovány do týdne nebo dvou. Po kompletní reindexaci se provádí druhý časový snímek návštěvnosti. Vyhodnocují se dopady provedené optimalizace na návštěvnost webu. Snaha o získání kvalitních zpětných odkazů je samozřejmě nutná neustále.

Pro praktickou ukázkou jsem zvolil dva weby, internetové knihkupectví mých rodičů [Jasminka.cz](http://www.jasminka.cz/)¹ a můj vlastní projekt internetový bazar [Bazos.cz](http://www.bazos.cz/)². Jako první popíšu [Bazos.cz](http://www.bazos.cz/).

6.1. Optimalizace serveru [Bazos.cz](http://www.bazos.cz/)

[Bazos.cz](http://www.bazos.cz/) měl před provedením SEO jedinou překážku optimalizace, a tou byly dynamické URL. Před vlastní optimalizací jsem zkontroloval kolik stránek má například Google ze serveru zaindexováno. Do vyhledávacího políčka Google jsem v květnu 2003 zadal `site:bazos.cz bazos`. Po zjištění, že Google indexuje pouhých 200 stránek jsem moc nadšený nebyl. Jak se převádí dynamické stránky na statistické je uvedeno v kapitole č.3.7 Problém s dynamickými stránkami. Místo URL adresy www.bazos.cz/index.php?kat=3&typ=1&cr=12 se nyní v URL zobrazuje jednoduše zapamatovatelná adresa ve tvaru www.bazos.cz/prodam/pc/. Při kontrole počtu zaindexovaných stránek v únoru 2004 měl Google zaindexováno již přes deset tisíc stránek. Celý problém byl v

¹ <http://www.jasminka.cz/>

² <http://www.bazos.cz/>

tom, že Google indexoval dynamické stránky pouze se dvěma proměnnými (v současnosti 3). Na Bazos.cz jsem původně používal tři proměnné, a tak Google většinu webu vůbec neindexoval. Úspěšnost převedení dynamických stránek na statické by se dala také měřit pomocí počtu stáhnutých stránek roboty z vyhledávačů a samozřejmě kontrolou jaké soubory roboti stahují.

Tabulka 7: Počet stáhnutých souborů vyhledávacími roboty ze serveru Bazos.cz

	Květen 2003 před optimalizací	Leden 2004 po optimalizaci
Celkový počet za měsíc	4875	16171
Denní průměr	157	522

Po optimalizaci URL jsem zkoumal jaká slova lidé ve vyhledávacích používají, aby se na Bazos.cz dostali. Jelikož mám log soubory archivovány, provedl jsem analýzu zpětně od ledna 2003 (datum registrace domény). Analýzu logu jsem provedl proto, abych při novém výběru klíčových slov nepřišel o nějaké dříve hodně vyhledávané. Zjištěné výsledky nebyly nijak přesvědčivé, viz. Tabulka č. 8.

Tabulka 8: Počet hledaných slov ve vyhledávacích květen 2003

1 digitální fotoaparáty	42
2 pc bazar	31
3 notebook bazar	23
4 notebooky	21
5 mobilní telefony	18
6 bazar notebook	14
7 digitalni fotoaparaty	13
8 bazar pc	11
9 základní desky	10
10 grafické karty	10

Na prvním místě byla nechtěná optimalizace na frázi digitální fotoaparáty, kde byl Bazos na prvním místě například v Googlu a Jyxu. Při letném pohledu na některá ostatní klíčová slova jsem neočekával, že tito návštěvníci hledali inzerci. Začal jsem tedy přemýšlet, jak někdo zadá dotaz do vyhledávače, pokud chce prodat počítačovou součástku. Na prvním místě mě napadla slova bazar a inzerce. Ta jsem odzkoušel pomocí nástrojů Adwords a Wordtracker viz. kapitola č. 4.1.3 Nástroje na vyhledávaná slova. Největší přínos mělo mít slovo bazar v nějaké kombinaci s počítačovou součástkou. Mezi nejjednodušší optimalizovatelné patřil pc bazar, mezi nejtěžší pak samostatně stojící slovo bazar.

Jelikož Bazos.cz byla v té době malá stránka a navíc věnující se pouze oblasti PC, upustil jsem od optimalizace samostatně stojícího slova bazar. Hlavní stránku jsem raději optimalizoval na frázi pc bazar. Jednotlivé kategorie pak vždy na slovo bazar a inzerce spojené s názvem dané kategorie. Takže například kategorie Notebooky je optimalizována na notebook bazar, kategorie Monitory na monitor bazar. Optimalizace obsahu kategorií byla provedena pomocí tagů title a description, což je úplný základ. Výsledkem bylo, že po měsíci jsem byl v hlavních vyhledávacích nejhůře na prvních třech místech na fráze název kategorie a bazar.

Do popředí se dostala klíčová slova, která lze jednoznačně spojovat se zaměřením serveru Bazos, tedy s řádkovou inzercí (viz tabulka č. 9). Samozřejmě výběr klíčových slov je neustálý proces, takže se je dále snažím ladit. Vždy dávám ale pozor, abych nepřišel o nějakou vyhledávanou frázi, která je spojená s inzercí (má dobrý konverzní poměr). Konverzní poměr v případě serveru Bazos chápu jako počet přidanych inzerátů lidmi co přišli z vyhledávačů / celkový počet návštěvníků z vyhledávačů.

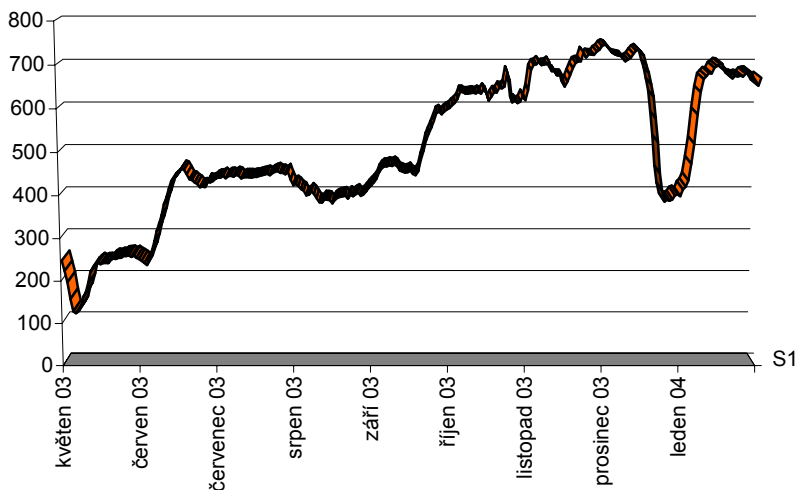
Tabulka 9: Vyhledávaná slova leden 2004

1 pc bazar	369
2 bazar	69
3 bazar pc	54
4 notebook bazar	49
5 inzerce	49
6 bazar notebook	39
7 grafické karty	33
8 pda bazar	27
9 bazos	26
10 pc bazary	25

Poslední uvedené jsou off page faktory, které však patří mezi ty nejdůležitější. Co to jsou off page faktory je uvedeno v kapitole č. 4.2 Off page faktory. Základem je získat co nejvíce odkazů z cizích stránek. To je více méně pracná záležitost. Zaregistrovat web do tří hlavních českých katalogů Seznam, Centrum a Atlas považuji za samozřejmost, a ne za SEO. Osobně doporučuji zaregistrovat web i do menších katalogů a samozřejmě do DMOZ, což je mezinárodní katalogový projekt. Neustálé hledání nových možností odkazů je nutností, i když je daná stránka na vrcholu ve vyhledávacích výsledcích. Za dobu této studie tedy od května 2003 do ledna 2004 vzrostl Google PageRank hlavní stránky serveru Bazos.cz ze 3 na 5.

Tabulka 10: Počet návštěvníků z vyhledávačů, server Bazos.cz

	Květen 2003 před optimalizací	Leden 2004 po optimalizaci
Celkový počet za měsíc	1580	7168
Denní průměr	51	231

Graf 7: Vývoj návštěvnosti Bazos.cz, 7 denní klouzavý průměr

Výsledky optimalizace byly více než uspokojivé. První fáze proběhla v květnu 2003. Vyhledávače na ni reagovaly o měsíc později v červnu. Druhá fáze pak koncem srpna, reakce byla opět za měsíc na konci září. Výkyv v prosinci je způsoben Vánoce. Díky SEO se ze serveru Bazos stal za dobu studie druhý nejnavštěvovanější server, zaměřený speciálně na počítačovou inzerci. Ačkoliv je to výborný výsledek musím přiznat, že je to způsobeno hlavně tím, že v době studie neměly konkurenční servery optimalizaci provedenu a pokud vím, ani se k ní ještě nechystaly.

Obrázek 21: Inzertní server Bazos.cz, 2004


The screenshot shows the Bazos.cz website in a Mozilla Firefox browser window. The page title is "Bazos.cz - Inzerce, PC bazar - Mozilla Firefox". The browser address bar shows "http://www.bazos.cz". The page content includes a navigation menu on the left, a main listing area with a table of PC items, and a right-hand sidebar with additional category links.

Navigation Menu (Left):

- Uživatelská
- CD - DVD ROM
- Digitální fotoaparáty
- Faxmodemy
- FDD, ZIPy
- Grafické karty
- Hard disky
- Herní zařízení
- Chladíče
- Klávesnice, myši
- Kopírovací stroje
- Mobilní telefony
- Monitory, LCD
- MP3 přehrávače
- Notebooky
- Paměti
- PC, Počítače
- PDA, Palm
- Síťové prvky
- Skáňovací
- SKříně, zdroje
- Software
- Tiskárny
- Wireless
- Základní desky
- Zdroje
- Zvukové karty
- Ostatní
- Uživatelská
- Knihkupectví Jasmínka

Main Listing Area:

PC bazar - smloumá inzerce zdarma - prodej i koupě
 Prodejte vaše staré PC, notebook, mobil nebo PDA, to vše zvládne náš Bazar.
 Akce s 2FC - sdílení inzerátů.



Přidat inzerát Prohledat inzeráty:

Typ	Typové	Cena	Datum a čas
P	sálec do HP L3 1100	700,-	29.4. - 19:05 X
P	Ibm 768l	nabídnete.-	29.4. - 19:31 X
N	Prosdím Compaq IPAQ H3970	9500,-	29.4. - 19:15 X
N	Prosdím digitál s min. 2 megapixely	1,-	29.4. - 19:15 X
N	Prosdím NOKIA 6110	1,-	29.4. - 19:14 X
P	Koupím monitor	do 500 Kč.-	29.4. - 18:55 X
N	P111 500 MHz za nízkou cenu. Posuďte	3300 Kč.-	29.4. - 18:08 X
N	Discman Panasonic SL-50510	1200,-	29.4. - 17:17 X
N	OCZ 256MB 400MHz CL2	dohoda.-	29.4. - 17:02 X
N	cisco 805 - 1600	dohoda.-	29.4. - 16:47 X
P	koupím klávesnici	do 500,-	29.4. - 16:38 X
N	Seagate 40GB	1400,-	29.4. - 16:11 X
N	Externí faxmodem WELL FM 56 Wt za	800,-	29.4. - 15:57 X
N	IBM ThinkPad 388D	9500,-	29.4. - 15:38 X
N	HDD Western Digital 40GB	1400,-	29.4. - 15:12 X

Počet inzerátů: 5269

Stránka: 1, 2, 3, 4, 5, 6, 7, 8

Right-hand Sidebar:

- Monitory
- CD - DVD ROM
- Digitální fotoaparáty
- Faxmodemy
- FDD, ZIPy
- Grafické karty
- Hard disky
- Herní zařízení
- Chladíče
- Klávesnice, myši
- Kopírovací stroje
- Mobilní telefony
- Monitory, LCD
- MP3 přehrávače
- Notebooky
- Paměti
- PC, Počítače
- PDA, Palm
- Procesory
- Síťové prvky
- Skáňovací
- SKříně, zdroje
- Software
- Tiskárny
- Wireless
- Základní desky
- Zdroje
- Zvukové karty
- Ostatní

6.2. Optimalizace serveru Jasminka.cz

Jako první ze všeho jsem zkontroloval, kolik stránek má Google ze serveru Jasminka.cz zaindexováno. Do vyhledávacího políčka Googlu jsem v září 2003 zadal site:jasminka.cz jasminka. Problém byl stejný jako u předchozího příkladu Bazos.cz, Google indexoval pouhých 30 stránek. Opět tři proměnné a pravděpodobně neochota Googlu zaindexovat proměnnou isbn (viz. kapitola č. 3.7 Problém s dynamickými stránkami). Po převodu na statické stránky ve tvaru <http://www.jasminka.cz/knihy/detska/> již Google a ostatní vyhledávače vše bez problému zaindexovaly. V lednu 2004 měl Google zaindexováno přes 500 stránek (přibližně počet prodávaných knih). Jasminka.cz neměla před začátkem studie jinou překážku optimalizace, která by zabraňovala úplné indexaci. Pomocným ukazatelem počtu zaindexovaných stránek může být počet stáhnutých souborů vyhledávacími roboty (zjištěný z log souboru). viz. Tabulka č. 11.

Tabulka 11: Počet stáhnutých souborů vyhledávacími roboty

	Září 2003 před optimalizací	Leden 2004 po optimalizaci
Celkový počet za měsíc	3661	9741
Denní průměr	122	314

Pomocí programku WebLog Expert jsem zanalyzoval log za měsíc září, abych zjistil jaká slova návštěvníci používají pro vyhledání tohoto knihkupectví. Výsledek nebyl nic moc viz. Tabulka č 12.

Tabulka 12: Vyhledávaná slova září 2003, Jasminka.cz

1	knihkupectví	8
2	dagmar kludská	6
3	chat	6
4	valja stýblová	5
5	levné knihy	5
6	james herriot	4
7	pavel frýbort	4
8	kludská	4
9	wilbur smith	4
10	devatero pohádek	3

Tyto slova jsem zkontroloval pomocí Adwords a pomocí Etargetu abych zjistil jejich potencionál. Jako nejlepší se jevílo knihkupectví a etargetem navrhnuté internetové knihkupectví. Zkusil jsem ve vyhledávačích na jakých pozicích se Jasmínka pohybuje. Na žádné z těchto klíčových slov nebyla v první desítce, což jistě byl problém. Takže jsem optimalizoval pomocí výše zmíněných technik hlavní stránku na výraz knihkupectví a každou jednotlivou knihu na název knihy a autora. Samozřejmě vyhledávačům to někdy trvá, než znovu zaindexují celý web, takže proto uvádím výsledky až z ledna 2004, kdy již měli dostatek času na novou indexaci.

Tabulka 13: Vyhledávaná slova leden 2004, Jasminka.cz

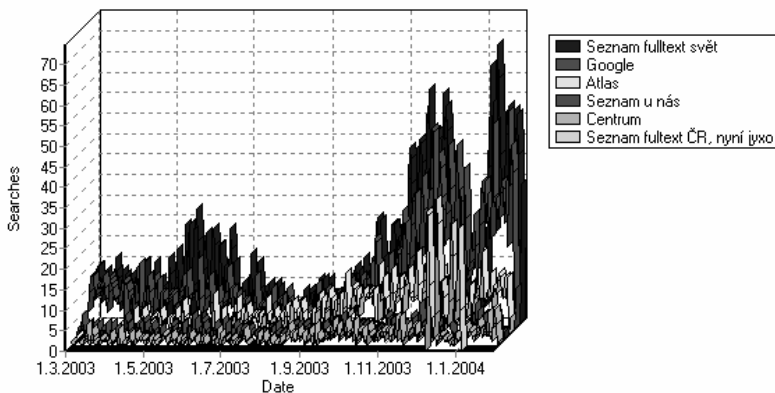
1	knihkupectví	122
2	michal viewegh	66
3	josef lada	52
4	vít chaloupka	48
5	harry potter a fénixův řád	39
6	internetové knihkupectví	38
7	harry potter	35
8	kludská	35
9	knihkupectvi	31
10	vít chaloupka	31

Všechna zmíněná slova se mi bez větších problémů podařilo dostat do první desítky do většiny vyhledávačů, a to jen pomocí základních technik (nadpisy, description a anchor text). Žádné další nebyly potřeba, protože konkurenční internetová knihkupectví neměla v této době zoptimalizované stránky. Nedostatky na serveru Jasminka jsou v současné době hlavně v on page faktorech. Je potřeba celý web předělat do kaskádových stylů, doplnit nadpisy H1, atd.

Tabulka 14: Počet návštěvníků z vyhledávačů, server Jasminka.cz

	Září 2003 před optimalizací	Leden 2004 po optimalizaci
Celkový počet za měsíc	896	2954
Denní průměr	30	95

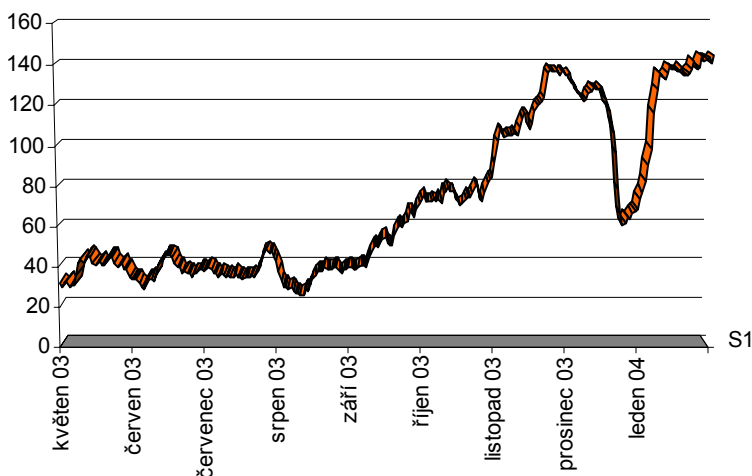
Graf 8: Vývoj počtu přístupů z vyhledávačů na server Jasminka.cz



Jak se vyvíjela návštěvnost z vyhledávačů lze krásně vidět z grafu č. 8. Nejvíce přístupů připadá na fulltextové vyhledávání Seznamu a na Google. Záměna technologie na Seznamu (Jyxo za Google) způsobila propad přístupů ze Seznam fulltext svět a vzestup Seznam fulltext ČR.

Nyní se dostáváme ke konverznímu poměru jednotlivých klíčových slov. Počítal by se jako počet objednávek z klíčového slova/celkový počet návštěvníků z vyhledávačů pomocí daného klíčového slova. Nejlepší konverzní poměr u knihkupectví Jasminka mají názvy jednotlivých knih. U některých se pohybuje až ve výši 100 procent. Není to však až tak překvapivé. Uživatel, který není schopen najít knihu v klasickém knihkupectví (v případě, že ji nemají na skladě), zadává název knihy do Googlu a poté co najde internetové knihkupectví Jasminka knihu koupí. Musím upozornit, že v době studie neměla ostatní internetová knihkupectví optimalizovány své stránky, takže například na Vltavě se nedala kniha najít ani v případě zadání do Googlu ve tvaru “autor název knihy Vltava”. Vltava užívala do ledna 2004 stejný nadpis pro všechny své stránky, stejný popis atd. Vltavu jsem vzal jako příklad, ostatní internetová knihkupectví na tom byla obdobně.

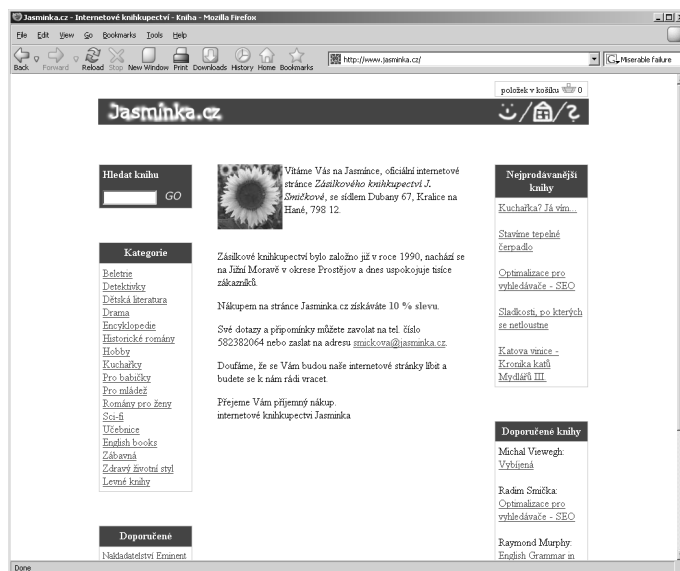
Graf 9: Vývoj návštěvnosti serveru Jasminka.cz 7 denní klouzavý průměr



Pro provedení důkazu úspěšnosti provedené optimalizace uvedu průměrný denní počet objednávek za den. Ty se zvýšily v lednu 2004 oproti září 2003 přibližně pětkrát a tento trend dále pokračuje, takže se nejedná o jednorázový

výkyv. V únoru zoptimalizovala své stránky konkurenční knihkupectví a bude nutné v optimalizaci dále pokračovat. Je to částečně problém celé optimalizace. SEO je neustálý proces, při kterém se musí sledovat pozice jednotlivých slov a v případě, že se nevyvíjí s očekáváním na to reagovat. Neviděl bych to však zase tak negativně. Uživatelé si zvyknou, že kniha se dá najít zadáním názvu do internetového vyhledávače a na prvních třech místech jim to vypíše tři různá internetová knihkupectví. Pro uživatele je to v tomto případě SEO určitě pozitivní věc. Pro knihkupectví to však může ve svém důsledku znamenat větší tlak na cenu.

Obrázek 22: Knihkupectví Jasminka



7. Závěr

Internetové vyhledávače jednoznačně hrají stále důležitější roli při získávání nových návštěvníků na internetové stránky. Přes devadesát procent lidí hledá právě pomocí vyhledávání a pouhých devět procent pomocí procházení katalogových sekcí. To vede každého majitele webových stránek k úvahám jak zvýšit pozici stránky ve vyhledávačích. Možnosti jak je možné toho dosáhnout je obsahem této knihy.

Ze všeho nejdůležitější je zjistit jestli je optimalizovaná stránka kompletně zaindexovaná. To je potřeba zjistit z důvodu, že uživatelé většinou nehledají informace, které jsou na úvodní stránce webu, ale většinou ty, ke kterým je potřeba se proklikat. Pokud webová stránka není dobře ve vyhledávačích zaindexovaná je nutné prověřit, zda nepoužívá některé překážky, přes které nejsou vyhledávače schopny přejít. Obecně by se dalo shrnout, že pokud dokáže optimalizovaný web dobře přečíst textový prohlížeč Lynx, dokáží ho přečíst a zaindexovat i vyhledávače. To znamená doporučení nepoužívat technologie typu Flash, Javascript, rámce atd.

Po bezchybném zaindexování všech stránek se pomocí SEO zlepšuje pozice každé stránky ve vyhledávačích. Je nutné se soustředit na všechny stránky webu, nikoliv jen na úvodní. Každá stránka na konkrétním webu musí mít unikátní obsah, ten je pak optimalizován pomocí on-page a off-page faktorů. Mezi základní on-page faktory patří titulek stránky a nadpisy. Mezi off-page faktory pak PageRank, neboli to, kolik webových stránek odkazuje na uvedenou stránku a anchor text.

Důkaz účinnosti SEO jsem provedl pomocí vlastního internetového bazaru Bazos.cz a internetového knihkupectví Jasminka. U serveru Bazos.cz došlo v důsledku optimalizace pro vyhledávače ke zvýšení návštěvnosti z 250 lidí za den na 750. Z toho 230 lidí denně je právě přes vyhledávače. Ze serveru Bazos

se tak po jednom roce od jeho vzniku stal jeden z hlavních serverů, který se v Česku zabývá počítačovou inzercí. Optimalizací serveru Jasminka bylo dosaženo nejenom zvýšení návštěvnosti webových stránek, ale i zvýšení počtu nových objednávek na zboží, které je zde nabízeno. Jejich počet se za dobu optimalizace zpětinásobil.

Z mojí studie jednoznačně vyplývá, že SEO je účinný marketingový nástroj. Jeho účinnost je samozřejmě vyšší pokud se podaří optimalizaci provést před konkurenčními servery. Je to určitá mezera na trhu, která ještě nebyla úplně zaplněna. Konkurenční knihkupectví v případě serveru Jasminka.cz a konkurenční internetové bazary v případě serveru Bazos.cz neměly a většina ani dosud nemá optimalizované své webové stránky. Projevovalo se to například tím, že pokud jsem v lednu 2004 zadal do Googlu název knihy a jejího autora, tak jsem dostal server Jasminka.cz jako jediný ve výsledcích vyhledávání. Současné výsledky vyhledávání však svědčí o tom, že se optimalizaci věnuje daleko více pozornosti než dříve. Některá konkurenční knihkupectví si již například změnila alespoň titulek u jednotlivých nabízených knih. Před tím používala stejný titulek na celém webu. Tím však také dochází ke zvýšení konkurence na všechna klíčová slova.

Bez provedení alespoň základní optimalizace se internetové obchody v budoucnu vůbec neobejdou. Pokud uživatel zná název produktu, tak je pro něj pohodlnější ho zadat do vyhledávače a pak jen zjišťovat, kde nabízí nejlepší podmínky. Pamatovat si jména deseti obchodů a až na jejich stránkách hledat příslušný produkt zabere hodně času, a ten jak známo stojí peníze.

Samozřejmě v současnosti již existují oblasti (cestovní ruch) ve kterých si není možné vystačit pouze s touto knihou. V takovém případě je nutné sledovat aktuální dění (některé důležité webové stránky jsou uvedeny ve zdrojích). Pro přehled o tom jak SEO funguje však tato kniha bohatě stačí.

Seznam použité literatury

Knihy:

- [1] Thurow S.: *Search Engine Visibility* - New York, New Riders, 2002, ISBN-07-3571-256-5
- [2] Page L., Brin S., Motwani R., Winograd T.: *The PageRank Citation Ranking: Bringing Order to the Web* - Stanford University, 1998
Dostupné na: <http://dbpubs.stanford.edu:8090/pub/1999-66>
- [3] Page L., Brin S.: *The Anatomy of a Large-Scale Hypertextual Web Search Engine* - Stanford University, 1998
Dostupné na: <http://dbpubs.stanford.edu:8090/pub/1998-8>
- [4] Page L., Brin S.: *Dynamic Data Mining: Exploring Large Rule Spaces by Sampling* - Stanford University, 1999
Dostupné na: <http://dbpubs.stanford.edu:8090/pub/1999-68>
- [5] Haveliwala H. T.: *Topic-Sensitive PageRank* - Stanford University, 2002
Dostupné na: <http://www2002.org/CDROM/refereed/127/>
- [6] Haveliwala H. T.: *Efficient computation of PageRank* - Stanford University, 1999
Dostupné na: <http://dbpubs.stanford.edu:8090/pub/1999-31>

Elektronické zdroje:

- [7] Dokumentace Apache: *Mod_rewrite*
Dostupné na: http://httpd.apache.org/docs/mod/mod_rewrite.html
- [8] Engelschall S. R.: *A Users Guide to URL Rewriting*
Dostupné na: <http://www.engelschall.com/pw/apache/rewriteguide/>

- [9] Sova v síti: Oborský L.: *Mod_rewrite pro lepší odkazy*
Dostupné na: http://www.sovavsiti.cz/2003/mod_rewrite.html
- [10] Root: Krause M.: *Což si tak trochu zapřepisovat?*
Dostupné na: <http://www.root.cz/clanek/542>
- [11] A list apart: Quack T.: *How to Succeed With URLs*
Dostupné na: <http://www.alistapart.com/articles/succeed/>
- [12] Interval: Málek V.: *SEO kontra ASP - za URL krásnější*
Dostupné na: <http://interval.cz/clanek.asp?article=2604>
- [13] Search Marketing: Wall A.: *History of Search Engines*
Dostupné na: <http://www.search-marketing.info/search-engine-history/>
- [14] Outer court: Lenssen P.: *Search Engine History*
Dostupné na: <http://blog.outer-court.com/history/>
- [15] SEOchat: Hrulbert W.: *Google PageRank Primer*
Dostupné na: <http://www.seochat.com/c/a/Google/Google-PageRank-Primer/>
- [16] eFactory: *The PageRank Algorithm*
Dostupné na: <http://pr.efactory.de/e-pagerank-algorithm.shtml>
- [17] SEOchat: Pfeiffer B.: *How to Optimize for Inktomi*
Dostupné na: <http://www.seochat.com/c/a/Search-Engine-Optimization/How-to-Optimize-for-Inktomi-in-2004/>
- [18] conBLOG: Bureš J.: *Co má být na chybové stránce*
Dostupné na: <http://blog.converter.cz/index.php?m=200403#450>

- [19] Jak psát web: Janovský D.: *ErrorDocument 404*
<http://www.jakpsatweb.cz/weblog/archiv/2004-03.html#150001>
- [20] BBC: *'Miserable failure' links to Bush*
Dostupné na: <http://news.bbc.co.uk/2/hi/americas/3298443.stm>
- [21] SEOGUY: *Channeling pagerank*
Dostupné na: <http://www.seo-guy.com/pagerank-articles/channeling-pagerank.html>
- [22] SearchEngineWatch: Sullivan D.: *comScore Media Metrix Search Engine Ratings*
Dostupné na: <http://searchenginewatch.com/reports/article.php/2156431>
- [23] Seochat: Wells T.: *What's Microsoft Gotten into Now? Search Me!*
Dostupné na: <http://www.seochat.com/c/a/News/Whats-Microsoft-Gotten-into-Now-Search-Me/1/>
- [24] chomat.net: Chomát J.: *Jyrobot nemusí být z Jyxa*
Dostupné na: <http://www.chomat.net/weboviny/archiv/2004-05.html#1083919019>

Webové stránky:

- [25] The Google Search Engine: <http://www.google.com/>
- [26] Sova v síti: <http://www.sovavsiti.cz/weblog/>
- [27] Yuhůův weblog o webu <http://www.jakpsatweb.cz/weblog/>
- [28] Česká konference o SEO: <http://seo.nawebu.cz/>
- [29] SEOchat: <http://www.seochat.com/>
- [30] Webmaster World: <http://www.webmasterworld.com/>
- [31] SEOGUY: <http://www.seo-guy.com/>
- [32] Robots.txt: <http://www.robotstxt.org/>

- [33] Search Engine Roundtable Weblog: <http://www.seroundtable.com/>
- [34] conBLOG | Jiří BUREŠ: <http://blog.converter.cz/>
- [35] Dlouhý Web: Weblog: <http://www.vitdlouhy.cz/weblog/>
- [36] About weblog: <http://www.zpravy.net/weblog/>
- [37] CubSEO: <http://cubseo.bloguje.cz/>
- [38] SearchEngineWatch: <http://searchenginewatch.com/>

Slovníček

A

Alt text

Text, který je umístěn v HTML značce obrázku. Pokud se neobjeví na obrazovce obrázek, zobrazuje se alt text.

Anchor text

Text, který je mezi párovou HTML značkou `<a>`. Běžně se mu říká také textový odkaz.

B

Browser (prohlížeč)

Software, který se používá na zobrazování webových stránek na internetu. Mezi nejrozšířenější prohlížeče patří Microsoft Internet Explorer, Opera a Mozilla.

C

CSS (Kaskádové styly)

Šablony pomocí nichž se formátuje webová stránka. Určují tedy jak mají vypadat různé elementy na stránce (nadpisy, odstavce, odkazy ...).

Click through rate (CTR)

Marketingový výraz, který udává jaké procento lidí klikne po shlédnutí na inzerát.

Click through popularity

Počet kolikrát uživatelé klikli na odkaz ve výsledcích vyhledávání a jak dlouho na webové stránce zůstali. Některé vyhledávače používají popularitu pro zvýšení relevantnosti vyhledávání.

Cloaking

Jedná se o vytváření speciálních stránek pro vyhledávací roboty. Robotu se podsouvá jiný obsah než běžnému uživateli.

Cookies

Malé textové soubory, které Web server posílá uživateli. Slouží k zapamatování specifických informací o uživateli.

Crawler

Jiné slovo pro vyhledávacího robota.

D**Doménové jméno**

Textové jméno, které identifikuje web. Například example.com.

Doorway pages

Znamenají vytvoření více stránek, které nemají jako takové jiný účel, než směřovat na domovskou stránku, a tak zajistit její lepší pozici ve vyhledávacích. Neobsahují téměř žádné užitečné informace pro uživatele.

Dynamické stránky

Stránky jejichž URL je ve tvaru www.example.com/abc.php?id=xyz&id=abc. Obsahují tedy znaky jako ?, =, & a jsou psány v PHP, ASP, CFM

F

Frames (rámce)

Technika, která umožňuje tvůrcům stránek rozdělit obrazovku prohlížeče do několika částí. Každá sekce je jednotlivá stránka.

Fulltextové vyhledávání

Jedná se o vyhledávání v databázi, která obsahuje kompletní obsah dokumentu. Ne tedy pouze nadpis nebo URL.

G

Getter

Část robota vyhledávače která stahuje webové stránky pro pozdější indexaci.

Google update (Google dance)

Větší změna algoritmu nebo dat ve vyhledávači Google. Změny algoritmů neprobíhají v pravidelných intervalech. Některé změny probíhají průběžně.

H

HTML

Zkratka pro Hypertext Markup Language. Jazyk pro tvorbu hypertextových dokumentů.

HTTP

Zkratka pro HyperText Transfer Protocol. Systém, který se používá pro přenos dat mezi webovým serverem a prohlížečem.

Hustota klíčového slova (keyword density)

Vyjadřuje podíl četnosti klíčového slova (počtu výskytů) ku celkovému počtu slov celého textu.

I

Indexer

Část vyhledávače, která upravuje (odstraňuje html značky ...) a ukládá webové stránky do databáze.

IP adresa

Unikátní číslo, které identifikuje počítač (zařízení) na internetu. Současný tvar je "255.255.255.255".

J

Javascript

Skriptovací jazyk vyvinutý společností Netscape. Používá se například pro kontrolu formulářů.

JyxoRank

Obdoba PageRanku. Číslo, které udává jak populární je webová stránka na internetu. Základem této popularity je kolik kvalitních odkazů vede na webovou stránku. JyxoRank je zjistitelný z Jyxo Toolbaru.

K

Katalog

Web, který obsahuje celou řadu odkazů na webové stránky. Odkazy jsou řazeny do tématických oblastí. Mezi nejznámější český katalogy patří Seznam, Centrum a Atlas, mezi zahraniční Yahoo! a ODP.

Keyword density (hustota klíčového slova)

Vyjadřuje podíl četnosti klíčového slova (počtu výskytů) ku celkovému počtu slov celého textu.

Keyword prominence

Udává jak vysoko se objevuje slovo na webové stránce (ve zdroji html stránky).

Klíčová slova (keywords)

Slova, která jsou vložena do vyhledávače.

Konverzní poměr

Udává poměr mezi určitou akcí a počtem návštěvníků. Pokud například na webový obchod přijde z vyhledávače 100 lidí a 10 si něco koupí je konverzní poměr 10 procent.

L**Link farma (odkazová farma)**

Webové stránky, které na sebe odkazují aby uměle zvýšily popularitu své stránky. Většinou se jedná o stránky co nemají nic společného.

Log soubor

Soubor, který vytváří web nebo proxy server, obsahuje všechny informace o požadavcích na daný server.

M**Mapa serveru**

Webová stránka, která odkazuje na podstatné stránky daného webu.

Meta značka (meta tag)

HTML značka, která je umístěna mezi <head> a </head>. Je v ní obsažen například popis obsahu webu, klíčová slova Meta značky nevidí uživatel v prohlížeči (pouze pokud si zobrazí zdroj stránky).

N

Návštěvnost (traffic)

Počet uživatelů kteří přijdou na webovou stránku.

Noframes

Obsah této značky se zobrazuje ve vyhledávači v případě, že web používá rámce.

O

Obrázková mapa

Obrázek, který je rozčleněn na části pomocí HTML kódu. Tyto části odkazují na rozlišné webové stránky.

Open Directory Project (ODP)

Mezinárodní katalog, který spravují dobrovolní editoři.

P

PPC (Pay Per Click)

Reklama u které se platí až za kliknutí uživatelem.

PFI (Pay For Inclusion)

Vyhledávač garantuje, že zahrne stránku do databáze po zaplacení poplatku, negarantuje však jak vysoko se stránka umístí.

PageRank

Číslo, které udává jak populární je webová stránka na internetu. Základem této popularity je kolik kvalitních odkazů vede na webovou stránku.

Poznámka (comment tag)

HTML značka<!-- a -->, která slouží pouze pro tvůrce webu, v prohlížeči se nezobrazuje. Většina vyhledávačů tuto značku ignoruje.

Prohlížeč (browser)

Software, který se používá na zobrazování webových stránek na internetu. Mezi nejrozšířenější prohlížeče patří Microsoft Internet Explorer, Opera a Mozilla.

R**Ranking system**

Systém hodnocení webových stránek. Většinou podle počtu a kvality odkazů, které směřují na danou stránku.

Referrer

URL webové stránky ze které přišel návštěvník webu. Zkoumá se jestli přišel přímo nebo z jiných stránek. Pokud návštěvník přišel z vyhledávače zkoumá se referrer pro zjištění použitých klíčových slov.

Relevantnost

Důležitost výsledků vyhledávání vzhledem k položenému dotazu uživatele. Často je velice subjektivní.

Robot

Softwarový program, který je součástí vyhledávače. Prochází web, následuje odkazy, stahuje text webových stránek a ukládá je do databáze. Někdy se používají také výrazy spider nebo crawler.

S

SERP

Zkratka pro Search Engine Results Page. Je to stránka s výsledky vyhledávání.

Server

Počítač, na kterém jsou umístěny webové stránky.

Spam

Všechny neetické postupy, které mají cíl dosažení vyšších pozic ve vyhledávačích. Jejich použitím se snižuje relevantnost vyhledávaného výrazu a tím se snižuje uživateli použitelnost vyhledávače.

Spider

Software, který automaticky prochází web aby získal webové stránky pro databázi vyhledávače.

Spam report

Nástroj jehož pomocí se hlásí neetické (spam) techniky vyhledávačům. Na základě Spam reportu, vyhledávače stránku mohou penalizovat.

Splash page

Vstupní stránka, která obsahuje jen velký obrázek nebo flashovou animaci.

Statické stránky

Stránky jejichž URL je ve tvaru www.example.com/xyz/abc. Neobsahují tedy znaky jako ?, =, &.

Stemming (skloňování)

Některé vyhledávače dokáží skloňovat slova. Odvozený tvar má, ale nižší váhu než přesný výraz.

T

Titulek (title)

Text v html značce `<title> </title>`. Nejdůležitější značka pro SEO.

U

URL

Zkratka pro Uniform Resource Locator, adresa souboru(stránky) na internetu.

V

Vyhledávač (Search engine)

Software který stahuje a indexuje webové stránky do databáze a v této databázi umožňuje uživatelům vyhledávat pomocí dotazů. Mezi nejznámější zahraniční vyhledávače patří Google, mezi české pak Jyxo a Morfeo.

W

Web copywriting

Psaní textu speciálně pro webové stránky. Jedna z nejdůležitějších částí SEO.

Webová stránka

Jednotlivá stránka která je součástí webu. Obsahuje text, grafiku a odkazy.

Web (website)

Soubor webových stránek, které mají společné téma, obvykle jedna doména.

Z

Zpětný odkaz

Odkazy z jiných webových stránek vedoucích na konkrétní stránku.

Přílohy

Seznam status kódů webového serveru:

HTTP status kód určuje zda přenos proběhl v pořádku nebo nikoliv a proč. Tyto kódy jsou definovány v RFC 2616 sekce 6.1.1

"Success" codes:

- 100 = Success:** Continue
- 101 = Success:** Switching Protocols
- 200 = Success:** OK
- 201 = Success:** Created
- 202 = Success:** Accepted
- 203 = Success:** Non-Authoritative Information
- 204 = Success:** No Content
- 205 = Success:** Reset Content
- 206 = Success:** Partial Content
- 300 = Success:** Multiple Choices
- 301 = Success:** Moved Permanently
- 302 = Success:** Found
- 303 = Success:** See Other
- 304 = Success:** Not Modified
- 305 = Success :** Use Proxy
- 307 = Success :** Temporary Redirect

"Failed" codes:

- 400 = Failed:** Bad Request
- 401 = Failed:** Unauthorized
- 402 = Failed:** Payment Required
- 403 = Failed:** Forbidden
- 404 = Failed:** Not Found
- 405 = Failed:** Method Not Allowed
- 406 = Failed:** Not Acceptable
- 407 = Failed:** Proxy Authentication Required
- 408 = Failed:** Request Time-out
- 409 = Failed:** Conflict
- 410 = Failed:** Gone
- 411 = Failed:** Length Required

412 = Failed: Precondition Failed
413 = Failed: Request Entity Too Large
414 = Failed: Request-URI Too Large
415 = Failed: Unsupported Media Type
416 = Failed: Requested range not satisfiable
417 = Failed: Expectation Failed
500 = Failed: Internal Server Error
501 = Failed: Not Implemented
502 = Failed: Bad Gateway
503 = Failed: Service Unavailable
504 = Failed: Gateway Time-out
505 = Failed: HTTP Version Not Supported

Konfigurace programu Weblog Expert pro české vyhledávače

Do souboru SearchEngines.cfg je nutné přidat:

```
Seznam fulltext svět = search1.seznam.cz; w
Seznam fultext ČR, nyní jyxo =
search2.seznam.cz; w
Seznam u nás = search.seznam.cz; w
Seznam celkem = seznam.cz; w
Idnes = idnes; q
Jyxo = jyxo; s
Atlas = atlas; q
Centrum = centrum; q
Morfeo.cz = morfeo.cz; q_req
Quick = quick; ftxt_query
zoohoo = zoohoo; q
Tiscali.cz = tiscali; query
Redbox = redbox; qs
Volny.cz = volny.cz; search
```

Do souboru Spiders.cfg:

```
holmes = Morfeo Holmes
jyxobot = Jyxobot
```

Bazos.cz+.sk, Jasminka.cz