

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÝCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER SYSTEMS

ON-PAGE A OFF-PAGE OPTIMALIZACE WEBOVÝCH STRÁNEK

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

MARTIN ŠIMKO

BRNO 2011



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÝCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER SYSTEMS

ON-PAGE A OFF-PAGE OPTIMALIZACE WEBOVÝCH STRÁNEK

ON-PAGE AND OFF-PAGE WEB PAGE OPTIMALIZATION

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

MARTIN ŠIMKO

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. PAVOL KORČEK

BRNO 2011

Abstrakt

Tato práce se zabývá principem fungování moderních fulltextových vyhledávačů a definuje jejich odlišnost od katalogů. Dále popisuje, které on-page a off-page faktory vyhledávače ve svých algoritmech zohledňují a jakou jim přiřkládají váhu. Vybrané faktory jsou následně prakticky otestovány na množině deseti webů ve vyhledávačích Seznam a Google. Přitom jsou měřeny pozice ve výsledcích vyhledávání, pomocí kterých probíhá také vyhodnocení.

Abstract

This bachelor's thesis deals with functioning of modern full-text search engines and defines their difference from directories. It also describes which on-page and off-page factors are more or less important for these search engines. Selected factors are then practically tested on a set of ten websites in search engines Seznam and Google. In doing so positions are measured in search engine result page. Through the use of these positions is made evaluation of factors.

Klíčová slova

seo, optimalizace pro vyhledávače, vyhledávač, katalog, Seznam, Google, on-page faktor, off-page faktor, serp

Keywords

seo, search engine optimization, search engine, directory, Seznam, Google, on-page factor, off-page factor, serp

Citace

Martin Šimko: On-Page a Off-Page optimalizace webových stránek, bakalářská práce, Brno, FIT VUT v Brně, 2011

On-Page a Off-Page optimalizace webových stránek

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením pana Ing. Pavla Korčeka. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....
Martin Šimko
11. května 2011

Poděkování

Rád bych touto cestou poděkoval vedoucímu mé bakalářské práce Ing. Pavlovi Korčekovi za podnětné rady a pomoc při vytváření této publikace. Naše konzultace se vždy vedly v přátelské atmosféře a byly mi inspirací v další tvorbě práce. V neposlední řadě bych také rád poděkoval tvůrcům vyhledávačů, kteří umožnili snazší dohledatelnost dokumentů na internetu.

© Martin Šimko, 2011.

Tato práce vznikla jako školní dílo na Vysokém učení technickém v Brně, Fakultě informačních technologií. Práce je chráněna autorským zákonem a její užití bez udělení oprávnění autorem je nezákonné, s výjimkou zákonem definovaných případů.

Obsah

1 Úvod	5
2 Vyhledávače	6
2.1 Jak pracují vyhledávače	6
2.1.1 Supplemental index	6
2.1.2 Periodicita procházení roboty	6
2.1.3 Identifikace robotů	7
2.1.4 Omezení vstupu robotů na stránky	7
2.2 Významné vyhledávače v českém prostředí	9
2.3 Hodnocení stránek vyhledávači	9
2.3.1 PageRank	10
2.3.2 S-rank	10
2.3.3 Další ranky	10
3 Katalogy	12
3.1 Schvalování nových záznamů	12
3.2 Řazení záznamů	13
3.3 Partnerské a placené zápisy	13
3.4 Nejznámější katalogy	13
4 On-page faktory	14
4.1 Definování cíle	14
4.2 Klíčová slova	14
4.2.1 Výběr klíčových slov	15
4.2.2 Hustota klíčových slov	15
4.3 Umístění klíčových slov na stránce	16
4.3.1 Text stránek	16
4.3.2 Tag title	16
4.3.3 Meta tag description	17
4.3.4 Meta tag keywords	17
4.3.5 Nadpisy h1 – h6	17
4.3.6 Zvýraznění textu	17
4.3.7 Doménové jméno a URL	18
4.4 Mapa stránek	18
4.4.1 HTML mapa	18
4.4.2 XML Sitemap Protocol	18
4.5 Duplicitní obsah	20
4.5.1 Zamezení duplicitnímu obsahu	20

4.6	Rychlost načítání stránek	22
5	Off-page faktory	23
5.1	Zpětné odkazy	23
5.1.1	Vlastnosti zpětných odkazů	23
5.1.2	Detekce zpětných odkazů	25
5.2	Odkazový profil	25
5.3	Linkbuilding – získávání zpětných odkazů	26
5.3.1	Linkbaiting	26
5.3.2	Záložkování	27
5.3.3	Sociální sítě	27
5.3.4	Registrace do katalogů	27
5.3.5	Další možnosti	27
6	Testování on-page a off-page faktorů	28
6.1	Metodika testování	28
6.1.1	Detekce přeindexování stránek	28
6.2	Měření pozic	29
6.2.1	Specifika měření pozic u vyhledávače Google	29
6.3	Testované webové stránky	30
6.3.1	Struktura stránek	30
6.3.2	Webhostingové řešení	31
6.4	On-page faktory	31
6.4.1	Titulek stránky – HTML tag <title>	32
6.4.2	Zvýrazněné písmo – HTML tag 	33
6.4.3	Nadpisy – HTML tagy <h1> až <h6>	34
6.4.4	Nejdůležitější on-page faktory	36
6.5	Off-page faktory	36
6.5.1	Záložkovací služba Linkuj.cz	37
6.5.2	Vzájemné odkazy mezi testovanými weby	38
6.5.3	Registrace do katalogu <i>odkazy.seznam.cz</i>	40
6.5.4	Nejdůležitější off-page faktory	41
7	Zhodnocení testů a vzorový web	43
7.1	On-page faktory	43
7.2	Off-page faktory	43
7.3	Vzorový web	44
8	Závěr	45
A	Naměřené pozice ve vyhledávačích	51
A.1	Vysvětlivky k tabulkám	51
B	Obsah CD	64

Seznam obrázků

2.1	Identifikace robotů pomocí programu AWStats [18]	7
2.2	Podíly vyhledávačů na českém trhu ke dni 9. 12. 2010	10
2.3	Zjednodušený výpočet PageRanku	11
4.1	Princip long tail klíčových slov	15
4.2	SERP českého vyhledávače Seznam	16
5.1	Analýza zpětných odkazů webu FIT VUT pomocí Yahoo! Site Explorer	26
6.1	Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po úpravách titulku	32
6.2	Pozice webů na dané klíčové fráze ve vyhledávači Google před a po úpravách titulku	33
6.3	Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po zvýraznění klíčových slov pomocí tagu strong	34
6.4	Pozice webů na dané klíčové fráze ve vyhledávači Google před a po zvýraznění klíčových slov pomocí tagu strong	35
6.5	Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po zvýraznění klíčových slov pomocí nadpisů	36
6.6	Pozice webů na dané klíčové fráze ve vyhledávači Google před a po zvýraznění klíčových slov pomocí nadpisů	37
6.7	Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po získání zpětného odkazu ze služby Linkuj.cz	38
6.8	Pozice webů na dané klíčové fráze ve vyhledávači Google před a po získání zpětného odkazu ze služby Linkuj.cz	39
6.9	Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po vytvoření zpětných odkazů mezi testovanými weby	40
6.10	Pozice webů na dané klíčové fráze ve vyhledávači Google před a po nastavení zpětných odkazů mezi testovanými weby	41
6.11	Pozice web kreta.hu.cz na dané klíčové fráze ve vyhledávači Seznam před a po vytvoření zpětného odkazu v katalogu odkazy.seznam.cz	42
6.12	Pozice web kreta.hu.cz na dané klíčové fráze ve vyhledávači Seznam před a po vytvoření zpětného odkazu v katalogu odkazy.seznam.cz	42

Seznam tabulek

4.1	Escape sekvence pro entity podle XML Sitemap protokolu	19
A.1	Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po úpravách titulku	52
A.2	Pozice webů na dané klíčové fráze ve vyhledávači Google před a po úpravách titulku	53
A.3	Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po zvýraznění klíčových frází pomocí tagu <code></code>	54
A.4	Pozice webů na dané klíčové fráze ve vyhledávači Google před a po zvýraznění klíčových frází pomocí tagu <code></code>	55
A.5	Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po zvýraznění klíčových slov pomocí nadpisů	56
A.6	Pozice webů na dané klíčové fráze ve vyhledávači Google před a po zvýraznění klíčových slov pomocí nadpisů	57
A.7	Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po získání zpětného odkazu ze služby Linkuj.cz	58
A.8	Pozice webů na dané klíčové fráze ve vyhledávači Google před a po získání zpětného odkazu ze služby Linkuj.cz	59
A.9	Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po vytvoření zpětných odkazů mezi testovanými weby	60
A.10	Pozice webů na dané klíčové fráze ve vyhledávači Google před a po vytvoření zpětných odkazů mezi testovanými weby	61
A.11	Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po vytvoření zpětného odkazu z katalogu odkazy.seznam.cz	62
A.12	Pozice webů na dané klíčové fráze ve vyhledávači Google před a po vytvoření zpětného odkazu z katalogu odkazy.seznam.cz	63

Kapitola 1

Úvod

S rozmachem internetu došlo k velké popularizaci webových stránek. Z počátku, kdy existovalo webů ještě málo, uživatelé nepotřebovali vyhledávače jak je známe dnes. Jejich úlohu zastávaly rozcestníky označované také jako internetové portály. Tyto portály byly jakousi vstupní branou do světa internetu – obsahovaly odkazy na webové stránky, které byly často tříděny do kategorií. S nárůstem počtu webů však bylo potřeba zajistit v rozumné míře vyhledávání. Uživatelům přestaly samotné odkazy na další weby stačit a začali požadovat relevantní odpovědi na svoje vyhledávací dotazy. Z tohoto důvodu se začaly objevovat vyhledávače, které podle svých algoritmů nabízejí uživatelům soubor webů, na kterých by se měla nacházet odpověď na položený dotaz. Provozovatelé webových stránek brzy zjistili, že vyhledávače jsou cenným zdrojem návštěvnosti, a proto se začali snažit svoje weby vytvářet tak, aby byly přístupné pro vyhledávače. Zrodil se tak nový internetový obor, který se označuje anglickým termínem *Search Engine Optimization*, zkráceně *SEO*. Do češtiny jej překládáme jako *optimalizace pro vyhledávače*.

Cílem této práce je představit fungování vyhledávačů (kapitola 2) a katalogů (kapitola 3). Pro pochopení toho, jakým způsobem je nutné stránky optimalizovat, aby byly dohledatelné na internetu, je důležité vysvětlit principy on-page a off-page optimalizace. On-page faktory jsou popsány v kapitole 4 a off-page v kapitole 5. Protože však algoritmy moderních vyhledávačů jako je například Google nebo Seznam nejsou veřejně známé, popisované skutečnosti jsou založeny na empirickém pozorování a výzkumu. Tyto informace byly čerpány z použitých zdrojů práce.

Počet stránek na internetu denně roste, a proto se tvůrci vyhledávačů snaží neustále inovovat indexovací a vyhledávací algoritmy. Jejich cílem je, aby uživatel vždy dostal co nejrelevantnější odpověď na vyhledávací dotaz. Z tohoto důvodu se váhy jednotlivých on-page a off-page faktorů webových stránek neustále mění. Tato práce si klade za cíl prakticky otestovat vybrané faktory a na základě měření pozic ve výsledcích vyhledávání odvodit, které jsou pro vyhledávače důležité více a které méně. Této problematice se věnuje kapitola 6. Zhodnocení naměřených dat je následně popsáno v kapitole 7, kde jsou také uvedena doporučení tvůrcům a správcům webových stránek.

Kapitola 2

Vyhledávače

2.1 Jak pracují vyhledávače

Dnešní vyhledávače používají pro procházení webových stránek tzv. *roboty*. Robot se také často označuje jako *crawler*, *prohledávač*, *webový pavouk* nebo pouze jako *bot*. V zásadě jde o program, který si přes HTTP protokol stahuje webové stránky. Na začátku je potřeba robotu zadat, která URL má procházet. Ta jsou mu dodána jednak na základě předcházejících prohledávání a také ze souboru *sitemap.xml* (více viz kapitola 4.4.2), pokud ho daný web obsahuje. Poté se robot v podstatě chová jako běžný uživatel, to znamená, že na stránce načítá textové informace a vyhledává hypertextové odkazy na další dokumenty (tj. obsah atributů *src* a *href* v HTML elementech). Tato URL si ukládá do tzv. *databáze vyhledávače*, aby je mohl později navštívit. Ze stránek, které již robot prošel, se poté vytváří *index vyhledávače*. Nutně však neznamená, že URL, která jsou v databázi, musí být i v indexu [35]. V tomto indexu se následně vyhledávají uživatelské dotazy. Reflektuje se zde především text, který je na dané stránce – mluvíme tedy o tzv. *fulltextových vyhledávačích* [45], dále pak různé faktory, které vyhledávač dané stránce přisoudil. Přesný algoritmus vyhledávání v indexu však není znám a vyhledávače si jej pečlivě chrání a neustále jej obměňují a vyvíjí.

2.1.1 Supplemental index

Google má kromě „normálního indexu“ také další, který se nazývá *supplemental index* – jde o doplňkový index, kam si robot ukládá stránky, které mu nepřijdou příliš kvalitní. To mohou být například takové, které mají velmi podobný obsah s jinými stránkami, mají nízký rank apod. V *supplemental indexu* se běžně nevyhledává, Google do něj sahá pouze v případě, že nemá dostatek výsledků pro aktuálně zadanou vyhledávací frázi. To jsou často víceslovné fráze nebo takové dotazy, které mají vylučovací operátor [33]. Zda je stránka v *supplemental indexu* se nedá zjistit.

2.1.2 Periodicita procházení roboty

Ne všechny stránky jsou roboty procházeny stejně často. Záleží především na tom, jak často se aktualizuje obsah daného webu. Pokud robot zjistí, že při každé jeho návštěvě nalezne na stránce nový obsah, zvyšuje periodicitu procházení. To platí především u zpravodajských webů apod. Neplatí však, že pokud web není aktualizovaný, roboti jej už nenavštíví. Právě naopak, roboti procházejí po určitém čase všechny již nalezené stránky a hledají, zda nedošlo ke změnám od poslední návštěvy.

2.1.3 Identifikace robotů

Jak už bylo zmíněno, roboti stahují jednotlivé webové stránky prostřednictvím protokolu HTTP. V HTTP hlavičce o sobě posílají informace – především jméno. Z toho plyne, že je lze na daném serveru jednoznačně identifikovat. Například robot českého vyhledávače Seznam.cz se identifikuje jako *SeznamBot*, crawler Google jako *Googlebot*, Yahoo jako *Yahoo Slurp* atd. Seznam robotů lze nalézt v [47]. Příklad identifikace robotů z programu *Advanced Web Statistics*, který dokáže rozpoznat 643 robotů [16], je zobrazen na obrázku 2.1.

Roboty (Top 10) - Úplný seznam - Poslední návštěvy			
26 různých robotů*	Hity	Přenesená data	Poslední návštěvy
Unknown robot (identified by 'robot')	89+2	84.82 KB	02 Kvě 2011 - 21:04
Unknown robot (identified by 'bot*')	34+55	256.28 KB	04 Kvě 2011 - 04:33
BSpider	32+12	207.13 KB	04 Kvě 2011 - 04:58
Yahoo Slurp	31+13	152.09 KB	04 Kvě 2011 - 05:43
Yandex bot	19+10	64.79 KB	04 Kvě 2011 - 06:21
Googlebot	10+11	70.86 KB	04 Kvě 2011 - 06:17
Unknown robot (identified by empty user agent string)	20	596.66 KB	04 Kvě 2011 - 01:39
MJ12bot	10+10	194.43 KB	04 Kvě 2011 - 04:26
Unknown robot (identified by 'checker')	19	0	04 Kvě 2011 - 05:06
BaiDuSpider	14+2	146.15 KB	04 Kvě 2011 - 04:38
Ostatní	25+24	320.21 KB	

* Zde uvedené roboty představují zátěž, která nebyla způsobena běžnými uživateli, a proto není zahrnuta v ostatních tabulkách. Čísla po + udávají počet úspěšných stažení souboru robots.txt.

Obrázek 2.1: Identifikace robotů pomocí programu AWStats [18]

2.1.4 Omezení vstupu robotů na stránky

Může se stát, že nechceme, aby nějaké části webu robot procházel. Typicky jde např. o tiskovou verzi stránek, přihlašovací stránky do administrace apod. Zamezení přístupu robotům se také používá pro odstranění duplicit napříč webem. Podrobnosti duplicitního obsahu jsou popsány v kapitole 4.5. Pro zamezení vstupu robotů na stránky máme několik možností.

Robots Exclusion Protocol

Robots Exclusion Protocol, někdy také označovaný jako *Robots Exclusion Standard*, je způsob, jak jednoduchým textovým souborem zamezit přístup vyhledávacím robotům na určité části webu. Příkazy tohoto protokolu se zapisují do souboru, který musí být pojmenován `robots.txt` a umístěn v kořenovém adresáři webu. Roboty lze buď omezit všechny, nebo nastavit různé podmínky pro různé roboty. Poté, co robot přijde na web, nejprve zkontroluje, zda existuje soubor `robots.txt`. Pokud ano, načte ho a dále se řídí příkazy, které jsou v tomto souboru uvedené. Pokud soubor neexistuje, robotu nejsou kladena omezení a prochází všechny stránky. Syntaxe protokolu je následující:

User-agent: <jméno robota>

Disallow: <složka, která nemá být procházena>

Jméno robota lze buď přesně specifikovat, nebo jej nahradit znakem *, který bude přikazovat níže napsaná pravidla všem robotům.

Ukázka souboru robots.txt, který zakazuje všem robotům přístup do složky *duverne* je uvedena v kódu 2.1.

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

Kód 2.1: Ukázka souboru robots.txt s jedním pravidlem

Ukázka souboru, která zakazuje přístup vyhledávacímu robotu Seznam.cz přístup do složek *tiskoveSestavy* a *faktury* je uvedena v kódu 2.2. V případě zákazu vstupu robotů do více složek, je nutné tyto složky vyjmenovat vždy samostatně s využitím klíčového slova `Disallow:`.

```
User-agent: SeznamBot  
Disallow: /tiskoveSestavy/  
Disallow: /faktury/
```

Kód 2.2: Ukázka souboru robots.txt s více pravidly

Meta tag robots

Meta tag robots je další způsob, jak zakázat indexaci dané stránky roboty. Meta tagy se umisťují přímo do HTML stránky a to konkrétně do její hlavičky (mezi tagy `<head>` a `</head>`). Atribut `name` nabývá hodnoty `robots`. Takto se budou pravidla aplikovat na všechny roboty, ale pravidla lze nastavit pouze pro specifické roboty vyhledávačů. V tomto případě se klíčové slovo `robots` nahradí názvem robota, viz výše.

Druhý atribut meta tagu se nazývá `content` a může nabývat následujících hodnot:

- `index`, `noindex` – povoluje, respektive zakazuje, danou stránku zobrazovat ve vyhledávání
- `follow`, `nofollow` – povoluje, respektive zakazuje, následovat odkazy směřující na další stránky
- `all` – vše povoleno, stejné jako `index` a `follow`

Googlebot navíc umí zpracovat následující hodnoty[34]:

- `snippet`, `nosnippet` – povoluje, respektive zakazuje, používat obsah stránky jako úryvky textu (snippets) ve vyhledávání
- `none` – ekvivalentní zápis k `noindex`, `nofollow`
- další hodnoty (které přímo nesouvisí se zákazem vstupu robotů) lze nalézt v nápovědě společnosti Google [32]

V jednom HTML souboru může být meta tagů robots i více než jeden a hodnoty atributu `content` lze seskupovat. Ukázka meta tagu robots je uvedena v kódu 2.3.

```
<meta name="robots" content="noindex , nofollow">  
<meta name="googlebot" content="nosnippet">
```

Kód 2.3: Ukázka použití meta tagu robots

X-Robots-Tag

`X-Robots-Tag` je element, který může být uveden v hlavičce odpovědi HTTP protokolu pro dané URL. Jako parametr tohoto elementu může být použita hodnota z atributu `content meta` tagu, přičemž elementů je povoleno uvádět i více. V současné době `X-Robots-Tag` podporuje např. vyhledávač Google [32], Bing [28] a Yahoo! [53]. Použití `X-Robots-Tag` je uvedeno v kódu 2.4.

```
HTTP/1.1 200 OK
Date: Tue, 7 Dec 2010 15:11:37 GMT
(...)
X-Robots-Tag: noarchive
X-Robots-Tag: unavailable_after: 24 Dec 2010 15:00:00 PST
(...)
```

Kód 2.4: Ukázka použití X-Robots-Tag

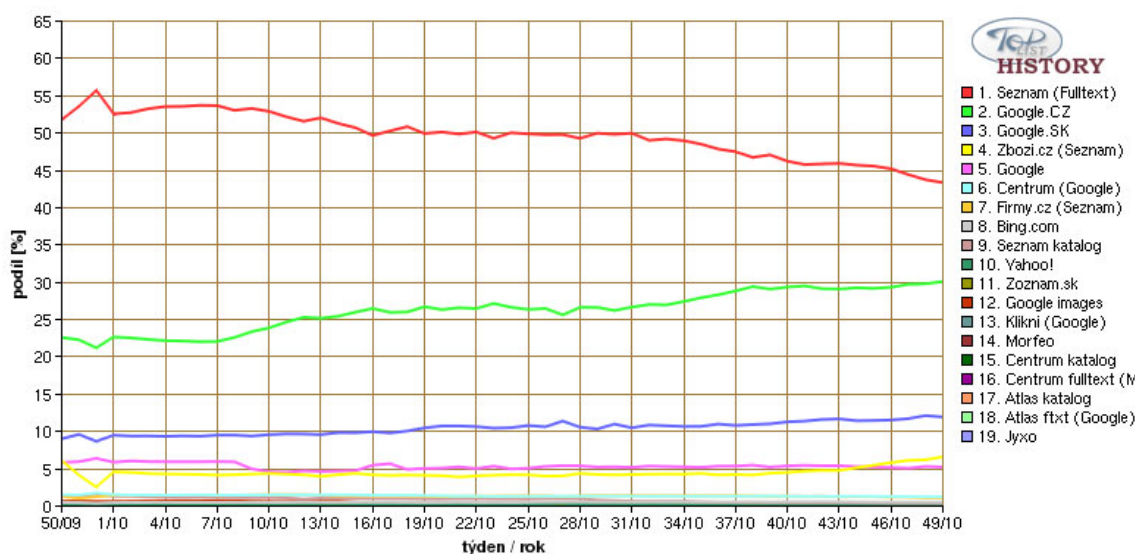
2.2 Významné vyhledávače v českém prostředí

Přesné měření struktury vyhledávačů není tak jednoduché, jak by se mohlo zdát [51]. Podíly vyhledávačů jsou většinou počítány společnostmi, které nabízejí měření návštěvnosti. Reflektuje se zde tzv. *HTTP referrer* – jde o URL stránky, ze které přišel návštěvník na daný web. Z těchto referrerů se poté vytvářejí statistiky podílů vyhledávačů na trhu. Tyto statistiky v českém prostředí vytváří například TopList.cz. Je však nutné připomenout, že jsou zde zahrnuti pouze ti návštěvníci, kteří přijdou z vyhledávače na web a ten má na svých stránkách umístěn měřicí kód TopListu. Samotné vyhledávače žádné statistiky nezveřejňují.

Jak je vidět na obrázku 2.2, na českém trhu podle historie služby TopList vede vyhledávač Seznam.cz. Druhý je Google.cz a třetí Google.sk. Za zmínku také stojí čtvrté místo, na kterém není fulltextový vyhledávač, ale *vyhledávač zboží* - Zbozi.cz. Podíly vyhledávačů se neustále mění, ale v posledních cca 5 letech se přední pozice již ustálily natolik, abychom mohli tvrdit, že mezi nejvýznamnější fulltextové vyhledávače v českém prostředí patří *Seznam* a *Google*.

2.3 Hodnocení stránek vyhledávači

Je běžné, že si vyhledávače podle svých algoritmů interně označují, jakou má daná stránka „sílu“. Často se tato síla vyjadřuje na stupnici, kde stránka s hodnocením 0 je nejméně autoritativní a s rostoucím hodnocením roste také její síla. Toto hodnocení není lineární a každý vyhledávač má vlastní algoritmus jeho výpočtu, název apod. Mluvíme zde o tzv. *ranku*, např. Google PageRank, Seznam S-rank apod. Primárně jde o interní hodnocení, největší vyhledávače je však zveřejňují. Zveřejněný rank ale nemusí plně korespondovat s aktuálním – tj. tím, se kterým v daném okamžiku pracují vyhledávače. K aktualizaci veřejných ranků totiž dochází periodicky po určité době a tato aktualizace není na vyhledávačích vynutitelná. Zveřejnění často probíhá pomocí tzv. *toolbarů*, které se instalují do webového prohlížeče – např. Google Toolbar [7], Seznam Lištička [9] apod. Dalším způsobem, jak zjistit aktuální ranky webů, jsou webové stránky [38] nebo pluginy do prohlížečů [48].



Obrázek 2.2: Podíly vyhledávačů na českém trhu ke dni 9. 12. 2010

2.3.1 PageRank

PageRank je algoritmus, podle kterého jsou ohodnocovány jednotlivé internetové stránky vyhledávačem Google. Jeho autory jsou *Sergej Brin* a *Larry Page*. PageRank je založen na hodnocení pomocí zpětných odkazů a předávání síly z jedné stránky na druhou. Ve své podstatě jde o podobný princip jako jsou citace ve vědeckých pracích, navíc se zde však také zohledňuje hodnocení odkazujících stránek [2]. Google Tool PageRank (GTPR), tedy ten rank, který se zobrazuje například pomocí Google Toolbaru, nabývá hodnot 0–10. Je však docela možné, že vnitřní algoritmus PageRanku pracuje s větším rozsahem hodnot. Každá stránka vystupuje autonomně (vyhledávače neznají pojem „web“) a získává tak svůj GTPR nezávisle na hodnotě GTPR ostatních stránek v rámci jednoho webu.

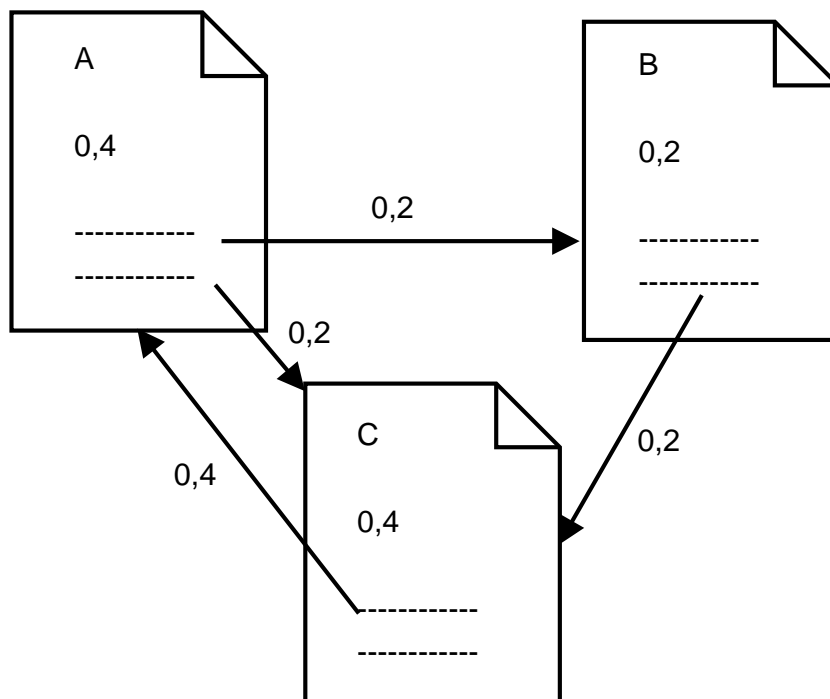
Zjednodušený výpočet PageRanku je znázorněn na obrázku 2.3. V obrázku nejsou však zohledněny další faktory, které se při výpočtu berou v úvahu. Takto by teoreticky pro dosažení PageRanku deset stačilo, aby na jednu stránku vedlo deset odkazů ze stránek, které mají PageRank roven jedné. Přesný algoritmus výpočtu není znám, základní myšlenka a obecný vzorec výpočtu je však uveden v [43].

2.3.2 S-rank

S-rank je hodnota, která udává důležitost dané stránky ve vyhledávači Seznam. Výpočet *S-ranku* je založen na algoritmu *Hubs & Authorities* [1] a nabývá hodnot 0 – 10. Do tohoto výpočtu jsou zahrnuty ještě další faktory, které však nejsou veřejně známé. Opět jako u PageRanku jsou však nejdůležitější off-page faktory [12]. *S-rank* mohou získat pouze české stránky.

2.3.3 Další ranky

Ve světovém měřítku se můžeme setkat i s jinými ranky. Například *Alexa rank*, který udává návštěvnost domény druhého řádu. Výsledný rank se zde počítá pouze z návštěvníků webu, kteří mají ve svém prohlížeči nainstalovaný Alexa toolbar. Z tohoto důvodu je tento rank



Obrázek 2.3: Zjednodušený výpočet PageRanku

v českém prostředí irelevantní, protože Alexa toolbar u nás používá mizivé množství uživatelů. Alexa rank nabývá cca hodnoty z intervalu $< 1; 30\,000\,000 >$, přičemž nižší hodnota znamená navštěvovanější web.

Dalším rankem je *Compete rank*, který je podobný Alexa ranku, sleduje však pouze milion nejnavštěvovanějších webů.

Kapitola 3

Katalogy

V minulé kapitole byla rozebrána struktura fulltextových vyhledávačů. Vedle nich existují ještě *katalogy*, které nabízí zcela něco jiného, i když jsou tyto dva pojmy často zaměňované. *Katalog* je web, který obsahuje seznam stránek. Stránky jsou zde tříděny do hierarchické struktury kategorií. Kategorie lze cíleně procházet a dostat se tak k vybraným tematickým webům. Většina katalogů také nabízí vlastní vyhledávání. Nejedná se tu však o vyhledávání jak jej známe z vyhledávačů; katalogy vyhledávají pouze ve své databázi webů.

Dalším podstatným rozdílem oproti fulltextovým vyhledávačům je také fakt, že do katalogů se musí webové stránky *ručně vkládat* – registrovat. Záznam v katalogu nejčastěji obsahuje následující informace:

- URL zaregistrované stránky (tzv. *zpětný odkaz*)
- nadpis
- krátký popis; nejčastěji do 255 znaků včetně mezer
- jméno a příjmení kontaktní osoby
- e-mail na kontaktní osobu
- klíčová slova

Některé katalogy také povolují vkládat další informace jako například logo webu o dané velikosti, otevírací dobu pobočky apod.

Tím, že je v katalogu umístěn zpětný odkaz na zaregistrovaný web, je tomuto webu předávána určitá *autoritativní síla* odkazujícího webu označovaná jako *link juice*[30]. Podrobnosti o zpětných odkazech a link juice jsou popsány v kapitole 5.1.

3.1 Schvalování nových záznamů

U některých katalogů je nutné, aby administrátoři schválili navržené weby. Kontrolují, zda je web funkční a splňuje podmínky katalogu a také jestli je zařazen do správné kategorie. Pokud dojde k porušení nějakých pravidel, administrátoři zaregistrovaný web z databáze katalogu odstraní. U takovýchto katalogů je tedy alespoň částečně zaručen kvalitní obsah. Právě tyto katalogy mají pro zaregistrované weby větší užitek než ty, u kterých není nutné web po zaregistrování schválit.

Posledně jmenované katalogy se často mohou stát útokem spammerů¹, kteří jej zahltní nekvalitními weby. Poté je jasné, že pro vyhledávače bude více autoritativní ten katalog, který má ve své databázi pořádek, neobsahuje *nepovolený obsah* jako je např. *warez*², stránky s hazardem³ apod. Při registraci do katalogů je tedy nutné myslet na to, zda bude registrace prospěšná nebo nikoli. Může se také stát, že si web zaregistrovaný do pochybného katalogu naruší svůj *odkazový profil* [40] a bude ho muset pracně získávat zpět. Více o odkazovém profilu je uvedeno v kapitole 5.2. Obecně se tedy dá tvrdit, že registrace do katalogů, kde neprobíhá schvalovací proces, se nedoporučuje provádět.

3.2 Řazení záznamů

Řazení v katalogích může být uzpůsobeno podle několika pravidel:

- abecední řazení
- weby, které byly zaregistrovány dříve se zobrazují na předních příčkách
- řazení podle ranků – Srank, PageRank apod. (problematice ranků se věnuje kapitola 2.3)
- kombinace těchto metod

3.3 Partnerské a placené zápisy

Není výjimkou, že katalogy nabízejí za úplaty, nebo umístění zpětného odkazu na katalog nějaké výhody. Často se jedná o přednostní zápis v katalogu, zvýraznění (např. barvou pozadí) nebo umožnění vložení dalších informací k zápisu. Vždy je potřeba si rozmyslet, zda se tato investice vyplatí.

3.4 Nejznámější katalogy

Mezi nejznámější světový katalog patří Open Directory Project (ODP), který je na adrese www.dmoz.org. Má také svoji českou jazykovou verzi [3]. Jde o jeden z nejautoritativnějších katalogů na internetu. Jak lze nalézt v literatuře [39], popisky, které jsou uvedené u registrovaných webů, mohou vyhledávače použít jako úryvky ve vyhledávání — mluvíme o tzv. *snippetech*. Tomuto chování lze zabránit přidáním meta tagu na danou stránku, příklad je uveden v kódu 3.1.

```
<meta name="robots" content="noindex">
```

Kód 3.1: Zákaz zobrazování popisků z ODP ve snippetech vyhledávání

V České republice jsou známé především katalogy tuzemských portálů Seznam (www.firmy.cz) a Centrum (www.najisto.cz). Za zmínku stojí také katalog pro nekomerční weby odkazy.seznam.cz, ve kterém jsou weby řazeny podle S-Ranku (více o S-Ranku naleznete v kapitole 2.3.2).

¹Označení uživatele, případně programu, který šíří nevyžádaná (často komerční) sdělení.

²Díla chráněná autorským zákonem, která jsou nelegálně distribuována.

³Weby, které nabízejí služby kasín, hazardních her, sázení apod.

Kapitola 4

On-page faktory

Termínem *on-page faktory* označujeme takové faktory, které můžeme ovlivnit přímo na dané stránce. Jedná se například o *HTML tagy*, *klíčová slova*, *interní odkazy*¹, *doménové jméno*² apod.

4.1 Definování cíle

Před samotnou tvorbou webové prezentace je nutné si stanovit, jakých cílů chceme dosáhnout. Tato fáze se skládá z průzkumu trhu, průzkumu potenciálních konkurentů, hledání slabín a silných stránek budoucí prezentace (tzv. *SWOT analýza*) apod. Cíle mohou být například: zvýšení prodeje v internetovém obchodě, zvýšení odběratelů noviněk pomocí newsletteru³ apod. Z předcházející věty by mělo být patrné, že primárně nejde o to zvýšit návštěvnost webu, ale také o to udělat z návštěvníků zákazníky a dosahovat vysokých konverzí.

4.2 Klíčová slova

Klíčová slova jsou jedním z nejdůležitějších faktorů, které může tvůrce stránek ovlivnit. Jsou to slova, případně fráze, které napíše uživatel do vyhledávacího pole na stránce vyhledávače, aby získal relevantní výsledky. Catherine Seda [44] rozděluje klíčová slova na dvě skupiny:

- *spojená s obchodní značkou*
- *všeobecná*

U prvně zmíněných jde především o taková slova a fráze, které se pojí s názvem firmy a produkty, jež daná společnost vyrábí. Všeobecná slova se hledají o poznání hůře, než ta spojená s názvem firmy a produktů. Při výběru je nutné také zohlednit fakt, že ne všechna slova a fráze uživatelé vyhledávají. Například název mladé a neznámé firmy bude zřejmě vyhledávat méně lidí, než obecnou frází, která se nepojí s konkrétní firmou.

¹Hypertextové odkazy, které jsou směřovány v rámci jednoho webu.

²Jednoznačný identifikátor webové stránky, např. www.fit.vutbr.cz.

³Informativní zprávy zasílané především pomocí e-mailu, které si uživatel zaregistruje k odběru.

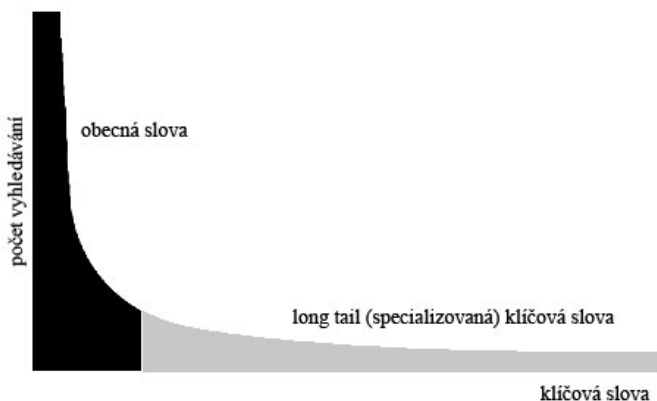
4.2.1 Výběr klíčových slov

Jak bylo již naznačeno výše, klíčová slova lze hledat ze dvou různých množin. Nejdříve je vhodné vybrat několik slov a frází, ze kterých se bude vycházet. Poté je potřeba ověřit, že tato slova jsou uživateli opravdu zadávána do vyhledávače a případně přidat slova a fráze, která nás dříve nenapadla. Pro tuto činnost existuje několik online nástrojů. Mezi nejznámější patří nástroj *Keyword Tool External* [4], který primárně slouží pro návrh klíčových slov pro PPC systém Google AdWords. Obecně lze PPC systémy použít pro návrh klíčových slov, protože obsahují množinu všech slov, které uživatelé vyhledávají. Další možností je hledání synonym. K tomuto účelu lze použít slovník synonym (tzv. *thesaurus*) dostupný například v [5].

Dále je také nutné podotknout, že pokud má web pouze regionální působitě (např. v rámci jednoho města, okresu apod.), je dobré tento region přidat do klíčové fráze. Tyto fráze jednak budou méně konkurenční a navíc se odfiltrují návštěvníci, kteří jsou z jiného regionu a stejně by nevyužili služeb webu.

Obecná vs. specializovaná klíčová slova

Na první pohled se může jevit, že použití obecnějších klíčových slov je lepší. Jsou vyhledávána více uživateli a musí tedy přinášet vyšší návštěvnost. Na druhou stranu je však také nutné myslet na to, že taková slova mohou být vysoce konkurenční. Proto je na ně optimalizace složitější a dražší. Pokud si však vybereme specializovaná slova, situace se obrací. Hledanost i konkurence je menší. Tento fakt označujeme termínem *long tail* a do češtiny se překládá často jako *dlouhý ocas*. Pojem *long tail* prezentoval *Chris Anderson* v časopise *The Wired* [21]. Z obrázku 4.1, kde je *long tail* znázorněn, je patrné, že *long tail* klíčová slova zde reprezentují specializovaná slova. Často jde o fráze složené z několika slov. Tato slova sice nepřivádějí tolik návštěvníků jako slova obecná, je u nich však vyšší konverzní poměr⁴.



Obrázek 4.1: Princip long tail klíčových slov

4.2.2 Hustota klíčových slov

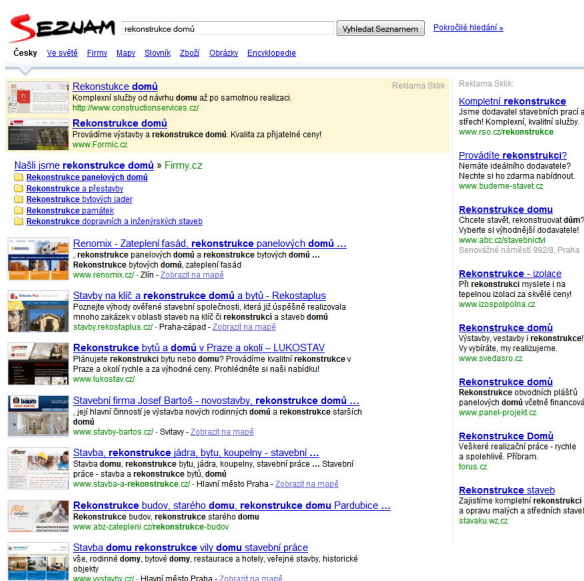
Pro vyjádření poměru klíčových slov k ostatnímu textu se používá termín *hustota klíčových slov*. Dříve algoritmy vyhledávačů nebyly tak dokonalé, jako tomu je dnes. Správci stránek

⁴Poměr návštěvníků webu ku zákazníkům vyjádřený v procentech.

brzy zjistili, že pokud zaplní svůj web pouze klíčovými frázemi, na které chtějí stránky optimalizovat, dostanou se jednoduše na přední příčky vyhledávání. Tento fakt byl však nezadržitelný a vyhledávače proti němu začaly z pochopitelných důvodů bojovat. Dnes se jako ideální hustota klíčových slov udává 7% [42]. Lepší než zahlcovat texty na stránce stejnými klíčovými slovy, je využívat synonym a různých tvarů slov. Texty budou jednak vypadat přirozeně a také se zvýší počet návštěvníků z *long tailu*.

4.3 Umístění klíčových slov na stránce

Pozice klíčových slov v rámci stránky je velmi důležitý faktor, který se podílí na pozici stránky ve vyhledávání – mluvíme o tzv. SERP. SERP je zkratka z anglického sousloví *Search Engine Result Page*. Jedná se o stránku s výsledky vyhledávání, příklad je uveden na obrázku 4.2.



Obrázek 4.2: SERP českého vyhledávače Seznam

4.3.1 Text stránek

Samotný obsah stránky je *velmi důležitým on-page faktorem*. V textu by měla být zmíněna klíčová slova vybraná pro danou stránku. Je nutné připomenout, že pokud se na stránce nevyskytuje žádný text, bude obtížnější ji optimalizovat. Platí zde nepsané pravidlo „obsah je král“.

4.3.2 Tag title

Obsah tohoto tagu se uvádí v HTML hlavičce mezi značky `<title>` a `</title>`. Na webu se poté zobrazuje v horní liště internetového prohlížeče. Jedná se o jeden z nejdůležitějších *on-page faktorů* na stránce. Je to především kvůli faktu, že se obsah tohoto tagu zobrazuje v SERPu jako popis webu. Tento text je vyhledávači zvýrazněn jako nadpis a zároveň slouží jako hypertextový odkaz na danou stránku. Každá stránka v rámci webu by měla mít unikátní titulek, který bude obsahovat nejdůležitější klíčová slova ze stránky. Není tedy

příliš vhodné, aby `<title>` tag obsahoval pouze název firmy bez další identifikace dané stránky. Podle tohoto titulu se totiž uživatelé často rozhodují, zda v SERPu kliknout právě na náš web. Ideální délka by se měla pohybovat mezi 10–70 znaky. Titulek stránky je považován za nejdůležitější on-page faktor [39].

4.3.3 Meta tag description

Obsah meta tagů (ať už jde o meta tag *description*, nebo *keywords*) se nezobrazuje přímo na stránce v prohlížeči, ale slouží jako pomocná informace vyhledávačům. Poté záleží na konkrétní implementaci algoritmu vyhledávače, zda daná data bude zohledňovat nebo ne. Meta tag *description* obsahuje popis stránky, na které je uveden. Nejčastěji se jedná o stručné shrnutí obsahu s tím, že se v textu používají vybraná klíčová slova. Vyhledávače mohou tento text použít pro snippety ve výsledcích vyhledávání, pokud se jim zdá relevantní a v samotném textu stránek nenaleznou lepší úryvek textu.

V souvislosti s tímto tagem můžeme také mluvit o jeho optimalizaci. Z důvodu, že text v HTML elementu může být použit jako snippet, se nedoporučuje do něj dávat pouze klíčová slova. Délka by měla být „přiměřená“, to znamená, že nemá cenu uměle prodlužovat obsah tohoto tagu. Může se totiž stát, že vyhledávač bude brát v úvahu pouze prvních pár vět a zbytek ignorovat. Pokud je to možné, měl by být meta tag *description* uváděn na každé stránce a být v rámci daného webu unikátní.

4.3.4 Meta tag keywords

Do meta tagu *keywords* se zapisují jednotlivá klíčová slova, která obsahuje text na dané stránce webu. Slova jsou od sebe oddělena čárkou (např. *slovo, klíčová fráze, další kw*). Opět se jedná o informace pro vyhledávací roboty, které mohou být zohledněny při samotném vyhledávání. V současné době však moderní vyhledávače jako je např. Google [25] nebo Seznam [8] tento tag ignorují. Pro optimalizaci webu tak není nutné meta tag *keywords* do kódu stránky uvádět.

4.3.5 Nadpisy h1 – h6

Pro označení nadpisů se ve značkovacím jazyce HTML používají tagy `<h1>` až `<h6>`. Jde o párové tagy a jejich formát je následující: `<h1>Text nadpisu</h1>`. Použití by mělo být především *logické*, tzn. používat je tak, jako se používají například v knihách, časopisech apod. Z toho plyne, že by se na stránce měl nejdříve objevit nadpis první úrovně a další až za ním. O použití klíčových slov v nadpisech platí to samé jako pro ostatní elementy na stránce – střídmost.

4.3.6 Zvýraznění textu

Pro zvýraznění důležitých věcí na stránce (např. klíčových frází, důležitých pojmů apod.) je možné používat tagy k tomu určené. Mluvíme zde o zvýraznění tučným řezem písma (párové tagy `` a ``) a kurzívou (párové tagy `` a `<i>`). Tato zvýraznění jsou důležitá především pro vizuální vzhled stránky – upoutání pozornosti návštěvníka. Pro vyhledávače nejsou takto zvýrazněná slova příliš zvýhodňována oproti ostatním.

4.3.7 Doménové jméno a URL

Pokud optimalizujeme web na málo konkurenční slova, jsou klíčová slova v doméně velmi cenná. Pravděpodobnost, že se na klíčové slovo, které uživatel vyhledává a zároveň je v naší doméně, zobrazí web na předních příčkách vysoce stoupá [36]. U konkurenčních slov však tato výhoda téměř neexistuje. Je to způsobeno především tím, že většina jednoslovných doménových jmen, která obsahují smysluplné slovo, je již zabraná. Vyhledávače proto nemohou dávat přednost doménám, které nemají v názvu klíčové slovo, protože by tak znevýhodňovaly všechny novější weby.

Vyhledávač Google bere v úvahu také stáří domény. Tato veličina se dá zjistit pomocí dvou metod:

- doba od první návštěvy vyhledávacího robota
- doba kdy se na doménu objeví první zpětný odkaz

Oba výše uvedené ukazatele jsou časově často velmi blízko sebe. Navíc podle Matta Cuttse příliš nezáleží, jestli je doména stará 6 nebo 12 měsíců a mnohem důležitější než stáří domény je především samotný obsah [27].

4.4 Mapa stránek

Mapa stránek, anglicky označovaná jako *sitemap*, je soubor všech URL na daném webu. Existuje několik různých formátů a typů těchto map. Např. HTML, XML, RSS/ATOM apod. Sitemapy jsou často rozdělovány na ty, které jsou určeny pro koncové uživatele a na ty pro vyhledávací roboty. Uživatelské mapy stránek slouží především pro jednodušší orientaci.

4.4.1 HTML mapa

HTML mapa je přehled všech podstránek webu zobrazených v jednom seznamu. Nejčastěji se používají HTML seznamy (tagy ``, ``, ``), aby se znázornilo zanoření stránek v rámci kategorií a sekcí webu. Každá položka takového seznamu by také měla být současně hypertextovým odkazem (párový tag `<a>`), aby zároveň sloužila jako navigace. HTML mapu je také vhodné umístit na chybovou stránku webu – *404 Nenalezeno*⁵. Odkaz na mapu by měl být také umístěn na každé podstránce webu například v patičce.

4.4.2 XML Sitemap Protocol

Jedná se o XML soubor, který napomáhá v procházení daného webu roboty vyhledávačů. Původní koncept pochází z roku 2005 a byl vytvořen společností Google pod názvem *Google Sitemap Protocol*. Ostatní vyhledávače později vyjádřily zájem tuto technologii také používat, proto se Google rozhodl protokol přejmenovat na *XML Sitemap Protocol*, aby byl považován za „univerzální“. V současné době je standard tohoto protokolu udržován na webové stránce www.sitemaps.org. Použití *XML Sitemap Protocol* je dobrovolné, a proto vyhledávače nepenalizují weby, které jej nepoužívají. Pro jeho použití je nutné vytvořit soubor `sitemap.xml` a umístit ho do kořenového adresáře webu. Je také dobré umístit direktivu o existenci tohoto souboru do `robots.txt`.

⁵Stavový HTTP kód, který udává, že požadovaná stránka neexistuje.

Struktura

Protokol *XML Sitemap Protocol* se skládá z XML tagů. Veškerá data (včetně URL) musí být nahrazena *escape sekvencemi* podle tabulky 4.1.

Znak	Escape sekvence
Ampersand (&)	&
Apostrof (')	'
Úvozovky (“	"
Menší než (>)	>
Menší než (<)	<

Tabulka 4.1: Escape sekvence pro entity podle XML Sitemap protokolu

Soubor `sitemap.xml` se poté skládá z jednotlivých atributů `log`, `lastmod`, `changefreq` a `priority`.

V atributu `loc` se uvádí URL dané stránky webu. Maximální délka by měla být 2048 znaků a URL musí být dle standardu RFC 3986.

Atribut `lastmod` obsahuje datum, kdy došlo k poslední změně na dané adrese. Tento atribut je volitelný, ale je vhodné jej používat, protože to je způsob, jak dát robotům vědět, že na stránce došlo ke změně. Datum zde může být uvedeno ve dvou tvarech, které jsou popsány v <http://www.w3.org/TR/NOTE-datetime>.

Dalším atributem je `changefreq`, který udává, jak často se daná stránka aktualizuje. Hodnota zde uvedená nemusí být zcela přesná. Vyhledávače web nebudou nijak postihovat, pokud hodnota `changefreq` nekoresponduje se skutečností. Tento atribut může nabývat následujících hodnot:

- `always` – pro stránky, které se mění při každém přístupu
- `hourly` – změna na stránce probíhá každou hodinu
- `daily` – změna na stránce probíhá každý den
- `monthly` – změna na stránce probíhá každý měsíc
- `yearly` – změna na stránce probíhá jednou ročně
- `never` – použití pro stránky v archivu

Je nutné poznamenat, že roboti vyhledávačů mohou procházet stránky častěji, případně méně často, než je uvedeno v atributu `changefreq`. Atribut `priority` může nabývat hodnot z intervalu $< 0, 0; 1, 0 >$, přičemž vyšší priorita udává vyšší pravděpodobnost, že bude stránka navštívena robotem. Jde opět o nepovinný atribut.

Jak může výsledný soubor `sitemap.xml` vypadat je zobrazeno v kódu 4.1.

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9" >
  <url>
    <loc>http://www.example.com/</loc>
    <lastmod>2010-12-24</lastmod>
    <changefreq>monthly</changefreq>
```



```
<priority >0.8</priority >
</url >
<url >
  <loc >
    http://www.example.com/strana?item=12&katgorie=1
  </loc >
  <lastmod >2004-12-23T18:00:15+00:00 </lastmod >
  <changefreq >weekly </changefreq >
</url >
</urlset >
```

Kód 4.1: Příklad souboru sitemap.xml

Pro vytváření XML sitemap je možné použít online generátory, např. na adrese www.xml-sitemaps.com.

4.5 Duplicitní obsah

Jako duplicitní obsah označujeme ten, který se na stránce vyskytuje několikrát. Přesněji řečeno – stejný obsah (především texty) je na více stránkách, které mají různá URL. Vyhledávače takový fakt postihují. Dochází například k rozproštění interních hodnocení vyhledávačů mezi tyto stránky. V rámci webu mohou vznikat duplicity, aniž by o nich správce stránek věděl. Například *www.example.cz*, *example.cz/index.html*, *www.example.cz/index.html* a *example.cz* jsou vyhledávači rozlišovány jako čtyři zcela autonomní stránky a přitom se na všech adresách většinou nachází stejný obsah. Tento problém však např. Google již řeší a nebere ji jako duplicitu. Přesto je vhodné tyto duplicity eliminovat například přesměrováním pouze na jednu stránku.

Další možností vzniku duplicit jsou parametry v URL adrese, případně stránky zobrazené po vyhledávání uživatelem. Nejčastěji jde například o `session id`, jméno kategorie, pomocné proměnné apod. Příklady duplicitních stránek jsou uvedeny v kódu 4.2.

```
www.example.cz/stranka.php?id=5
www.example.cz/stranka.php?id=5&zdroj=seznam
www.example.cz/stranka.php?id=5&sessionId=2341GE57FAF195DE34027A95
```

Kód 4.2: Ukázka duplicitního obsahu v rámci webu

4.5.1 Zamezení duplicitnímu obsahu

Duplicitnímu obsahu je nutné v co největší míře zamezit. Nejlepší je se mu vyhnout již při generování dynamických stránek v redakčním systému. U menších a statických webů se duplicitní obsah většinou nevyskytuje. Jednotlivé techniky, jak se proti duplicitnímu obsahu bránit, jsou popsány níže.

Přesměrování stránek

Přesměrování stránek je vhodné zejména pro statické stránky a také pro případy, kdy chceme řešit problém prefixu `www` (viz výše). Pro tento účel lze použít HTTP stavového kódu *301 Moved Permanently*, který je definován v RFC 2616 [31]. Pokud klient obdrží na svůj HTTP dotaz v odpovědi od serveru kód 301 s novou adresou, bude tuto adresu

následovat. Jde tedy o přesměrování na straně serveru a pro tento způsob přesměrování lze u serveru Apache použít `.htaccess`. Jde o textový soubor, podle kterého se server řídí. Stejně příkazy lze zadat případně také do konfiguračního souboru `httpd.conf`, pokud k tomu máme příslušná práva. Přesměrování pomocí stavového kódu 301 je vhodné použít například pokud vytváříme nový web se zcela odlišnou strukturou kategorií oproti původnímu webu. Poté můžeme přesměrovat původní stránky na nové (v rámci zachování kontextu stránek), abychom nepřišli o body, které dané stránce už vyhledávače přidělily. Pomocí tohoto přesměrování se totiž autoritativní body přenáší. Je nutné dát pozor, aby se pomocí přesměrování nevytvořila nekonečná smyčka – stránka A přesměrovává na stránku B a ta opět na A. Tímto bychom zamezili přístup jednak robotům, ale také uživatelům, protože v obou případech by se stránka samozřejmě ze serveru nestáhla. Příklad přesměrování pomocí `.htaccess` je uveden v kódu 4.3. Se souborem `.htaccess` lze nastavovat i další věci jako přesměrování v případě nenalezené stránky (stavový HTTP kód 404), vytváření zaheslované složky, blokování IP adres pro přístup na web apod. Tyto funkce jsou popsány např. v [6].

```
# presmerovani v pripade HTTP kodu 404 stranka nenalezena
ErrorDocument 404 /chybova-stranka.php

# presmerovani pomoci stavoveho HTTP kodu 301 stranky
# stranka1.php na stranku stranka2.php
Redirect 301 /stranka1.php http://example.cz/stranka2.php

# presmerovani vseh adres bez www prefixu na www
# napr. example.cz na www.example.cz
RewriteEngine On
RewriteCond %{HTTP_HOST} ^example\.cz [nc]
RewriteRule (.*) http://www.example.cz/$1 [R=301,L]
```

Kód 4.3: Přesměrování pomocí `.htaccess`

HTML canonical link element

Proti duplicitám lze také bojovat pomocí HTML elementu `<link rel='canonical' href='kanonicka-stranka'>`. Jde o to, určit autoritativní stránku, která je primárním nositelem informace a tím pádem i privilegována. Matt Cutts na svém blogu [24] popisuje kanonizaci takto: „*Kanonizace je proces vybírání nejlepší URL v případě, že existuje několik možností pro výběr.*“.

Dejme tomu, že máme stránku `www.example.cz/produkt.php`, kterou jsme určili jako primární. Poté máme např. stránky `www.example.cz/vyhledavani.php?title=produkt` a `www.example.cz/katalog.php?strana=156`, které mají stejný obsah jako URL uvedená výše. Do HTML kódu těchto dvou stránek přidáme canonical link element, který se bude odkazovat na primární stránku. HTML element `<link>` se umísťuje do hlavičky webu a má následující podobu: `<link rel='canonical' href='http://www.example.cz/produkt.php'>`. Tímto říkáme vyhledávačům, že na stránce, která je takto označená, se vyskytuje stejný text jako na `www.example.cz/produkt.php` a ta je kanonická, tedy primární.

Kanonické odkazy podporuje v současné době (leden 2011) např. vyhledávač Google, Yahoo a Bing. Naopak český vyhledávač Seznam je zatím nepodporuje.

4.6 Rychlost načítání stránek

V posledních několika týdnech se velmi diskutuje o rychlosti načítání stránek jako jednom z faktorů, který také ovlivňuje pozici v SERPu. Potvrdil to i sám Google na svém blogu [46]. Podle tohoto článku rychlost načítání stránky v současné době ovlivňuje méně než 1 % výsledků. Optimalizovat stránky na rychlost však doporučuje Matt Cutts i v jednom z videí na Google Webmaster Central [26]. Vliv rychlosti načítání stránek v SERPu Google testoval například Geoff Kenyon [37].

V době psaní této práce není známo, že by rychlost načítání stránek byla jedním z faktorů ovlivňující výsledky vyhledávání na českém vyhledávači Seznam.

Kapitola 5

Off-page faktory

V minulé kapitole byly probrány faktory, které má tvůrce stránek pod absolutní kontrolou – sám rozhoduje jaká klíčová slova použije, jaké texty na dané stránky umístí apod. V této kapitole budou popsány faktory, které provozovatelé webových stránek nemohou přímo ovlivňovat. Respektive ne tak snadno, jako on-page faktory. Existují však postupy, jak alespoň částečně usměrňovat *off-page faktory* a využít je tak pro prospěch webu.

5.1 Zpětné odkazy

Vyhledávače se snaží hodnotit jednotlivé weby podle jejich síly a relevance k danému vyhledávacímu dotazu. Problémem však je, jak vypočítávat tuto „sílu stránky“. Jednou z možností, a dnes nejvíce používanou, je zohledňování kvality a kvantity *zpětných odkazů*.

Pojmem „zpětný odkaz“ označujeme hypertextové odkazy, které vedou na danou stránku z jiné stránky. (V literatuře se zpětné odkazy často také označují jako *zpětné linky*, *back-linky* apod.) Tato stránka může být buď v rámci jednoho webu, v tomto případě mluvíme o tzv. interních zpětných odkazech, nebo na zcela jiném webu – tzv. externí zpětné odkazy. Je nutné zdůraznit, že vyhledávače vnímají každou stránku webu jako autonomní, neznají pojem jako je web ve smyslu jak jej vnímají uživatelé, tedy jako množinu stránek umístěných většinou na jedné doméně druhého řádu (např. www.example.cz). Z tohoto faktu tedy plyne, že odkazy bychom měli vytvářet i v rámci webu, tedy odkazovat z jedné stránky na další. Zpětný odkaz můžeme chápat také jako hlas, jenž je udělen stránce, na kterou se odkazuje.

5.1.1 Vlastnosti zpětných odkazů

Zpětné odkazy mají svá specifika. Kvalitativní hledisko zde převyšuje kvantitativní. V praxi to znamená, že pokud na naši stránku odkazuje jeden vysoce autoritativní web, má to větší přínos, než kdyby na nás odkazovalo deset malých webů. Není tedy možné podle absolutního počtu zpětných odkazů vyvozovat jejich význam, pokud neznáme, odkud odkazy pocházejí.

Anchor text

Další významnou vlastností odkazů je tzv. *anchor text*. Tímto pojmem označujeme text, který je uzavřen v párové HTML značce *a* (anchor) a zároveň se zobrazuje jako aktivní („klikatelný“) text na dané stránce v prohlížeči. Anchor text je popsán v kódu [5.1](#).

```
<a href="http://www.example.cz">Anchor text</a>
```

Kód 5.1: Anchor text u hypertextových odkazů

Anchor text by se měl vytvářet tak, aby zahrnoval klíčová slova, na která chceme posilovat odkazovaný web. Samozřejmě musí také korespondovat s okolním obsahem webu. Dále je také nutné, aby ve všech anchor textech nebyla stejná slovní spojení. Mohlo by totiž dojít také k tomu, že web budou vyhledávače za tuto věc penalizovat. Anchor text je pro vyhledávače vysoce důležitý. Svědčí o tom například tzv. *Google bomby*. Jedná se o snahu podpořit odkazovaný web na klíčová slova, která se však přímo na webu vůbec nenacházejí. V České republice je nejznámější Google bomba s frází „senilní ješita“, kdy se na tento dotaz zobrazovaly ve vyhledávání jako první webové stránky Václava Klause. Samozřejmě, že se tento text na stránkách vůbec nevyskytoval. Vše bylo způsobeno pouze tím, že na stránky Václava Klause vedl velký počet odkazů s anchor textem „senilní ješita“. V současné době se Google snaží podobným bombám zabránit. Seznam podobným neduhem netrpí, protože zobrazuje ve výsledcích vyhledávání pouze takové stránky, které mají danou frázi přímo na stránce nebo v URL [39].

Tematická příbuznost odkazovaných stránek

Vyhledávače také zohledňují tematickou relevanci jednotlivých odkazů. Pokud například web, který se věnuje zubní technice, odkazuje na stránku o jaderné fyzice, bude tento odkaz méně hodnotný než ten, který by vedl například ze stránek fyzikální fakulty. Proto je nutné myslet na to, že obsahová příbuznost webů je velmi důležitá.

Link juice

Pojmem *link juice* označujeme „přelévání“ autoritativní síly stránky pomocí odkazů. Daná stránka získá sílu (*juice*) tak, že na ni odkazují jiné stránky. Tuto nakumulovanou sílu může prostřednictvím odkazů předat dalším stránkám. Přitom se *juice* rozloží rovnoměrně na všechny stránky, na které se odkazuje. Pochopení link juice je nutné zejména pro vnitřní prolínání webu. Například z hlavní strany e-shopu je vhodné odkazovat na kategorie a až z kategorií na jednotlivé stránky produktů. Pokud bychom z hlavní strany odkazovali na všechny stránky produktů, rozlijeme všechnu *juice* a sílu dostanou i méně důležité stránky – např. produkty, které jsou již v archivu apod. Nehledě na to, že takováto hlavní strana by obsahovala stovky až tisíce odkazů a mohla by být považována za link farmu.

Atribut rel='nofollow'

Podobně jako u meta tagu robots, lze i odkazům nastavit atribut `nofollow`. Říkáme tak vyhledávačům, aby tento odkaz nenásledovaly. Vyhledávač tak odkazovanou stránku nebude zahrnovat do svého indexu a Google mu ani nepřenesou část PageRanku. Podle nápovědy společnosti Seznam [20] je tímto atributem vhodné označovat takové odkazy, u kterých si nejsme jisti jejich důvěryhodností. Jde typicky o komentáře u blogů, návštěvní knihy, diskuzní fóra apod. Ukázka odkazu s atributem `rel='nofollow'` je zobrazena v kódu 5.2. Další informace lze dohledat v Centrále pro webmastery [19].

```
<a href="http://www.example.cz" rel="nofollow">Anchor text</a>
```

Kód 5.2: Atribut nofollow u hypertextových odkazů

5.1.2 Detekce zpětných odkazů

Během životního cyklu webu by měl mít počet zpětných odkazů rostoucí tendenci. Důležité je však také to, aby nedocházelo k nárazovitému prudkému růstu zpětných odkazů, poté k utlumení apod. Je tedy nutné mít přehled, které weby na náš web odkazují, v jaké je to souvislosti apod. Jde především o to zjistit, kde je potenciál dalšího růstu zpětných odkazů, ale také vědět, zda náš web někdo nepomlouvá, šíří nepravdy apod. Většina dnešních vyhledávačů nabízí možnost zjištění, jaké zpětné odkazy na náš web směřují. Pro tuto detekci se používá operátor `link:` ve vyhledávání. Pokud například do vyhledávacího pole na Seznamu zadáme `link:example.cz`, dostaneme ve výsledcích vyhledávání seznam stránek s odkazem na tuto doménu. S operátorem `link:` je možné uvádět doménu s `www` i bez `www`, druhá možnost však vrací více výsledků [11]. Operátor `link:` v současné době podporují například vyhledávače Seznam, Google a Yahoo!. Každý vyhledávač však dává odlišné výsledky. Je to jednak z toho důvodu, že každý má svůj index a může tedy mít i jiný reálný počet zaindexovaných zpětných odkazů na náš web. Dále například Google zobrazuje pouze ty stránky, které považuje za hodnotné.

Yahoo! Site Explorer

Dobrým nástrojem pro zjištění největšího počtu zpětných odkazů je *Yahoo! Site Explorer* [10], který umí zobrazit jak externí zpětné odkazy, tak odkazy v rámci webu. Výsledek z analýzy zpětných odkazů pomocí nástroje Yahoo! Site Explorer je zobrazen na obrázku 5.1.

Open Site Explorer

Dalším nástrojem pro zjištění zpětných odkazů, které vedou na daný web, je *Open Site Explorer* [13] od společnosti SEOmoz. Nástroj je zdarma, je však omezený počtem denních reportů na tři a počtem zobrazených zpětných odkazů na 200. Po registraci zdarma však nástrojem můžeme provádět neomezený počet analýz a je nám zobrazeno až 1000 odkazů. V obou případech můžeme navíc zpětné odkazy řadit podle anchor textu, autority odkazované domény, nebo si nechat vypsát tzv. plnou metriku zpětných odkazů, která nám zobrazí, kolik procent odkazů jsou nofollow a follow a také kolik procent odkazů je interních a kolik externích. Popis funkcí nástroje Open Site Explorer nalezneme například v článku [29].

5.2 Odkazový profil

Každá stránka webu si již od svého vzniku buduje tzv. *odkazový profil*. Ten se skládá ze všech odkazů, které na danou stránku vedou. Je proto velmi důležité dávat pozor na to, kam zpětné odkazy umísťujeme. Může se totiž stát, že si svůj odkazový profil poškodíme a jen těžko jej budeme dostávat do původního stavu. Jako nevhodné odkazy se berou například ty z *link farm*¹, *warez* webů apod. Je tedy pochopitelné, že vyhledávače odkazový profil důsledně kontrolují a v případě nevhodných odkazů mířících na daný web jej mohou také penalizovat.

¹Weby, které obsahují velké množství zpětných odkazů a snaží se tak předat svoji sílu odkazovaným webům a obelhat tak algoritmy vyhledávačů.

The screenshot shows the Yahoo! Site Explorer interface. At the top, the URL 'www.fit.vutbr.cz' is entered in the search bar, with an 'Explore URL' button to its right. Below the search bar, the 'SITE EXPLORER' header is visible. On the left side, there is a sidebar with various site management options: 'Add to MySites', 'My Sites', 'Submit Your Site', 'Preferences', 'Blog', 'Badge', 'Web Service API', and 'Feedback'. The main area is titled 'Results' and shows a list of 17 search results. Each result includes a number, a title, a URL, and an 'Explore' button. The results are as follows:

Rank	Title	URL	Cache Size
1.	Faculty of Information Technology, Brno University of Technology	text/html www.fit.vutbr.cz	7k - cache
2.	Faculty of Information Technology, Brno University of Technology	text/html www.fit.vutbr.cz/index.php.en	7k - cache
3.	Doc. Ing. Tomáš Vojnar, Ph.D.	text/html www.fit.vutbr.cz/~vojnar	7k - cache
4.	ICES 2008	text/html www.fit.vutbr.cz/events/ices2008	9k - cache
5.	Ing. Pavel Ti'novský	text/html www.fit.vutbr.cz/~tisnovpa	5k - cache
6.	ECBS 2004	text/html www.fit.vutbr.cz/events/ECBS2004	7k - cache
7.	Information Systems Modelling 2002	text/html www.fit.vutbr.cz/events/ism	8k - cache
8.	INFINITY 2008	text/html www.fit.vutbr.cz/events/infinity08	5k - cache
9.	Faculty of Information Technology, Brno University of Technology	text/html www.fit.vutbr.cz.en	7k - cache
10.	Ing. Rudolf ějka	text/html www.fit.vutbr.cz/~cejkar	6k - cache
11.	MEMICS 2006	text/html www.fit.vutbr.cz/events/memics06	9k - cache
12.	Bible kralick??	text/html www.fit.vutbr.cz/~michal/kr	32k
13.	Home Speech Processing Group	text/html www.fit.vutbr.cz/speech	11k - cache
14.	Doc. Ing. Luká' Sekanina, Ph.D.	text/html www.fit.vutbr.cz/~sekanina	7k - cache
15.	Doc. RNDr. Pavel Smr'ka, Ph.D.	text/html www.fit.vutbr.cz/~smrz	6k - cache
16.	Prof. Ing. Miroslav @véda, CSc.	text/html www.fit.vutbr.cz/~sveda	7k - cache
17.	DDECS 2002	text/html www.fit.vutbr.cz/events/ddec02	9k - cache

Obrázek 5.1: Analýza zpětných odkazů webu FIT VUT pomocí Yahoo! Site Explorer

5.3 Linkbuilding – získávání zpětných odkazů

Vytváření zpětných odkazů, tzv. *linkbuilding*, je jednou z nejdůležitějších částí optimalizace pro vyhledávače. Protože často nemůžeme zpětné odkazy na náš web vytvářet přímo (za předpokladu, že nemáme rozsáhlou síť webů), musíme hledat jiné cesty, jak zpětné odkazy získat. Nejlepším způsobem je vytvářet kvalitní obsah. Při vytváření zpětných odkazů je potřeba myslet na to, abychom odkazy rovnoměrně rozložili na celý web. Není vhodné odkazovat pouze na hlavní stranu webu se všemi klíčovými slovy v anchor textu.

5.3.1 Linkbaiting

Pojmem *linkbaiting* označujeme činnost, která vede k získávání zpětných odkazů formou „lákání“ (český výraz pro linkbaiting se zatím neustálil). Nejčastěji jde o tvorbu zajímavého obsahu jako jsou například on-line hry, zajímavá videa, aplikace apod. Pokud se nám podaří takový obsah vytvořit, ostatní stránky na něj budou odkazovat samy a informace o obsahu se budou rychle šířit po internetu. Mluvíme zde o tzv. *virálním obsahu* a k jeho šíření dnes přispívají především sociální sítě jako jsou Facebook, Twitter, Youtube apod.

5.3.2 Záložkování

Pro vytváření zpětných odkazů lze také využít služeb tzv. *záložkovacích webů*. Jedná se o webové stránky, kam je možné vkládat URL naší stránky společně s nadpisem a krátkým textem. Záložkování je primárně vytvořeno pro ukládání zajímavých článků a u každé záložky mohou návštěvníci hlasovat, tímto se liší od registrací do katalogů (téma katalogů bylo diskutováno v sekci 5.3.4). Záložky s nejvíce hlasy se poté zobrazují na úvodní straně a tím pádem i předávají větší *link juice* odkazované stránce. Platí zde tedy, že čím je obsah zajímavější a čím lépe vytvoříme nadpis a popisek, tím více návštěvníků můžeme získat. Mezi nejznámější české záložkovací služby patří:

- www.linkuj.cz
- www.jagg.cz
- www.topclanky.cz

5.3.3 Sociální sítě

V současné době se velmi rozvíjí tzv. *sociální sítě*. Na těchto webech se soustřeďuje velký počet lidí, kteří jsou navzájem propojeni. K šíření obsahu zde tedy dochází velmi interaktivním způsobem. Proto je vhodné zapojit do tvorby odkazů i sociální sítě typu Facebook, Twitter, LinkedIn, MySpace, Youtube a další.

5.3.4 Registrace do katalogů

Registrací do vhodných katalogů lze také získat cenné zpětné odkazy. Je však důležité volit správné katalogy, abychom si neničili odkazový profil odkazy z link farem. Téma katalogů bylo diskutováno v kapitole 3.

5.3.5 Další možnosti

Každý web je svým způsobem specifický, proto je vždy důležité hledat nové možnosti zisku zpětných odkazů. Jako inspirace může sloužit článek „How to Build Links Fast: 101 Tips & Strategies“ [52].

Kapitola 6

Testování on-page a off-page faktorů

Tato práce si klade za cíl otestovat jednotlivé on-page a off-page faktory a jejich vliv na pozice ve výsledcích vyhledávání. Vyhodnocení probíhá na základě měření pozic ve vyhledávacích Seznam a Google. Ty mají v České republice majoritní podíl mezi vyhledávači, a proto není nutné uvažovat ostatní vyhledávače, jejichž celkový podíl je v řádech jednotek procent.

6.1 Metodika testování

Pro možnosti testování bylo vytvořeno deset jednoduchých webových prezentací. Měření pozic ve vyhledávacích probíhá vždy před úpravou on-page nebo off-page faktoru a po přeindexování jednotlivých stránek. Naměřené pozice se poté zaznamenávají.

Jednotlivé změny on-page faktorů a posilování off-page faktorů probíhá vždy na stejné množině pěti webů. V rámci této množiny je u všech webů upraven jeden on-page faktor, případně posílen off-page faktor. Po každé změně on-page nebo off-page faktoru se pozice opět změní a zaznamenají. V rámci množiny, u které byly testovány jednotlivé faktory, se naměřené pozice zprůměrují. Tím se určí, zda měl daný faktor kladný nebo záporný vliv na pozice ve výsledcích vyhledávání. Na základě těchto hodnot se odvodí ty faktory, které mají největší vliv na posilování pozic ve vyhledávacích. Dále bude vytvořen jeden modelový web, kde budou zahrnuty nejdůležitější on-page faktory.

6.1.1 Detekce přeindexování stránek

On-page faktory

Aby bylo možné zjistit, kdy došlo k novému přeindexování stránky po změně on-page faktorů, je po každé změně upraven také titulek stránky (tag `<title>`). Protože se titulek objevuje ve výsledcích vyhledávání, dá se jednoduše odvodit, kdy už je v indexu vyhledávače stránka nová a kdy stará. Všechny weby měly na počátku testování titulek „Titulek ’!’“.

Off-page faktory

Při testování off-page faktorů je nutné postupovat jinak. Do testování byly zahrnuty zpětné odkazy z jiných webových stránek, které vedou na stránky testované. To, že byl odkaz již

vyhledávači zaindexován bylo testováno pomocí operátoru `site:`. Jako modelový příklad uvedme zpětný odkaz ze záložkovací služby Linkuj.cz. Zpětný odkaz, který zde byl vytvořen, měl anchor text „Informace o Krétě“. Pokud tedy chceme detekovat, zda byl již odkaz vyhledávači zaindexován, zadáme do vyhledávacího pole `site:linkuj.cz ‘‘Informace o Krétě’’`. Následné výsledky v SERPu je nutné manuálně projít, protože se může stát, že v indexu je již nějaký web se stejným anchor textem. Předchozí postup lze uplatnit jak ve vyhledávači Seznam, tak Google.

6.2 Měření pozic

Pro jednodušší měření pozic ve vyhledávačích byl vytvořen PHP skript, který podle zadané klíčové fráze zjistí pozice všech testovacích webů. Tento skript využívá open source knihovnu Simple HTML DOM [23]. Nejdříve se do proměnné načte stránka s výsledky vyhledávání, ve které se následně prochází DOM struktura a testuje se, zda se na stránce vyskytuje některá z URL testovaných webů. Aby se omezil počet HTTP dotazů, je pomocí URL parametrů nastaveno u vyhledávače Seznam i Google 20 výpisů na stránku (standardní počet výsledků je 10). U Seznamu je tento parametr nazván `count` a u Google `num`. Ve vyhledávači Seznam se kontrolují výsledky až do 1000. pozice, u Google do 600. Tyto hodnoty byly vybrány na základně testování skriptu. Seznam vrací výsledky pouze do 1000. pozice, i když jich má v indexu i více. Pokud by se i u Google vyhledávalo do 1000. pozice, začíná Google po několika měřeních vracet místo výsledků pouze HTTP chybu `503 - Service unavailable`¹.

Protože se na jedné stránce zobrazuje pouze 20 výsledků, je nutné vyřešit také přechod na další stránky. Ve vyhledávači Seznam je nutné nastavit v URL parametr `from` a v Google parametr `start` na hodnotu, od které se mají zobrazovat následující výsledky. V kódu 6.1 jsou ukázány URL pro vyhledávače Seznam a Google, které zobrazí výsledky vyhledávání od pozice 21 do pozice 40 na klíčovou frázi „FIT VUT“.

```
// vyhledavac Seznam
http://search.seznam.cz/searchScreen?q=FIT+VUT&count=20&from=21

// vyhledavac Google
http://www.google.cz/search?hl=cs&num=20&source=hp&q=FIT+VUT
&as_qdr=all&start=20
```

Kód 6.1: URL pro zobrazení druhé strany výsledků vyhledávání v Seznamu a Google

6.2.1 Specifika měření pozic u vyhledávače Google

Vyhledávač Google je po technické stránce vyspělejší než Seznam. Nabízí mimo jiné také *personalizované vyhledávání*; to znamená, že každému uživateli mohou být na stejný dotaz nabídnuty jiné výsledky vyhledávání. Personalizace vychází jednak z vyhledávací historie, která se uchovává 180 dní, dále také z historie navštívených stránek. Tato data se uchovávají u uživatele v souborech *cookies*. Další informace o personalizovaném vyhledávání na Google lze nalézt v [49] a [50].

Z výše zmíněných důvodů je zřejmé, že je nutné, aby se měření pozic u vyhledávače Google vždy provádělo ve stejném prohlížeči se smazanými cookies a webovou historií. Je

¹Stavový HTTP kód, který udává, že server je přetížený nebo na něm právě probíhá údržba (podle RFC 2616).

důležité, aby měření probíhalo ze stejné IP adresy, protože Google i podle tohoto faktoru upravuje výsledky vyhledávání.

6.3 Testované webové stránky

Všechny testované webové stránky jsou zaměřeny na jedno téma a hustota klíčových slov je na všech cca stejná. Důležité je také to, že každá stránka má unikátní texty. Pokud by se použily u všech stránek texty stejné, mohlo by dojít k tomu, že je vyhledávače vyhodnotí jako duplicitní a vůbec nepřidají do svých indexů. Pro testování bylo nutné vybrat téma webových stránek, u kterých je možné použít klíčová slova, která mají vyšší konkurenci. Jinak by se mohlo stát, že se všechny stránky umístí na předních pozicích a testování jednotlivých faktorů by se tak znesnadnilo.

Z tohoto důvodu byl jako hlavní téma vybrán řecký ostrov Kréta. O Krétě existuje velké množství webů a je zde také potenciál pro výběr klíčových frází. Pro testovací účely bylo vybráno deset klíčových frází. Několik z nich je vysoce konkurenčních, jiné jsou konkurenční méně. Výběr probíhal za pomoci nástroje návrhu klíčových slov ve webové aplikaci Google AdWords [4]. Vybrané klíčové fráze jsou následující:

- doprava na Krétě
- tržiště v Chanii
- krétské tržiště
- krétská letiště
- krétské jídlo
- krétské autobusy
- taxi na Krétě
- krétská dálnice
- letiště v Chanii
- krétská kuchyně

6.3.1 Struktura stránek

Všechny webové stránky používají stejnou HTML a CSS šablonu, aby se předešlo možnému zkreslení výsledků v důsledku jinak uspořádaného HTML kódu apod. Kvůli rychlejší „přeindexaci“ stránek vyhledávači byla zvolena taková struktura, kde každý web má pouze jednu stránku. Na ní se nachází kompletní texty a v horní levé části je pro snazší navigaci návštěvníka umístěno kontextové menu. Pokud se na webu upraví on-page faktor a je potřebné stránku znovu posoudit vyhledávači, stačí, aby robot navštívil právě tuto jednu stránku. Takovéto indexování zkracuje čas pro následné měření pozic, kdy robot nemusí procházet x webových stránek.

6.3.2 Webhostingové řešení

Weby jsou umístěné na freehostingu společnosti Internet Centrum (*www.ic.cz*) na doménách třetího řádu. Z toho důvodu, že hosting i domény jsou zdarma, je podle podmínek nutné mít umístěn na každé stránce webu v její horní části pruh s několika reklamami ve formě hypertextových odkazů. Přítomnost těchto odkazů může mít jistý vliv na hodnocení stránek vyhledávači, ale protože se stejná reklama nachází na všech testovacích stránkách, nebude ovlivňování výsledků reklamou uvažováno. Domény testovacích webů jsou následující:

- www.kreta.hu.cz
- www.kreta.ezin.cz
- www.kreta.yc.cz
- www.kreta.kx.cz
- www.kreta.byl.cz
- www.kreta.own.cz
- www.kreta.chytry.cz
- www.kreta.webovka.eu
- www.kreta.vyjimecny.cz
- www.kreta.tym.cz

6.4 On-page faktory

V této práci byly testovány následující on-page faktory:

- zvýrazněné písmo pomocí HTML tagu ``
- titulek stránky (HTML tag `<title>`)
- nadpisy (HTML tagy `<h1>` až `<h6>`)

Další on-page faktory — jako je například meta tag description — testovány nebyly, protože je zřejmé, jak fungují. Záměrně také nebyla testována například hustota klíčových slov. Při vysoké hustotě bychom totiž mohli narazit na problém keyword stuffingu („tapetování klíčovými slovy“) [14] a je možné, že by vyhledávače web penalizovali, případně ho vyřadili z indexu. Tento fakt by poté znesnadnil další testování.

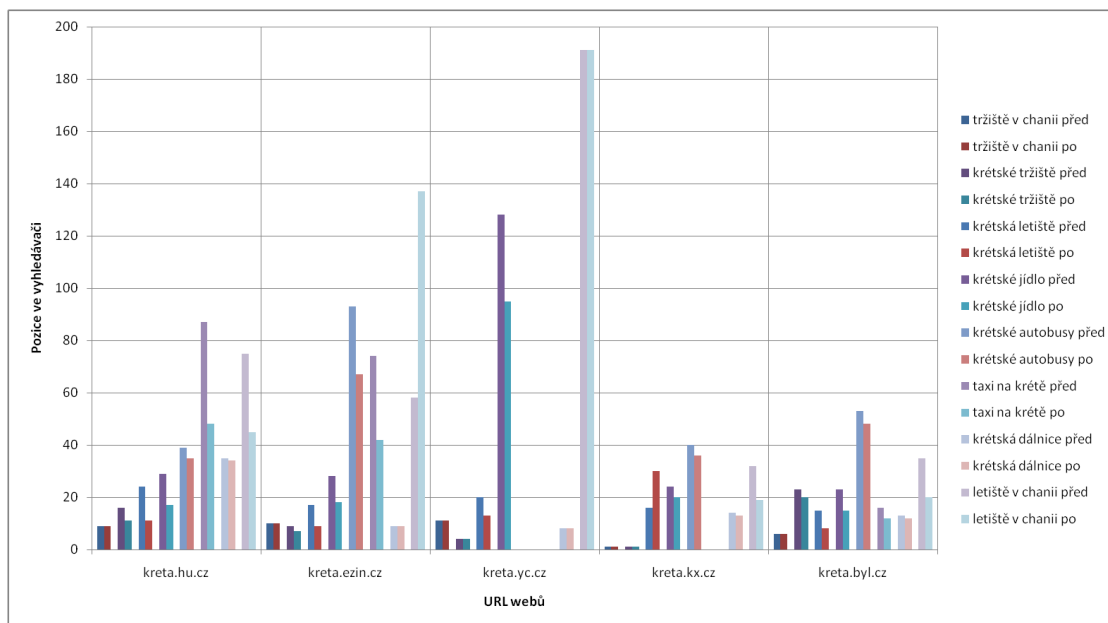
Během testování on-page faktorů nebyl vytvořen ani na jeden web žádný zpětný odkaz. Je to především z toho důvodu, aby nemohlo dojít k nekontrolovatelnému nárůstu odkazů na testované weby, pokud by se na ně autoři jiných stránek rozhodli odkazovat. Webhosting, kde jsou testované weby umístěny, ale všechny stránky zařazuje do svého interního katalogu. Byla zde tedy možnost, že i přesto vznikne nějaký nechtěný zpětný odkaz. Proto bylo vždy před samotným měřením pozic kontrolováno pomocí nástroje *Yahoo! Site Explorer* 5.1.2, zda takové odkazy nevznikly.

6.4.1 Titulek stránky – HTML tag <title>

Titulek stránky je jeden z nejdůležitějších on-page faktorů, proto byl testován jako první. Do titulku je důležité vložit klíčová slova, která se nacházejí v textu dané stránky. Z tohoto důvodu byl pro celou množinu stránek zvolen titulek „== Informace o Krétě – letiště, doprava, jídlo“. Dvě rovnítka na začátku titulku jsou zde proto, aby mohla být při dalších změnách on-page faktorů detekována nová indexace ze strany vyhledávačů, jak to bylo popsáno v sekci 6.1.1. Tento titulek se při následných úpravách webu měnil pouze v prvních dvou znacích (v rovnítkách), aby byla zachována klíčová slova.

Seznam

Průměrný posun ve výsledcích vyhledávání po změně titulku byl ve vyhledávači Seznam +5,21. To znamená, že weby se na testované klíčové fráze posunuly v průměru o tuto hodnotu vpřed, jak je vidět z obrázku 6.1. V tomto měření nebyly brány v úvahu klíčové fráze „doprava na Krétě“ a „krétská kuchyně“, protože se ani jedna z nich vůbec neumístila v SERPu. Z tabulky A.1 je také patrné, že u webu *kreta.yc.cz* se neumísťovaly ve vyhledávání i klíčové fráze „krétské autobusy“ a „taxi na Krétě“. Ke stejnému problému došlo i u webu *kreta.kx.cz* v případě klíčové fráze „taxi na Krétě“. Z toho důvodu nebyly žádné z těchto klíčových frází započítávány do výsledného posunu pozic.

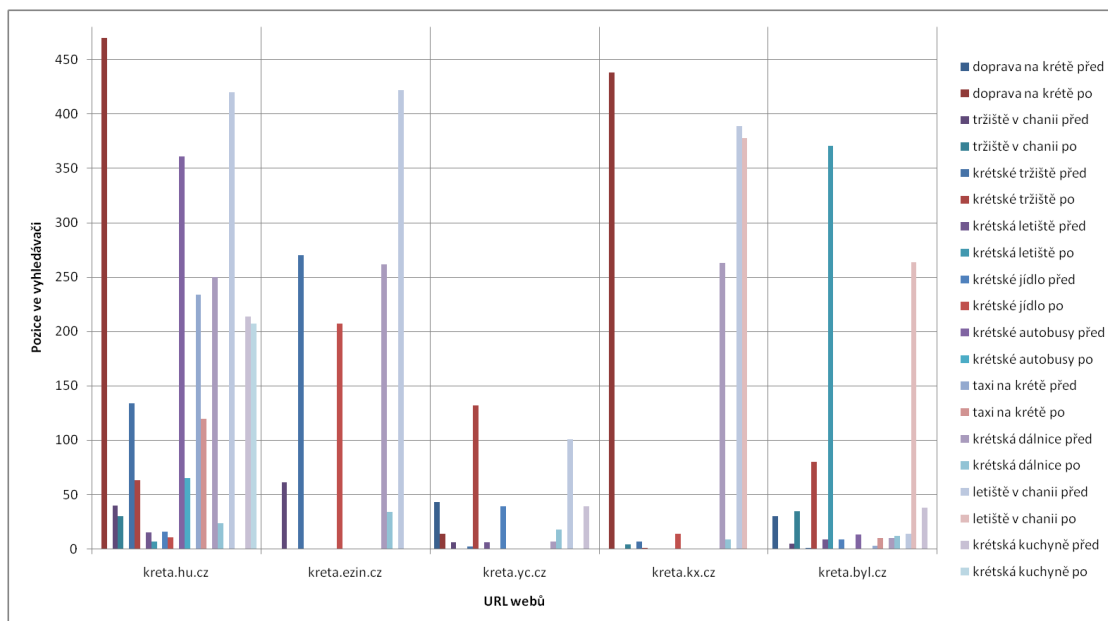


Obrázek 6.1: Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po úpravách titulku

Google

Průměrný posun pozic ve vyhledávači Google v závislosti na úpravě <title> tagu je +50,29. Zde je však nutné zohlednit fakt, že ve 29 případech (viz tabulka A.2) se jednotlivé klíčové fráze nezapočítávaly do celkového průměru. Jak je patrné z obrázku 6.2, jde buď o případy, kdy se při měření pozic před úpravou titulku web ve vyhledávání neumístil, nebo se před úpravou sice umístil, ale po změně z vyhledávání vypadl. Aby nebyly celkové posuny webů

příliš zkresleny, jsou tyto případy z výpočtů vynechány. Je ale nutné na zmíněný fakt pamatovat.



Obrázek 6.2: Pozice webů na dané klíčové fráze ve vyhledávači Google před a po úpravách titulku

Zhodnocení

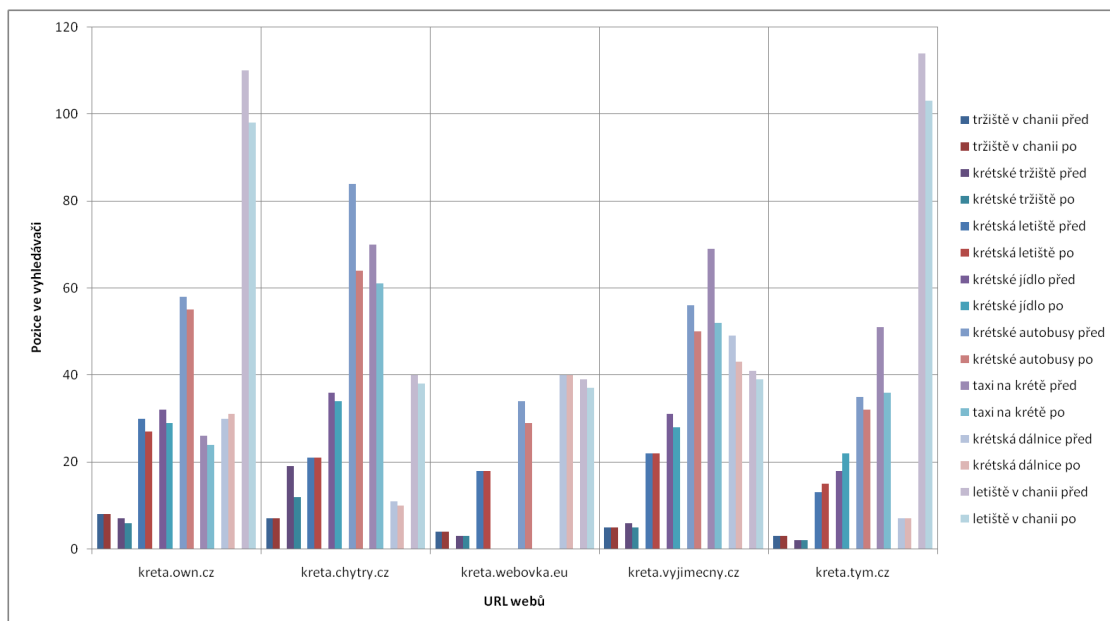
Jak je z naměřených pozic vidět, titulek ovlivňuje pozice u obou vyhledávačů, co se posunu na vyšší příčky týče. Během testování bylo také zjištěno, že pokud Google nepovažuje uvedený titulek za relevantní obsahu stránky, dynamicky ho v SERPu upraví tak, aby dával uživatelům větší smysl. Například u testovaného webu, který měl titulek „Titulek ’!’“, Google zobrazoval na některé klíčové fráze v SERPu text „Kréta“. O tomto faktu lze nalézt další informace např. v [41]. U Seznamu nebylo takové chování zjištěno.

6.4.2 Zvýrazněné písmo – HTML tag

Zvýraznění klíčových slov pomocí tagu bylo paralelně testováno s titulkem stránky na jiné množině webů. Takto byly zvýrazněny všechny klíčové fráze, na které se weby optimalizovaly a u kterých byly také měřeny jednotlivé pozice ve vyhledávačích.

Seznam

Jak je patrné z obrázku 6.3, růst pozic není tak výrazný jako u titulku stránky. Což značí, že zvýraznění klíčových slov pomocí tučného písma není pro vyhledávače tolik důležité. U testovaných webů, kde byly klíčové fráze zvýrazněny tagem , byl posun u vyhledávače Seznam +3,28. Poznamenejme, že z daného měření byly opět — jako u titulku stránky — vynechány klíčové fráze „doprava na Krétě“ a „krétská kuchyně“, které se ve výsledcích vyhledávání neumísťovaly. Dále pak byly vynechány fráze „krétské jídlo“ a „taxi na Krétě“ u webu *kreta.webovka.eu* a to ze stejných důvodů, jak je vidět v tabulce A.3.



Obrázek 6.3: Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po zvýraznění klíčových slov pomocí tagu **strong**

Google

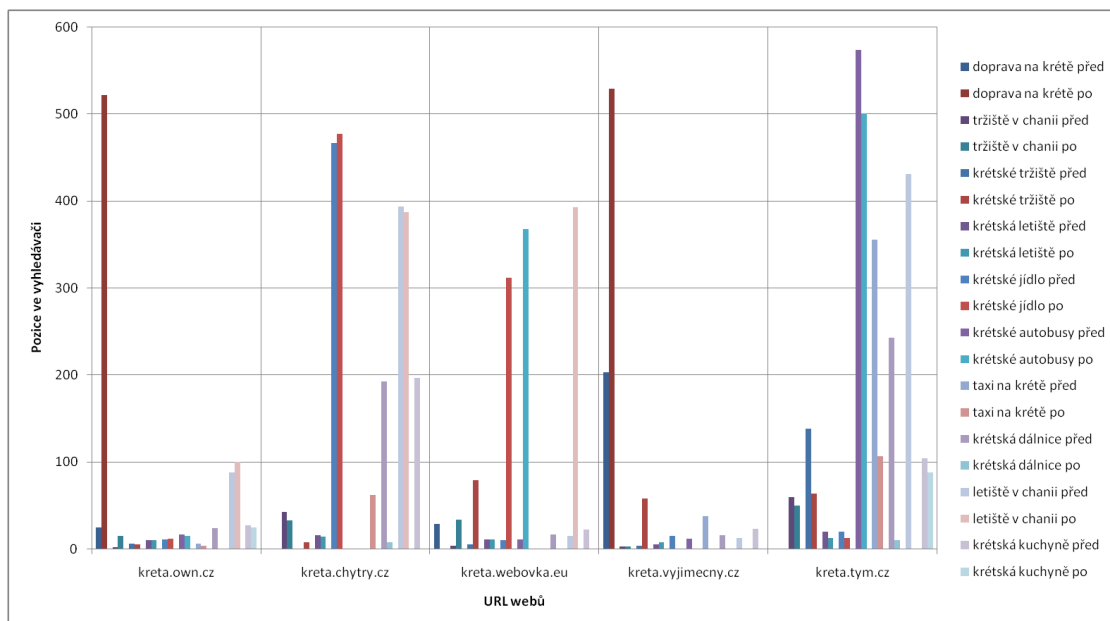
U vyhledávače Google je situace poněkud odlišná. Z naměřených dat vyplývá, že po označení klíčových frází pomocí tagu **** poklesla pozice průměrně o $-44,17$. Z výše zmíněného však nelze jednoznačně soudit, že zvýrazňování klíčových frází má nutně špatný vliv na umístění v SERPu u vyhledávače Google. Data jsou zkruslena tím, že ne všechny klíčové fráze byly po úpravě tohoto on-page faktoru ve výsledcích nalezeny. V tomto případě se jedná o 18 výskytů, jak je vidět na obrázku 6.4 a v tabulce A.4. Dále také tím, že u několika klíčových frází — zejména „doprava na Krétě“ — byl pokles v řádech stovek pozic: u webu *kreta.own.cz* -497 pozic a u *kreta.vyjimecny.cz* -326 pozic. Výsledek nejvíce ovlivňuje web *kreta.webovka.eu*, který má průměrný pokles pozic $-190,17$. Naopak druhým směrem — největším růstem pozic — se projevil web *kreta.tym.cz*, který měl průměrný růst pozic $+83,63$.

Zhodnocení

Jak plyne z měření, vliv zvýraznění klíčových frází pomocí tagu **** není příliš velký v porovnání s titulkem. Pokud navíc zprůměrujeme průměrné posuny webů na klíčové fráze z vyhledávačů Seznam a Google, dostáváme se na $-20,45$ pozic. V tomto duchu se také vyjádřilo i 72 SEO expertů při průzkumu mínění, který byl pořádán v roce 2009 společností SEOMoz [15]. Označení klíčových slov pomocí tagů **** nebo **** má podle tohoto průzkumu *minimální význam*.

6.4.3 Nadpisy – HTML tagy **<h1>** až **<h6>**

Nadpisy jsou logickou částí struktury webových stránek. Jak bylo uvedeno v sekci 4.3.5, klíčová slova v nadpisech by se měla používat především s rozumem a ne s cílem oklamat vyhledávače. Na testovaných webech byly jako nadpisy označeny logické části textů, které



Obrázek 6.4: Pozice webů na dané klíčové fráze ve vyhledávači Google před a po zvýraznění klíčových slov pomocí tagu **strong**

uvozují jednotlivé odstavce. Pokud tedy například odstavec na testované webové stránce pojednává o dopravě, v nadpisu je uveden text „Doprava na Krétě“, „Jak vypadá doprava na Krétě“ apod. Použity byly pouze nadpisy první a druhé úrovně, tedy tagy `<h1>` a `<h2>`. K použití nadpisů dalších úrovní nebyl vzhledem k povaze webových stránek důvod.

Seznam

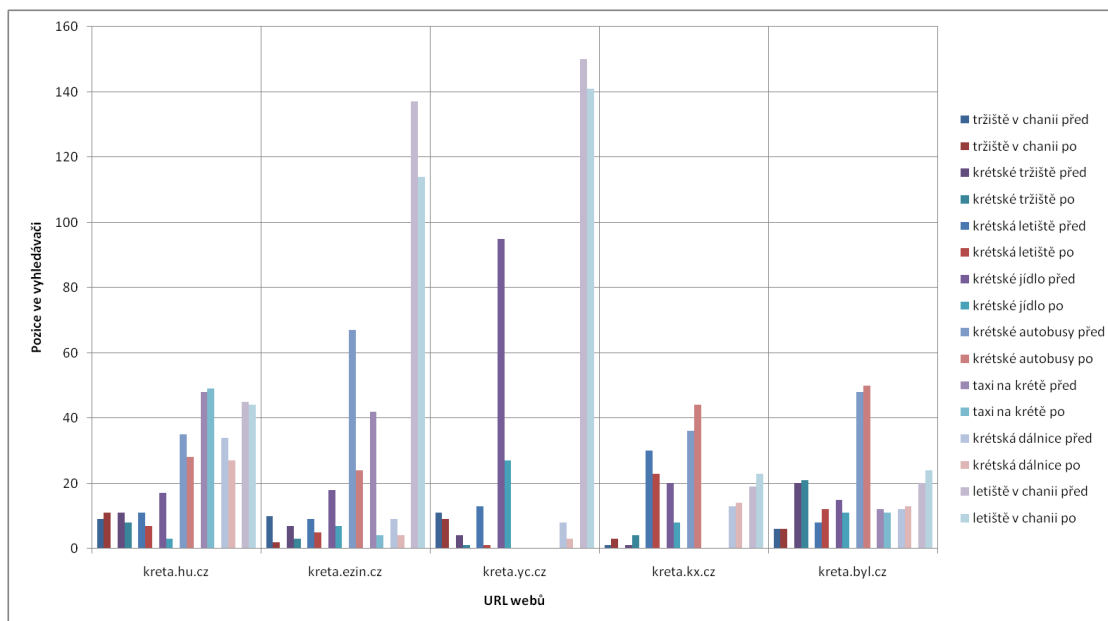
Při vyhodnocování pozic v závislosti na HTML nadpisech nebyly uvažovány klíčové fráze „doprava na Krétě“ a „krétská kuchyně“ (viz tabulka A.5). Opět z důvodu, že se na dané fráze weby před nastavením nadpisů ve vyhledávání na Seznamu neumísťovaly. Dále jsou z hodnocení vyjmuty fráze „taxi na Krétě“ u webů *kreta.yc.cz* a *kreta.kx.cz*, dále fráze „krétské autobusy“ u webu *kreta.yc.cz* ze stejného důvodu. Jednotlivé posuny pozic klíčových frází lze vysledovat z obrázku 6.5. Průměrný posun byl +7,38.

Google

U Google je vyhodnocení pozic obtížnější. Ve 31 případech — což je více než polovina — se weby ve vyhledávání neumísťovaly, a proto jsou vyřazeny z hodnocení (viz tabulka A.6). Pokud bychom však počítali se všemi ostatními změnami pozic, průměrný posun je +21,61. Z obrázku 6.6 je ale patrné, že u webu *kreta.hu.cz* nastal daný problém pouze u jedné klíčové fráze – „letišť v Chanii“. Uvažujme tedy pouze tento web. Poté je průměrný posun pozic ve vyhledávání +81,89. Výsledek nejvíce ovlivňuje klíčová fráze „doprava na Krétě“, u které byl posun +456 pozic.

Zhodnocení

Z naměřených výsledků vyplývá, že označení klíčových frází pomocí HTML nadpisů `<h1>` a `<h2>` má pozitivní vliv na pozice v SERPu. Průměrný posun pozic ve vyhledávacích



Obrázek 6.5: Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po zvýraznění klíčových slov pomocí nadpisů

Seznam a Google je +14,5.

6.4.4 Nejdůležitější on-page faktory

Z testovaných on-page faktorů je patrné, že největší vliv na pozice ve vyhledávání má titulek stránky – tag `<title>`. Po úpravě tohoto faktoru byl růst průměrně +27,75 pozic.

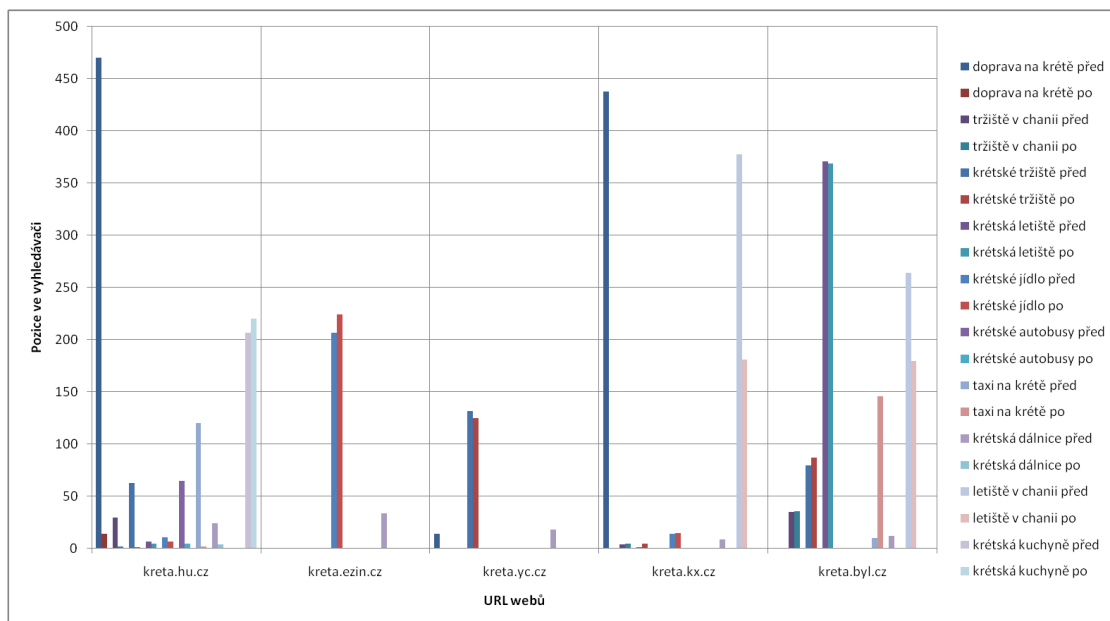
Druhým nejdůležitějším on-page faktorem je zvýraznění klíčových frází pomocí nadpisů – tagy `<h1>` a `<h2>`. Průměrný růst byl +14,5 pozic.

U zvýraznění klíčových frází pomocí tagu `` byl naměřen naopak pád pozic a to průměrně o –40,89 pozic. Důvody propadu byly diskutovány v sekci 6.4.2.

6.5 Off-page faktory

Jak vyplývá z kapitoly 5, mezi off-page faktory patří zejména zpětné odkazy. Není výjimkou, že tvorba a získávání zpětných odkazů je pro celou řadu webů problémem. To samé platí i pro testované weby. Tento fakt také stěžuje skutečnost, že je nutné sehnat stejné zpětné odkazy pro pět webů, aby bylo možné změřit jejich vliv na pozice ve výsledcích vyhledávání. Proto byly pro testovací účely vybrány takové techniky linkbuildingu, které je snadné ovlivnit samotným autorem zpětných odkazů. Jde zejména o hypertextové odkazy ze záložkovacích služeb (viz sekce 5.3.2). Dále byly také otestovány zpětné odkazy, které vedly z jednoho testovaného webu na druhý.

Nastává ale další problém, který je nutné řešit. Poté co se vytvoří zpětný odkaz na testovaný web, je možnost, že se nějakému uživateli daný web zalíbí natolik, že vytvoří další zpětný odkaz. Tento fakt by samozřejmě ovlivnil výsledky, protože platí, že čím více odkazů na daný web vede, tím lépe jej vyhledávače hodnotí (v této chvíli neuvažujeme kvalitu jednotlivých zpětných odkazů, která je ovšem důležitější než kvantita). Proto byl



Obrázek 6.6: Pozice webů na dané klíčové fráze ve vyhledávači Google před a po zvýraznění klíčových slov pomocí nadpisů

během testování kontrolován počet zpětných odkazů na jednotlivé weby pomocí nástroje *Yahoo! Site Explorer* 5.1.2.

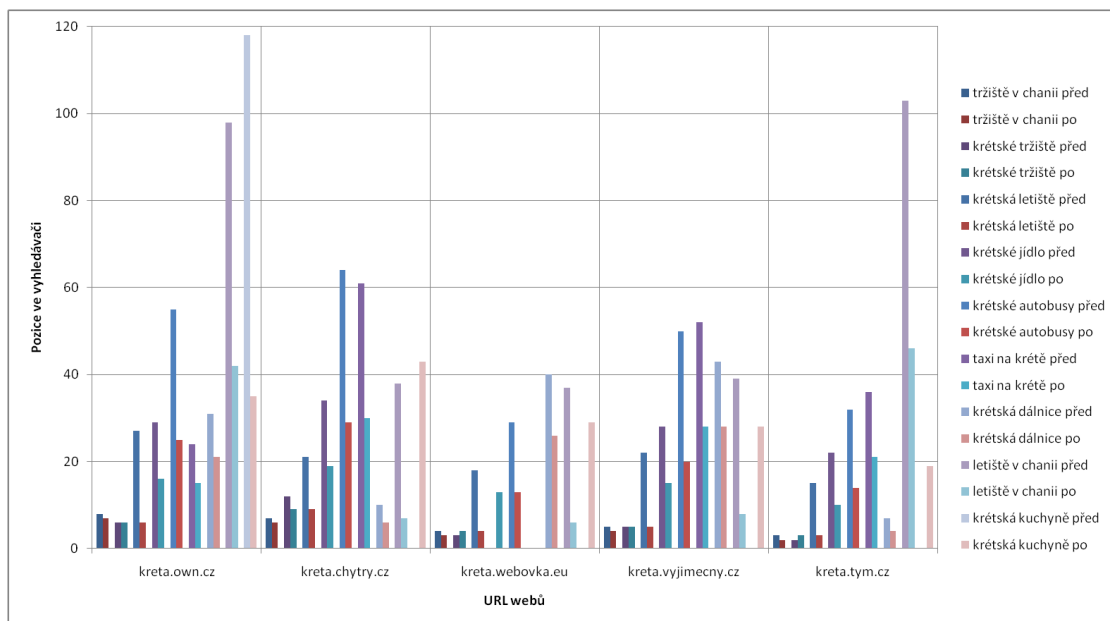
6.5.1 Záložkovací služba Linkuj.cz

Web *www.linkuj.cz* funguje na jednoduchém principu – registrovaní uživatelé mohou do systému vkládat odkazy na webové stránky, které považují za zajímavé. Ostatní návštěvníci webu mohou pro tyto odkazy hlasovat. Vložené stránky, které mají nejvíce hlasů se následně zobrazují na hlavní straně webu Linkuj.cz.

Seznam

Do záložkovací služby Linkuj.cz byla vkládána stejná množina webů, u kterých byl testován vliv pozic vzhledem k tagu ****. Proto i v tomto případě jsou z výsledného hodnocení vyjmuty klíčové fráze „doprava na Krětě“ a „krétská kuchyně“, které se ve výsledcích vyhledávání neumísťovaly a to až na jednu výjimku – web *kreta.own.cz* a frázi „krétská kuchyně“, která se nyní již v SERPu umístila na 35. pozici, což je posun o +83 pozic oproti minulému stavu. Vyřazeny jsou také fráze „krétské jídlo“ a „taxi na Krětě“ u webu *kreta.webovka.eu* a to z obdobných důvodů (viz tabulka A.7). Před měřením pozic nebyly zjištěny žádné další zpětné odkazy, které by vedly na testované weby. Není tedy třeba při vyhodnocování počítat i s tímto faktorem.

Jak je patrné z obrázku 6.7, u téměř všech testovaných frází došlo ve vyhledávači Seznam k posunu na vyšší pozice. Poklesly pouze pozice u klíčové fráze „krétské tržičiště“ a to u webů *kreta.webovka.eu* a *kreta.tym.cz*. Jednalo se však pouze o posun –1 pozice, což je zanedbatelné. Průměrný posun pozic po zaindexování zpětných odkazů ze služby Linkuj.cz je u vyhledávače Seznam +16,96.



Obrázek 6.7: Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po získání zpětného odkazu ze služby Linkuj.cz

Google

U vyhledávače Google narážíme s touto množinou webů na stejný problém, jaký byl u testování tagu ``. Z celkového počtu sledovaných klíčových frází u všech webových stránek, kterých je 50, se jich 20 vůbec neumístilo ve vyhledávání. O které fráze přesně jde, je patrné z obrázku 6.8 a tabulky A.8. Po vynechání těchto klíčových frází dostáváme průměrný posun pozic +39,78.

Zhodnocení

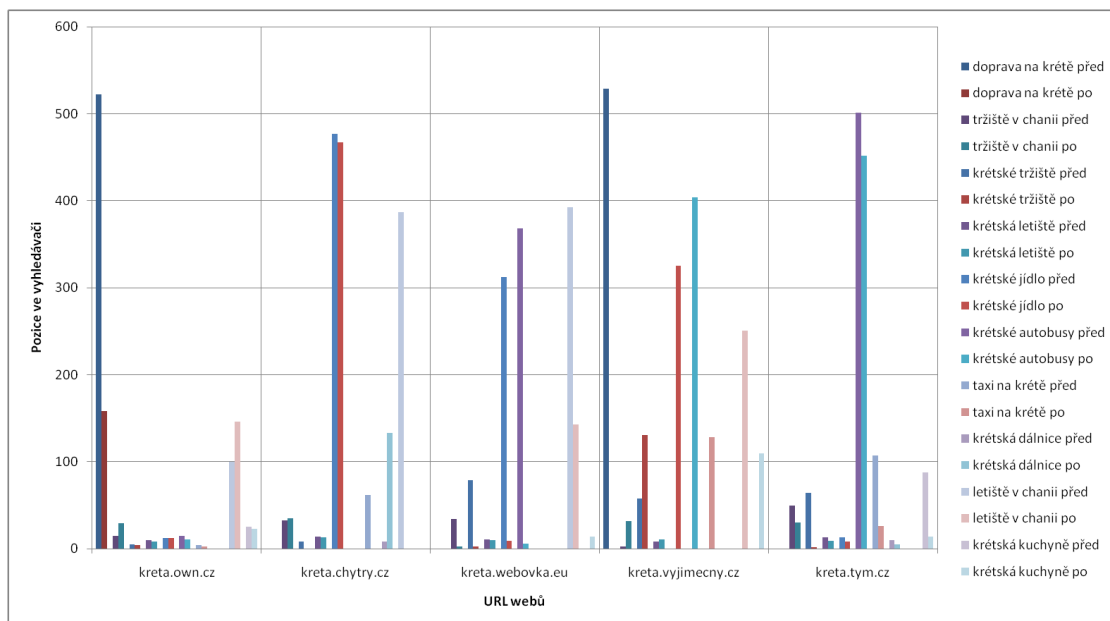
Měřením bylo dokázáno, že zpětné odkazy ze záložkovací služby Linkuj.cz mají pozitivní vliv na růst pozic. Průměrně se weby posunuly ve vyhledávání o +28,37 pozic. Z toho lze tedy usuzovat, že je vhodné do záložkovacích serverů vkládat webové stránky. Posilují se tím jednak pozice ve výsledcích vyhledávání, jednak je zde také velký potenciál pro získání přímých návštěvníků webu.

6.5.2 Vzájemné odkazy mezi testovanými weby

Zpětné odkazy bylo možné také vytvořit vzájemným odkazováním mezi testovanými weby. Weby na sebe odkazují „kruhovým způsobem“, tj. následovně: *kreta.own.cz* → *kreta.chytry.cz* → *kreta.webovka.eu* → *kreta.vyjimecny.cz* → *kreta.tym.cz* → *kreta.own.cz*. Jako anchor texty byly uvedeny testované klíčové fráze.

Seznam

Během tohoto testování byla u vyhledávače Seznam vynechána fráze „taxi na Krétě“ u webu *kreta.webovka.eu* a to opět proto, že se nevyskytovala v SERPu. Z naměřených hodnot, které jsou zobrazeny v tabulce A.9, vyplývá, že vzájemné prolínání webů má na pozice



Obrázek 6.8: Pozice webů na dané klíčové fráze ve vyhledávači Google před a po získání zpětného odkazu ze služby Linkuj.cz

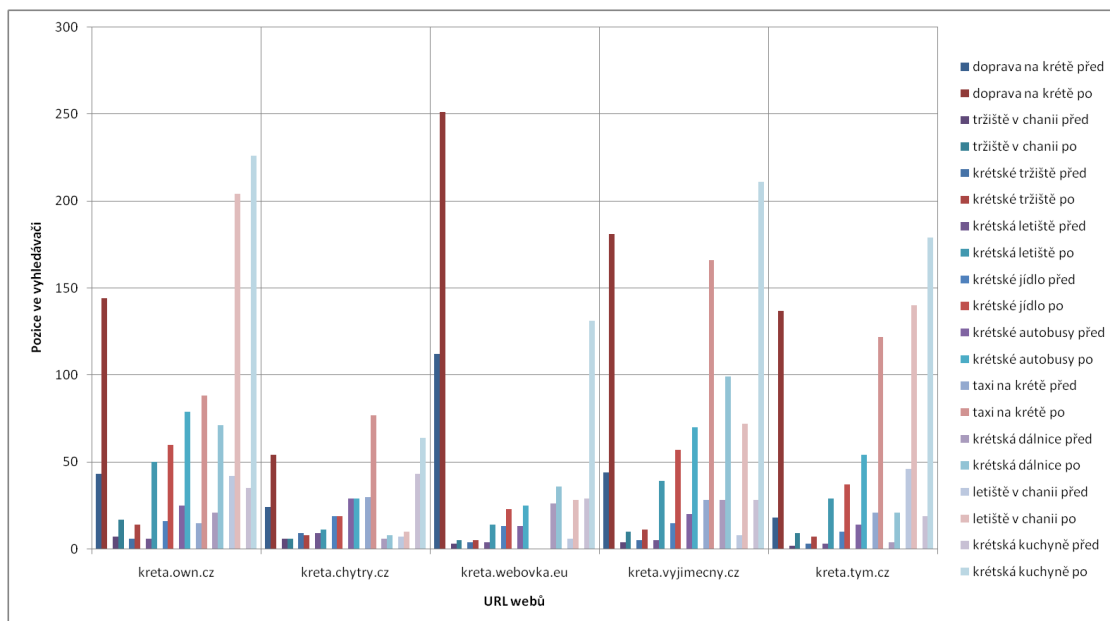
negativní vliv. Pouze v jednom případě došlo k růstu pozic – u fráze „krétské tržiště“ a webu *kreta.chytry.cz*. Šlo ovšem pouze o růst o jednu pozici. Ve třech dalších případech pozice stagnovaly (jedná se opět o web *kreta.chytry.cz*). U ostatních webů a frází však došlo k pádům. Nejvýraznější pád byl naměřen u fráze „krétská kuchyně“ (web *kreta.own.cz*) a to –191 pozic. Tento web má také nejvýraznější průměrný pád pozic na všechny klíčové fráze: –73,7 pozic. Posuny pozic vzhledem k tomuto off-page faktoru jsou zobrazeny na obrázku 6.9.

Google

U tohoto faktoru jsou z pohledu vyhledávače Google hodnoceny pouze weby *kreta.chytry.cz* a *kreta.tym.cz*. Je to způsobeno tím, že nedošlo k přeindexování všech webů po úpravě stránek ani po 25 dnech, kdy byl tento faktor testován. Mezi dvěma testovanými weby jsou však velké rozdíly. Z tabulky A.10 je vidět, že v případě stránky *kreta.chytry.cz* byl průměrný posun pozic +136,5. Do testování byly zapojeny však pouze čtyři fráze. Naopak u webu *kreta.tym.cz* se do měření promítlo šest frází, byl zde však výrazný pád pozic. Pouze u fráze „krétské tržiště“ došlo k růstu +1 pozice, tato fráze se tak umístila na prvním místě v SERPu. Celkový průměrný posun webů je +27,58 a je ovlivněn především frází „krétské jídlo“, která si polepšila ve vyhledávání o +450 pozic. Posuny pozic lze pozorovat také v obrázku 6.10.

Zhodnocení

Vzájemné prolinkování webů pomocí hypertextových odkazů se projevilo u testovaných webů především negativně. K poklesu pozic došlo u vyhledávače Seznam u všech stránek, u Google byl pokles u jednoho webu ze dvou měření. Celkový průměrný posun u obou vyhledávačů je –11,3 pozic.



Obrázek 6.9: Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po vytvoření zpětných odkazů mezi testovanými weby

6.5.3 Registrace do katalogu *odkazy.seznam.cz*

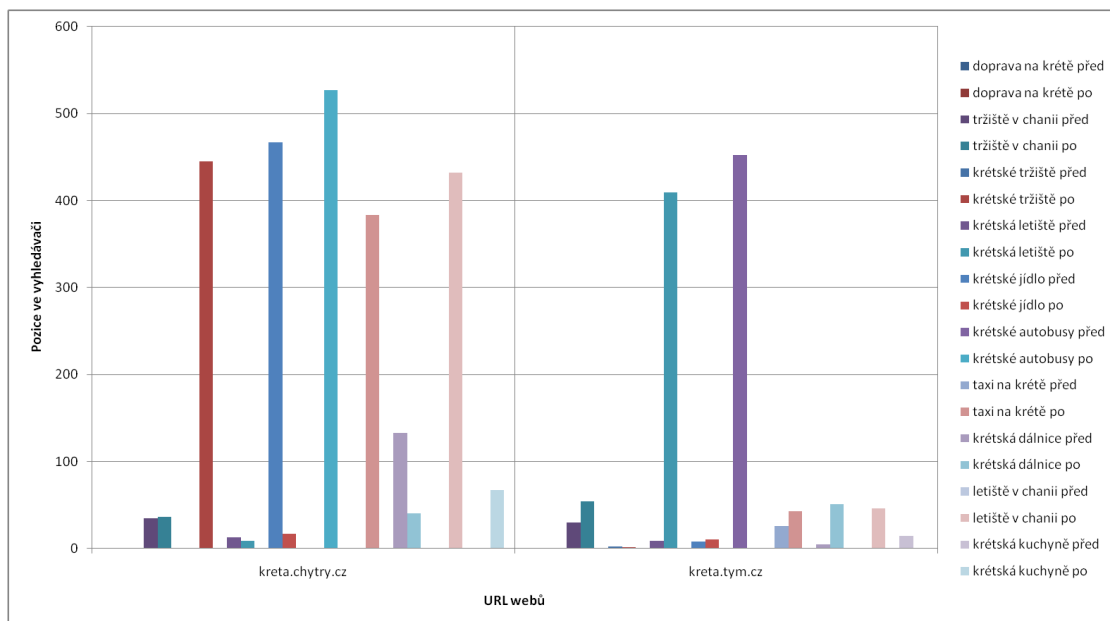
Katalog *odkazy.seznam.cz*, jak už samotné URL napovídá, je katalog www odkazů společnosti Seznam. Jedná se o jeden z nejdůležitějších katalogů na českém internetu. Žádosti o zařazení jsou schvalovány administrátory. Tím je zajištěno, že v katalogu budou pouze hodnotné weby a nestane se z něj link farma. Do katalogu bylo vloženo pět webů, naneštěstí byl však schválen pouze jeden – *kreta.hu.cz*. Ostatní byly zamítnuty z důvodu „téměř duplicitního obsahu s webem *kreta.hu.cz*“. Proto je toto testování omezeno jenom na tento web a výsledky nemusí být tolik průkazné jako u ostatních testů. Web je v katalogu umístěn v kategorii *Cestopisy z Kréty* [17].

Seznam

U testování tohoto faktoru byla vynechána fráze „krétská kuchyně“, protože se před vytvořením odkazu z katalogu v SERPu web *kreta.hu.cz* na tuto frázi nevyskytoval. Po zaindexování odkazu z katalogu byl však web na tuto frázi nalezen na 116 pozici. I z tohoto faktu můžeme soudit, že zpětné odkazy mají pozitivní vliv na pozice ve výsledcích vyhledávání. Z tabulky A.11 je patrné, že k růstu pozic došlo v sedmi případech, zatímco k poklesu je u dvou frází. Změny pozic jsou také znázorněny v obrázku 6.11. Průměrný posun klíčových frází po vytvoření zpětného odkazu z katalogu *odkazy.seznam.cz* je +10 pozic.

Google

U vyhledávače Google je situace odlišná. Z devíti případů, které se započítávají do celkové změny pozic, došlo v šesti k poklesu pozic. Jak je vidět z tabulky A.12, poklesy nejsou příliš velké, ve většině případů jde o propad o jednu nebo dvě pozice. Pouze u fráze „krétská dálnice“ je pokles −43 pozic. Celková průměrná změna pozic je však +3,7. Průměrný růst



Obrázek 6.10: Pozice webů na dané klíčové fráze ve vyhledávači Google před a po nastavení zpětných odkazů mezi testovanými weby

pozic ovlivňuje především fráze „krétská kuchyně“, u které došlo k růstu +85 pozic. Poklesy a růsty pozic jsou také zobrazeny v obrázku 6.12.

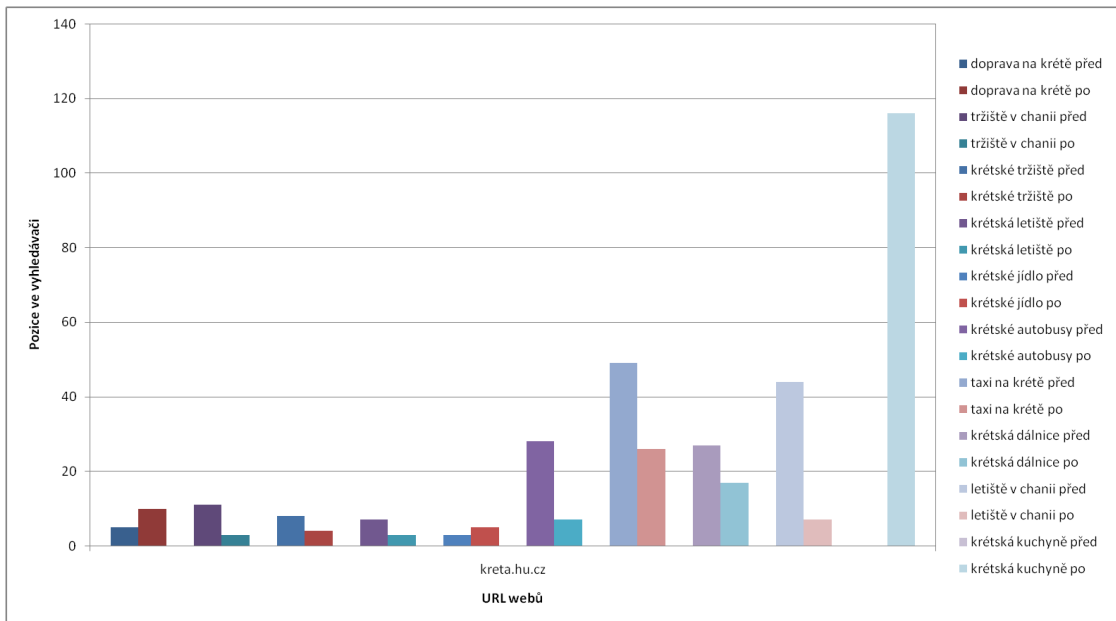
Zhodnocení

Vytvoření zpětného odkazu z katalogu *odkazy.seznam.cz* mělo u obou vyhledávačů pozitivní vliv na růst pozic. Průměrný růst pozic byl +6,85. U Seznamu byl vliv tohoto odkazu 2,7krát větší než u vyhledávače Google. Tento jev je možné vysvětlit tím, že katalog, kam byl testovaný web zaregistrován, je spravován přímo Seznamem a tak mu přiřkládá vyšší váhu.

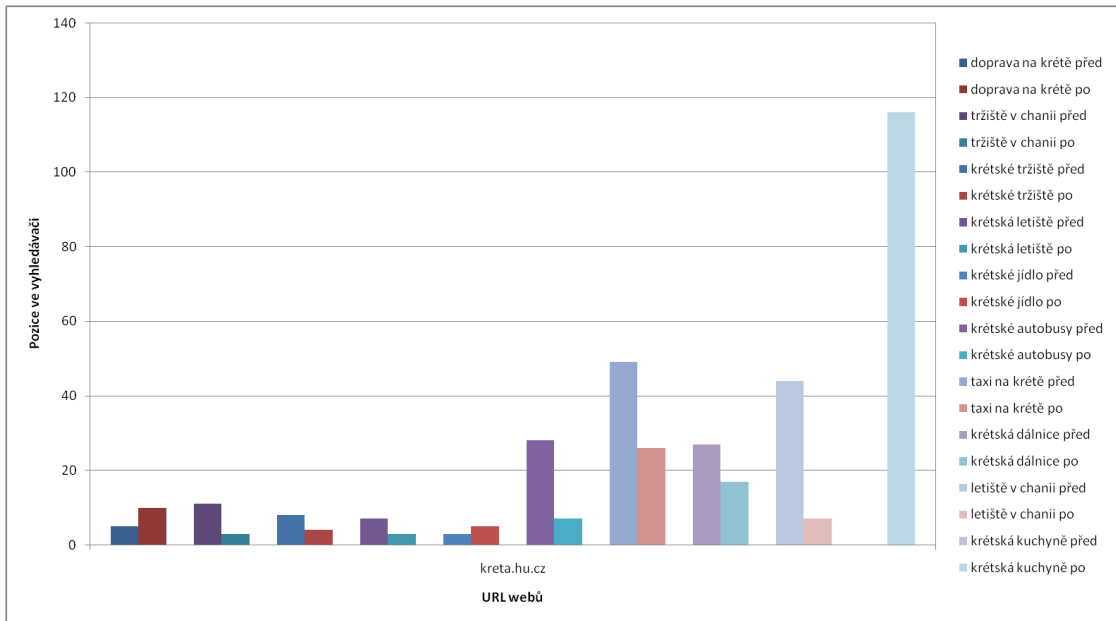
6.5.4 Nejdůležitější off-page faktory

Protože off-page faktory jsou především zpětné odkazy na optimalizované weby, je žádoucí se při hodnocení zaměřit na jejich kvalitu. Jak bylo popsáno v sekci 5.2, každý web si buduje tzv. *odkazový profil*. Proto se testované zpětné odkazy v této práci zaměřovaly pouze na ty, které jsou vedeny z webů s dobrou pověstí a mají podobné téma. Z testů vyšly nejlépe odkazy ze záložkovací služby *Linkuj.cz*, které posílily pozice webů v průměru o +28,37. O podstatně méně — +6,85 pozic — byly posíleny pozice díky odkazu z katalogu *odkazy.seznam.cz*. Naopak v testu odkazování se jednoho webu na druhý se stránky propadly v průměru o −11,3 pozic. Z toho vyplývá, že pro vyhledávače jsou pravděpodobně hodnotnější odkazy, které vedou z autoritativních webů a nejsou reciproční².

²Reciproční odkazy jsou takové, kdy web A odkazuje na web B a web B odkazuje zpět na web A.



Obrázek 6.11: Pozice web kreta.hu.cz na dané klíčové fráze ve vyhledávači Seznam před a po vytvoření zpětného odkazu v katalogu odkazy.seznam.cz



Obrázek 6.12: Pozice web kreta.hu.cz na dané klíčové fráze ve vyhledávači Seznam před a po vytvoření zpětného odkazu v katalogu odkazy.seznam.cz

Kapitola 7

Zhodnocení testů a vzorový web

Tato kapitola shrnuje jednotlivé testy prováděné v práci a vybírá ty on-page faktory, které mají pozitivní vliv na pozice v SERPu. Ty jsou následně aplikovány na vzorový web. Je tak dán návod tvůrcům stránek a administrátorům, jak správně tvořit weby, aby byly co nejlépe optimalizovány pro vyhledávače. Ne všechny faktory jsou pro vyhledávače stejně důležité. Důležité je vědět, že jeden totožný faktor může být jinak posuzován mezi jednotlivými vyhledávači. Protože algoritmy vyhledávačů nejsou veřejně známé, nezbyvá nic jiného, než sadou testů podobných jako v této práci, ověřovat faktor po faktoru.

V praxi není možné vytvářet dvě a více verzí webů tak, aby jedna byla určená pro vyhledávač A, další pro vyhledávač B apod. Z tohoto důvodu je důležité udělat kompromis a při tvorbě webu zohledňovat také věci jako je přístupnost a použitelnost a nehnát se slepě pouze za pozicemi ve vyhledávání.

7.1 On-page faktory

Testováním bylo zjištěno, že nejdůležitějším on-page faktorem je titulek (HTML tag `<title>`). Do titulku je vhodné dát nejdůležitější klíčová slova z dané stránky. To stejné platí také pro nadpisy (HTML tagy `<h1>` až `<h6>`). Podle testů nadpisy posilují pozice ve vyhledávání. Dále zpřehledňují strukturu textu, ale jsou také důležité například pro nevidomé návštěvníky, kteří používají pro prohlížení stránek hlasové čtečky. U zvýraznění klíčových frází pomocí tagu `` (případně ``, `` nebo `<i>`) byly u vyhledávačů Seznam a Google rozdílné výsledky. V Seznamu se zvýraznění projevilo mírně pozitivně, u Googlu naopak došlo k pádu pozic. Přesto je do vzorového webu aplikován i tento on-page faktor a to především kvůli tomu, že zvýrazněné písmo upoutá návštěvníkovu pozornost a pokud tuto frázi hledá, je větší pravděpodobnost, že na webu zůstane.

7.2 Off-page faktory

Co se týče off-page faktorů, lze pouze obecně doporučit, aby tvořené odkazy pocházely především z tematicky podobných webů. U registrací do katalogů je důležité pečlivě vybírat jednak samotné katalogy, ale také správné kategorie.

7.3 Vzorový web

Na základě zhodnocení testů byl vytvořen vzorový web, který aplikuje následující on-page faktory:

- klíčová slova v titulku
- nastavení nadpisů
- zvýraznění slov pomocí tagu ``

Jako vzorový web byl vybrán z testovaných ten, který se v součtu nejlépe umisťoval ve výsledcích vyhledávání. Jde o web *kreta.hu.cz*. Na něj byly aplikovány faktory uvedené výše. Takto vytvořený web je umístěn na přiloženém CD v adresáři *vzorovy-web*.

Kapitola 8

Závěr

V této bakalářské práci byly shrnuty nejdůležitější faktory v procesu optimalizace webových stránek pro vyhledávače. Vybrané on-page a off-page faktory byly následně prakticky otestovány na množině deseti webů. Po každé změně faktoru probíhalo měření pozic ve výsledcích vyhledávání vyhledávačů Seznam a Google. Díky tomu bylo zjištěno, které faktory nejvíce ovlivňují pozice v SERPu, a proto je důležité se na ně zaměřit při tvorbě a správě www stránek.

Testování jednotlivých faktorů je jediný možný způsob, jak zjistit, jakou váhu jim přiřkládají vyhledávače. V práci byly diskutovány a testovány vybrané faktory a následně porovnány mezi sebou. Pokud se budou autoři stránek řídit doporučeními z této práce, budou jejich webové stránky kvalitně optimalizovány pro vyhledávače, alespoň co se on-page faktorů týče. Protože SEO je proces, nikoli stav, je nutná neustálá práce jak na samotném webu (tvorba nového obsahu, správná navigační struktura apod.), tak i na off-page faktorech (zpětné odkazy). Je však důležité myslet na to, že tvůrci vyhledávačů se snaží inovovat indexovací a vyhledávací algoritmy tak, aby dávaly co nejrelevantnější odpovědi na položené dotazy. Například vyhledávač Google v roce 2010 testoval více než 6000 změn v algoritmu. Pouze 500 z nich se však promítlo do algoritmu natrvalo [22]. O které změny se jedná, ale Google neprozradil. Je tedy pravděpodobné, že naměřené výsledky z této práce mohou být za určitý čas považovány za zastaralé.

Z tohoto důvodu je pro tvůrce a správce webů, kteří se zabývají optimalizací pro vyhledávače, nutné, aby neustále sledovali vývoj a reagovali na změny. V tomto směru je otevřen prostor i pro další vývoj práce. Je možné testovat další faktory ovlivňující pozice ve výsledcích vyhledávání. Zajímavé by bylo zaměření se čistě na off-page faktory, tedy zpětné odkazy, jejich tematickou příbuznost, kvalitu vs. kvantitu apod. Další možností je provedení identických testů jako v této práci o několik měsíců později a jejich vzájemné porovnání.

Literatura

- [1] Hubs and Authorities [online].
<http://nlp.stanford.edu/IR-book/html/htmledition/hubs-and-authorities-1.html>, 2009-07-04 [cit. 2011-01-22].
- [2] PageRank – Wikipedie [online]. <http://cs.wikipedia.org/wiki/PageRank>, 2010-08-03 [cit. 2010-12-22].
- [3] Open Directory – World: Český. <http://www.dmoz.org/World/%C4%8Cesky/>, 2010 [cit. 2010-12-10].
- [4] Google AdWords: Návrhy klíčových slov [online].
<https://adwords.google.com/select/KeywordToolExternal>, 2010 [cit. 2010-12-15].
- [5] Synonyma online Tezaurus – slovník synonym [online].
<http://www.synonyma-online.cz/tezaurus.php>, 2010 [cit. 2010-12-15].
- [6] Apache Tutorial: .htaccess files [online].
<http://httpd.apache.org/docs/1.3/howto/htaccess.html>, 2010 [cit. 2010-12-22].
- [7] Google Toolbar [online].
<http://www.google.com/intl/cs/toolbar/ff/index.html>, 2010 [cit. 2010-12-22].
- [8] Půjdou moje stránky zaindexovat? [online].
<http://napoveda.seznam.cz/cz/hledani-fulltext-pujdou-moje-stranky-zaindexovat-pridat-do-vyhledavani.html>, 2010 [cit. 2010-12-22].
- [9] Seznam Lištička — Seznam software [online].
<http://software.seznam.cz/listicka>, 2010 [cit. 2010-12-22].
- [10] Site Explorer – Yahoo! Site Explorer [online].
<http://siteexplorer.search.yahoo.com>, 2010 [cit. 2010-12-22].
- [11] Vaše stránky v indexu Google [online].
<http://www.google.com/support/webmasters/bin/answer.py?hl=cs&answer=35256>, 2010 [cit. 2010-12-22].
- [12] Algoritmus — Seznam Náповěda [online].
<http://napoveda.seznam.cz/cz/hledani-fulltext-algoritmus-vyhledavani-razeni-vysledku-faq-dotazy.html>, 2010 [cit. 2011-01-22].
- [13] Open Site Explorer: Link Popularity & Backlink Analysis Tool [online].
<http://www.opensiteexplorer.org>, 2011 [cit. 2011-03-24].

- [14] Tapetování klíčovými slovy [online].
<http://www.google.com/support/webmasters/bin/answer.py?answer=66358>,
2011 [cit. 2011-03-24].
- [15] Search Engine Ranking Factors — SEOmoz [online].
<http://www.seomoz.org/article/search-ranking-factors#ranking-factors>,
2011 [cit. 2011-04-13].
- [16] AWStats – Free advanced log file analyzer for web, ftp or mail statistics (GNU GPL).
<http://awstats.sourceforge.net/>, 2011 [cit. 2011-05-04].
- [17] Cestopisy z Kréty – Seznam.
<http://odkazy.seznam.cz/Cestovani/Cestopisy-a-informace-o-cestovani/Cestopisy-z-Evropy/Cestopisy-z-Recka/Cestopisy-z-Krety/>, 2011 [cit. 2011-05-04].
- [18] Statistics for destailleur.fr (2011-05) – main.
<http://www.nltechno.com/awstats/awstats.pl?config=destailleur.fr>, 2011
[cit. 2011-05-04].
- [19] Informace o atributu rel=“nofollow” – Náповěda Nástroje pro webmastery.
<http://www.google.com/support/webmasters/bin/answer.py?hl=cs&answer=96569>,
2011 [cit. 2011-05-08].
- [20] Nofollow — Seznam Náповěda. <http://napoveda.seznam.cz/cz/fulltext-hledani-v-internetu/komunikace-s-vyhledavacim-robotem/nofollow/>,
2011 [cit. 2011-05-08].
- [21] Anderson, C.: The Long Tail. *The Wired*, ročník 12.10, 10 2004, ISSN 1059-1028.
- [22] Chapman, G.: Google hones search edge to stay sharp. <http://www.google.com/hostednews/afp/article/ALeqM5gYLg8pB1TXX2F3UWJx5UHqxy2Rzw?docId=CNG.887a901975c3f70a5f64f725a13a9937.121>, 2011-04-13 [cit. 2011-05-07].
- [23] Chen, S.: PHP Simple HTML DOM Parser [online].
<http://simplehtmldom.sourceforge.net/>, 2011 [cit. 2011-04-13].
- [24] Cutts, M.: SEO advice: url canonicalization [online].
<http://www.matcutts.com/blog/seo-advice-url-canonicalization/>,
2006-01-04 [cit. 2010-12-22].
- [25] Cutts, M.: Google does not use the keywords meta tag in web ranking [online].
<http://googlewebmastercentral.blogspot.com/2009/09/google-does-not-use-keywords-meta-tag.html>, 2009-09-21 [cit. 2010-12-22].
- [26] Cutts, M.: If you were an SEO of a large company, what would you include in your 2011 strategy? [online databáze videí].
<http://www.youtube.com/user/GoogleWebmasterHelp#p/u/5/vLp9Qf99DCI>,
2010-10-26 [cit. 2010-03-24].
- [27] Cutts, M.: How does Google determine domain age, and is it important for ranking? [online databáze videí]. <http://www.youtube.com/watch?v=-pnpg00FWJY>,
2010-10-26 [cit. 2010-12-22].

- [28] DeJarnette, R.: Prevent a bot from getting “lost in space” (SEM 101) [online].
http://www.bing.com/community/site_blogs/b/webmaster/archive/2009/08/21/prevent-a-bot-from-getting-lost-in-space-sem-101.aspx, 2008-01-11 [cit. 2010-12-09].
- [29] Dover, D.: Introducing New Features for Open Site Explorer [online].
<http://www.seomoz.org/blog/introducing-new-features-for-open-site-explorer>, 2010-03-04 [cit. 2011-03-24].
- [30] Enge, E.; et al.: *The Art of SEO*, kapitola 4. Sebastopol: O’Reilly Media, Inc., první vydání, 2009, ISBN 978-0-596-51886-8, str. 578.
- [31] Fielding, R.; et al.: RFC 2616 - Hypertext Transfer Protocol – HTTP/1.1 [online].
<http://tools.ietf.org/html/rfc2616>, 1999 [cit. 2010-12-22].
- [32] Google, I.: Robots meta tag and X-Robots-Tag HTTP header specifications – Controlling Crawling and Indexing – Google Code [online].
http://code.google.com/intl/cs/web/controlcrawlindex/docs/robots_meta_tag.html, 2010 [cit. 2010-12-09].
- [33] Janovský, D.: Jak zjistit míru zaindexování [online].
<http://www.jakpsatweb.cz/seo/zaindexovani.html#supplemental>, 2010 [cit. 2010-03-24].
- [34] Janovský, D.: O Google: algoritmy, vlastnosti Google, jak optimalizovat [online].
<http://www.jakpsatweb.cz/google.html>, 2010 [cit. 2010-12-04].
- [35] Janovský, D.: Co je databáze a co index [online].
<http://www.jakpsatweb.cz/seo/zaindexovani.html#databaze>, 2010 [cit. 2010-12-09].
- [36] Jerkovic, J. I.: *SEO Warrior*, kapitola 14. Sebastopol: O’Reilly, druhé vydání, 2010, ISBN 978-0-596-15707-4, str. 472.
- [37] Kenyon, G.: Site Speed – Are You Fast? Does it Matter? [online].
<http://www.seomoz.org/blog/site-speed-are-you-fast-does-it-matter>, 2010-02-20 [cit. 2011-03-24].
- [38] Klír, J.: PageRank, S-Rank, Alexa, Compete, JyxoRank Checker – pagerank.jklir.net [online]. <http://pagerank.jklir.net>, 2010 [cit. 2010-12-22].
- [39] Kubíček, M.: *Velký průvodce SEO: Jak dosáhnout nejlepších pozic ve vyhledávacích*. Brno: Computer Press, a.s., první vydání, 2008, ISBN 978-80-251-2195-5, str. 320.
- [40] Lešetický, M.: Katalogománie. Aneb rizika bezhlavé registrace do všemožných katalogů [online]. <http://plavacek.posterous.com/katalogomanie-aneb-rizika-bezhlave-registrace>, 2010-20-02 [cit. 2010-12-10].
- [41] Malmborg, D.: Evidence Google is Ignoring the Title Tag [online].
<http://www.seo.com/blog/google-ignoring-title-tag/>, 2009-02-17 [cit. 2011-04-12].

- [42] Moran, M.; Hunt, B.: *Search Engine Marketing, Inc.*, kapitola 2. Upper Saddle River: Pearson plc, první vydání, 2006, ISBN 0-13-185292-2, str. 560.
- [43] Page, L.; et al.: The PageRank Citation Ranking: Bringing Order to the Web. – Stanford InfoLab Publication Server [online].
<http://ilpubs.stanford.edu:8090/422/>, 2001-10-30 [cit. 2011-01-22].
- [44] Seda, C.: *Search Engine Advertising: Buying Your Way to the Top to Increase Sales*, kapitola 3. Indianapolis: New Riders Publishing, první vydání, 2004, ISBN 0-7357-1399-5, str. 347.
- [45] Seznam: Fulltext hledání v internetu [online].
<http://napoveda.seznam.cz/cz/hledani-fulltext-vyhledani-v-internetu.html>, 2010 [cit. 2010-12-10].
- [46] Singhal, A.; Cutts, M.: Using site speed in web search ranking [online].
<http://googlewebmastercentral.blogspot.com/2010/04/using-site-speed-in-web-search-ranking.html>, 2010-04-09 [cit. 2011-03-24].
- [47] Staeding, A.: List of User-Agents (Spiders, Robots, Browser) [online].
<http://www.user-agents.org>, 2010 [cit. 2010-12-03].
- [48] Stibůrek, M.; Blahoš, P.: SEO Profesionál Toolbar — SEO – Profesionál.cz [online].
<http://www.seo-profesional.cz/toolbar/>, 2010 [cit. 2010-12-22].
- [49] Sullivan, D.: Google Search History Expands, Becomes Web History [online].
<http://searchengineland.com/google-search-history-expands-becomes-web-history-11016>, 2007-04-19 [cit. 2011-04-14].
- [50] Sullivan, D.: Google Now Personalizes Everyone's Search Results [online].
<http://searchengineland.com/google-now-personalizes-everyones-search-results-31195>, 2009-12-04 [cit. 2011-04-14].
- [51] Tichý, J.: Jaké jsou skutečné podíly vyhledávačů? [online].
<http://blog.medio.cz/podily-vyhledavacu>, 2010-01-29 [cit. 2010-12-09].
- [52] Wall, A.; Hagans, A.: How to Build Links Fast: 101 Tips & Strategies [online].
<http://www.seobook.com/archives/001792.shtml>, 2008-08-16 [cit. 2011-01-25].
- [53] Yahoo!, I.: How to Prevent Search Engines from Indexing Specific Pages [online].
<http://help.yahoo.com/l/us/yahoo/search/indexing/slurp-04.html>, 2010-08-18 [cit. 2010-12-09].

Seznam příloh

Dodatek A – tabulky s naměřenými pozicemi ve výsledcích vyhledávání

Dodatek B – obsah CD

Dodatek A

Naměřené pozice ve vyhledávačích

Testování on-page a off-page faktorů probíhalo na základě měření absolutních pozic ve výsledcích vyhledávání vyhledávačů Seznam a Google. Níže jsou k dispozici tabulky s naměřenými daty.

A.1 Vysvětlivky k tabulkám

V prvním řádku tabulek jsou uvedeny domény webových stránek. Kvůli úspoře místa jsou zde pouze domény druhého řádu, místo domén třetího řádu je uveden znak * (u všech webů je doména třetího řádu „kreta“). V prvním sloupci jsou zobrazena jednotlivá klíčová slova. V dalším sloupci jsou uvedeny tři řádky:

- *před* – pozice daného webu ve vyhledávání před úpravami stránky
- *po* – udává pozici daného webu ve vyhledávání po úpravách stránky
- *posun* – relativní změna pozice ve výsledcích vyhledávání; kladné číslo udává růst pozic, záporné pád pozic

Každý faktor byl testován na množině pěti webů, jednotlivé pozice jsou zaznamenány do tabulek. Pokud je místo čísla uveden u pozice znak „x“, daný web nebyl na uvedené klíčovou frázi ve výsledcích vyhledávání nalezen. Prohledávání u vyhledávače Seznam je nastaveno do 1000. pozice a u vyhledávače Google do 600. pozice. Znak „x“ v řádku „posun“ značí, že alespoň v jednom případě nebyl web na danou klíčovou frázi nalezen ve vyhledávání. Tato fráze je následně vynechána z celkového hodnocení. Vyhodnocení probíhá na základě zprůměrování jednotlivých posunů ve výsledcích vyhledávání. Průměry jednotlivých webů jsou následně zprůměrovány do výsledného čísla, které je označeno jako „celkový průměrný posun“. Tato hodnota udává průměrný růst nebo pád daného on-page nebo off-page faktoru.

Klíčová fráze		*.hu.cz	*.ezin.cz	*.yc.cz	*.kx.cz	*.byl.cz
doprava na Krétě	před	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	po	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	posun	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
tržiště v Chanii	před	9	10	11	1	6
	po	9	10	11	1	6
	posun	0	0	0	0	0
krétské tržiště	před	16	9	4	1	23
	po	11	7	4	1	20
	posun	5	2	0	0	3
krétské letiště	před	24	17	20	16	15
	po	11	9	13	30	8
	posun	13	8	7	-14	7
krétské jídlo	před	29	28	128	24	23
	po	17	18	95	20	15
	posun	12	10	33	4	8
krétské autobusy	před	39	93	<i>x</i>	40	53
	po	35	67	<i>x</i>	36	48
	posun	4	26	<i>x</i>	4	5
taxi na Krétě	před	87	74	<i>x</i>	<i>x</i>	16
	po	48	42	<i>x</i>	<i>x</i>	12
	posun	39	32	<i>x</i>	<i>x</i>	4
krétská dálnice	před	35	9	8	14	13
	po	34	9	8	13	12
	posun	1	0	0	1	1
letiště v Chanii	před	75	58	191	32	35
	po	45	137	191	19	20
	posun	30	-79	0	13	15
krétská kuchyně	před	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	po	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	posun	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>

Celkový posun	104	-1	40	8	43
Průměrný posun na web	13	-0,125	6,667	1,143	5,375

Celkový průměrný posun	5,21
-------------------------------	------

Tabulka A.1: Pozice webů na dané klíčové fráze ve vyhledávací Seznam před a po úpravách titulku

Klíčová fráze		*.hu.cz	*.ezin.cz	*.yc.cz	*.kx.cz	*.byl.cz
doprava na Krétě	před	<i>x</i>	<i>x</i>	43	<i>x</i>	30
	po	470	<i>x</i>	14	438	<i>x</i>
	posun	<i>x</i>	<i>x</i>	29	<i>x</i>	<i>x</i>
tržiště v Chanii	před	40	61	6	<i>x</i>	5
	po	30	<i>x</i>	<i>x</i>	4	35
	posun	10	<i>x</i>	<i>x</i>	<i>x</i>	-30
krétské tržiště	před	134	270	2	7	1
	po	63	<i>x</i>	132	1	80
	posun	71	<i>x</i>	-130	6	-79
krétské letiště	před	15	<i>x</i>	6	<i>x</i>	9
	po	7	<i>x</i>	<i>x</i>	<i>x</i>	371
	posun	8	<i>x</i>	<i>x</i>	<i>x</i>	-362
krétské jídlo	před	16	<i>x</i>	39	<i>x</i>	9
	po	11	207	<i>x</i>	14	<i>x</i>
	posun	5	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
krétské autobusy	před	361	<i>x</i>	<i>x</i>	<i>x</i>	13
	po	65	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	posun	296	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
taxi na Krétě	před	234	<i>x</i>	<i>x</i>	<i>x</i>	3
	po	120	<i>x</i>	<i>x</i>	<i>x</i>	10
	posun	114	<i>x</i>	<i>x</i>	<i>x</i>	-7
krétská dálnice	před	250	262	7	263	10
	po	24	34	18	9	12
	posun	226	228	-11	254	-2
letiště v Chanii	před	420	422	101	389	14
	po	<i>x</i>	<i>x</i>	<i>x</i>	378	264
	posun	<i>x</i>	<i>x</i>	<i>x</i>	11	-250
krétská kuchyně	před	214	<i>x</i>	39	<i>x</i>	38
	po	207	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	posun	7	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
Celkový posun		737	228	-112	271	-730
Průměrný posun na web		92,125	228	-37,333	90,333	-121,667
Celkový průměrný posun		50,29				

Tabulka A.2: Pozice webů na dané klíčové fráze ve vyhledávači Google před a po úpravách titulku

Klíčová fráze		*.own.cz	*.chytry.cz	*.webovka.cz	*.vyjimecny.cz	*.tym.cz
doprava na Krétě	před	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	po	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	posun	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
tržiště v Chanii	před	8	7	4	5	3
	po	8	7	4	5	3
	posun	0	0	0	0	0
krétské tržiště	před	7	19	3	6	2
	po	6	12	3	5	2
	posun	1	7	0	1	0
krétské letiště	před	30	21	18	22	13
	po	27	21	18	22	15
	posun	3	0	0	0	-2
krétské jídlo	před	32	36	<i>x</i>	31	18
	po	29	34	<i>x</i>	28	22
	posun	3	2	<i>x</i>	3	-4
krétské autobusy	před	58	84	34	56	35
	po	55	64	29	50	32
	posun	3	20	5	6	3
taxi na Krétě	před	26	70	<i>x</i>	69	51
	po	24	61	<i>x</i>	52	36
	posun	2	9	<i>x</i>	17	15
krétská dálnice	před	30	11	40	49	7
	po	31	10	40	43	7
	posun	-1	1	0	6	0
letiště v Chanii	před	110	40	39	41	114
	po	98	38	37	39	103
	posun	12	2	2	2	11
krétská kuchyně	před	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	po	118	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	posun	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
Celkový posun		23	41	7	35	23
Průměrný posun na web		2,875	5,125	1,167	4,375	2,875

Celkový průměrný posun	3,28
-------------------------------	------

Tabulka A.3: Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po zvýraznění klíčových frází pomocí tagu

Klíčová fráze		*.own.cz	*.chytry.cz	*.webovka.cz	*.vyjimecny.cz	*.tym.cz
doprava na Krétě	před	25	<i>x</i>	29	203	<i>x</i>
	po	522	<i>x</i>	<i>x</i>	529	<i>x</i>
	posun	-497	<i>x</i>	<i>x</i>	-326	<i>x</i>
tržiště v Chanii	před	2	43	4	3	60
	po	15	33	34	3	50
	posun	-13	10	-30	0	10
krétské tržiště	před	6	<i>x</i>	5	4	138
	po	5	8	79	58	64
	posun	1	<i>x</i>	-74	-54	74
krétské letiště	před	10	16	11	5	20
	po	10	14	11	8	13
	posun	0	2	0	-3	7
krétské jídlo	před	11	467	10	15	20
	po	12	477	312	<i>x</i>	13
	posun	-1	-10	-302	<i>x</i>	7
krétské autobusy	před	17	<i>x</i>	11	12	574
	po	15	<i>x</i>	368	<i>x</i>	501
	posun	2	<i>x</i>	-357	<i>x</i>	73
taxi na Krétě	před	6	<i>x</i>	<i>x</i>	38	356
	po	4	62	<i>x</i>	<i>x</i>	107
	posun	2	<i>x</i>	<i>x</i>	<i>x</i>	249
krétská dálnice	před	24	193	17	16	243
	po	<i>x</i>	8	<i>x</i>	<i>x</i>	10
	posun	<i>x</i>	185	<i>x</i>	<i>x</i>	233
letiště v Chanii	před	88	394	15	13	431
	po	100	387	393	<i>x</i>	<i>x</i>
	posun	-12	7	-378	<i>x</i>	<i>x</i>
krétská kuchyně	před	27	197	22	23	104
	po	25	<i>x</i>	<i>x</i>	<i>x</i>	88
	posun	2	<i>x</i>	<i>x</i>	<i>x</i>	16
Celkový posun		-516	194	-1141	-383	669
Průměrný posun na web		-57,333	38,8	-190,167	-95,75	88,625

Celkový průměrný posun	-44,17
-------------------------------	--------

Tabulka A.4: Pozice webů na dané klíčové fráze ve vyhledávači Google před a po zvýraznění klíčových frází pomocí tagu

Klíčová fráze		*.hu.cz	*.ezin.cz	*.yc.cz	*.kx.cz	*.byl.cz
doprava na Krétě	před	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	po	5	13	6	4	8
	posun	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
tržiště v Chanii	před	9	10	11	1	6
	po	11	2	9	3	6
	posun	-2	8	2	-2	0
krétské tržiště	před	11	7	4	1	20
	po	8	3	1	4	21
	posun	3	4	3	-3	-1
krétské letiště	před	11	9	13	30	8
	po	7	5	1	23	12
	posun	4	4	12	7	-4
krétské jídlo	před	17	18	95	20	15
	po	3	7	27	8	11
	posun	14	11	68	12	4
krétské autobusy	před	35	67	<i>x</i>	36	48
	po	28	24	<i>x</i>	44	50
	posun	7	43	<i>x</i>	-8	-2
taxi na Krétě	před	48	42	<i>x</i>	<i>x</i>	12
	po	49	4	<i>x</i>	<i>x</i>	11
	posun	-1	38	<i>x</i>	<i>x</i>	1
krétská dálnice	před	34	9	8	13	12
	po	27	4	3	14	13
	posun	7	5	5	-1	-1
letiště v Chanii	před	45	137	150	19	20
	po	44	114	141	23	24
	posun	1	23	9	-4	-4
krétská kuchyně	před	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	po	<i>x</i>	44	18	<i>x</i>	<i>x</i>
	posun	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
Celkový posun		33	136	99	1	-7
Průměrný posun na web		4,125	17	16,5	0,143	-0,875

Celkový průměrný posun	7,38
-------------------------------	------

Tabulka A.5: Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po zvýraznění klíčových slov pomocí nadpisů

Klíčová fráze		*.hu.cz	*.ezin.cz	*.yc.cz	*.kx.cz	*.byl.cz
doprava na Krétě	před	470	<i>x</i>	14	438	<i>x</i>
	po	14	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	posun	456	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
tržiště v Chanii	před	30	<i>x</i>	<i>x</i>	4	35
	po	2	<i>x</i>	<i>x</i>	5	36
	posun	28	<i>x</i>	<i>x</i>	-1	-1
krétské tržiště	před	63	<i>x</i>	132	1	80
	po	1	<i>x</i>	125	5	87
	posun	62	<i>x</i>	7	-4	-7
krétské letiště	před	7	<i>x</i>	<i>x</i>	<i>x</i>	371
	po	5	<i>x</i>	<i>x</i>	<i>x</i>	369
	posun	2	<i>x</i>	<i>x</i>	<i>x</i>	2
krétské jídlo	před	11	207	<i>x</i>	14	<i>x</i>
	po	7	224	<i>x</i>	15	<i>x</i>
	posun	4	-17	<i>x</i>	-1	<i>x</i>
krétské autobusy	před	65	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	po	5	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	posun	60	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
taxi na Krétě	před	120	<i>x</i>	<i>x</i>	<i>x</i>	10
	po	2	<i>x</i>	<i>x</i>	<i>x</i>	146
	posun	118	<i>x</i>	<i>x</i>	<i>x</i>	-136
krétská dálnice	před	24	34	18	9	12
	po	4	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	posun	20	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
letiště v Chanii	před	<i>x</i>	<i>x</i>	<i>x</i>	378	264
	po	<i>x</i>	<i>x</i>	<i>x</i>	181	180
	posun	<i>x</i>	<i>x</i>	<i>x</i>	197	84
krétská kuchyně	před	207	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	po	220	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	posun	-13	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
Celkový posun		737	-17	7	191	-58
Průměrný posun na web		81,889	-17	7	47,75	-11,6

Celkový průměrný posun	21,61
-------------------------------	-------

Tabulka A.6: Pozice webů na dané klíčové fráze ve vyhledávači Google před a po zvýraznění klíčových slov pomocí nadpisů

Klíčová fráze		*.own.cz	*.chytry.cz	*.webovka.cz	*.vyjimecny.cz	*.tym.cz
doprava na Krétě	před	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	po	43	24	112	44	18
	posun	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
tržiště v Chanii	před	8	7	4	5	3
	po	7	6	3	4	2
	posun	1	1	1	1	1
krétské tržiště	před	6	12	3	5	2
	po	6	9	4	5	3
	posun	0	3	-1	0	-1
krétské letiště	před	27	21	18	22	15
	po	6	9	4	5	3
	posun	21	12	14	17	12
krétské jídlo	před	29	34	<i>x</i>	28	22
	po	16	19	13	15	10
	posun	13	15	<i>x</i>	13	12
krétské autobusy	před	55	64	29	50	32
	po	25	29	13	20	14
	posun	30	35	16	30	18
taxi na Krétě	před	24	61	<i>x</i>	52	36
	po	15	30	<i>x</i>	28	21
	posun	9	31	<i>x</i>	24	15
krétská dálnice	před	31	10	40	43	7
	po	21	6	26	28	4
	posun	10	4	14	15	3
letiště v Chanii	před	98	38	37	39	103
	po	42	7	6	8	46
	posun	56	31	31	31	57
krétská kuchyně	před	118	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	po	35	43	29	28	19
	posun	83	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
Celkový posun		223	132	75	131	117
Průměrný posun na web		24,778	16,5	12,5	16,375	14,625

Celkový průměrný posun	16,96
-------------------------------	-------

Tabulka A.7: Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po získání zpětného odkazu ze služby Linkuj.cz

Klíčová fráze		*.own.cz	*.chytry.cz	*.webovka.cz	*.vyjimecny.cz	*.tym.cz
doprava na Krétě	před	522	<i>x</i>	<i>x</i>	529	<i>x</i>
	po	158	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
	posun	364	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>
tržiště v Chanii	před	15	33	34	3	50
	po	29	35	3	32	30
	posun	-14	-2	31	-29	20
krétské tržiště	před	5	8	79	58	64
	po	4	<i>x</i>	3	131	2
	posun	1	<i>x</i>	76	-73	62
krétské letiště	před	10	14	11	8	13
	po	8	13	10	11	9
	posun	2	1	1	-3	4
krétské jídlo	před	12	477	312	<i>x</i>	13
	po	12	467	9	325	8
	posun	0	10	303	<i>x</i>	5
krétské autobusy	před	15	<i>x</i>	368	<i>x</i>	501
	po	11	<i>x</i>	6	404	452
	posun	4	<i>x</i>	362	<i>x</i>	49
taxi na Krétě	před	4	62	<i>x</i>	<i>x</i>	107
	po	3	<i>x</i>	<i>x</i>	128	26
	posun	1	<i>x</i>	<i>x</i>	<i>x</i>	81
krétská dálnice	před	<i>x</i>	8	<i>x</i>	<i>x</i>	10
	po	<i>x</i>	133	<i>x</i>	<i>x</i>	5
	posun	<i>x</i>	-125	<i>x</i>	<i>x</i>	5
letiště v Chanii	před	100	387	393	<i>x</i>	<i>x</i>
	po	146	<i>x</i>	143	251	<i>x</i>
	posun	-46	<i>x</i>	250	<i>x</i>	<i>x</i>
krétská kuchyně	před	25	<i>x</i>	<i>x</i>	<i>x</i>	88
	po	23	<i>x</i>	14	110	14
	posun	2	<i>x</i>	<i>x</i>	<i>x</i>	73
Celkový posun		314	-116	1023	-105	300
Průměrný posun na web		34,889	-29	170,5	-15	37,5

Celkový průměrný posun	39,78
-------------------------------	-------

Tabulka A.8: Pozice webů na dané klíčové fráze ve vyhledávači Google před a po získání zpětného odkazu ze služby Linkuj.cz

Klíčová fráze		*.own.cz	*.chytry.cz	*.webovka.cz	*.vyjimecny.cz	*.tym.cz
doprava na Krétě	před	43	24	112	44	18
	po	144	54	251	181	137
	posun	-101	-30	-139	-137	-119
tržiště v Chanii	před	7	6	3	4	2
	po	17	6	5	10	9
	posun	-10	0	-2	-6	-7
krétské tržiště	před	6	9	4	5	3
	po	14	8	5	11	7
	posun	-8	1	-1	-6	-4
krétské letiště	před	6	9	4	5	3
	po	50	11	14	39	29
	posun	-44	-2	-10	-34	-26
krétské jídlo	před	16	19	13	15	10
	po	60	19	23	57	37
	posun	-44	0	-10	-42	-27
krétské autobusy	před	25	29	13	20	14
	po	79	29	25	70	54
	posun	-54	0	-12	-50	-40
taxi na Krétě	před	15	30	<i>x</i>	28	21
	po	88	77	<i>x</i>	166	122
	posun	-73	-47	<i>x</i>	-138	-101
krétská dálnice	před	21	6	26	28	4
	po	71	8	36	99	21
	posun	-50	-2	-10	-71	-17
letiště v Chanii	před	42	7	6	8	46
	po	204	10	28	72	140
	posun	-162	-3	-22	-64	-94
krétská kuchyně	před	35	43	29	28	19
	po	226	64	131	211	179
	posun	-191	-21	-102	-183	-160
Celkový posun		-737	-104	-308	-731	-595
Průměrný posun na web		-73,7	-10,4	-34,222	-73,1	-59,5

Celkový průměrný posun	-50,18
-------------------------------	--------

Tabulka A.9: Pozice webů na dané klíčové fráze ve vyhledávací Seznam před a po vytvoření zpětných odkazů mezi testovanými weby

Klíčová fráze		*.own.cz	*.chytry.cz	*.webovka.cz	*.vyjimecny.cz	*.tym.cz
doprava na Krétě	před	neměřeno	x	neměřeno	neměřeno	x
	po	neměřeno	x	neměřeno	neměřeno	x
	posun	neměřeno	x	neměřeno	neměřeno	x
tržiště v Chanii	před	neměřeno	35	neměřeno	neměřeno	30
	po	neměřeno	36	neměřeno	neměřeno	54
	posun	neměřeno	-1	neměřeno	neměřeno	-24
krétské tržiště	před	neměřeno	x	neměřeno	neměřeno	2
	po	neměřeno	445	neměřeno	neměřeno	1
	posun	neměřeno	x	neměřeno	neměřeno	1
krétské letiště	před	neměřeno	13	neměřeno	neměřeno	9
	po	neměřeno	9	neměřeno	neměřeno	409
	posun	neměřeno	4	neměřeno	neměřeno	-400
krétské jídlo	před	neměřeno	467	neměřeno	neměřeno	8
	po	neměřeno	17	neměřeno	neměřeno	10
	posun	neměřeno	450	neměřeno	neměřeno	-2
krétské autobusy	před	neměřeno	x	neměřeno	neměřeno	452
	po	neměřeno	527	neměřeno	neměřeno	x
	posun	neměřeno	x	neměřeno	neměřeno	x
taxi na Krétě	před	neměřeno	x	neměřeno	neměřeno	26
	po	neměřeno	383	neměřeno	neměřeno	43
	posun	neměřeno	x	neměřeno	neměřeno	-17
krétská dálnice	před	neměřeno	133	neměřeno	neměřeno	5
	po	neměřeno	40	neměřeno	neměřeno	51
	posun	neměřeno	93	neměřeno	neměřeno	-46
letiště v Chanii	před	neměřeno	x	neměřeno	neměřeno	x
	po	neměřeno	432	neměřeno	neměřeno	46
	posun	neměřeno	x	neměřeno	neměřeno	x
krétská kuchyně	před	neměřeno	x	neměřeno	neměřeno	14
	po	neměřeno	67	neměřeno	neměřeno	x
	posun	neměřeno	x	neměřeno	neměřeno	x
Celkový posun		neměřeno	546	neměřeno	neměřeno	-488
Průměrný posun na web		neměřeno	136,5	neměřeno	neměřeno	-81,333

Celkový průměrný posun	27,58
-------------------------------	-------

Tabulka A.10: Pozice webů na dané klíčové fráze ve vyhledávači Google před a po vytvoření zpětných odkazů mezi testovanými weby

Klíčová fráze		*.hu.cz	*.ezin.cz	*.yc.cz	*.kx.cz	*.byl.cz
doprava na Krétě	před	5	neměřeno	neměřeno	neměřeno	neměřeno
	po	10	neměřeno	neměřeno	neměřeno	neměřeno
	posun	-5	neměřeno	neměřeno	neměřeno	neměřeno
tržiště v Chanii	před	11	neměřeno	neměřeno	neměřeno	neměřeno
	po	3	neměřeno	neměřeno	neměřeno	neměřeno
	posun	8	neměřeno	neměřeno	neměřeno	neměřeno
krétské tržiště	před	8	neměřeno	neměřeno	neměřeno	neměřeno
	po	4	neměřeno	neměřeno	neměřeno	neměřeno
	posun	4	neměřeno	neměřeno	neměřeno	neměřeno
krétské letiště	před	7	neměřeno	neměřeno	neměřeno	neměřeno
	po	3	neměřeno	neměřeno	neměřeno	neměřeno
	posun	4	neměřeno	neměřeno	neměřeno	neměřeno
krétské jídlo	před	3	neměřeno	neměřeno	neměřeno	neměřeno
	po	5	neměřeno	neměřeno	neměřeno	neměřeno
	posun	-2	neměřeno	neměřeno	neměřeno	neměřeno
krétské autobusy	před	28	neměřeno	neměřeno	neměřeno	neměřeno
	po	7	neměřeno	neměřeno	neměřeno	neměřeno
	posun	21	neměřeno	neměřeno	neměřeno	neměřeno
taxi na Krétě	před	49	neměřeno	neměřeno	neměřeno	neměřeno
	po	26	neměřeno	neměřeno	neměřeno	neměřeno
	posun	23	neměřeno	neměřeno	neměřeno	neměřeno
krétská dálnice	před	27	neměřeno	neměřeno	neměřeno	neměřeno
	po	17	neměřeno	neměřeno	neměřeno	neměřeno
	posun	10	neměřeno	neměřeno	neměřeno	neměřeno
letiště v Chanii	před	44	neměřeno	neměřeno	neměřeno	neměřeno
	po	7	neměřeno	neměřeno	neměřeno	neměřeno
	posun	37	neměřeno	neměřeno	neměřeno	neměřeno
krétská kuchyně	před	<i>x</i>	neměřeno	neměřeno	neměřeno	neměřeno
	po	116	neměřeno	neměřeno	neměřeno	neměřeno
	posun	<i>x</i>	neměřeno	neměřeno	neměřeno	neměřeno
Celkový posun		100	neměřeno	neměřeno	neměřeno	neměřeno
Průměrný posun na web		10	neměřeno	neměřeno	neměřeno	neměřeno

Celkový průměrný posun	10
-------------------------------	----

Tabulka A.11: Pozice webů na dané klíčové fráze ve vyhledávači Seznam před a po vytvoření zpětného odkazu z katalogu odkazy.seznam.cz

Klíčová fráze		*.hu.cz	*.ezin.cz	*.yc.cz	*.kx.cz	*.byl.cz
doprava na Krétě	před	14	neměřeno	neměřeno	neměřeno	neměřeno
	po	16	neměřeno	neměřeno	neměřeno	neměřeno
	posun	-2	neměřeno	neměřeno	neměřeno	neměřeno
tržiště v Chanii	před	2	neměřeno	neměřeno	neměřeno	neměřeno
	po	3	neměřeno	neměřeno	neměřeno	neměřeno
	posun	-1	neměřeno	neměřeno	neměřeno	neměřeno
krétské tržiště	před	1	neměřeno	neměřeno	neměřeno	neměřeno
	po	3	neměřeno	neměřeno	neměřeno	neměřeno
	posun	-2	neměřeno	neměřeno	neměřeno	neměřeno
krétské letiště	před	5	neměřeno	neměřeno	neměřeno	neměřeno
	po	3	neměřeno	neměřeno	neměřeno	neměřeno
	posun	2	neměřeno	neměřeno	neměřeno	neměřeno
krétské jídlo	před	7	neměřeno	neměřeno	neměřeno	neměřeno
	po	6	neměřeno	neměřeno	neměřeno	neměřeno
	posun	1	neměřeno	neměřeno	neměřeno	neměřeno
krétské autobusy	před	5	neměřeno	neměřeno	neměřeno	neměřeno
	po	7	neměřeno	neměřeno	neměřeno	neměřeno
	posun	-2	neměřeno	neměřeno	neměřeno	neměřeno
taxi na Krétě	před	2	neměřeno	neměřeno	neměřeno	neměřeno
	po	3	neměřeno	neměřeno	neměřeno	neměřeno
	posun	-1	neměřeno	neměřeno	neměřeno	neměřeno
krétská dálnice	před	4	neměřeno	neměřeno	neměřeno	neměřeno
	po	47	neměřeno	neměřeno	neměřeno	neměřeno
	posun	-43	neměřeno	neměřeno	neměřeno	neměřeno
letiště v Chanii	před	<i>x</i>	neměřeno	neměřeno	neměřeno	neměřeno
	po	12	neměřeno	neměřeno	neměřeno	neměřeno
	posun	<i>x</i>	neměřeno	neměřeno	neměřeno	neměřeno
krétská kuchyně	před	220	neměřeno	neměřeno	neměřeno	neměřeno
	po	135	neměřeno	neměřeno	neměřeno	neměřeno
	posun	85	neměřeno	neměřeno	neměřeno	neměřeno
Celkový posun		37	neměřeno	neměřeno	neměřeno	neměřeno
Průměrný posun na web		3,7	neměřeno	neměřeno	neměřeno	neměřeno

Celkový průměrný posun	3,7
-------------------------------	-----

Tabulka A.12: Pozice webů na dané klíčové fráze ve vyhledávači Google před a po vytvoření zpětného odkazu z katalogu odkazy.seznam.cz

Dodatek B

Obsah CD

- soubor *xsimko06.pdf* – tato bakalářská práce ve formátu PDF
- adresář *tex* – zdrojové texty této bakalářské práce
- adresář *position-checker* – zdrojové kódy PHP skriptu pro měření pozic testovaných webů na klíčové fráze
- adresář *weby* – jednotlivé verze webových stránek, tak jak byly v průběhu práce testovány
- adresář *vzorovy-web* – vzorový web s aplikovanými on-page faktory otestovanými v této práci