

TECHNICKÁ UNIVERZITA V LIBERCI
Fakulta mechatroniky, informatiky a mezioborových studií

Studijní program: B2646 Informační technologie

Studijní obor: Informační technologie

**Webová aplikace pro online SEO analýzu webových
stránek**

Web application for online SEO website analysis

Bakalářská práce

Autor:

Tomáš Brzek

Vedoucí práce:

Ing. Mojmír Volf

V Liberci 12. 5. 2012

Prohlášení:

Byl jsem seznámen s tím, že na mou diplomovou práci se plně vztahuje zákon č. 121/200 Sb. o právu autorském, zejména § 60 – školní dílo.

Beru na vědomí, že Technická univerzita v Liberci (TUL) nezasahuje do mých autorských práv užitím mé bakalářské práce pro vnitřní potřebu TUL.

Užiji-li bakalářskou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědom povinnosti informovat o této skutečnosti TUL; v tomto případě má TUL právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Diplomovou práci jsem vypracoval samostatně s použitím uvedené literatury a na základě konzultací s vedoucím bakalářské práce.

Datum

Podpis

Abstrakt

Tato bakalářská práce se zabývá optimalizací stránek pro internetové vyhledávače. Metodika, která se zabývá touto problematikou, se nazývá Search Engine Optimization (SEO). První část se věnuje uvedením do problematiky, stanovením faktorů pro optimalizaci stránek a jejich rozdělení. Další část se věnuje sestavením postupu optimalizace na základě stanovených faktorů. Tento postup se v poslední části implementuje jako samostatná webová aplikace, která analyzuje internetovou stránku a zhodnotí míru optimalizace pro vyhledávače, případně sestaví doporučený postup pro optimalizování stránky.

Abstract

This bachelor thesis is concerned with optimizing sites for search engines. The methodology, which deals with this issue is called Search Engine Optimization (SEO). The first part is dedicated to state the problem, and to identify factors in order to optimize pages and their distribution. Second part is dedicated to optimize the assembly process based on defined factors. This procedure is implemented in the last section as a separate web application that analyzes the website and evaluates the degree of search engine optimization, or establish a recommended procedure for optimizing websites.

Obsah

Prohlášení:	3
Abstrakt	4
Abstract	4
Obsah.....	5
Seznam obrázků	7
1 Teoretická část	8
1.1 Internetový vyhledávač	8
1.1.1 Jak funguje vyhledávač	8
Procházení webových stránek	8
Databáze výskytu slov	9
Indexování	9
1.1.2 Vyhledávače v ČR a ve světě	9
1.2 Internetový katalog.....	10
1.3 Optimalizace stránek pro vyhledávací služby.....	11
1.3.1 Faktory významné z hlediska SEO.....	11
On-page faktory.....	11
Titulek stránky.....	11
Meta tag description	12
Meta tag keywords	12
Struktura URL	12
Nadpisy.....	13
Popis obrázků	13
Navigace a odkazy.....	13
Off-page faktory	13
Zpětné odkazy	13
Anchor text.....	14

Zakázané metody.....	14
Skrytí obsahu	14
Cloaking	15
Link farms	15
Doorway page.....	15
Duplicitní stránky	16
2 Cíl práce.....	17
3 Praktická část – webová aplikace	18
3.1 Použité technologie	18
3.2 Návrh databáze.....	18
3.3 Postup analýzy	20
3.3.1 Načtení stránky	21
3.3.2 Analýza URL.....	22
3.3.3 Kódování stránky.....	22
3.3.4 Načtení a zpracování důležitých elementů	23
Titulek	23
Nadpisy.....	24
Meta značky	24
Odkazy.....	25
Odstavce	25
Obrázky	25
Ostatní	25
3.3.5 Robots.txt.....	26
3.3.6 Sitemap.....	26
3.4 Hodnocení	26
3.4.1 Procentuální hodnocení	27
3.4.2 Bodové hodnocení	28

3.5	Popis aplikace.....	28
3.5.1	Vstupní formulář	28
3.5.2	Stránka s vyhodnocením.....	29
4	Vyhodnocení řešení	31
4.1	Časová a paměťová náročnost.....	31
4.2	Možnosti rozšíření.....	32
5	Srovnání již existujících aplikací	34
5.1	Seo-servis	34
5.2	SEO Workers	35
6	Shrnutí práce	37
	Citovaná literatura	39

Seznam obrázků

Obrázek 1.1:	Podíl vyhledávačů na světě ()	10
Obrázek 1.2:	Mnoho stránek nepoužívá titulek stránky	12
Obrázek 3.1:	ER diagram databáze	20
Obrázek 3.2:	Ukázka vstupního formuláře aplikace)	29
Obrázek 3.3:	Tabulka základních informací stránky při vyhodnocení analýzy	29
Obrázek 3.4:	Ukázka části vyhodnocení analýzy	30
Obrázek 5.1:	Ukázka vyhodnocení nástroje analýzy zdrojového kódu	34
Obrázek 5.2:	Výsledek analýzy síly webové stránky	35
Obrázek 5.3:	Ukázka nástroje SEO Workers	36

1 Teoretická část

1.1 Internetový vyhledávač

Internetový vyhledávač je webová aplikace (služba), která umožňuje vyhledat internetové stránky a dokumenty, které obsahují požadované informace. Vyhledává se pomocí klíčových slov, které charakterizují hledanou informaci, a vyhledávač vypíše seznam odkazů stránek, které hledané informace obsahují. Tyto informace jsou udržovány převážně automaticky, kde si je každý internetový vyhledávač obstarává a indexuje sám.

Cílem internetového vyhledávače je uživateli poskytnout co nejrelevantnější informace, proto jsou stránky v databázi různě hodnoceny.

1.1.1 Jak funguje vyhledávač

Vyhledávač pracuje z velké části automaticky, k tomu využívá desítky až statisíce počítačů. Kvalita vyhledávače je závislá na kvalitě odpovědi, to znamená, že vyhledávač zobrazí co nejrelevantnější odpovědi na hledanou informaci. Proto každý vyhledávač musí měřit kvalitu stránek, které má ve své databázi.

Většina internetových vyhledávačů pracuje následovně:

1. Procházení webových stránek
2. Vytvoření databáze výskytu slov
3. Indexování

Procházení webových stránek

Vyhledávač prochází webové stránky pomocí programu, tzv. vyhledávacího robota (pavouk, crawler, fish atp.), který se pomocí hypertextových odkazů snaží navštívit co největší počet webových stránek na internetu.

Robot začíná s procházením ve vlastní databázi, kde si vybere odkaz stránky, kterou následně navštíví. Každou navštívenou stránku si robot stáhne do svého úložiště, poznamená si její URL adresu, aby stránku neprocházel v jednom cyklu znovu. Ve zdrojovém kódu načtené popř. uložené stránky hledá odkazy ukryté v párových značkách <a> a . Pokud takový odkaz najde, přejde na tuto stránku a celý proces se rekurzivně opakuje.

Jakmile se informace o stránce dostane do hlavní databáze, nastane načítání stránek, tzv. fetching. To obstarává jiný typ robota, který denně zpracuje okolo dvou milionů stránek. Robot odešle dotaz HTTP, odpověď přečte a zpracuje. Mezi nejdůležitější informace patří:

- Hlavička dokumentu
- Redirekce, přesměrování
- Stavové kódy
- Informace o poslední změně na stránce
- Čas, za který je stránka zpracována

Databáze výskytu slov

Stránky, které robot uložil na pevný disk, je nutné zpracovat a vytvořit z ní databázi. Tato databáze v sobě uchovává data o klíčových slovech a URL adresách, ke kterým se tato klíčová slova vážou. Problémem databáze bývá její velikost, protože sekvenční prohledání by trvalo neúměrně dlouho. K optimalizaci hledání v databázi slouží tzv. indexování

Indexování

Indexování databáze napomáhá k rychlému prohledávání databáze potažmo vyhledání požadované informace. Index je tvořen tak, aby na prvních místech odpovědi poskytoval stránky s nejvyšší váhou. Pro výpočet relevance se používají různé algoritmy, které zkoumají stránky z různých úhlů. Významné faktory pro analýzu popíšu později.

1.1.2 Vyhledávače v ČR a ve světě

Na internetu je velké množství internetových vyhledávačů. Uvedu proto ty nejdůležitější.

V České republice je nejpoužívanějším internetovým vyhledávačem Seznam, který dlouhodobě používá více než 60% uživatelů. Významně se však zvyšuje podíl uživatelů vyhledávače Google s 30% podílem vyhledávání v ČR. Třetím největším vyhledávačem je Centrum s pěti procentním podílem, je však na ústupu a používá systém od společnosti Google. Posledním vyhledávačem, který stojí za zmínku je Jyxo, který nemá ani jedno procento.

Ve světě jednoznačně dominuje Google, na druhém místě Yahoo!, který používá vyhledávací systém Bing, třetí je čínský vyhledávač Baidu. Dalšími vyhledávači jsou Bing, Ask a AOL.



Obrázek 1.1: Podíl vyhledávačů na světě ()

1.2 Internetový katalog

Katalog je seznam odkazů na webové stránky, které jsou roztrženy do kategorií a podkategorií. Na rozdíl od vyhledávačů, které stránky do své databáze dávají automaticky, do katalogů převážně vkládají webové stránky jejich majitelé. Editoři katalogů návrhy kontrolují a schvalují k zařazení.

Kvalita katalogu je závislá na počtu odkazů, jejich kvalitě, správného třídění do kategorií nebo počtu kategorií.

Mezi nejznámější katalogy ve světě patří Yahoo!, v Česku například Seznam, Atlas, nebo Centrum.

1.3 Optimalizace stránek pro vyhledávací služby

Metodika pro úpravu webových stránek tak, aby jejich forma a obsah byly pro vyhledávací služby co nejvhodnější, se nazývá Search Engine Optimization (dále jen SEO). Cílem této metodiky je získat ve výsledku hledání ve vyhledávačích pro danou stránku a daná klíčová slova nejvyšší pozici.

1.3.1 Faktory významné z hlediska SEO

Přesné metody a algoritmy pro optimalizaci stránek jsou obvykle tajemstvím každé firmy. Základní mechanismy jsou však veřejně známé nebo odpozorované.

Faktorů pro SEO je velké množství. Základní rozdělení je na on-page a off-page faktory. On-page faktory se nacházejí na stránce a jsou s ní přímo spojeny. Autor stránky může tyto faktory přímo ovlivnit. Naopak off-page faktory se na stránce přímo neobjevují. Autor je sice nemůže přímo ovlivnit, mají však přímý vliv na relevanci stránky.

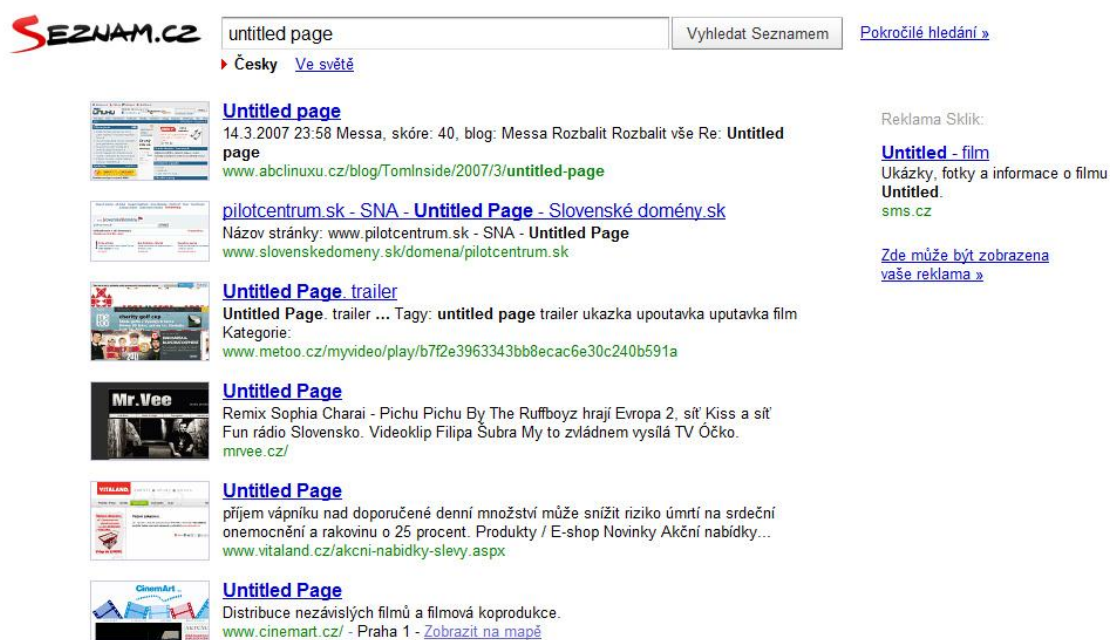
On-page faktory

Jak jsem popsal výše, tyto faktory jsou spojeny přímo se stránkou, lze je tedy ovlivňovat přímo editací webu.

Titulek stránky

Titulek stránky je umístěn v hlavičce stránky v párovém tagu <title>. Titulek je z hlediska vyhledávání nejvýznamnější on-page faktorem po několik let. Je to v podstatě to nejdůležitější, co webová stránka obsahuje. Dalším faktorem, proč je důležité vyplnit titulek, je že mnoho vyhledávačů zobrazují jako titulek vyhledávaného spojení právě slova z titulku stránky. To je například vidět na obrázku 1.3.2. Google se chová tak, že pokud se na stránce titulek nenachází, jako titulek výsledku zobrazí jiný text, často nadpis úrovně h1 (obrázek 1.3.1). Optimální délka titulku je 10-70 znaků. Důležité je také, aby každý titulek v rámci jednoho webu byl unikátní, vždy v závislosti na obsahu.

Internet [Firmy](#) [Mapy](#) [Slovník](#) [Zboží](#) [Obrázky](#) [Videa](#) [Encyklopedie](#)



Obrázek 1.2: Mnoho stránek nepoužívá titulek stránky

Meta tag description

Meta tag description slouží ke stručnému popisu obsahu stránky. Z hlediska SEO není až tak významný, ale jako v případě titulku, především Google používá tento meta tag jako krátký popis pro danou stránku ve výsledcích hledání. Předpokladem pro správné používání tagu description je také jeho unikátnost v rámci jednoho webu.

Meta tag keywords

Tento HTML tag je podobný, tagu description, měl by však obsahovat pouze klíčová slova uvedená v textu stránky. Poslední dobou sice nemá žádnou (Google), nebo malou (Seznam) váhu, přesto ho zde uvádím, protože by každá stránka měla tento tag obsahovat a existují některé vyhledávače, které do svých algoritmů tento tag započítávají.

Struktura URL

URL, celým názvem Uniform Resource Locator („jednoznačný lokátor zdrojů“), definuje doménovou adresu serveru, umístění zdroje na serveru a protokol, kterým je možné zdroj zpřístupnit.

Ideální URL je takové, které se obsahově shoduje se zaměřením webu a je identická s klíčovým slovem. Mělo by být co nejjednodušší a nejsrozumitelnější.

Například URL adresa `http://example.com/show_example.php?category=2&id=145` je pro člověka, nesrozumitelná a neobsahuje klíčová slova. Optimalizovaná adresa by mohla vypadat třeba takto: `http://example.com/seo_examples/search_engine_friendly`. Slova použitá v této URL jsou relevantní k obsahu stránky a struktuře webu.

Nadpisy

Tak jako v knihách, nadpisy jsou velice důležité i na webových stránkách. Vyhledávače jim dávají velkou váhu a jsou důležité pro logické rozdělení stránek. V HTML se používá šest úrovní nadpisů a sice `<h1>` až `<h6>`. `<h1>` dávají vyhledávače nejvyšší váhu, naopak `<h6>` nejnižší. Měli by se používat hierarchicky, tzn. Nadpis `<h2>` by měl být následován `<h3>` a ne `<h4>`.

Popis obrázků

Obrázky jsou nedílnou součástí webových stránek. Při vkládání obrázků je důležité používat atribut `alt`, který by měl obsahovat stručný popis obsahu obrázku. Navíc je tento atribut povinný z hlediska validity.

Navigace a odkazy

Navigace s odkazy velice úzce souvisí. Navigace není nic jiného, než odkazy směřující na stránky v rámci své domény. K tomu, aby robot mohl projít každý ...

Off-page faktory

Tyto faktory označujeme jako vnější vlivy působící na danou stránku. Mezi off-page faktory patří především zpětné odkazy.

Zpětné odkazy

Zpětnými odkazy (tzv. back-links) rozumíme odkazy směřující na daný web z jiných stránek na internetu. Vyhledávače považují zpětné odkazy za pozitivní reference, které jsou doporučovány odkazovaným webem.

Zpětné odkazy se dají získat několika způsoby. Přirozeným způsobem, což znamená, že lidé na stránky odkazují sami díky kvalitnímu obsahu, propagací stránek na sociálních sítích, nebo vložením stránek do katalogů. Obecně se získáváním zpětných odkazů říká linkbuilding.

Anchor text

Anchor text je viditelná část odkazu, na kterou lze kliknout a tím se odkázat na příslušnou webovou stránku nebo dokument. Slova obsažená v anchor textu mohou zvýšit pozici odkazující i odkazované stránky.

Do anchor textu se obvykle zadávají slova, které krátce popisují obsah odkazované stránky. Z toho důvodu začali internetové vyhledávače anchor text využívat k hodnocení stránek. Má to však své nevýhody. Pokud by se na internetu vyskytlo mnoho odkazů na web s anchor textem, který nesouvisí s obsahem stránky, mohlo by to cílený web poškodit. Například společnost Google musela změnit své vyhodnocování právě z cíleného poškozování webů, a dnes se musí anchor text tematicky shodovat s odkazovanou stránkou a pokud se neshoduje, není tento text brán v potaz pro vyhodnocení.

Zakázané metody

Tato část je nazývána Black hat SEO, která je brána jako podvodná. Tyto metody mohou krátkodobě zvýšit rank stránek, ale pokud je vyhledávač objeví, penalizuje tyto stránky. Takto penalizované stránky mohou být dlouhodobě vyřazeny z výsledků vyhledávání.

Mezi zakázané metody patří:

- skrytí obsahu
- cloaking
- link farms
- doorway page

Skrytí obsahu

Jako skrytí obsahu je považována technika, jejímž cílem je zvýšit hustotu klíčového slova v textu stránky. Může mít mnoho podob: malé písmo, stejná barva písma a pozadí apod. Dalším způsobem může být překrytí textu obrázkem.

Tato praktika se často špatně odhaluje, pokud internetový vyhledávač zjistí tuto praktiku, dotyčný web je penalizovaný.

Cloaking

Mezi další zakázané metody patří tzv. cloaking. Jde o metodu, kdy robotům vyhledávačů je nabízen jiný obsah, než klasickým návštěvníkům.

Pokud robot jakéhokoli vyhledávače navštíví webové stránky, je mu podstrčena kvalitní stránka z hlediska SEO faktorů. Tuto stránku si vyhledávač zaindexuje a přiřadí příslušný rank, který je mnohem vyšší. Pokud stránku navštíví běžný uživatel, zobrazí se mu stránka odlišná, často reklamního charakteru s vysokým konverzním poměrem.

Link farms

Za link farmy se považují stránky, které obsahují veliké množství externích odkazů, které jsou na stránce pouze za účelem zvýšení indexu stránky. Linkovou farmou se může stát stránka, která obsahuje i kvalitní obsah, ale jenom její část (například pata stránky) obsahuje nadměrné množství externích odkazů.

Výhodou pro tuto praktiku je, že stránky, pro které jsou link farmy vytvářeny, si mohou pouze polepšit, protože při odhalení této praktiky, je penalizována pouze stránka s odkazy. Kdyby se penalizovali i stránky, na které se odkazuje, začala by se tato technika využívat pro cílené poškozování stránek.

Doorway page

Tato praktika je podobná cloakingu, s drobnými rozdíly. Doorway page je stránka, jejíž obsah je velice optimalizován pro roboty vyhledávačů. Měla by se proto umístit na vysokém místě ve výsledcích vyhledávání.

Taková stránka je však optimalizována pro robota, ne pro uživatele. Využívá se proto přesměrování (většinou pomocí JavaScriptu, který vyhledávače neumí dobře rozpoznat) na cílovou stránku. Občas také tato stránka obsahuje výraznou reklamu, která má cílového návštěvníka na web přivést.

Díky JavaScriptu je velice obtížné na tuto metodu strojově přijít. Občasné reklamy přes celou stránku mohou být pouze obrázky, které roboti nejsou schopni analyzovat. Jedná se však o zjevnou manipulaci s výsledky vyhledávání, proto za její použití hrozí taktéž penalizace.

Duplicitní stránky

Duplicitní obsah vzniká tehdy, pokud více různých adres odkazují na totožný nebo velice podobný obsah. Častým typem duplicitních stránek bývá, když na jeden obsah odkazuje následující čtveřice adres:

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

Z výše uvedených adres si internetový vyhledávač obvykle vybere jen jednu a zbylé ve výsledku vyhledávání nezobrazuje. Někdy navíc ignoruje i zpětné odkazy na ostatní varianty URL a stránka tak ztrácí nárok na body z těchto odkazů.

2 Cíl práce

Cílem bakalářské práce bylo stanovit faktory významné z hlediska optimalizace stránek pro vyhledávací služby. Tyto faktory se mohou u každé vyhledávací službu více či méně odlišovat, obecně se dá však říci, že existují obecné postupy pro získání vyšších pozic ve vyhledávacích službách na internetu.

Tyto postupy bylo potřeba sjednotit a vytvořit z nich univerzální postup pro analýzu faktorů na webových stránkách. Na základě takového postupu vytvořit samostatnou webovou aplikaci, která dokáže tento postup implementovat a analyzovat jakoukoli internetovou stránku.

Aplikace je nástroj, který pomocí algoritmů analyzuje stránku (její zdrojový kód) a odhalí nedostatky, které jsou potřeba změnit. K tomu slouží základní databáze doporučených postupů pro optimalizace analyzované webové stránky.

Existuje sice mnoho nástrojů, které mají velice podobnou funkcionalitu, cílem však nebylo vytvořit takový nástroj, který by přinášel něco nového, ale potvrdit schopnost řešit danou problematiku a dostat se k správnému výsledku řešení.

3 Praktická část – webová aplikace

3.1 Použité technologie

Aplikace je psaná v programovacím jazyce PHP s databázovým systémem MySQL. Použil jsem knihovny z PHP frameworku Nette a databázové abstraktní knihovny Dibi.

3.2 Návrh databáze

Pro návrh databáze byl použit software MySQL Workbench. Tento program od společnosti Oracle je vizuální nástroj pro návrh, správu, vývoj a údržbu SQL databází. Použitá verze 5.2.35CE.

Databáze se skládá z pěti tabulek:

- seo_site
- seo_page
- percentages
- seo_recommendation
- seo_recommendation_site

První tabulka s názvem seo_page obsahuje informace analyzované stránky. Sloupce title, description a keywords obsahují obsah hlavičky stránky, které se zobrazí v horní části vyhodnocení, jako základní informace o stránce. Path, query a hostName popisují url adresu stránky. Dalšími sloupci jsou pRatio a bodyRatio, ve kterých je uložen záznam o poměru výskytu klíčových slov a počtu všech slov v odstavcích respektive v celém těle dokumentu. Zbylé sloupce uchovávají údaje o počtu odkazů na stránce, dále pak sémantické a validační chyby nebo počet některých dalších elementů. Seo_page obsahuje cizí klíč na tabulku seo_site.

Seo_site je tabulka, která se váže k celé doméně, ve které se nachází analyzovaná stránka. Robots je sloupec, ve kterém je informace, zda doména obsahuje textový dokument robots.txt. Sitemap nese informaci, zda jsou na doméně xml dokumenty s mapou celého webu, popřípadě jejich url adresa.

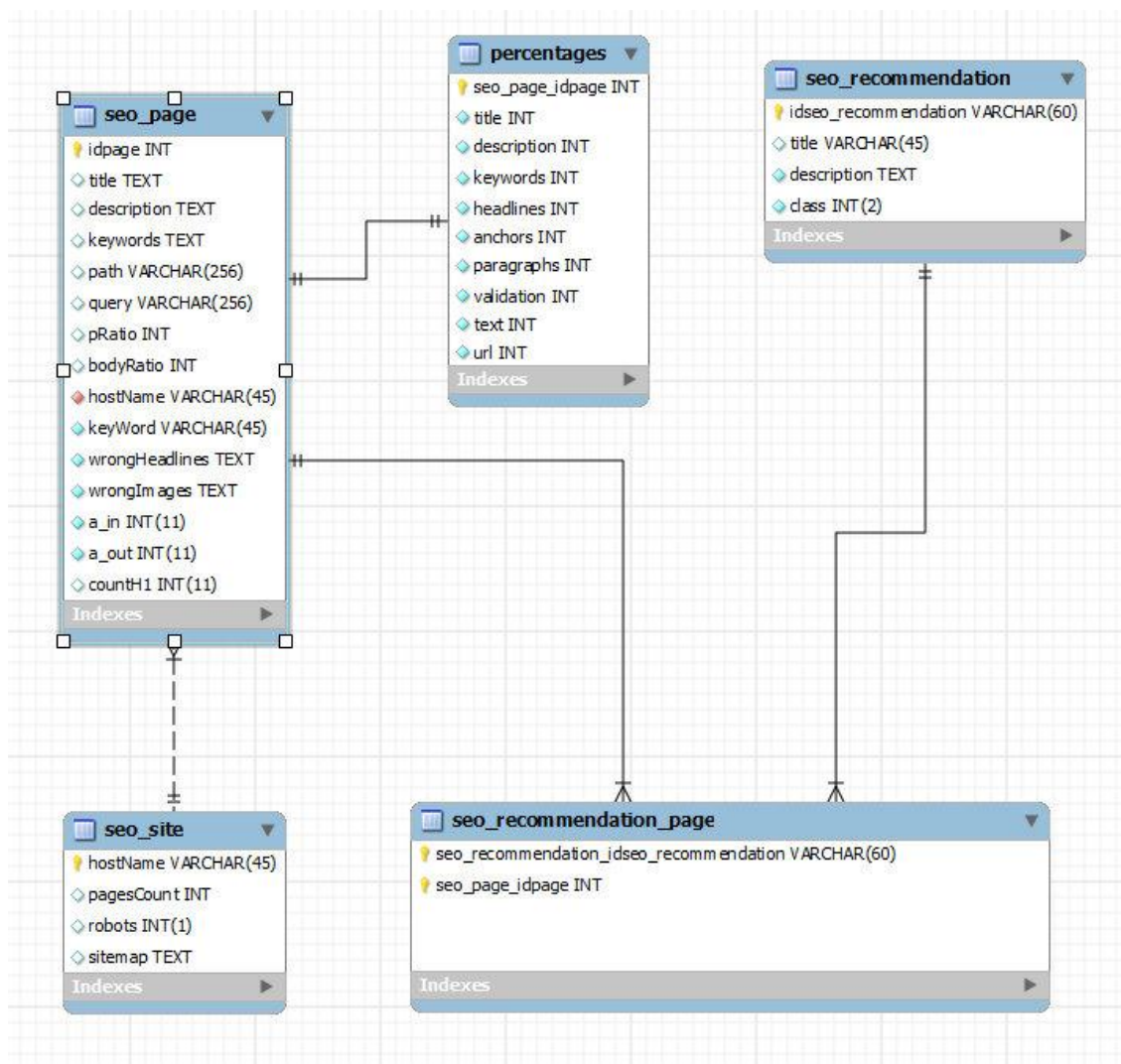
Percentages je tabulka, která nese informace o procentuálním vyjádření hodnocení stránky, uspořádané do následujících skupin:

- title
- description
- keywords
- headlines
- anchors
- paragraphs
- validation
- text
- url
- other

Pro základní databázi doporučení slouží tabulka seo_recommendation. Každé doporučení obsahuje titulek (název), popis, ve kterém je stručně popsáno, jak postupovat, při daném problému, k jeho odstranění. Sloupec class obsahuje třídu (skupinu), do které doporučení spadá.

Poslední tabulkou je relační tabulka seo_recommendation_page, která nese pouze informace, ke které stránce patří některé doporučení.

Schéma databáze s typy sloupců je na následujícím ER diagramu.



Obrázek 3.1: ER diagram databáze

3.3 Postup analýzy

Analýza stránky probíhá s ohledem na faktory, které jsou popsány v teoretické části. Nejdříve bylo potřeba stanovit základní strukturu analýzy, která se postupně rozšiřovala na konkrétní body. Tato základní struktura vypadá následovně:

1. Načtení stránky
2. Analýza URL
3. Analýza kódu
 - a) Vyhledání a uložení důležitých elementů z pohledu SEO
 - b) Zpracování elementů
 - i) Hledání klíčového slova v textu
 - ii) Dodržení HTML pravidel
 - c) Validita kódu

4. Uložení výsledků

Při analyzování stránky je hodnocení rozděleno na dvě části, které se provádějí společně. První je procentuální vyjádření, kde se hodnotí kvalita stránky jako celku i s ohledem na klíčové slovo, druhá je pouze bodové vyhodnocení síly klíčového slova v dokumentu.

3.3.1 Načtení stránky

Pro načtení a zpracování dokumentu bylo na výběr mezi dvěma možnostmi. Vytvoření vlastní knihovny pomocí regulárních výrazů, nebo implementace již hotové knihovny. První možnost pomocí regulárních výrazů by byla časově velice náročná a robustní, takže byla zavrhnuta jako nevyhovující pro tuto práci.

Zbývala možnost implementace externí knihovny. Byla zvolena open source knihovna PHP Simple HTML DOM Parser. Tato knihovna využívá DOM, což je objektové orientovaná reprezentace dokumentu, v tomto případě HTML. Výhodou DOM je v přístupu k dokumentu jako ke stromu. Tato technologie, nazývaná grove, vyžaduje nahrání celého parsovaného dokumentu do paměti, z čehož plyne, že její optimální použití je tam, kde je k jednotlivým elementům dokumentu přistupováno v náhodném pořadí nebo opakovaně. Dokumentace a zdrojové kódy jsou dostupné na adrese <http://simplehtmldom.sourceforge.net/>.

Samotné načtení stránky probíhá tak, že pomocí parseru se stránka načte do proměnné, ve které je uložen ve stromové struktuře. Příklad dokumentu stránky www.seznam.cz:

```
simple_html_dom(21) {
  root => simple_html_dom_node(9) {
    nodetype => 5
    tag => "root" (4)
    attr => array(0)
    children => array(2) [
      0 => simple_html_dom_node(9) { ... }
      1 => simple_html_dom_node(9) { ... }
    ]
    nodes => array(4) [
      0 => simple_html_dom_node(9) { ... }
      1 => simple_html_dom_node(9) { ... }
      2 => simple_html_dom_node(9) { ... }
      3 => simple_html_dom_node(9) { ... }
    ]
  ]
}
```

```
parent => NULL
_ => array(2) [
  0 => -1
  1 => 1859
]
tag_start => 0
dom private => simple_html_dom(21) { *RECURSION* }
}
.
.
.
}
```

Po načtení stránky přichází na řadu samotná analýza dokumentu. K analýze je kromě dokumentu také potřeba klíčové slovo, ke kterému se vztahuje. Toto klíčové slovo zadává, stejně jako adresu stránky, uživatel. Pokud se dále v práci bude vyskytovat pojem klíčové slovo, je tím myšleno právě toto uživatelsky definované.

3.3.2 Analýza URL

Prvním krokem, který aplikace po načtení dokumentu provede, je analýza URL. Tato část není moc složitá, po načtení URL adresy je adresa rozdělena na tři části.

- a) Doména
- b) Adresářová cesta
- c) Dotaz

V každé z těchto částí se vyhledává klíčové slovo. Pokud se klíčové slovo v url nevyskytuje ani jednou, je stránka penalizována 5 procenty. Ideální stav je, pokud se klíčové slovo vyskytuje v URL jednou, a to buď v doménové části, nebo v části s adresářovou cestou, popřípadě jeden výskyt v každé z obou dvou. Další penalizace nastává v případě, že se klíčové slovo vyskytuje v některé části vícekrát.

3.3.3 Kódování stránky

Při vytváření aplikace vznikl problém při načítání dokumentu a jeho dalším zpracováním, pokud dokument nebyl reprezentován ve znakové sadě typu UTF-8. Z toho důvodu byla do aplikace implementována metoda, která vyhledá, ve které znakové sadě byl dokument napsán.

K správnému přiřazení znakové sady slouží meta značka s atributem charset. Tato značka se zapisuje do hlavičky html souboru a může vypadat například takto:

```
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
```

Pokud analyzovaná stránka tuto značku neobsahuje, je zvolena znaková sada UTF-8.

3.3.4 Načtení a zpracování důležitých elementů

Toto je stěžejní část analýzy. Prohledává celý dokument a ukládá si všechny elementy stránky ke zpracování. Dokument se prohledává od začátku, ke konci, kde pro každý element jsou společně uchovávány následující informace:

- Typ elementu
- Originální text
- Počet slov v textu
- Pozice elementu v dokumentu
- Atributy elementu

Pro tyto společné vlastnosti elementů je společná třída s názvem *HtmlElement*. Tato třída obsahuje také metody, které jsou pro všechny elementy společné.

Nejdůležitější z nich je metoda pro hodnocení výskytu klíčového slova v textu elementu.

Titulek

Titulek, jak už bylo popsáno v teoretické části, je jeden z nejdůležitějších on-page faktorů z hlediska SEO. Proto má jako element největší váhu při hodnocení. V procentuálním měřítku se na něj vyhradilo dvacet procent a v bodovém hodnocení tento element dostal váhu 200 bodů. V hodnocení se v titulku stránky zohledňují následující faktory:

- Výskyt elementu
- Délka textu
- Výskyt klíčových slov

Pokud se v dokumentu tento element vůbec nevyskytuje, penalizace je celých 20%. Pokud ano, ale neobsahuje ani jedno klíčové slovo penalizace je 10%. Optimální délka titulku je 10 – 70 znaků, pokud se tato délka nedodrží, stránka je penalizována dalšími pěti procenty. Zbýlých pět procent je vyhrazeno pro opakující se počet

klíčových slov v textu, konkrétně při větším výskytu, než dvou, je strženo dalších pět procent.

Nadpisy

Tato část analýzy je trochu komplikovanější, než hledání klíčových slov v elementu. Důležitý není pouze výskyt klíčových slov, ale také sémantika nadpisů.

Prvním krokem, při analýze nadpisů, je zjištění počtu hlavních nadpisů h1. Tento nadpis by se měl na stránce vyskytovat právě jednou. Pokud analyzovaný dokument tuto podmínku nesplní, dostává penalizaci pět procent.

V sémantické části se kontroluje správné pořadí nadpisů. To znamená, že umístování nadpisů do stránky musí být postupné. Za nadpisem první úrovně by se měli vyskytovat pouze nadpisy úrovně druhé nebo vyšší, v žádném případě ale ne nižší, atp. Správné pořadí nadpisů se kontroluje tak, že se procházejí všechny nadpisy na stránce a pro každý nadpis se kontroluje úroveň následujícího nadpisu, kde rozdíl úrovní nesmí být větší, než plus jedna.

Při nedodržení tohoto pravidla se stránka penalizuje. Penalizace je přímo úměrná počtu špatně umístěných nadpisů, nejvíce však deset procent.

Meta značky

Aplikace se zaměřuje pouze na některé meta značky, konkrétně na dvě. Meta description a meta keywords. Na ostatní meta značky nebere aplikace ohled, protože jsou pouze informativní a z hlediska SEO nejsou důležité.

Description, jak už bylo popsáno v teoretické části, není z hlediska hodnocení stránky vyhledávači tak důležitý, ale vzhledem k tomu, že je často používán jako popis stránky ve výsledku vyhledávání, byl zahrnut do analýzy a bylo na něj vyhrazeno deset procent. Penalizace stránky za meta značku typu description nastává ve třech případech. Pokud stránka tento element neobsahuje, stránka je penalizována deseti procenty. V případě že se na stránce vyskytuje, ale není v něm obsaženo klíčové slovo, penalizace je tři procentní. Poslední případ penalizace je za nedodržení doporučené délky textu, konkrétně při překročení 160 znaků, dalšími pěti procenty.

Meta značka keywords se hodnotí rozdílněji oproti ostatním elementům. Vyhledávače dnes převážně pouze kontrolují při výskytu tohoto elementu na stránce

jednu věc. Všechny slova se musí vyskytovat aspoň jednou v textu dokumentu. V opačném případě je stránka penalizována.

Tato metoda byla implementována do analýzy a při porušení tohoto pravidla je stránka penalizována pěti procenty. Mimo to byla do algoritmu zahrnuta ještě jedna podmínka, maximální počet klíčových slov v elementu. Ten je nastaven na deset, po překročení je další penalizace dvě procenta.

Odkazy

Odkazy se v analýze nehodnotí. Pouze se sleduje výskyt na stránce a uživatel se zobrazí pouze přehled množství odkazů rozdělené na interní a externí zdroje.

Odstavce

V tomto elementu se počítá počet klíčových slov a jejich poměr k celkovému počtu slov v odstavcích.

Poměr se počítá tak, že pro každý odstavec se spočítá počet klíčových slov, který se vydělí počtem všech slov v odstavci. Tyto hodnoty se ze všech odstavců sečtou a vydělí jejich počtem. Tím se získá výsledná hodnota, která je dále hodnocena. Pokud je výsledný průměr menší než pět procent, je penalizace za odstavce v rozmezí od dvou do desíti procent v závislosti hodnotě poměru.

Obrázky

Co se týče obrázků, konkrétně elementů ``, je u nich kontrolováno, zda obsahují povinný atribut `alt`. Tento postup se provádí na základě dvou faktorů.

Prvním z nich je validita stránek. U elementu `` jsou dva povinné elementy, a to `src`, který udává zdroj obrázku, a atribut `alt`, který specifikuje alternativní text. Pokud aspoň jeden z těchto atributů neobsahuje, je tato část kódu nevalidní a internetový vyhledávač validitu kódu sleduje a hodnotí.

Druhým faktorem je, že tento atribut může být jediným zdrojem informací o tom, co se na daném obrázku vyskytuje.

Ostatní

Ve všech ostatních elementech se pouze hledá výskyt klíčových slov, který se započítává do bodového hodnocení stránky.

3.3.5 Robots.txt

Při analyzování stránky se také hledají dokumenty, které se nalézají na jiné adrese, jsou ale společné pro všechny stránky na dané doméně. Prvním z nich je dokument robots.txt. Soubor robots.txt by se měl podle normy vyskytovat na adrese /robots.txt, kde je důležité dodržet malá písmena (např.: www.example.com/robots.txt). Z toho důvodu bylo jednoduché hledat tento soubor pro jakoukoli hledanou stránku. Vždy se hledá v rámci dané domény na adrese /robots.txt. Pokud se na této adrese soubor vyskytuje, načte se a je dále zpracován, v opačném případě se poznamená, že soubor neexistuje.

V souboru robots.txt aplikace hledá odkazy na xml dokumenty sitemap, a zjišťuje, zda není zakázán přístup pro vyhledávače (roboty) pro danou stránku. Algoritmus pro vyhledávání v dokumentu byl použit pomocí regulárních výrazů.

3.3.6 Sitemap

Sitemap je dalším dokumentem, který je společný pro celou doménu. Na rozdíl od robots.txt není dáno, na jaké adrese by se měl vyskytovat. Navíc sitemap nemusí být jeden, ale může jich být několik.

V aplikaci byl zvolený postup pro hledání následující:

- Hledání v dokumentu robots.txt
- Hledání na adrese /sitemap.xml

O prvním způsobu byla zmínka již v kapitole o robots.txt. Odkaz na každý dokument typu sitemap je vždy na novém řádku a začíná řetězcem „sitemap: “. Za tímto řetězcem se nalézá adresa na tento dokument. Pokud se v dokumentu robots.txt odkaz na sitemap vyskytuje, je tato adresa (popř. všechny adresy) uložena do databáze pro další použití.

3.4 Hodnocení

Jak již bylo zmíněno, výsledkem celé analýzy je hodnocení stránky. Udává, jak je optimalizována pro internetové vyhledávače. Je důležité si uvědomit, že nelze jednoznačně určit, zda je stránka plně optimalizována pro konkrétní vyhledávač. V případě, že výsledek analyzované stránky se blíží nebo dosahuje 100%, neznamená to, že již na stránce není z hlediska SEO nic zlepšovat. Každý internetový vyhledávač

má vlastní algoritmy pro hodnocení stránek, které si chrání a jsou pro veřejnost neznámé. Výše popsany postup je zhotoven z doporučení, které poskytují vývojáři těchto vyhledávačů. Z výsledku hodnocení se uživatel dozví, do jaké míry tyto doporučení splňuje.

3.4.1 Procentuální hodnocení

Toto hodnocení se vztahuje k celému dokumentu jako celku. Jsou v něm zahrnuty všechny faktory analýzy. Každá část analýzy má vyhrazenou procentuální hodnotu, ve které se pohybuje. Konkrétní hodnoty pro každou část jsou zobrazeny v následující tabulce.

Tabulka 3.1: Procentuální rozdělení analýzy

Hodnocená část	Hodnota v procentech
Titulek stránky	20
Meta Description	10
Meta Keywords	7
Nadpisy	15
Odkazy	5
Odstavce	10
Validace	10
Text	10
Adresa stránky	10
Ostatní	3

Hodnocení bylo rozděleno podle informací z doporučení o důležitosti jednotlivých částí stránky při SEO analýze. Nejvíce, dvacet procent, získal titulek stránky, který je po několik let nejvýznamnějším on-page faktorem a je proto vyžadováno, aby byl tento element na stránce správně obsažen a nejlépe, aby obsahoval klíčové slovo. I u ostatních částí analýzy byly přidělovány procenta podle významnosti jednotlivých faktorů.

3.4.2 Bodové hodnocení

Bodové hodnocení je vyjádření síly klíčového slova v dokumentu. Hodnotí se podle výskytu klíčového slova v jednotlivých elementech. Každý element má určenou váhu, a podle té se pro každý element na stránce vypočítá tzv. rating. Tento rating se sečte a výsledná hodnota je bodové hodnocení. V následující tabulce jsou vypsány váhy elementů.

Tabulka 3.2: bodové rozdělení hodnocení jednotlivých elementů

Element	Váha
<title>	250
<h1>	100
<h2>	80
<h3>	50
<h4>	30
<h5>	10
<h6>	5
<a>	10
<p>	2
<meta> (keywords)	10
<meta> (description)	20
	3
ostatní	1

Stejně jako při procentuálním hodnocení byla váha elementů sestavena na základě SEO faktorů a doporučení.

3.5 Popis aplikace

Aplikace je klasický webový portál, který se skládá ze dvou stránek. První stránka obsahuje vstupní formulář, druhá vyhodnocení.

3.5.1 Vstupní formulář

Vstupní formulář je jednoduchý html formulář se dvěma vstupními textovými poli a odesílacím tlačítkem.

SEO analyzer



Adresa webu:

Klíčové slovo:

Odeslat

Obrázek 3.2: Ukázka vstupního formuláře aplikace)

Do prvního textového pole se zadává adresa analyzované stránky, do druhého klíčové slovo.

3.5.2 Stránka s vyhodnocením

Po odeslání formuláře se zobrazí pohled s vyhodnocením stránky. V horní části se zobrazuje procentuální výsledek, po něm následuje hodnocení bodové.

Výsledek

78 %

234 bodů

Adresa stránky:	http://trampolinypatrman.cz
Zadané klíčové slovo:	trampoliny
Titulek:	Novinky Trampoliny Patrman, oddíl skoků na trampolíně
Popis stránky:	Trampoliny Patrman, nejen oddíl skoků na trampolíně v Liberci, ale i show, exhibice a vše spojené s trampolínou.
Klíčová slova:	Trampolína, Patrman, Miroslav, Liberec, Skoky, oddíl, skoky na trampolíně, oddíl skoků na trampolíně
Sitemap:	Neexistuje Základní informace
Robots.txt:	Existuje

Obrázek 3.3: Tabulka základních informací stránky při vyhodnocení analýzy

Pod částí s hodnocením se nacházejí základní informace o analyzované stránce:

- Adresa stránky – odkaz na analyzovanou stránku
- Zadané klíčové slovo – klíčové slovo, pro které se provádí analýza
- Titulek – text atributu <title>
- Popis stránky – text meta značky Description
- Klíčová slova – text meta značky Keywords

- Sitemap – informace o xml dokumentu sitemap
- Robots.txt – informace o existenci textového souboru robots.txt

Pod touto tabulkou se nachází výsledek analýzy stránky. Výsledek je rozdělen do částí, kde každá část interpretuje jeden typ elementů, popřípadě část stránky. U každé části je popsáno, do jaké míry je tato část optimalizována. V případě nějakého nedostatku se vypíše doporučený postup pro dosažení lepšího výsledku.

Titulek

Titulek: Novinky | Trampolíny Patrman, oddíl skoků na trampolíně

- ✔ Počet znaků v titulku má doporučený počet.
- ✘ V titulku stránky se nenachází klíčové slovo.

Meta Description

Popisek stránky: Trampolíny Patrman, nejen oddíl skoků na trampolíně v Liberci, ale i show, exhibice a vše spojené s trampolínou.

- ✘ Meta značka description neobsahuje klíčové slovo.

Meta Keywords

Klíčová slova: Trampolína, Patrman, Miroslav, Liberec, Skoky, oddíl, skoky na trampolíně, oddíl skoků na trampolíně

- ✘ V meta značce keywords je mnoho klíčových slov. Maximální doporučené množství je 10 slov.
- ✘ Klíčové slovo není obsaženo v meta značce Keywords.

Obrázek 3.4: Ukázka části vyhodnocení analýzy

4 Vyhodnocení řešení

Aplikace plní úlohu nástroje, který má analyzovat internetovou stránku a na základě SEO faktorů vyhodnotit, jakým způsobem a do jaké míry je optimalizovaná pro internetové vyhledávače.

Pro aplikaci bylo zvoleno webové rozhraní, které je nejjednodušší a nejsnazší řešení pro vytvoření aplikace. Byl použit PHP Framework Nette a pro parsování stránky Simple HTML DOM Parser.

Zaměření aplikace je na on-page faktory, které jsou přímo spojeny se stránkou a dají se ovlivnit editací stránky. Analýza se provádí pro všechny tyto faktory, které jsou popsány v teoretické části.

Byla snaha zaměřit aplikaci také na některé off-page faktory, konkrétně udávat počet zpětných odkazů na danou internetovou stránku. Existuje nástroj od společnosti Google, který vrací počet indexovaných zpětných odkazů ve formátu xml, po podrobném nastudování bylo zjištěno, že tento nástroj je placený a cena je pro účely bakalářské práce vysoká.

4.1 Časová a paměťová náročnost

Po webové aplikace je vyžadována co nejmenší časová a paměťová náročnost z důvodu zátěže serveru. Z toho důvodu se provedly testy.

Byly vybrány čtyři internetové stránky, které mají stejné tematické zaměření. Pro tuto čtveřici byla provedena analýza pro stejné klíčové slovo – gymnastika. V následující tabulce je přehled analyzovaných stránek včetně velikosti každé stránky. Velikostí stránky je myšlena velikost zdrojového kódu stránky na příslušné adrese bez externích souborů, jako jsou kaskádové styly, obrázky apod.

Tabulka 4.1: Přehled testovaných stránek

Adresa stránky	Velikost stránky
http://www.gymlib.cz	10 097 B
http://cs.wikipedia.org/wiki/Gymnastika	44 931 B
http://gymnastikareporyje.cz/	63 047 B
http://gymnastika.cstv.cz/page/100.aktuality/	34 264 B

Testy se prováděly ve dvou prostředích, vývojovém a produkčním. Produkční prostředí obsahuje mnohonásobně vyšší výkon a vyskytuje se na místě s menší internetovou odezvou. Testován byl čas, za který se provedla analýza a velikost alokované paměti. Čas analýzy je čistý čas potřebný k načtení a analyzování stránky bez sestavení a poslání pohledu s vyhodnocením. Alokovaná paměť je část paměti, které si interně vyhradilo PHP pro svoje výpočty a meziukládání. První tabulka ukazuje výsledky měření na vývojovém serveru, druhá na serveru produkčním.

Tabulka 4.2: Měření časové a paměťové náročnosti na vývojovém serveru

Adresa stránky	Čas analýzy	Alokovaná paměť
http://www.gymlib.cz	1 314.9 ms	7.79 MB
http://cs.wikipedia.org/wiki/Gymnastika	1 598.3 ms	9.71 MB
http://gymnastikareporyje.cz/	2 336.0 ms	9.17 MB
http://gymnastika.cstv.cz/page/100.aktuality/	2 169.5 ms	9.28 MB

Tabulka 4.3: Měření časové a paměťové náročnosti na produkčním serveru

Adresa stránky	Čas analýzy	Alokovaná paměť
http://www.gymlib.cz	378.2 ms	13.52 MB
http://cs.wikipedia.org/wiki/Gymnastika	889.1 ms	18.64 MB
http://gymnastikareporyje.cz/	949.3 ms	17.14 MB
http://gymnastika.cstv.cz/page/100.aktuality/	865.7 ms	17.67 MB

Z výsledků je patrný rozdíl ve výkonu obou serverů. V případě vývojového serveru se čas pohyboval v rozmezí 1 – 2,5 vteřin, na produkčním serveru se čas zmenšil v průměru na třetinu. Dále je z výsledků testu patrné, že časová i paměťová náročnost má souvislost s velikostí analyzované stránky. V případě paměti je rozdíl minimální, v případě času analýzy je velikost stránky hlavním faktorem pro dobu výpočtu.

4.2 Možnosti rozšíření

Tato práce je zaměřena na on-page faktory (především praktická část). V teoretické části jsou však popsány další faktory, které obsahuje metodika SEO. Tyto

faktory by bylo vhodné implementovat do analýzy, popřípadě implementovat některé již vytvořené nástroje, které se na tuto problematiku zaměřují.

Další možností je zaměřením se na zakázané metody, za které dávají internetové prohlížeče penalizaci, a tudíž ovlivňují výsledek ve vyhledávání.

5 Srovnání již existujících aplikací

Nástrojů, které se zabývají analýzou webových stránek, je na internetu veliké množství. Některé jsou placené, většina z nich je však bezplatných. V této práci se zaměřím na dva, konkrétně na jeden český a jeden zahraniční.

5.1 Seo-servis

Seo-servis je webový portál zabývající se SEO problematikou. Obsahuje sadu nástrojů pro optimalizaci webových stránek pro vyhledávače. Portál je rozdělen na následující služby:

- Analýza zdrojového kódu stránky
- Analýza klíčových slov stránky
- Pozice ve vyhledávacích
- Analýza síly webové stránky
- Linkbuilding – výměna odkazů
- SEO školení

Z pohledu srovnání s aplikací této práce je nejbližší služba analýza zdrojového kódu. Analýza stránky se zaměřuje na informace v hlavičce, chyby HTML kódu, sémantiku obsahu a přístupnost.

Úvodní strana » Síla webu » Výsledek testu

Síla webu

Adresa: trampolinypatrman.cz

- Datum testování: 11. 05. 2012
- Celková síla webu: 45 %

Webová stránka

Titulek	Novinky Trampoliny Patrman, oddíl skoků na trampolíně
Popis	Trampoliny Patrman, nejen oddíl skoků na trampolíně v Liberci, ale i show, exhibice a vše spojené s trampolínou.
Obsah	Webová stránka obsahuje celkem 312 slov .
Validita	Webová stránka není XHTML 1.1 validní - obsahuje 4 html chyby .



Obrázek 5.1: Ukázka vyhodnocení nástroje analýzy zdrojového kódu

Za zmínku také stojí analýza síly webové stránky. Tato služba je komplexní analýza OnPage i Offpage faktorů, která se zaměřuje na HTML kód, zpětné odkazy a jejich potenciál, pozice ve vyhledávačích, návštěvnost a potenciál domény.



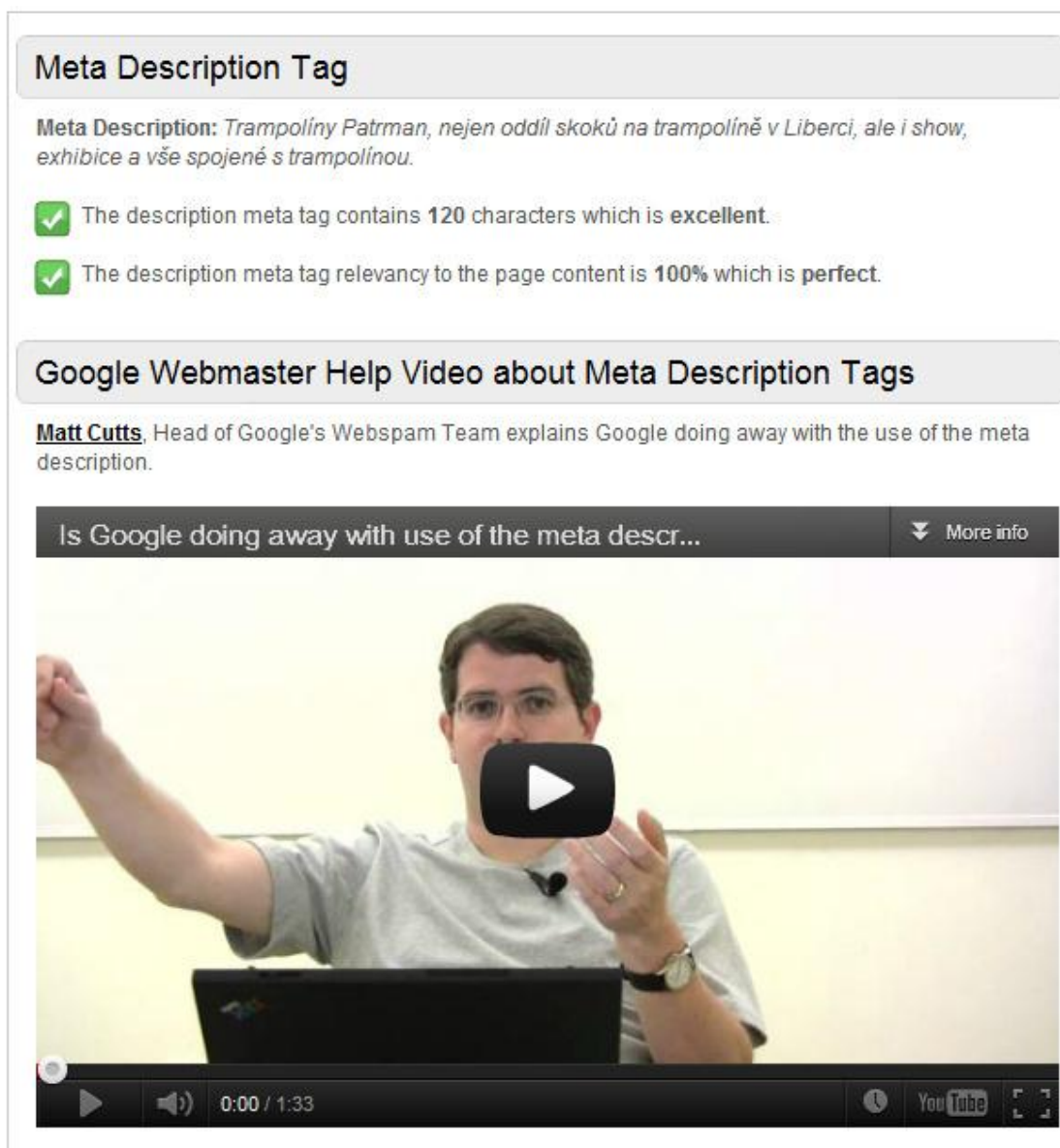
Obrázek 5.2: Výsledek analýzy síly webové stránky

5.2 SEO Workers

Jeden ze zahraničních nástrojů, které se zabývají online SEO analýzou, je SEO Analysis Tool od společnosti SEO Workers. Tento nástroj se zaměřuje pouze na on-page faktory, konkrétně na obsahovou část stránky. Rozbor stránky je rozdělen na následující sekce:

- Kontrola HTTP hlavičky
- Analýza meta tagů
- Relevantnost klíčových slov
- Nalezené odkazy na stránce
- Klíčová slova v odkazech
- Klíčová slova v atributu alt
- Analýza nadpisů a odstavců

Každá část je podrobně vysvětlena a jsou přiložena i videa, která se vztahují k dané problematice.



Obrázek 5.3: Ukázka nástroje SEO Workers

Nevýhodou tohoto nástroje je, že nepodporuje české znaky, což je nepříjemné především v analýze klíčových slov na stránce. Seo analysis tools se vyskytuje na internetové adrese <http://www.seoworkers.com/tools/analyzer.html>.

6 Shrnutí práce

Bakalářská práce se věnuje tématu SEO (Search Engine Optimization), což je metodika vytváření a upravování webových stránek tak, aby jejich forma a obsah byly vhodné pro automatizované zpracování v internetových vyhledávačích. Cílem metodiky je zajistit co nejvyšší příčku ve výsledcích vyhledávačů a tím přivést na webové stránky co nejvíce zákazníků.

V teoretické části jsou definovány základní pojmy v oblasti internetových vyhledávačů a metodice SEO. Na základě prostudovaných literárních zdrojů je zde seznámení s významnými SEO faktory, jejich rozdělení, popř. příklady správného a špatného použití. Tímto tématem se zabývá poměrně mnoho knih, jsou ale často zaměřené na marketingovou část, např.: správný výběr klíčových slov, na kterou se tato práce nesoustředí. Teoretická část tyto faktory seskupuje do celku, který podává komplexní přehled o této metodice a jejích postupech.

Praktická část navazuje na danou problematiku. Zaměřuje se především na on-page faktory, které jsou přímo spojeny se zdrojovým kódem internetové stránky a dají se jednoduše ovlivnit editací stránky. V praktické části je načtena internetová stránka, která je analyzována a zpracována právě na základě faktorů popsanych v teoretické části práce.

Pro analýzu stránky je použita knihovna PHP Simple HTML DOM Parser, která na základě objektového modelu dokumentu dokáže provést rozbor stránky pro další zpracování. Po rozboru stránky jsou načteny a zpracovány důležité elementy. Každý element je zpracován zvlášť a každý typ elementu má jiná pravidla. Stránka je pak hodnocena jako celek, hodnocení se dělí na dvě skupiny, procentuální a bodové. Procentuální se vztahuje k celé stránce v závislosti na klíčovém slově. Udává, v jaké míře jsou splněny on-page faktory ve stránce. Oproti tomu bodové hodnocení udává sílu klíčového slova. Každý atribut je ohodnocen váhou a podle výskytu klíčového slova v různých attributech se vypočítává bodové hodnocení.

Po analýze stránky a vypočítání hodnocení se uživateli zobrazí stránka s vyhodnocením analýzy. Tato stránka obsahuje hodnocení s rozepsáním nejdůležitějších částí stránky. Dále je zde seznam doporučení pro optimalizaci stránky a zlepšení hodnocení.

Na práci lze dále navázat zaměřením na off-page faktory, nebo zakázané metody, které mají vliv na výsledek vyhledávání.

Závěrem lze říci, že výsledky práce lze považovat za poměrně úspěšné a přínosné v oblasti metodiky pro zdokonalování stránek.

Citovaná literatura

Google Inc. 2010. Search Engine Optimization Starter Guide. *Google*. [Online] 2010.

[Cited: Březen 19, 2012.]

http://static.googleusercontent.com/external_content/untrusted_dlcp/www.google.com/en//webmasters/docs/search-engine-optimization-starter-guide.pdf.

Janovský, Dušan. SEO, optimalizace pro vyhledávače. *Jak psát web*. [Online] [Citace:

15. Březen 2012.] <http://www.jakpsatweb.cz/seo/seo.html>.

Kubíček, Michal. 2008. *Velký průvodce SEO*. Praha : COMPUTER PRESS, 2008. 978-80-251-2195-5.

Search Engine Market Share. *Net Market Share*. [Online] [Citace: 16. Leden 2012.]

<http://www.netmarketshare.com/search-engine-market-share.aspx?qprid=4>.

Seznam. Optimalizace webu. *Seznam nápověda*. [Online] [Citace: 22. Březen 2012.]

<http://napoveda.seznam.cz/cz/optimalizace-webu.html>.

(Janovský) (Google Inc., 2010) (Seznam)