

Jihočeská univerzita v Českých Budějovicích

Pedagogická fakulta

Katedra informatiky

**Relevance vyhledávání českých a světových
vyhledávačů a jejich hodnotící algoritmy**

**Web-searching relevance of Czech and World search
engines and their evaluative algorithms**

Bakalářská práce

Miroslav Pešťák

Vedoucí práce:

PaedDr. Petr Pexa

ČESKÉ BUDĚJOVICE 2011

JIHOČESKÁ UNIVERZITA V ČESKÝCH BUDĚJOVICÍCH

Pedagogická fakulta

Katedra informatiky

Akademický rok: 2008/2009

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Miroslav PEŠTÁK**

Studijní program: **B1802 Aplikovaná informatika**

Studijní obor: **Výpočetní technika**

Název tématu: **Relevance vyhledávání českých a světových vyhledávačů a jejich hodnotící algoritmy**

Z á s a d y p r o v y p r a c o v á n í :

Cílem bakalářské práce bude popsat algoritmy hodnocení relevantnosti webové stránky z pohledu vyhledávače a představit problematiku hodnocení stránek v podobě tzv. ranků (Page-Rank, S-rank, JyxoRank, BrowseRank) s hlavním zaměřením na vyhledávače Seznam, Google a Live. Představení dalších vyhledávačů, specializovaných na určitý segment trhu. Porovnání rozdílů vyhodnocování relevantnosti výsledků vyhledávání a to jak pro česky psané webové stránky tak pro celosvětový internet. Hlubší seznámení s praktikami SEO optimalizace pro konkrétní vyhledávače, nastínění profesionálních technik získání vysokého hodnocení od vyhledávačů.

Rozsah grafických prací:
Rozsah pracovní zprávy: 60
Forma zpracování bakalářské práce: tištěná

Seznam odborné literatury:

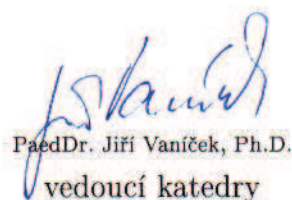
1. Kubíček, Michal. Velký průvodce SEO. Jak dosáhnout nejlepších pozic ve vyhledávačích. Computer Press, 2008.
2. Lapáček, Jiří. Seznam.cz, vyhledávání, e-mail a další služby portálu. Computer Press, 2004.
3. Iskra, Jiří. Google. Vyhledávání, Gmail, Google Talk a další služby. Computer Press, 2006.
4. Smička, Radim. Optimalizace pro vyhledávače-SEO. Jasmínka, Dubany 2006.
5. Müller, Václav. Optimalizace webových stránek. Computer, Press 2007.
6. Domes, Martin. Tvorba WWW pro pokročilé i začátečníky. Computer Press, 2008.
7. Petrželka, Jiří. Web Site Optimisation. Computer Press, 2007.
8. Šustáček, Josef. W3C. Computer Press, 2006.

Vedoucí bakalářské práce: PaedDr. Petr Pexa
Katedra informatiky

Datum zadání bakalářské práce: 1. dubna 2009
Termín odevzdání bakalářské práce: 30. dubna 2010



doc. PhDr. Alena Hošpesová, Ph.D.
děkanka



PaedDr. Jiří Vaníček, Ph.D.
vedoucí katedry

V Českých Budějovicích dne 1. dubna 2009

Prohlášení

Prohlašuji, že jsem předloženou bakalářskou práci vypracoval samostatně, s použitím citovaných literárních pramenů. Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své bakalářské práce, a to v nezkrácené podobě elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách.

V Českých Budějovicích dne 26. dubna 2011

Miroslav Pešťák

Poděkování

Rád bych poděkoval panu PaedDr. Petru Pexovi za odborné konzultace, čas, cenné rady, trpělivost, připomínky a ochotu, s nimiž mě při psaní této práce vedl.

Dále bych rád vyslovil díky všem, kteří mě při psaní této práce všestranně podporovali.

Anotace

Cílem této bakalářské práce bylo porovnat a zhodnotit výsledky vyhledávání českých a světových vyhledávačů na českém internetu. Dále bylo cílem představit porovnávané vyhledávače, jejich funkce, hodnotící algoritmy a tzv. ranky, které napomáhají jednotlivým vyhledávacím robotům k řazení výsledků vyhledávání v SERP, tj. na stránce s výsledky vyhledávání.

Při hodnocení vyhledávání na českém internetu byla použita data z českých vyhledávačů Seznam, Jyxo a Morfeo. Světové vyhledávače byly zastoupeny trojicí vyhledávačů Google, Bing a Ask. Do porovnávání bylo také zařazeno několik menších lokálních a specificky zaměřených vyhledávačů.

Abstract

The goals of my bachelor thesis were the evaluation and comparison of results of internet searching through the Czech and international browsers. The other goal was an introduction of commonly used internet browsers, their functions and evaluation of their algorithms, as well as an introduction of "ranks" whose help browser robots with sorting of results of internet searching in SERP, i.e. at the page with searching results.

For the evaluation of the searching results in the Czech internet were used data originated from the national internet browsers Seznam, Jyxo and Morfeo and international internet browsers Google, Bing and Ask, as well as from few small local very specific internet browsers.

Osnova

1	Úvod.....	7
2	Jak fungují vyhledávače	8
2.1	<i>Crawling – Procházení webů.....</i>	8
2.2	<i>Indexing – Indexování.....</i>	8
2.3	<i>Ranking – Hodnocení.....</i>	9
3	Známé faktory ovlivňující výpočtu relevantnosti	10
3.1	<i>Off-page faktory</i>	10
3.1.1	<i>Zpětné odkazy</i>	10
3.2	<i>On-page faktory</i>	11
3.2.1	<i>Titulek</i>	11
3.2.2	<i>Meta tag keywords a description.....</i>	11
3.2.3	<i>URL adresa webu.....</i>	12
3.2.4	<i>Keyword density</i>	13
3.2.5	<i>Sémantika zdrojového kódu.....</i>	13
3.2.6	<i>Kvalita obsahové části.....</i>	14
3.3	<i>Stáří webu</i>	14
3.4	<i>Dostupnost stránky.....</i>	14
4	Google.....	15
4.1	<i>Historie.....</i>	16
4.1.1	<i>Vznik názvu Google.....</i>	17
4.2	<i>Operátory a speciální znaky ve vyhledávači Google.....</i>	17
4.3	<i>Příkazy vyhledávacího řádku.....</i>	18
4.4	<i>Další skryté funkce Google vyhledávání.....</i>	19
4.5	<i>Google PageRank.....</i>	21
4.6	<i>BadRank.....</i>	23

5	Seznam	24
5.1	<i>Historie.....</i>	24
5.2	<i>Operátory a speciální znaky ve vyhledávači Seznam.....</i>	25
5.3	<i>Příkazy vyhledávacího řádku</i>	25
5.4	<i>S-rank</i>	26
6	Jyxo.....	27
6.1	<i>Historie.....</i>	28
6.2	<i>Operátory a speciální znaky ve vyhledávači Jyxo.....</i>	28
6.3	<i>Příkazy vyhledávacího řádku</i>	28
6.4	<i>JyxoRank</i>	29
7	Bing.....	30
7.1	<i>Operátory a speciální znaky vyhledávače Bing</i>	30
7.2	<i>Příkazy vyhledávacího řádku Bing.....</i>	31
7.3	<i>BrowseRank</i>	32
8	Morfeo	33
8.1	<i>Operátory a speciální znaky v českém vyhledávači Morfeo</i>	33
8.2	<i>Příkazy vyhledávacího řádku</i>	33
9	Ask.....	35
9.1	<i>Příkazy vyhledávacího řádku Ask</i>	35
9.2	<i>Doplňkové funkce vyhledávání Ask.....</i>	36
9.3	<i>ExpertRank.....</i>	36
10	Další vyhledávače	37
10.1	<i>AOL.....</i>	37
10.2	<i>Altavista</i>	37
10.3	<i>DuckDuckGo.....</i>	38

10.4	<i>Lycos</i>	39
10.5	<i>WolframAlpha</i>	39
10.6	<i>Yahoo</i>	42
10.7	<i>Yandex</i>	42
10.8	<i>Zoohoo</i>	43
11	Test vybraných fulltextových vyhledávačů.....	44
11.1	<i>Testování relevantnosti</i>	44
11.1.1	Kritéria testování relevantnosti	46
11.1.2	Kompletní přehled testovaných frází	48
11.1.3	Vyhodnocení testování relevantnosti	51
11.2	<i>Test rychlosti indexování</i>	52
11.2.1	Vyhodnocení rychlosti indexování	55
11.3	<i>Celkové vyhodnocení testování fulltextových vyhledávačů</i>	57
12	Zpracování výzkumu	60
12.1	<i>Reklama ve vyhledávačích</i>	63
13	Závěr.....	67
	Použité zdroje	68
	Přílohy	71

1 Úvod

Hodnotit co je relevantním a co není relevantním výsledkem vyhledávání je složité hlavně z důvodu, že každý vidí relevantnost trochu jinak. V první řadě by neměli relevantnost porovnávat majitelé nebo provozovatelé posuzovaných webů. Zajisté každý provozovatel webu o Praze by chtěl mít svůj web při zadání klíčového slova Praha na prvním místě. V tu chvíli začíná rozepře provozovatelů webů, proč právě jejich web by měl být na prvním místě v daném vyhledávači a o co jsou jejich webové stránky lepší než konkurenční. Do toho vstupuje internetový vyhledávač, který je ten, kdo určuje, který web je relevantnější na zadané klíčové slovo. Toto na první pohled jednoduché rozhodnutí v sobě skrývá složitý proces algoritmizace, který si každý vyhledávač pečlivě střeží. Každý z vyhledávačů má svůj algoritmus, do kterého vstupují stovky proměnných a těm jsou přidělovány různě velké váhy. To celé je poté součástí vzorců každého z vyhledávačů při posuzování relevantnosti jednotlivých stránek a následné řazení výsledků vyhledávání v SERP¹.

Téma relevance vyhledávání českých a světových robotů jsem si vybral, protože mi je hodně blízké. Již několik let se zabývám tvorbou webových stránek jako hobby a poslední dobou se také blíže zajímám o problematiku SEO². Právě díky snaze prohloubit své znalosti SEO optimalizace jsem se snažil více pochopit a porozumět samotným vyhledávačům, ale také uživatelům, pro které jsou webové stránky připravovány.

Cílem této práce je blíže seznámit čtenáře s jednotlivými vyhledávači, představit některé méně známé vyhledávače a následně je podrobit testování. Hlavním prvkem testování bude relevantnost, okrajovou částí hodnocení v testu bude také rychlost indexování, přehlednost a další doplňková funkce jednotlivých vyhledávačů.

¹ SERP – Search engine results page = stránka s výsledky vyhledávání

² SEO – Search Engine Optimization = optimalizace webových stránek pro vyhledávače

2 Jak fungují vyhledávače

Samotnému vyhledávání předchází tři důležité procesy, které musí vyhledávače provést, než mohou zařadit stránku do výsledku vyhledávání a to je crawling, indexing a ranking.

2.1 Crawling – Procházení webů

Robot každého fulltextového vyhledávače prochází přirozeným způsobem internetovou sítí stránek, obsah dané stránky načte (uloží do indexu) a dále přechází z odkazů na stránce na další stránky, ať už na daném webu (interní odkazy) nebo na jiném (externí odkazy).

Díky přirozenému propojení webů pomocí zpětných odkazů (prolinkování) se robot na stránku po nějaké době vrací (doba návratu na stránku je závislá na počtu zpětných odkazů směřujících na danou stránku). Při další návštěvě robot zkontroluje obsah a porovná ho s obsahem, který zaindexoval při poslední návštěvě. Robot hledá změny na současné stránce oproti původnímu obsahu a o každé další návštěvě si vede záznam. Pokud nastaly změny v obsahu robot je zaindexuje, pokud ne, pouze přidá záznam do tabulky o poslední návštěvě stránky. Pokud z nějakého důvodu robot nemůže načíst obsah stránky, ať už byla stránka smazána nebo je jen nedostupný webhostingu, nemohou být stránky indexovány a hodnoceny, a tedy se ani nebudou zobrazovat ve výsledcích vyhledávání.

2.2 Indexing – Indexování

Během procesu procházení webu se každá navštívená stránka robotem ukládá do databáze vyhledávače, probíhá tzv. indexování. Obsah stránek se analyzuje a shromažďuje k následnému rychlému vyhledávání.

Z toho plyne, že při vyhledávání není zpracováván okamžitý obsah stránky, ale poslední dostupný (indexovaný) obsah. U velkých portálů a v poslední době také u sociálních služeb, dochází k tak častému indexování (v řádech minut), že můžeme nabývat představ, že dochází k online vyhledávání.

2.3 Ranking – Hodnocení

Konečný kritický aspekt vyhledávání je způsob, jakým vyhledávač rozhodne o nejrelevantnějších výsledcích vyhledávání. Pokud hledáme aktuální informace, zajisté bude mít větší hodnocení článek aktuální jen několik hodin, než článek několik měsíců starý. U obecných témat bude zase kladen větší nárok na kvalitu webu, na němž je stránka umístěna, tématické zaměření celého webu, množství externích i interních odkazů nebo délka textu.

Hodnocení je neveřejná a utajovaná část celého algoritmu každého z vyhledávačů. Přičemž některé faktory jsou přesto známé, nebo praxí prověřené. Mezi známé faktory patří množství zpětných odkazů z nejlépe hodnocených stránek s příbuzným tématickým zaměřením.

3 Známé faktory ovlivňující výpočtu relevantnosti

Fulltextové vyhledávače k řazení výsledků vyhledávání používají jednak své hodnotící algoritmy (např. PageRank nebo S-rank), ale také kombinaci off-page a on-page faktorů každé stránky, jejichž přesné použití a váhy žádný z fulltextových vyhledávačů veřejně neuvádí.

3.1 Off-page faktory

Off-page faktory jsou elementy ovlivňující kvalitu webové stránky, které majitel z větší části nemůže ovlivnit. Poznámka „z větší části“ je myšlena tak, že provozovatel nebo majitel webu může mít rozsáhlé portfolio webových stránek (až v řádech tisíců), a tak je schopen si sám vybudovat stovky zpětných odkazů s nejlepší možnou relevantností obsahové části.

3.1.1 Zpětné odkazy

Zpětné odkazy jsou nejstarším známým faktorem, který ovlivňuje výpočet relevantnosti daného webu. Zpětné odkazy berou v potaz všechny známé fulltextové vyhledávače, ať české nebo zahraniční. Každý ovšem dává zpětným odkazům různé váhy. A tak je známo, že například Google postupem času dává množství zpětných odkazů menší váhu, více se dívá na kvalitu odkazující stránky a také na umístění odkazu na stránce, která na daný web odkazuje. Touto skutečností se snaží zamezit automatickému a hromadnému vyměňování zpětných odkazů, v patičkách webových stránek. Webové stránky plné „patičkových odkazů“ jsou hlavně z důvodu penalizací od Googlu minulostí.

Například vyhledávač Seznam má s odhalením těchto nekalých SEO praktik stále problém. A tak pořízení velkého množství zpětných odkazů, v řádech stovek až tisíců odkazů, lze dostat webovou stránku na první místa v Seznam vyhledávání během relativně krátké doby. Tento fakt dokazují webové stránky, které se v některých případech objevují ve výsledcích vyhledávání. Takové stránky jsou tvořeny za účelem prodeje zpětných odkazů, nemají přínosné informace pro uživatele, přesto jsou z pohledu vyhledávače kvalitní, jelikož mají vybudovaný velký počet zpětných odkazů a kvalitní vnitřní odkazovou síť.

U zpětných odkazů hraje roli jejich počet, text odkazu, atribut title, umístění odkazu na stránce, tematičnost textu obklopující odkaz, počet všech odkazů na stránce, délka umístěného odkazu i stáří webu, tématická shodnost nebo příbuznost odkazující stránky a v neposlední řadě také možná penalizace webu. Z toho plyne jednoduché pravidlo, buďte zpětné odkazy, které povedou z kvalitních webových stránek zaměřených na stejné téma jako vaše stránka, s klíčovým slovem v názvu odkazu i jeho popisu. Odkaz by měl být umístěn v textu (nikoli v patičce nebo v postranním sloupci o samotě) co možná nejvýše na stránce a ze stránky by mělo vést co možná nejméně dalších odkazů.

3.2 On-page faktory

On-page faktory jsou faktory webové stránky, které ovlivňují vyhledávače a webmaster tyto faktory může ovlivňovat. On-page faktory patří mezi neefektivnější a vyhledávače jim dávají vysokou váhu, proto jsou také vnímány odbornou veřejností jako nejdůležitější prvky SEO optimalizace.

3.2.1 Titulek

Titulek je jedním z nejdůležitějších SEO prvků webové stránky. Právě v titulcích by mělo být obsaženo klíčové slovo nebo fráze, pro kterou mají být stránky optimalizovány. Vyhledávače v SERP zobrazují titulek každé nalezené stránky a ten musí být tedy nejen vhodně optimalizovaný pro vyhledávače, ale také věcný a stručný, aby uživatel na první pohled mohl odhadnout, co se na dané stránce nachází. Každá stránka na webu by měla mít unikátní titulek, nikoli stejný titulek pro více stránek, například název společnosti ve všech titulcích webu. Jeho délka by neměla překročit 64 znaků.

3.2.2 Meta tag keywords a description

Meta tagy byly důležitými prvky pro vyhledávače počátkem 90 let, od té doby hodnota meta tagů klesla na minimum. Meta tag keywords již nemá u většiny vyhledávačů žádnou váhu a to především díky možnému snadnému ovlivňování (spamování) výsledků vyhledávání.

Meta tag description má také menší váhu, ale stále je to jeden z faktorů, který ovlivňují vyhledávače. Například vyhledávač Google v SERP zobrazuje právě

description jako popis výsledku zobrazených stránek. Parametrem tagu description by měl být výstižný popis jednotlivé stránky s doporučenou délkou maximálně 160 znaků včetně mezer (ve výsledcích vyhledávání se delší text nezobrazuje). Description musí být pro každou stránku originální, nikoliv univerzální pro celý web podobně jako titulek stránky. Každý popis by měl obsahovat daná klíčová slova na stránce, pro kterou je optimalizována. Popisek může v podstatě nahrazovat tag keywords s tím rozdílem, že klíčová slova jsou „zabalena“ v kontextu, nikoli jen jednotlivě oddělena čárkou. [1]

„The meta description tag serves the function of advertising copy, drawing readers to a website from the results and thus, is an extremely important part of search marketing. Crafting a readable, compelling description using important keywords can draw a much higher click-through rate of searchers to the given web page. To maximize click-through rates on search engine result pages, it's important to note that Google and other search engines bold keywords in the description when they match search queries.“ [1]

3.2.3 URL adresa webu

Jasně stanovisko k URL adresám webů není, každý vyhledávač se k tomuto faktoru staví jiným způsobem. U převážné většiny vyhledávačů se doporučuje používat klíčová slova v URL adrese, ovšem efekt tohoto faktoru je již u každého vyhledávače rozdílný. Pokud se webová adresa skládá z více slov, je vhodné je oddělit pomlčkou, protože ne všechny vyhledávače jsou schopny rozeznat slova v adrese bez jakéhokoliv oddělení správně. Nedoporučují se oddělovat víceslovné adresy podtržítkem, protože dvě slova spojená podtržítkem považuje většina vyhledávačů jako celek, tedy jako jedno slovo.

Jistou roli při určování relevantnosti hraje také délka url adresy, struktura a doména prvního řádu. To vychází ze skutečnosti, že například domény .cz, .com nebo .net jsou komerční a může si je zaregistrovat každý, ale generické domény jako .edu nebo .gov může vlastnit pouze ten, kdo má speciální práva (vzdělávací instituce v USA používají .edu, vláda USA doménu .gov), a proto weby na doménách prvního řádu tohoto typu mohou být vyhledávači považovány za důvěryhodnější, a tedy lépe hodnocené. Česky psané webové stránky na doméně .cz (která je pro české webové stránky určena) tak mají přednost před ostatními doménovými koncovkami.

3.2.4 Keyword density

Keyword density nebo-li hustota klíčových slov na stránce vyjadřuje procento četnosti výskytu klíčového slova nebo fráze na webové stránce v porovnání s celkovým počtem slov na stránce. Hustota (četnost) klíčového slova na stránce je neoptimálnější 3-4%, častější výskyt klíčových slov na stránce může být vyhledávači brán jako spam, hodně nízký výskyt klíčového slova zase jako klíčové slovo s nerelevantním s obsahem.

Například webová stránka o tiskárnách by měla v textu obsahovat několikrát slovo tiskárna, slova příbuzná či odvozená (například „náplně do tiskáren“, „inkoustové tiskárny“ atd.). Vše by mělo být v kontextu s obsahem celého webu, nikoliv jen umístěno na jedné stránce bez dalšího obsahu.

3.2.5 Sémantika zdrojového kódu

Sémantika se často spojuje s tvorbou webů pro zrakově postižené uživatele. Hlavním důvodem jsou nároky softwarových čteček pro nevidomé uživatele. Podobně jako čtečka pro nevidomé uživatele vidí webovou stránku také indexovací robot. Z tohoto důvodu je nevhodné používat pouze divů nebo spanů k zvýrazňování struktury webů. Naopak je doporučováno využívat všech standardních tagů pro formátování a zobrazování prvků na webové stránce.

Pravidlem pro sémantický web je dodržování standardů W3C³. Pro každý prvek stránky je vhodné volit takový tag, který je pro danou část textu vhodný. Různé úrovně nadpisů se musejí vyskytovat v tagách h1 až h6 a ne využívat formátování divů pomocí css, které následně vypadají jako nadpisy. Zvýrazňování v textu provádíme pomocí tagu nebo nikoli pomocí tagu text rozdělujeme do přiměřeně velkých odstavců pomocí tagu <p>, pro odsazení textů se nemá používat
. Tabulkové tagy jsou určeny pro tabulky nikoli pro rozvržení stránky, proto používání tabulkového layoutu není správné. Všechny parametry týkající se vzhledu by měly být umístěny odděleně od zdrojového kódu v kaskádových stylech. Sémantikou se více zabývá například český web

³ World Wide Web Consortium – mezinárodní konsorcium, které vyvíjí webové standardy

www.semantika.name, kde je uvedeno množství příkladů a značek, které by měl v sémantický web obsahovat.

3.2.6 Kvalita obsahové části

Kvalita obsahové části patří také mezi důležité faktory, které vyhledávače při určování relevantnosti a hodnocení využívají. Pro webmastera je to nejspíše ovlivnitelný faktor. Nejen Google ve svých návodech a doporučení pro webmastery uvádí: „Vytvořte užitečný a informačně bohatý web...“ [2]. Jak je samotná kvalita webových stránek určována již neuvádí.

3.3 Stáří webu

V reálném životě dáváme za pravdu starším, kteří mají více zkušeností, jsou důvěryhodnější a spolehlivější. Z této zažité zvyklosti čerpá i řada vyhledávačů a při hodnocení je toto další z faktorů. Stáří domény si mohou nejen vyhledávače zjistit dle záznamů WHOIS⁴, kde je kromě data registrace a expirace uveden také majitel a registrátor domény.

Starší weby jsou tak kvalitnějším zdrojem stálých a neměnných informací například historických dat. Nové webové stránky zase mohou být aktuálnější tedy relevantnější na časově citlivé dotazy, například na vyhledávání aktuálních sportovních výsledků.

3.4 Dostupnost stránky

Častá nedostupnost webu, nutnost instalace plug-inu či některý z chybových kódů jsou důvodem vyřazení stránky z indexu, a tím i její nedostupnosti ve výsledcích vyhledávání.

„Nedostupnost může být také způsobena používáním URL redirektů, které vyhledávací roboti nemohou následovat, skrýváním obsahu za select formuláře, javascript nebo další pro roboty nepřekonatelnou formu navigace.“ [3]

⁴ Z anglického „who is“ – je databáze informací o doménových jménech

4 Google

V současné době je Google „jedničkou“ na trhu v celosvětovém měřítku vyhledávání. V České Republice je pak na místě druhém za českým fulltextovým vyhledávačem Seznam.cz. Google umožňuje vyhledávání na internetu jak v psané podobě, tak i mezi obrázky, videi, ve zprávách a v reálném čase mezi mikroblogy⁵. Mimo vyhledávání nabízí také více než desítku služeb, ve kterých z části uplatňuje své vyhledávací technologie. Patří sem například následující služby:

- **Gmail** – Nabízí emailovou schránku o velikosti více než 7,5GB zdarma, v rámci vyhledávání v gmailu můžete využít vyhledávacích služeb Google.
- **Google maps** – Vyhledávání v mapách, propojeno také se službou StreetView, která umožňuje virtuální prohlídky měst a Google Earth, který umožňuje prohlížení Země formou virtuálního glóbusu.
- **Google překladač** – Internetový překladač umožňující překlad mezi 57 různými jazyky, umí překládat jednotlivá slova, dlouhé texty i celé webové stránky.
- **Google kalendář** – Osobní webový kalendář, umožňuje přidávání jednorázových úkolů i úkolů s opakováním. Upozornění lze zasílat na email i na mobilní telefon prostřednictvím textové zprávy.
- **Picasa** – Webová fotoalba, které je možné spravovat přes freewarový počítačový program s možností editace jednotlivých fotek, samotný software pak nabízí řadu dalších funkcí a služeb.
- **Dokumenty Google** – Umožňují tvořit online dokumenty, tabulky, prezentace, kresby nebo formuláře. Dokumenty mohou být následně sdíleny pro prohlížení nebo pro možnou spolupráci při jejich vytváření a upravování. Online dokumenty je možné stáhnout do PC a editovat

⁵ Mikrblog je obdobou klasického blogu s rozdílem omezené délky, nejčastěji na 160 znaků, mezi neznámější mikrblogovací systém patří Twitter.

v kancelářských aplikacích a také naopak nahrát dokumenty vytvořené v kancelářských aplikacích do online dokumentů Google.

- **YouTube** – Služba umožňuje sdílet video, které může být nahráváno v plné HD kvalitě o délce 15 minut. Experimentálně je povoleno nahrávání i delších videí. YouTube nabízí také od počátku dubna 2011 živé přenosy vybraných příležitostí (sportovních utkání, prezentací atd.). Do budoucna se počítá s masovějším obsahem živého vysílání z více zdrojů.
- **Google Store** – Internetový obchod převážně s reklamními produkty společnosti Google.
- **Google Scholar** – Speciální vyhledávání v obsahu akademických prací, prohledává plné texty dokumentů, umožňuje omezit výsledky vyhledávání podle autora, data publikace a dalších možností vyhledávání.
- **AdSense / AdWords** – reklamní systémy, které spojují inzerenty (AdWords) a majitele stránek (AdSense). Inzerentům umožňují inzerovat na velkém počtu webových stránek, které jsou zapojeny do reklamního programu AdSense a majitelům webových stránek v systému AdSense umožňují získat příjem z kontextové i bannerové reklamy zobrazené na webu.
- **Blogger** – Systém pro tvorbu jednoduchých webových stránek bez znalosti programování. K psaní je využit WYSIWYG⁶ editor, který převede text a obrázky od uživatele do formy webové stránky.

4.1 Historie

Za vznikem společností Google od počátku stojí Larry Page a Sergey Brin, studenti ze Standfordské univerzity, kteří v roce 1995 přišli s nápadem organizovat informace v rámci World Wide Web. Zprovoznili první verzi vyhledávače s názvem BackRub, pojmenování Google vzniklo později (1997).

⁶ WYSIWYG – je zkratka anglické věty „What you see is what you get.“.

4.1.1 Vznik názvu Google

Slovo Google vzniklo překlepem ze slova "googol", což je číslo zapisované pomocí číslice 1, následované stovkou 0 (10^{100}). Tento název vybral Larry Page a v září roku 1997 byla zaregistrována doména google.com. Larry Page při dotazu, proč vybral jméno Google, řekl:

" 10^{100} (a gigantic number) is a googol, but we liked the spelling "Google" better. We picked the name "Google" because our goal is to make huge quantities of information available to everyone. And it sounds cool and has only six letters. ".[4]

4.2 Operátory a speciální znaky ve vyhledávači Google

Díky operátorům a speciálním znakům lze dosáhnout přesnějších výsledků vyhledávání, přesněji specifikujeme vyhledávači to, co hledáme. Jednotliví operátoři se dají také kombinovat, a tím ještě přesněji specifikovat vyhledávanou frázi. Google podporuje následující operátory:

- " " uvozovky ohraničují přesnou frázi, která musí být obsažena ve výsledku hledání
- + slovo za znakem plus musí být obsaženo ve výsledku hledání
- - slovo za znaménkem mínus nesmí být obsaženo ve výsledku hledání
- * hvězdička může nahradit slovo, které v přesné frázi neznáte, nebo může být různé
- .. dvě tečky slouží ke stanovení číselného rozsahu, například letopočtu
- ~ vyhledává synonyma k zadanému slovu nebo frázi. Operátor fungoval pouze v anglickém vyhledávání.
- **OR** pokud mezi slovy necháme mezeru, Google upřednostní stránky s oběma výrazy. Použijeme-li logickou spojku OR nebo znak |, seřadí se stránky podle relevance samostatných výrazů.
- **AND** logická spojka mezi slovy způsobí, že se zobrazí jen ty stránky, na kterých budou nalezena obě slova. Dá se nahradit umístěním znaku + před obě hledaná slova.

4.3 Příkazy vyhledávacího řádku

Googlu můžete také přikazovat, kde má vyhledávat, a tím ještě přesněji specifikovat vyhledávání. Můžete k tomu používat následující příkazy, některé však v českém vyhledávání nepracují zcela správně.

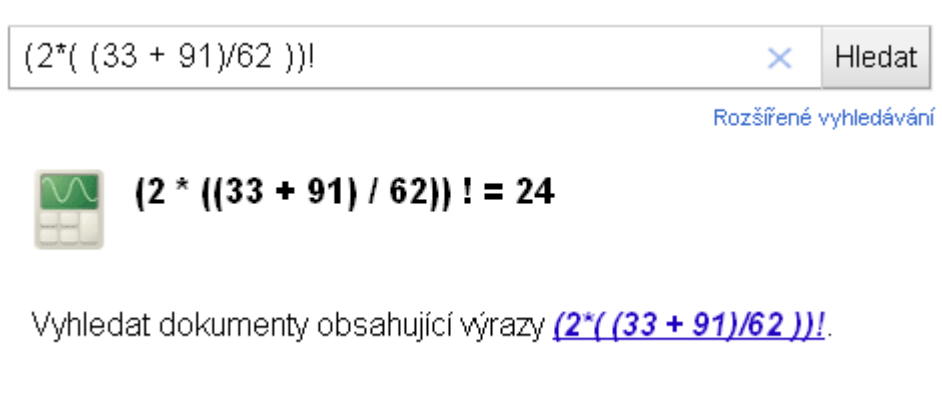
- **allintitle:** *hledaný výraz* – Hledání pouze v titulcích stránek bez ohledu na pořadí slov.
- **cache:** *url_adresa* – Nalezení poslední archivované kopie webu (stejně funguje i odkaz „Archiv“ ve výsledcích vyhledávání).
- **define:** *hledaný výraz* – Nalezne definici hledaného výrazu.
- **filetype:** *koncovka typu souboru* – Vyhledává relevantní soubory zadaného typu dle koncovky souboru.
- **inanchor:** *hledaný výraz* – Vyhledává pouze v textech odkazů.
- **info:** *url_adresa* – Google nabídne další vyhledávání související se zadanou url adresou.
- **intext:** *hledaný výraz* – Vyhledává se jen v textu stránky, titulek stránky je ignorován.
- **intitle:** *hledaný výraz* – Vyhledávání pouze v titulku stránek.
- **inurl:** *hledaný výraz* – Hledání výrazu v URL adresách stránek.
- **link:** *url_adresa* – Vyhledává stránky, ze kterých je odkazováno na zadanou url adresu.
- **related:** *url_adresa* – Vyhledá podobné stránky jako je zadaná adresa. Při testování jsem nebyl o jeho funkci zcela přesvědčen. Výsledky byly podle mého názoru dobré při hledání podobných stránek velkých jasně zaměřených portálů, například při zadání „related:ihned.cz“ byly zobrazeny stránky podobných zpravodajských serverů. Při testování hledání podobných webů pro menší webové stránky již byly výsledky nerovnoměrné. Tato funkce lze také vyvolat kliknutím na odkaz „Podobné“ ve výsledcích vyhledávání.

- **site:** *url_adresa* hledaný výraz – Hledání probíhá jen na stránkách zadané url adresy.
- **-site:** *url_adresa* hledaný výraz – Do výsledků vyhledávání nejsou zahrnuty stránky, které jsou na zadaném webu.

4.4 Další skryté funkce Google vyhledávání

Vyhledávání Google v sobě ukrývá mnohem více funkcí než jen samotné vyhledávání. Pomocí slovních příkazů nebo způsobu zadání můžete ve vyhledávání vyvolat funkce, nebo chcete-li doplňkové aplikace.

- **Kalkulačka** – Zadáním pouhého matematického výrazu a jeho vyhledáním se aktivuje funkce kalkulačky a tak bude příklad ve vyhledávání vypočten. Google provádí matematické úkony (sčítání, odečítání, násobení, dělení, umocňování, druhá odmocnina, goniometrické funkce, logaritmické funkce, procenta a faktoriál). Bezchybně vypočte i složitěji strukturované příklady s více závorkami.



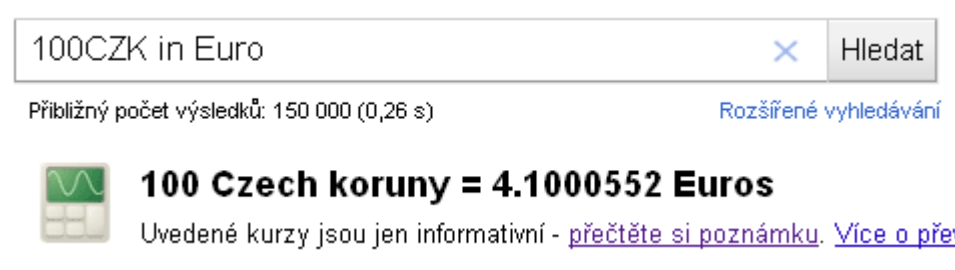
Obrázek 1: Funkce kalkulačky v hledání Google

- **Převodník jednotek** – Vyhledávání Google lze také při správném zadání parametrů použít jako převodník jednotek. Lze převádět délky, objem, váhu a teplotu. Lze použít více způsobů jak samotný převod aktivovat, Google nemá striktně daná pravidla zápisu a tak lze užít několik možností, například: „1km in mile“, „1km na mile“, „1km v mile“, ale i opačně „mile in 1km“ bude vždy výsledek stejný.



Obrázek 2: Funkce převodu délky v hledání Google

- **Převodník měn** – podobně jako převod jednotek je i převod měn intuitivní, není tedy striktně dáno, jak má vypadat zadaný příkaz pro převod. Lze tedy použít například následující příkazy se stejným výsledkem: „100 korun na euro“, „100CZK v EUR“.



Obrázek 3: Funkce převodu měn v hledání Google

- **Předpověď počasí** – Napsáním příkazu „počasí“ případně jeho anglického ekvivalentu „weather“ a jména města (v originálním znění nebo v anglickém překladu) v libovolném pořadí bude zobrazena informace o aktuální teplotě, vlhkosti vzduchu a větru. Funkce je v ČR dostupná pro všechny okresní města



Obrázek 4: Funkce předpověď počasí v hledání Google

- **Sportovní výsledky** – Zobrazení sportovních výsledků je méně známá funkce, vyvolá se zadáním jména týmu, ligy nebo soutěže. Zobrazen je

výsledek posledního zápasu a datum zápasu následujícího, pokud není znám výsledek poslední zápasu, jsou zobrazeny dva následující zápasy.



Obrázek 5: Funkce sportovních výsledků v hledání Google

- **Programy kin** – Příkazem „program kin“ a názvem města uvedete ve funkčnosti další z ukrytých funkcí vyhledávače Google. Tímto příkazem se vám zobrazí 3 filmy, které se budou vysílat v nejkratším možném čase. Odkazem „Více filmů“ si můžete zobrazit seznam všech filmů v jednotlivých kinech až na 4 dny dopředu.



Obrázek 6: Funkce programy kin v hledání Google

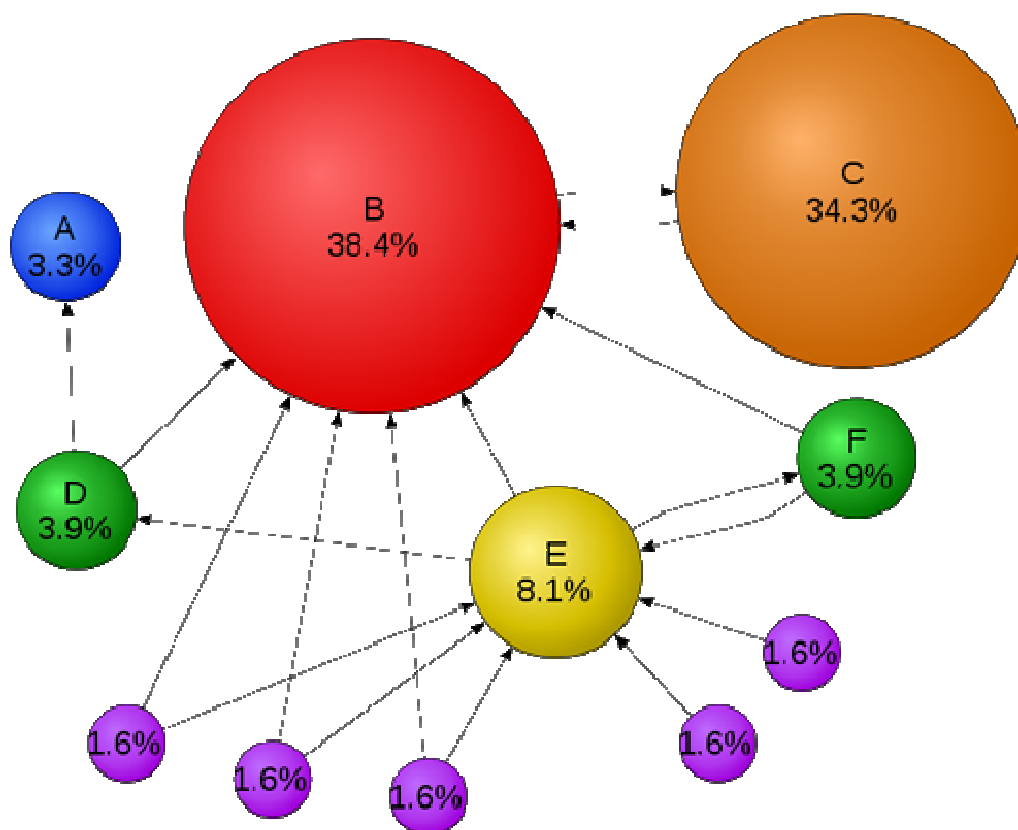
4.5 Google PageRank

PageRank používá Google jako veličinu pro hodnocení kvality webové stránky. Každá webová stránka, kterou Google indexuje má nějaký PageRank. Minimální PageRank (někdy se také používá označení zdrojový nebo přirozený PageRank) je získán při zaindexování stránky. Další nárůst PageRank, který nabývá hodnoty od 0 do 10, ovlivňuje odkazová síť, kvalita obsahu a mnoho dalších faktorů, které nejsou veřejné. Do rovnice výpočtu PageRank vstupuje více než 500 milionů proměnných se 2 miliardami termínů.

„Místo počítání přímých odkazů interpretuje technologie PageRank odkaz ze stránky A na stránku B jako hlas pro stránku B od stránky A. Technologie PageRank

poté vyhodnotí důležitost stránky podle počtu získaných hlasů. Zohledňuje také důležitost každé stránky, která udělila hlas. Hlasy od některých stránek mají větší hodnotu, a odkazovaná stránka tak získá vyšší ohodnocení. Důležité stránky obdrží vyšší ohodnocení PageRank a zobrazí se na začátku výsledků vyhledávání. Technologie společnosti Google používá kurčení důležitosti stránky souhrnné informace webu. Vyhledávač Google analyzuje také obsah stránky. Avšak místo prostého procházení textu na stránkách analyzuje celý obsah stránky a zohledňuje faktory, jako jsou typy a fonty písma, odstavce a přesné umístění každého slova.“ [5]

Složitost výpočtu PageRanku může demonstrovat obrázek 7, kde je vidět vícenásobné přeposílání PageRanku nejen mezi stránkou, která odkazuje na odkazovanou stránku, ale také přeposílání PageRanku ze stránky E na stránku A přes stránku D. (Obrázek 7)



Obrázek 7: Názorný graf přeposílání hodnoty PageRank⁷

⁷ <http://en.wikipedia.org/wiki/PageRank>

4.6 BadRank

BadRank je prakticky sub výpočtem PageRank, jeho výpočet se provádí také podobně jako výpočet PageRanku. Počítá se z odkazové sítě, pro každou webovou adresu zvlášť s určitým útlumem. [6]



Obrázek 8: Znázornění směru přenášení hodnoty BadRank⁸



Obrázek 9: Znázornění směru přenášení hodnoty PageRank⁸

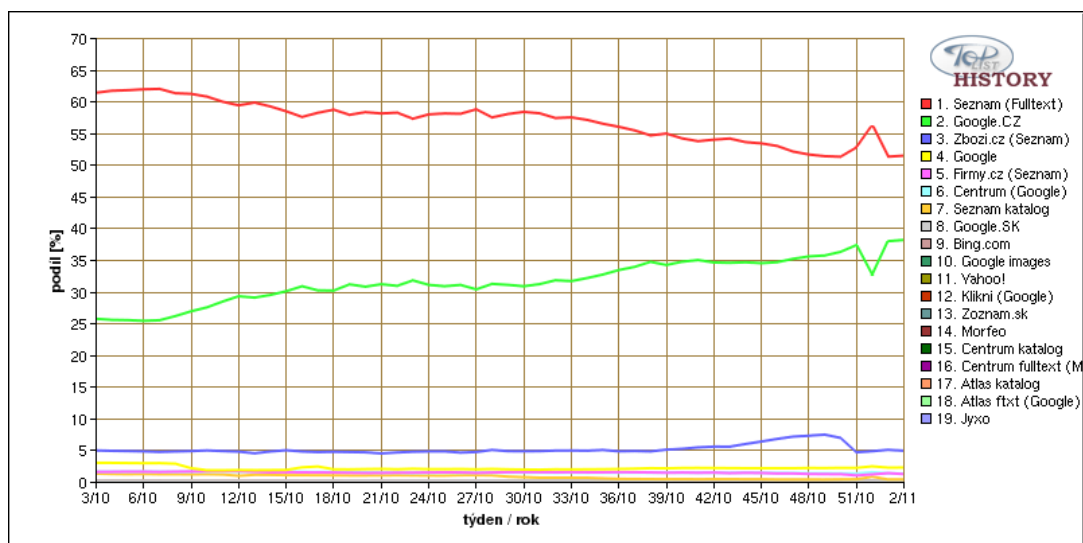
Z obrázku výše vyplývá, že je důležité, na jakou stránku je z webu odkazováno, odkazuje-li webová stránka na takovou stránku, kterou Google označil za špatnou, tj. stránka která, nějakým způsobem manipuluje s výsledky vyhledávání například využitím metody cloaking, nebo nákupem irelevantních zpětných odkazů nejčastěji patičkového typu. (Obrázek 8, Obrázek 9)

Z obrázků je také patrný rozdílný princip a logika algoritmu Google. Pokud odkazujete na kvalitní stránku je část vašeho PageRanku přeposlána na odkazovanou stránku. Odkazujete-li na závadnou stránku, její hodnocení se nezlepší, naopak vaše hodnocení bude horší, protože odkazovaná (závadná) stránka přepoše na vaši stránku BadRank.

⁸ <http://weblog.jakpsatweb.cz/d/1225209900-badrank-seznamte-se.html>

5 Seznam

Seznam.cz je nejstarším českým internetovým vyhledávačem, který v současné době poskytuje více jak desítku doplňkových služeb, nejen samotné vyhledávání. Koncem roku 2008 Seznam dle Financial Times [7] zprostředkovával 63% lokálního vyhledávání v České republice. Podle posledních statistik serveru Toplist.cz (Graf 1) je vyhledávač Seznam.cz stále jedničkou lokálního vyhledávání v České republice. Za poslední rok však ztratil, především díky vyhledávači Google, víc jak 10% z celkového objemu vyhledávání v ČR.



Graf 1: Postavení dominantních vyhledávačů na českém trhu⁹

5.1 Historie

Zakladatelem společnosti Seznam.cz je Ivo Lukačovič, který v roce 1996 spustil, první katalogový vyhledávač Seznam.cz. Seznam zpočátku vyhledával jen ve svém katalogu stránek, který přetrvává dodnes. Byla z něj však odštěpena služba firmy.cz jako katalog firem.

Z počátku využíval Seznam pro fulltextové vyhledávání vlastní technologii zvanou Kompas, následně ji vystřídala outsourcovaná služba od společnosti Empyreum. V roce 2007 byl partnerem a dodavatelem výsledků vyhledávání pro světové vyhledávání Google a Jyxo.cz pro vyhledávání na českém internetu. V roce 2005 spustil Seznam vlastní fulltextový vyhledávač, orientovaný na české

⁹ <http://www.toplist.cz/stat/?a=history&type=4>

a slovenské vyhledávání, ten přetrvává dodnes v pravidelných obměnách. Pro světové vyhledávání používal Seznam stále vyhledávač Google. V září roku 2009 rozvázal Seznam také spolupráci s Google a pro zahraniční vyhledávání začal využívat fulltextový vyhledávač Bing od Microsoftu.

5.2 Operátory a speciální znaky ve vyhledávači Seznam

Díky operátorům a speciálním znakům lze dosáhnout přesnějších výsledků vyhledávání. Jednotliví operátoři se dají kombinovat a tím přesněji specifikovat vyhledávanou frázi. Vyhledávač seznam podporuje následující operátory: [24]

- " " Uvozovky ohraničující přesnou frázi v zadaném pořadí, která musí být obsažena na stránkách ve výsledcích vyhledávání.
- + Znaménko plus před vyhledávaným slovem udává nutnost obsažení daného slova ve výsledcích vyhledávání.
- - Slovo za znaménkem mínus naopak ve výsledcích vyhledávání nesmí být obsaženo.

5.3 Příkazy vyhledávacího řádku

Také Seznam využívám příkazů pro upřesnění vyhledávání, které jdou kombinovat s povolenými operátory. K upřesnění je možno použít následující příkazy: [24]

- **intitle:** *hledaný výraz* – Hledaná fráze nebo slovo bude obsaženo v titulku stránek, které budou vyhledány.
- **inurl:** *hledaný výraz* – Hledání výrazu musí být obsažen v URL adrese webové stránky.
- **intext:** *hledaný výraz* – V případě vyhledávání jednoho slova nemá smysl, v případě víceslovného dotazu slouží k vyhledání zadaného slova přednostně v obsahu stránky.
- **site:** *url_adresa* *hledaný výraz* – Příkazem site s parametrem url adresy následováno hledaným výrazem, vyhledá hledaný výraz pouze na stránkách zadaného webu. Samotným příkazem site: se zadanou url adresou se vypíší všechny indexované stránky zadané domény.

- **-site:** *url_adresa* hledaný výraz – Opak předchozího příkazu, hledaná fráze nebo slovo bude prohledáno v celém indexu, ale stránky ze zadané domény budou vyjmuty z výsledků vyhledávání.
- **link:** *url_adresa* – Příkaz vyhledá všechny stránky, ze kterých je odkazováno na zadanou url adresu.
- **filetype:** *koncovka typu souboru* – Příkaz filetype umožňuje vyhledávat dokumenty podle zadaného formátu. Seznam indexuje formáty typu: html, doc, rtf, pdf, ppt a txt. Příkaz filetype s parametrem typu souboru musí následovat až po zadání hledaného výrazu. Funkčnost tohoto příkazu značně pokulhává a víceslovná vyhledávání konkrétního typu souboru se mi nepodařilo ověřit.

5.4 S-rank

S-rank stránky je veličina, která by měla vyjadřovat důležitost každé stránky na českém webu. Výpočet S-ranku je vysoce závislý na počtu zpětných odkazů, a tak není problém hodnotu S-ranku ovlivnit až do hodnoty 70, kterou lze dosáhnout pořízením většího množství zpětných odkazů. Až od hodnoty 80 se můžeme bavit o hodnotách, které vyjadřují důležitost stránky na českém internetu.

Počítá se zejména z odkazové sítě algoritmem, který zohledňuje jednak odkazy, které na stránku míří ale i skutečnost, kam odkazy ze stránky vedou. Z hodnoty S-ranku nelze odvozovat předpokládané pořadí ve výsledcích vyhledávání. Výsledná relevance výsledků vyhledávání se počítá z mnoha dalších kritérií a S-rank je jen jedním z nich. Přesný výpočet S-ranku není veřejný. [8]

S-rank stránky může nabývat hodnoty od 0 do 100 a není vázán na doménu, ale na každou jednotlivou webovou stránku. Hlavní stránka tak můžete nabývat hodnotu S-rank 60, jedna podstránka hodnotu 70 a u ostatních stránek webu se S-rank může pohybovat nejčastěji v rozmezí 0 až 30.

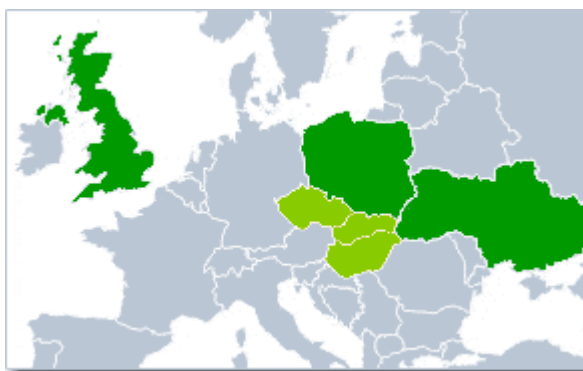
6 Jyxo

Jyxo je původem český internetový vyhledávač postavený na vlastní technologii. Jeho hlavními třemi pilířem je vyhledávání, kontextová reklama a publikační systém pro blog.

Internetové vyhledávání Jyxo.cz umožňuje vyhledávat na webových stránkách, ale také v souborech formátu pdf a doc. Také umí vyhledávat v obrázcích, v hudbě a videu, ale také například v sortimentu internetových obchodů. Pro vyhledávání ve světě využívá Jyxo vyhledávač Google, na který je uživatel přesměrován po zadání dotazu a zvoleným vyhledáváním ve světě.

Vyhledavač Jyxo obsahuje lingvistický modul, který umožňuje skloňování a časování českých slov a jakožto ryze český vyhledávač nabízí stejně jako Seznam kontrolu pravopisu zadávaných dotazů uživatelem, vyhledavač v případě překlepu nebo špatně napsaného slova nabídne nejpravděpodobnější správný tvar. Mezi nedostatky lze zařadit například chybějící „našeptávač“ [25]

Při vyhledávání jsem často narazil na pomalé zpracovávání výsledků, kterým odpovídaly desítky tisíc výsledků. Na vyhledávací dotaz „životopis format:doc“ jsem čekal v rozmezí 10 – 15 vteřin. Toto je jeden z velkých nedostatků, který dle mého mínění poukazuje na již nedostatečný hardware a pomalý vývoj celého fulltextového vyhledávání Jyxo. Také kvalita některých výsledků vyhledávání je nedostatečná, to dokazují výsledky testů relevantnosti v kapitole 11.1.2.



Obrázek 10: Dostupnost služeb Jyxo vyznačen na mapě.¹⁰

¹⁰ <http://jyxo.vybereme.cz/d/jyxo>

6.1 Historie

Společnost Jyxo.cz mimo jiné provozuje také blogovací systém Blog.cz, online fotogalerii Galerie.cz či přehled článků z různých webových serverů na adrese clanky.jyxo.cz. Pro společnost Seznam do roku 2010 vyvíjel PPC systém Sklik. Svou vyhledávací technologii poskytuje mnoha různým subjektům. Kromě češtiny se společnosti Jyxo specializuje také na slovenský a maďarský trh z části své služby nabízí také v Polsku, Ukrajině a ve Velké Británii. (Obrázek 10). V polovině roku 2008 koupila společnost CME (provozovatel TV stanice Nova) 100% podíl ve společnosti Jyxo s.r.o.

6.2 Operátory a speciální znaky ve vyhledávači Jyxo

- “ “ Uzavřením hledaných slov nebo fráze do uvozovek budou vyhledány pouze stránky a dokumenty, které obsahují zadaná slova v přesně uvedeném tvaru a pořadí.
- + Slovo za znakem plus musí být obsaženo ve výsledcích vyhledávání.
- - Slovo za znakem mínus se naopak ve výsledcích nesmí vyskytovat.
- # Výskyt slova za znakem # ve výsledcích vyhledávání je výhodou, ale není nezbytně nutný.
- **OR** Při použití logického operátoru OR bude nalezeno jedno či druhé slovo. OR lze také nahradit znakem |, doporučováno je místo operátoru OR volit raději operátor #.

6.3 Příkazy vyhledávacího řádku

- **domain:** *url_adresa* – Tímto příkazem lze omezit výsledky vyhledávání jen na zadanou doménu druhé nebo vyššího řádu. Lze kombinovat se znakem - (mínus), tedy naopak zakázat stránky ze zadané domény. Příkaz domain: lze nahradit také příkazy **host:** nebo vyhledávací častěji používaným příkazem **site:**, které fungují stejně jako příkaz domain.
- **title:** *hledaný výraz* – Hledaná fráze nebo slovo musí být obsaženo v titulku stránky zobrazené ve výsledcích vyhledávání.
- **url:** *hledaný výraz* – Hledání výrazu musí být obsažen v URL adrese webové stránky nebo dokumentu.

- **format:** *koncovka typu souboru* hledaný výraz – Budou vyhledávány relevantní soubory zadaného formátu, Jyxo podporuje formáty: txt, pdf a doc.
- **link:** *url_adresa* – Tento příkaz vyhledá stránky, kterých odkazují na zadanou url adresu.

6.4 JyxoRank

„JyxoRank je hodnota vyjadřující jak je stránka v českém internetu známá, důležitá, populární. Počítá se z odkazů stránky, na které odkazuje mnoho lidí, získají větší JyxoRank.“ [9]

JyxoRank nabývá hodnot od 0 do 220, přičemž vyšší hodnota je lepší. Také JyxoRank podobně jako PageRank nebo S-Rank má nepřímý vliv na řazení výsledků vyhledávání. Pro stanovení hodnoty JyxoRank jsou zohledněny zpětné odkazy a také jejich zdroj. Větší váha je při výpočtu JyxoRank přikládána stránkám, na které vedou odkazy z více nezávislých zdrojů. Ani v případě JyxoRank nejsou pouze zpětné odkazy hlavním faktorem ovlivňujícím výpočet hodnoty ranku. Patří sem také on-page faktory a další neveřejné hodnoty. Přepočet JyxoRanku na rozdíl od přepočtů S-ranku a Google PageRanku je častý a pravidelný, probíhá přibližně v rozmezí jednoho týdne.

7 Bing

Bing je nejmladší z posuzovaných fulltextový vyhledávačů (oficiálně byl představen v květnu 2009 a začátkem června téhož roku veřejně spuštěn), stojí za ním společnost Microsoft. Předchůdci vyhledávače Bing byli vyhledávač Live Search a MSN Search. Oba zmíněné vyhledávače jsou nyní nahrazeny technologií vyhledávání Bing. Kromě fulltextového vyhledávání nabízí také vyhledávání v obrázcích, videích, mapách a v dalších vyhledávacích službách, které však nejsou lokalizované do češtiny, například srovnávač a vyhledávač zboží, vyhledávač letenek či hotelů. V současné době je Bing kromě češtiny lokalizovaný do dalších 42 světových jazyků.

Bing je „světovou dvojkou“ ve vyhledávání („jedničkou“ je Google). Druhou pozici si však vybudoval převážně akvizicí s vyhledávačem Yahoo, pro který Bing dodává výsledky vyhledávání.

7.1 Operátory a speciální znaky vyhledávače Bing

Díky operátorům a speciálním znakům lze dosáhnout přesnějších výsledků vyhledávání také u vyhledávače Bing. Jednotlivé operátory se dají kombinovat a tím přesněji specifikovat vyhledávanou frázi.

- " " Uvozovky ohraničují přesnou frázi, která musí být obsažena ve výsledku vyhledání.
- + Slovo za znaménkem plus musí být obsaženo ve výsledcích vyhledávání.
- - Slovo nebo fráze v uvozovkách za znaménkem mínus nesmí být obsaženo ve výsledcích hledání, znaménko mínus lze nahradit logickým operátorem **NOT**.
- **OR** Použijeme-li logickou spojku OR, znak | nebo dva znaky ||, seřadí se stránky podle relevance samostatných výrazů.
- **AND** Logická spojka AND mezi slovy způsobí, že se zobrazí jen stránky, na kterých budou nalezena obě slova. Dá se nahradit umístěním znaku + před všechny hledaná slova.

7.2 Příkazy vyhledávacího řádku Bing

- **contains:** *koncovka typu souboru* – Vyhledány stránky, které obsahují zadané klíčové slovo a zároveň se na stránce nachází, odkazují na zadaný typ souboru. Například „ceník škoda auto contains:pdf“ vyhledá stránky relevantní na slova ceník, škoda, auto a zároveň je z webu odkazováno na nějaký pdf soubor.
- **define:** *hledaný výraz* – Vyhledá definici zadaného výrazu.
- **domain:** *url_adresa* – Příkaz vyhledá všechny subdomény zadané url adresy.
- **filetype:** *koncovka typu souboru* hledaný výraz – Vyhledá relevantní soubory v zadaném formátu.
- **imagesize:** *small/medium/large* – Příkaz pro vyhledávání předem specifikované velikosti obrázku, příkaz vyhledá obrázky v zadané velikosti. Příklad použití: „mapa ČR imagesize: large“
 - **small** – obrázky menší než 200px x 200px
 - **medium** – obrázky větší než 200px a menší než 500px
 - **large** – obrázky o rozměrech větších než 500px na šířku i výšku
- **inanchor:** *hledaná fráze* – Budou vyhledány stránky, na které je odkazováno z jiných webů se zadaným textem v textu odkazu např.:
inanchor:"Miroslav Pešťák" – budou vyhledány weby na které odkazováno například v tomto tvaru
`Miroslav Pešťák`.
- **inbody:** *hledaná fráze* – Vrátí webové stránky, které obsahují zadaný výraz v metadatech nebo v těle webové stránky.
- **intitle:** *hledaná fráze* – Ve výsledcích vyhledávání budou pouze stránky, které obsahují vyhledávaný dotaz v title tagu webové stránky.
- **ip:** *ip adresa* – Tento příkaz vypíše stránka nebo stránky, které jsou na serveru s danou IP adresou. Tento užitečný nástroj je unikátní u Bing, jiné vyhledávače ho nenabízejí. Můžete si tak například zobrazit všechny weby, které hostují na stejném hostingu jako je váš web. Například příkaz

„ip:46.28.105.7“ vypíše všechny weby (které Bing indexuje) a hostují na jednom ze serverů společnosti Wedos, který má IP adresu 46.28.105.7..

- **language:** *kod jazyku* – Vyhledává pouze ve stránkách, které jsou psané v zadaném jazyku. Např.: „robot language:es“ pro vyhledání relevantní stránky slovu robot ve španělštině.
- **msite:** *url_adresa* – Tento příkaz vyhledá obrázky a videa na zadaném webu podle zadaného klíčového slova např.: msite:youtube.com simpsons – vyhledá obrázky i videa které obsahují relevantní obsah simpsons.
- **site:** *url_adresa* – Vypíše všechny stránky na zadané doméně, pokud se přidá hledaná fráze, budou vyhledány výsledky pouze ze zadaného webu.

7.3 BrowseRank

BrowseRank je relativně nový hodnotící algoritmus, který se objevil s vyhledávačem Bing. BrowseRank je odlišný především v tom, že na rozdíl od Google algoritmu PageRank nevychází pouze z odkazů, ale převážně z chování uživatelů. Hodnotí také, jak dlouho se návštěvníci webu na stránkách zdržují. Předpokladem Microsoftu je, že uživatelé na stránkách s užitečnějším obsahem tráví více času. [10]

O tom zda se Bing ve výsledcích vyhledávání BrowseRank uplatňuje, jsem v žádném z literárních pramenů nenašel zmínku. Podle mého názoru může být BrowseRank užitečný. Neumím si však představit, jak může Bing získávat data o dalších činnostech návštěvníka na webu (například když návštěvník ihned po přístupu na vyhledaný web, nepokračuje na web jiný), pokud tedy nevyužije data z prohlížeče Internet Explorer. To však podle mého názoru není reálné ani etické. Z tohoto důvodu se možná o BrowseRanku přestalo mluvit a jeho vývoj nepokračuje, nebo je jen pouze nevýznamným faktorem při řazení výsledků vyhledávání v SERP v Bing.

8 Morfeo

Fulltextový vyhledávač Morfeo je výhradně český vyhledávač, který v současnosti indexuje přes 162 milionů stránek (počet indexovaných stránek odpovídá datu 8. 3. 2011), umožňuje hledání slov odvozených, synonym hledaných výrazů a také opravuje překlepy v českém jazyce.

Technologie vyhledávače Morfeo je založena na Open Source technologii Sherlock Holmes (Sherlock Holmes Search Engine). Vyhledávač Morfeo byl hlavním vyhledávačem portálu Centrum.cz. Později ale provozovatel portálu, společnost NetCentrum, rozhodla o odsunutí vyhledávače Morfeo do pozadí. Jako primární fulltextový vyhledávač je od té doby na hlavní stránce Centrum.cz využíván vyhledávač Google . Vyhledávač Morfeo se díky této skutečnosti přestal dále vyvíjet.

8.1 Operátory a speciální znaky v českém vyhledávači Morfeo

- " " Uvozovky umožňují vyhledat přesnou frázi v zadaném pořadí slov.
- - Hledané slovo za znaménkem mínus se nebude vyskytovat ve výsledcích vyhledávání.
- * Jako nahrazující znak lze využít hvězdičku, ta může nahradit libovolné počty slov i písmen, podmínkou je zadat minimálně 3 znaky, další znaky a slova je možné nahradit.

8.2 Příkazy vyhledávacího řádku

- **site:** *url_adresa* hledaný výraz – Hledaný výraz bude vyhledáván jen na zadané doméně a jejích subdoménách.
- **link:** *url_adresa* – Tento příkaz vyhledá všechny weby, které odkazují na zadanou url adresu.
- **hdr:** *hledaný výraz* – Vyhledá stránky, které obsahují hledaný výraz na stránce v nadpise libovolné úrovně (H1 až H6).
- **keyword:** *hledaný výraz* – Vyhledá stránky, které mají v meta tagu keywords zadaný hledaný výraz nebo slovo. Tento příkaz patří mezi ty, které se v dnešní době již nedají přínosně využít.

- **alt:** *hledaný výraz* – Dle manuálu by měl tento příkaz vyhledat takové stránky, které mají zadaný výraz v popisu obrázku zobrazeném na stránce. Podle mého testování však jsou vyhledány všechny stránky, které mají v parametru title (nikoli v parametru alt) hledaný výraz a to ať se jedná o title obrázku, nebo klasického textového odkazu.
- **filetype= "format souboru"** *hledaný výraz* – Příkaz pro vyhledání souborů zadaného typu, které obsahují hledaný výraz. Vyhledávač Morfeo podporuje formáty typu pdf, text, html, msword, excel. Na rozdíl od jiných příkazů, je zde nutno použít znak rovnítka místo obvyklé dvojtečky.

9 Ask

Fulltextový vyhledávač Ask je využíván v největší míře v Anglii, do češtiny není lokalizován. Řadou českých uživatelů je však tento vyhledávač dále využíván, protože některý software nainstaluje násilnou formou Ask.com Toolbar a zároveň nastaví Ask jako hlavní vyhledávač i domovskou stránku. Z vlastní zkušenosti jsem se setkal s tím, že mnoho uživatelů tento toolbar neumí odstranit, a tak využívají vyhledávač, který je v jejich internetovém prohlížeči nastaven jako výchozí.

Vyhledávač Ask umožňuje vyhledávat jak v textovém obsahu webu, tak v obrázcích, videích, obsazích i v mapách. Jelikož není lokalizovaný do češtiny, má s vyhledáváním v česky psaných stránkách menší problémy, přesto je však většina výsledků vyhledávání českých slov a frází uspokojivá.

Vývoj fulltextového vyhledávače Ask byl zastaven koncem minulého roku. „*Ask Networks se místo vyhledávání bude věnovat vývoji „Q&A Service“ – tedy něčemu, co také dlouhodobě nepatří mezi služby, které na Internetu nějak výrazně profitují a fungují.*“ [11] Zkratka Q&A znamená Question and Answer doslovně přeloženo otázka a odpověď, na otázky uživatelů budou odpovídat jiní vybraní uživatelé na základě svých znalostí a zkušeností. Tyto odpovědi budou indexovány a zobrazí se při dalším vyhledání stejného nebo podobného dotazu.

9.1 Příkazy vyhledávacího řádku Ask

Vyhledávač Ask podobně jako ostatní v tomto směru drží krok s vyhledávačem Google, a tak také nabízí základní příkazy pro zpřesnění vyhledávání. Vlastnosti funkcí, které jsou stejné jako u vyhledávače Google nepopisuji, najdete je v kapitole 4.3. Zde popisuji pouze funkce a příkazy odlišné.

- **site: inurl: intitle: inlink fines: OR** – viz kapitola 4.3
- **country:** *kod jazyku* – Do výsledků vyhledávání budou zahrnuty jen stránky, které jsou v zadaném jazyce.
- **+:** *hledaný výraz* – Slovo nebo výraz musí být ve výsledcích vyhledávání obsaženo. Tento příkaz je obdobný jako samotné znaménko plus ve vyhledávači Google, v případě vyhledávače Ask musí za znaménkem plus následovat dvojtečka.

- **-:** *hledaný výraz* – Slovo nebo výraz nesmí být obsažen ve výsledcích vyhledávání. Také tento příkaz je obdobou samotného znaménka mínus ve vyhledávači Google opět s přidáním dvojtečky.

9.2 Doplnkové funkce vyhledávání Ask

Také doplnkové funkce má vyhledávač Ask hodně podobné vyhledávači Google, s tím rozdílem, že nejsou lokalizované do češtiny.

- **Předpověď počasí** – Funkce se vyvolá klíčovým slovem **weather** nebo **forecast** s přidáním příslušného města. Velká česká města jsou podporována, pokud jsou zadána v mezinárodním tvaru, například „weather Pilsen“.
- **Převodník** – funkci převodu jednotek se vyvolá klíčovým slovem **conver** (může být také vynecháno) a zadáním vstupních, poté klíčového slova in nebo to a výstupních jednotek. Umožňuje převod měny, hmotnosti, objemu, délky. Například „2 Megabytes to bytes“ nebo „convert 10 kilometer to miles“.

9.3 ExpertRank

Algoritmus hodnocení webů ExpertRank je založen na technologii Teoma a k řazení výsledků vyhledávání ho využívá vyhledávač Ask. Hodnota ExpertRanku se vypočítává převážně z odkazů vedoucích na stránku. Hlavní roli při výpočtu hraje samotná analýza každého odkazu v kontextu odkazující webové stránky. Téměř nulovou hodnotu má odkaz směřující z webu jiného zaměření, naopak odkaz z webové stránky se stejným zaměřením má vysokou hodnotu, která se na odkazovanou stránku přenáší.

„Teoma, pronounced chawmuh, was an Internet search engine founded in 2000 by Professor Apostolos Gerasoulis and his colleagues at Rutgers University in New Jersey. Professor Tao Yang from the University of California, Santa Barbara co-led technology R&D. Their research grew out of the 1998 DiscoWeb project.“ [12]

10 Další vyhledávače

Mnoho velkých světových fulltextových vyhledávačů postupně ustoupilo nebo ustupuje velkým hráčům ve fulltextovém vyhledání a nabízí výsledky vyhledávání od největších rivalů internetového vyhledávání, kterými jsou v celosvětovém měřítku Google a Bing.

Výsledky vyhledávání Google nabízí například AOL, ICQ search, z českých vyhledávačů jsou to například Centrum nebo Atlas. Výše zmíněné vyhledávače vyplnilo v mém dotazníku několik tazatelů s tím, že to jsou jejich hlavní vyhledávače. O skutečnosti, že používají vyhledávání Google, nevěděli. To dokazuje, že používají jako alternativní vyhledávač právě samotný Google. Výsledky vyhledávání dodávané vyhledávačem Bing jsou například na serveru Conduit nebo na největším portálu v USA Yahoo, v Čechách ho pro vyhledávání ve světě implementuje například Seznam.

10.1 AOL

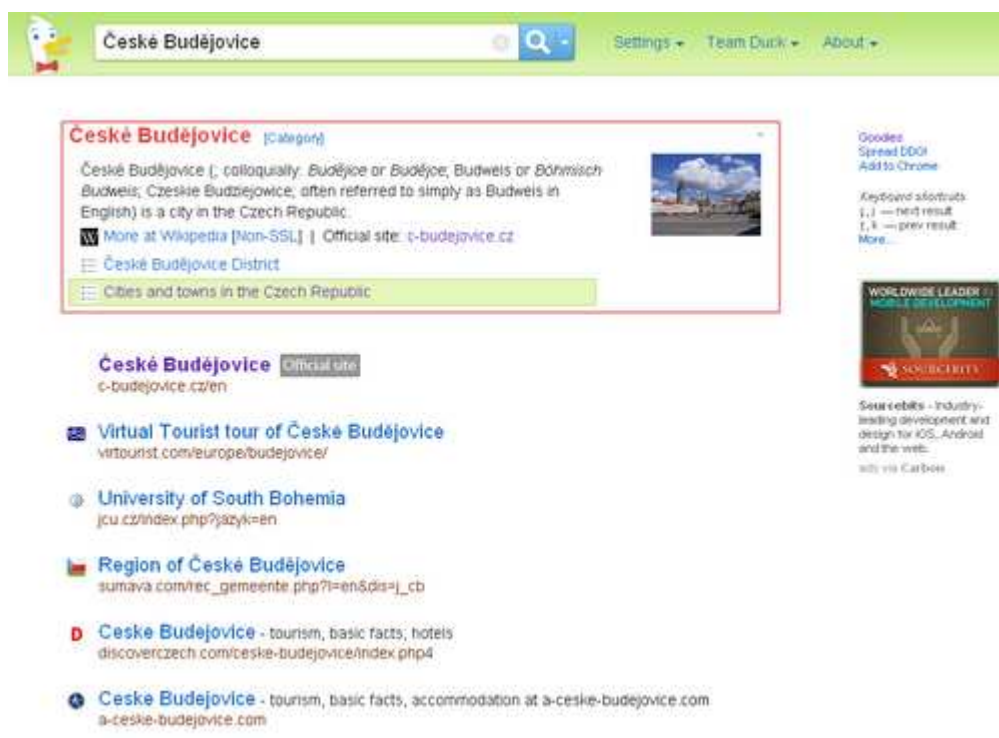
Společnost AOL (America Online) patřila v devadesátých letech mezi největší poskytovatele internetových služeb v USA a v nejlepší období měla téměř 30 milionů zákazníků. Od roku 2003 počet uživatelů vyhledávání AOL postupně klesá. V roce 2005 ukončila společnost AOL vývoj fulltextového vyhledávače a integrovala vyhledávání Google, které ve službách AOL přetrvává stále. AOL se tak spíše než vyhledávačem stalo zpravodajským a zábavním portálem s několika přidanými službami jakou je například emailová schránka zdarma.

10.2 Altavista

Altavista byla hlavním poskytovatelem výsledku vyhledávání pro portál Yahoo od počátku založení (1995). Později se vyhledávač Altavista osamostatnil a byl velkým konkurentem ve fulltextovém vyhledávání pro Google. V roce 2003 se historie obrátila a Yahoo koupilo vyhledávač Altavista a výsledky vyhledávání dodával vyhledávací engine Yahoo právě pro Altavistu. Vyhledávání přes Altavistu je v současné době stále směřováno na servery Yahoo, avšak ani Yahoo už nemá vlastní výsledky vyhledávání, ale zobrazuje výsledky vyhledávače Bing. Tak i Altavista nepřímo zobrazuje výsledky vyhledávání vyhledávače Bing.

10.3 DuckDuckGo

DuckDuckGo je unikátní fulltextový vyhledávač, který se kromě vlastního indexovacího robota spoléhá na dalších 30 zdrojů vyhledávání, přesněji na dalších 30 vyhledávačů, od kterých přebírá výsledky, jež není sám schopen dodat, a vybírá ty nejlepší dle svého algoritmu. Zdrojem vyhledávání tak mohou být výsledky od vyhledávače Google, Bing nebo články z internetové encyklopedie wikipedie, a v případě striktně položených dotazů jsou výsledky přebírány z kontroverzního vyhledávače WolfrmaAlpha (o tomto vyhledávači více v kapitole 10.5).



Obrázek 11: Náhled stránky s výsledky vyhledávání DuckDuckGo

Vyhledávač DuckDuckGo je vyvíjen od roku 2008. Samotným vzhledem a stylem zobrazení výsledků je patrné, že se nesnaží kopírovat Google ani Bing. Číslo stránek pro možnost stránkování výsledků nezobrazuje nestandardně, ale na konci výsledků vyhledávání jako odkaz pro zobrazení dalších výsledků, které jsou načteny dynamicky technologií jQuery. Celkový vzhled a rozložení vyhledávače působí velmi přehledně, nevyskytují se zde žádné rušivé reklamy a před výpisem nalezených stránek je vždy stručná odpověď na vyhledávanou frázi, často dostačující, takže není třeba následovat odkazy na nalezené stránky (Obrázek 11).

DuckDuckGo není lokalizovaný do češtiny, přesto jeho výsledky jsou z větší části uspokojivé i pro české vyhledávání a to převážně díky jeho kvalitním zdrojům

informací. DuckDuckGo také nabízí rozsáhlé funkce a usnadnění vyhledávání, nechybí integrovaná kalkulačka, která po zadání vypočte příklad. Ve složitějších případech, např. při výpočtu integrálu, přebírá výsledky z vyhledávače WolframAlpha (včetně grafu funkce). Dále umožňuje DuckDuckGo převod jednotek, zobrazit informace o doménovém jménu, zobrazení časových pásem, vyhledání mapy podle jména města nebo směrovacího čísla. Je také možné pokládat faktické otázky, v tomto případě jsou výsledky přebírány převážně od vyhledávače WolframAlpha (kapitola 10.5).

10.4 Lycos

Lycos patří mezi řadu dříve úspěšných, prosperujících a kvalitních fulltextových vyhledávačů, kteří však tlak konkurence neustáli. Lycos patřil se svými 25 jazykovými mutacemi mezi první desítku vyhledávačů ve světě.

V roce 2008 zaznamenal Lycos obrovský úpadek svých služeb, a tak prodal veškeré své portály a nadřazené služby a nyní nabízí primárně jen vyhledávání, jak ve psaném obsahu internetu, tak ve videích, obsazích a obrázcích. Nabízí také možnost vyhledávání lidí ve žlutých stránkách (pouze pro země, kde je Lycos lokalizovaný do místního jazyka), dále vyhledávač zboží a vyhledávání v inzerátech a pracovních nabídkách vlastního serveru.

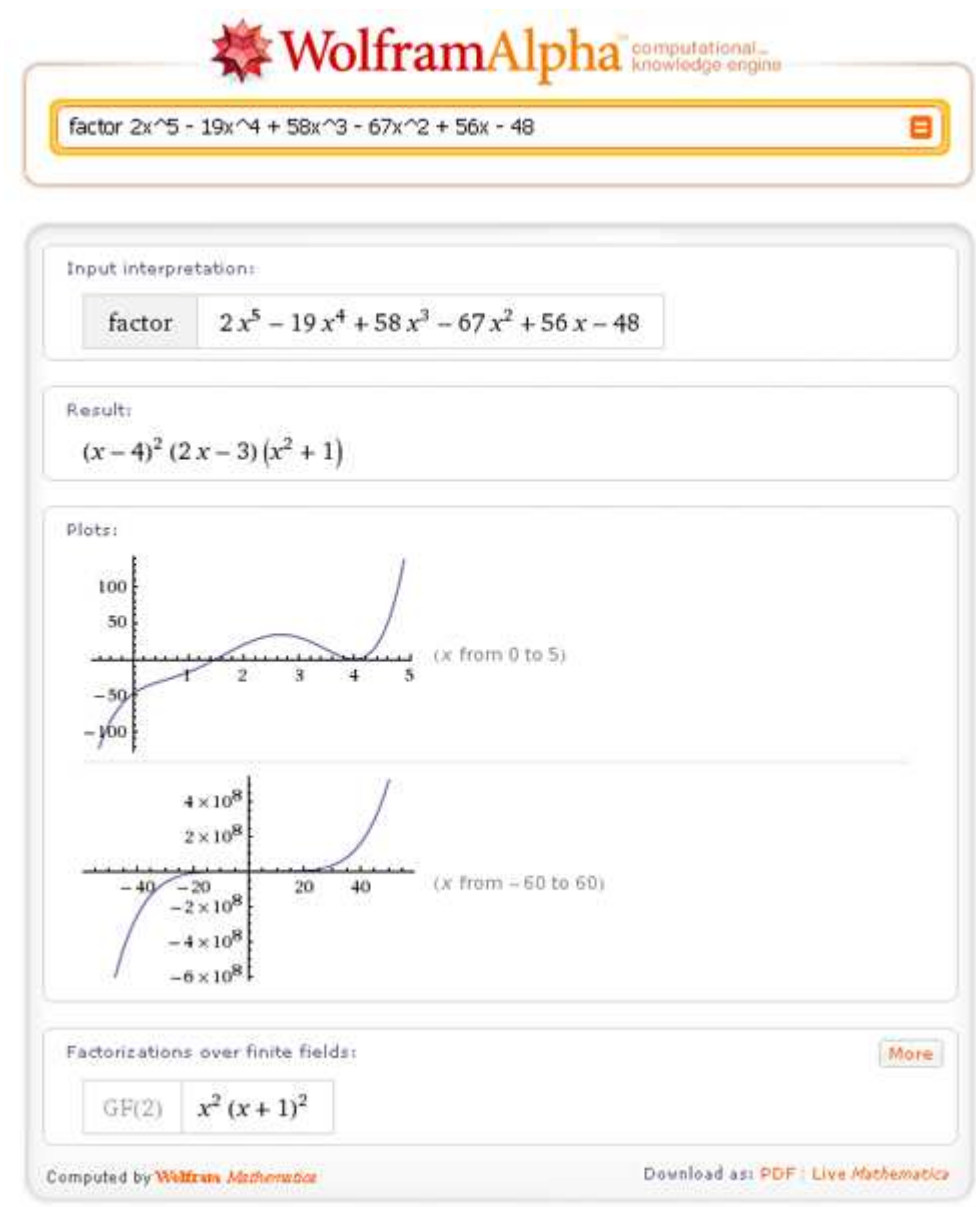
K dispozici je oprava překlepů (jen pro podporovaných 25 jazykových mutací – české překlepy nejsou rozpoznány). Při testování jsem ale postrádal našeptávač, nebo jiný způsob urychlení vyhledávání.

O podpoře operátorů nebo klíčových slov nebylo v nápovědě vyhledávací služby Lycos nic napsáno. Věděl jsem s jistotou, že někteří operátoři, například uvozovky, zde fungují, a tak jsem postupně zkoušel, zda jsou ve vyhledávači Lycos také podporováni operátoři jiných vyhledávačů. Mohu potvrdit funkčnost operátorů: site:, intitle:, filetype: (s podporovanými typy formátů doc pdf rtf a xls), uvozovky, AND (lze nahradit znakem plus) a poslední parametr NOT, který lze také nahradit a to znaménkem mínus.

10.5 WolframAlpha

WolframAlpha je vědomostní vyhledávač, který na rozdíl od fulltextových vyhledávačů prohledává předem zformátované báze dat, nikoli běžné stránky na

internetu. Výsledky vyhledávání se také liší od běžných fulltextových vyhledávačů. Výsledky vyhledávání nejsou odkazy na stránky, kde se hledaný problém vyskytuje nebo řeší, ale nabízí se rovnou odpověď na zadanou otázku nebo výpočet příkladu. Za vznikem vyhledávače stojí známý Britský fyzik a matematik Steven Wolfram, který je mimo jiné autorem softwaru Mathematica.



Obrázek 12: Náhled stránky s výsledky vyhledávání WolframAlpha

Systém se snaží k vyhledávaným heslům najít co nejvíce možných alternativ, aby měl uživatel komplexní přehled informací o hledaném heslu. Toho využijí ti, kteří nevědí přesně, co hledají, řeší úkol na dané téma nebo si chtějí rozšířit informace v dané problematice. WolframAlpha je převážně matematický a fyzikální nástroj. Tuto skutečnost naznačuje také samotné jméno vývojáře a hlavního zakladatele.

Poradí si také s elementárním výpočtem operací, převody mezi číselnými soustavami a jednotkami, dále pak zvládá základní aritmetické operace, integrály, derivace. Poradí si také s většinou rovnic (diferenciální, kvadratické, kubické a další). Kromě konečných výsledků zadaných příkladů jsou k dispozici grafy, alternativní formy výsledků, postup jakým bylo výsledku dosaženo nebo různé další užitečné aproximace. Není pro něj problémem řešení symbolické, ani numerické a je zde možné využívat také konstanty.

Mimo matematické a fyzikální výpočty si vyhledávač WolframAlpha poradí také se zeměpisnými daty, chemickými vzorci, informacemi o potravinách, dopravě, počasí. Obsahuje rozsáhlou historii teplotních dat vázaných na města a obce. Můžete si tak zobrazit například výkyvy teplot ve vašem městě za posledních několik let. Ale poradí si také například s ekonomickými informacemi. Veškerá tato data pak dokáže při jediném vyhledávání využít a nabídnout tak komplexní informace o hledaném heslu. Například při vyhledávání názvu státu obdrží uživatel komplexní data, tedy mezinárodní název, zkratky, vlajku, mapu, geografické informace, sousední státy, kompletní informace o populaci (počet obyvatel, průměrný věk, hustotu obyvatel), informace o hlavním městě a dalších velkých nebo významných městech, nechybí ani ekonomické informace o HDP, inflaci nebo o nezaměstnanosti. Většina dat jde dále na vyžádání (kliknutím na tlačítko) rozšířit. Kompletní výsledek vyhledávání jde exportovat do PDF, v případě matematických výpočtů je možný export do softwaru Mathematica. Jednotlivé skupiny informací jdou také ukládat zvlášť, buďto formou exportu do obrázku (ve formátu JPG) nebo čistého textového dokumentu.



Demographics: [Show rates](#) [Show distribution](#) [Show non-metric](#)

population	67.8 million people (world rank: 20 th) (2010 estimate)
population density	29.9 people/km ² (people per square kilometer) (world rank: 182 nd) (2008 estimate)
population growth	2.77% per year (world rank: 25 th) (2008 estimate)
life expectancy	54.4 years (world rank: 197 th) (2009 estimate)

[Save as image >](#) | [Copyable plaintext >](#)

Obrázek 13: Jedna z oblastí výsledku vyhledávání WolframAlpha s možností ukládání jako obrázek nebo zkopírování čistého textu.

Veškeré vyhledávání je opravdu komplexní a vyhledávač dodá všechny relevantní informace k zadanému dotazu. Pokud vyhledáte název firmy, obdržíte opět komplexní informace, historii, jméno nebo jména zakladatelů, počet zaměstnanců, finanční výsledky za poslední známé období zobrazené v tabulkách i grafech, případně mohou být dostupné i burzovní informace.

Přes veškerou komplexnost a užitečnost je do jisté míry omezený, díky pouhé anglické mutaci. V případě matematických výpočtů je to zanedbatelný nedostatek, ale pokud chceme využívat jeho plnou sílu, musíme zadávat dotazy v anglickém jazyce. V takovém případě je schopný zpracovat i víceslovné dotazy, které mají svou návaznost, je tak možné se například dotázat, jaké bylo počasí v Praze, když se narodil Václav Klaus. Dotaz upravený pro vyhledávač v anglickém jazyce pak vypadá následovně: „weather in Prague when Vaclav Klaus was born“.

10.6 Yahoo

Fulltextový vyhledávač Yahoo byl druhým největším vyhledávačem (světového měřítko), co se kvantity vyhledávaných dotazů týká. V roce 2009 zaujímal 6,4% trhu (první byl Google s 85.4%). [13]

Od 29. července 2009 přestal Yahoo vyvíjet své fulltextové vyhledávání a začal ve výsledcích vyhledávání zobrazovat výsledky vyhledávače Bing. Touto spoluprací si společnost Microsoft koupila své druhé místo mezi světovými vyhledávači.

Yahoo i nadále provozuje svůj portál, který patří k nejnavštěvovanějším v USA, nabízí široké spektrum služeb, sportovní i finanční zpravodajství, aktuality, ale také vyhledávač aut, letenek, nebo zboží.

10.7 Yandex

Yandex je Ruský internetový vyhledávač, který na domácí (Ruském) trhu v internetovém vyhledávání drží 60% podíl na trhu. [14] V tomto směru se mu hodně podobá český Seznam, který si zatím také na domácím trhu drží nadpoloviční podíl před globálním vyhledávačem Google, který má takřka monopol ve vyhledávání ve většině Evropských zemí.

Globálním vyhledávačem se Yandex stal začátkem května 2010 [15], kdy začal vyhledávat také mezi webovými stránky, které jsou psané v latině. Yandex

nezaostává za konkurencí, nabízí vyhledávání obrázků, videí a také dalších typů formátů (pdf, rtf, doc, xls, ppt, swf). Podporuje také řadu operátorů (např. uvozovky, znaménko plus a mínus) i několik příkazů vyhledávacího řádku (například site: nebo title:). Dostupná je i řada skrytých funkcí, například vepsáním „my ip“ do pole pro vyhledávání bude včetně výsledků vyhledávání zobrazena také IP adresa vašeho PC. Další skrytou funkcí je například zobrazení chemického prvku z periodické tabulky prvků. Pokud vyhledáváte název chemického prvku, bude včetně výsledků vyhledávání zobrazena chemická značka prvku, atomová hmotnost, řád a skupina.

10.8 Zoohoo

Česko-slovenský fulltextový vyhledávač, který je již „mrtvý“ (není vyvíjen ani aktualizován), obsahuje velké množství reklamy, která je přidávána majitelem vyhledávače, společností Etarget, jenž provozuje stejnojmenný reklamní systém.

S výsledky vyhledávání jsem nebyl spokojen téměř ani v jednom ze složitějších dotazů, nenabízí žádné přidané hodnoty ani funkcionalitu. Pro vyhledávání ve světě (v jiném jazyce než českém nebo slovenském) je použit vyhledávač Google.

11 Test vybraných fulltextových vyhledávačů

Na základě praktických zkušeností a poznatků z teoretické části mé bakalářské práce jsem prováděl testování vybraných fulltextových vyhledávačů s konkrétními příklady vyhledávání podle níže uvedených kritérií, které jsem pro testování zvolil.

K testování jsem vybral 9 výše popsaných vyhledávačů, jmenovitě: Google, Seznam, Jyxo, Bing, Morfeo, Ask, DuckDuckGo, Lycos a Yandex. Vyhledávače, které přebírají výsledky vyhledávání od konkurenčních vyhledávačů, tedy například AOL, Altavista, Yahoo nebo české Centrum či Atlas, jsem do testování nezařazoval, stejně tak jsem do testování nezařadil český Zoohoo, který není delší dobu vyvíjen a je silně nedostačující. Dále jsem mezi testované vyhledávače nezařazoval vyhledávač WolframAlpha, který není fulltextový vyhledávač, jeho funkce a vlastnosti jsou popsány v kapitole 10.5.

11.1 Testování relevantnosti

K testování relevantnosti výsledků vyhledávání ve fulltextových vyhledávačích jsem si zvolil dvacet slov a slovních spojení. V první části testovaných dotazů jsem se zaměřil na překlepy, které jsou uživateli často zadávány ať už nevědomě nebo z neznalosti oboru, o kterém chtějí informace vyhledat. Sledoval jsem, zda vyhledávač rozezná překlep, ale také jak se liší výsledky vyhledávání při správném a při špatném zadání hledaného výrazu. Testoval jsem několik vybraných překlepů, které uživatelé při vyhledávání často zadávají. (Tabulky 2, 3, 4, 5)

V další části jsem použil slova s českou diakritikou, která jsem vyhledával se zadáním české diakritiky i bez ní v různých kombinacích. Nejzajímavější byly výsledky vyhledávání slova měď, které při zadání bez české diakritiky má jiný význam a to med. Snad nejhorší možný výsledek ze všech testů měl vyhledávač Bing, který při zadání „Včelí med“ na první místo zařadil naprosto irelevantní stránku s erotickým videem a na druhém místě byl web med.cz, který ale neměl žádný obsah kromě nadpisu med.cz v H1 tagu. V ostatních vyhledávačích se web med.cz nenacházel, což bylo správně, jelikož na tomto webu se doposud obsah tematicky příbuzný medu nenacházel (soudě dle archivu webové stránky Google). (Tabulky 6, 7, 8)

Otestoval jsem také možnost vyhledávání v obrázcích. Všechny testované vyhledávače tuto službu nabízejí. Seznam nemá tuto službu vlastní, ale přebírá výsledky od společnosti picsearch, která se na vyhledávání obrázků specializuje. Výsledky ve vyhledávání obrázků byly kromě upadajících vyhledávačů Jyxo a Morfeo dostačující. (Tabulka 13)

Jeden z vyhledávacích dotazů byl položen podle manuálů pokročilého vyhledávání. Snahou bylo vyhledat manuál pro zařízení iphone 4 ve formátu pdf. V tomto případě jsem použil jiné hodnocení, vyhledávač mohl dostat 0, 5 nebo 10 bodů. 10 bodů obdržel pouze vyhledávač Google, který vyhledal požadovaný manuál v češtině, 5 bodů obdržely vyhledávače, které našli manuál v anglickém jazyce a ty vyhledávače které manuál v pdf nenašly žádný, měly bodů nula. (Tabulka 1)

Vyhledání manuálu ve formátu pdf pro iphone 4 dle možností vyhledávače		
Vyhledávač	body	dotaz
Google	10	filetype:pdf iphone 4 manuál
Seznam	0	iphone 4 manuál filetype:pdf
Jyxo	0	iphone 4 manuál format:pdf
Bing	5	filetype:pdf iphone 4 manuál
Morfeo	0	filetype="pdf" iphone 4 manuál
Ask	0	filetype:pdf iphone 4 manuál
DuckDuckGo	5	filetype:pdf iphone 4 manuál
Lycos	5	filetype:pdf iphone 4 manuál
Yandex	0	iphone 4 manuál pdf

Tabulka 1: Test vyhledávání s požadavkem na manuál ve formátu pdf

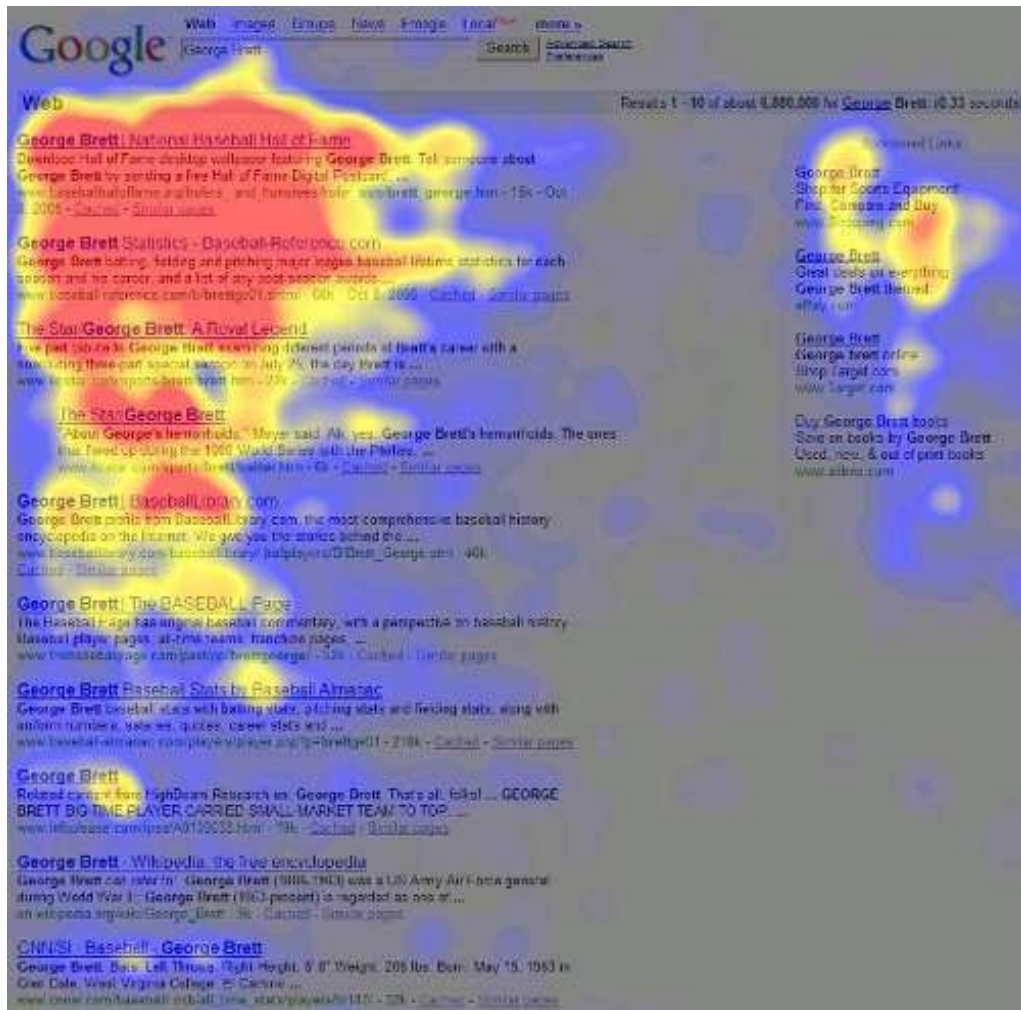
Dvacítku testovacích vyhledávacích frází jsem doplnil o vyhledávání softwaru Total Commander (Tabulka 17), vyhledáváním informací o osobách (Tabulky 14, 15) nebo vyhledáváním stránek menší obce Žichovice. (Tabulka 10) Dále jsem vyhledával informace o webovém standardu HTML5 (Tabulka 16), užitečné informace pro turisty jsem vyzkoušel při hledané frázi Pražský hrad (Tabulka 9) a nakonec všeobecné vyhledávání informací o Jihočeské univerzitě. (Tabulka 12) Nechybělo vyhledávání informací o aktuálně probíhajících událostech (Sčítání lidu 2011, Matějská pouť, Majáles 2011). (Tabulky 18, 19, 20)

11.1.1 Kritéria testování relevantnosti

Pro hodnocení relevantnosti výsledku jednotlivých vyhledávačů jsem zvolil bodové hodnocení od 0 do 10 bodů, kde 10 je maximální možný počet bodů.

První výsledek přirozeného vyhledávání (pokud se na prvních místech ve výsledcích vyhledávání zobrazovala reklama, nebyla v hodnocení relevantnosti brána v úvahu) měl hodnocení od 0 do 5 bodů, a maximálního počtu bodů dosáhl výsledek vyhledávání, který obsahoval informace o hledaném slovu. V případě dostupných oficiálních stránek byl plně relevantní, pouze web oficiální, přesto weby s informativní hodnotou daného tématu mohly dostat až 4 body.

Druhý a třetí výsledek vyhledávání dosahovaly 0 až 2 bodů, opět dle relevantnosti. Čtvrtý a pátý výsledek mohl být ohodnocen počtem bodů 0 nebo 0,5 bodu. Celkově tedy mohl každý vyhledávač získat maximum 10 bodů při vyhledávání jednoho dotazu. Další výsledky vyhledávání, tedy 6 a další výsledek na první straně a na dalších stránkách jsem nezohledňoval. Vycházel jsem z teplotní mapy výsledků vyhledávání Google (Obrázek 14), která dokazuje, že jen přibližně 1 - 2% uživatelů využije výsledků vyhledávání na páté a další pozici ve výsledcích vyhledávání.



Obrázek 14: Teplotní mapa (heat map) výsledků vyhledávání Google¹¹

¹¹ <http://promediablog.com/where-to-position-for-google-search-results-serp-heat-map-study-for-search-marketers/>

11.1.2 Kompletní přehled testovaných frází

Vyhledávaný dotaz - Hary Potter - překlep	
Vyhledávač	body
Google	7,5
Seznam	2
Jyxo	0
Bing	9
Morfeo	0,5
Ask	7,5
DuckDuckGo	5
Lycos	9,5
Yandex	9,5

Tabulka 2: Dosažené body při vyhledávání
překlepové fráze 1

Vyhledávaný dotaz - Mozilla Firefox	
Vyhledávač	body
Google	9
Seznam	3
Jyxo	3
Bing	8
Morfeo	1,5
Ask	9,5
DuckDuckGo	6
Lycos	9,5
Yandex	9,5

Tabulka 4: Dosažené body při vyhledávání
překlepové fráze 2

Vyhledávaný dotaz - Harry Potter - hodnocení rozdílů s překleповou verzí	
Vyhledávač	body
Google	10
Seznam	6
Jyxo	1,5
Bing	10
Morfeo	0,5
Ask	10
DuckDuckGo	9
Lycos	9,5
Yandex	9,5

Tabulka 3: Dosažené body při vyhledávání

Vyhledávaný dotaz - Rádio Impulz	
Vyhledávač	body
Google	10
Seznam	0,5
Jyxo	0
Bing	7,5
Morfeo	6
Ask	5,5
DuckDuckGo	7
Lycos	7,5
Yandex	0

Tabulka 5: Dosažené body při vyhledávání
překlepové fráze 3

Vyhledávaný dotaz - měď	
Vyhledávač	body
Google	10
Seznam	10
Jyxo	0
Bing	3,5
Morfeo	9
Ask	2,5
DuckDuckGo	4
Lycos	0
Yandex	0

Tabulka 6: Dosažené body při vyhledávání

Vyhledávaný dotaz - med	
Vyhledávač	body
Google	9
Seznam	5
Jyxo	4
Bing	0
Morfeo	4
Ask	0
DuckDuckGo	0
Lycos	0
Yandex	0

Tabulka 7: Dosažené body při vyhledávání

Vyhledávaný dotaz - Žichovice	
Vyhledávač	body
Google	10
Seznam	8,5
Jyxo	7
Bing	2
Morfeo	9,5
Ask	7,5
DuckDuckGo	0,5
Lycos	5
Yandex	0

Tabulka 10: Dosažené body při vyhledávání informací o obci

Vyhledávaný dotaz - Včeli med	
Vyhledávač	body
Google	10
Seznam	10
Jyxo	8
Bing	3
Morfeo	5
Ask	9
DuckDuckGo	1,5
Lycos	0,5
Yandex	3

Tabulka 8: Dosažené body při vyhledávání

Vyhledávání v mapách - Národní park Šumava	
Vyhledávač	body
Google	10
Seznam	8
Jyxo	4
Bing	8
Morfeo	5
Ask	10
DuckDuckGo	8
Lycos	8,5
Yandex	8

Tabulka 11: Dosažené body při vyhledávání v mapách

Vyhledávaný dotaz - Pražský hrad	
Vyhledávač	body
Google	10
Seznam	7,5
Jyxo	0,5
Bing	9
Morfeo	7
Ask	8
DuckDuckGo	6,5
Lycos	7,5
Yandex	3,5

Tabulka 9: Dosažené body při vyhledávání

Vyhledávaný dotaz - Jihočeská univerzita katedra informatiky	
Vyhledávač	body
Google	10
Seznam	9
Jyxo	7
Bing	1
Morfeo	1
Ask	9
DuckDuckGo	0
Lycos	0
Yandex	1

Tabulka 12: Dosažené body při vyhledávání

Vyhledávání v obrázcích - iPad	
Vyhledávač	body
Google	10
Seznam	8
Jyxo	9,5
Bing	9,5
Morfeo	9,5
Ask	5
DuckDuckGo	0
Lycos	10
Yandex	10

Tabulka 13: Dosažené body při vyhledávání v obrázcích

Vyhledávaný dotaz - HTML 5 elementy	
Vyhledávač	body
Google	9,5
Seznam	7
Jyxo	3,5
Bing	7
Morfeo	0
Ask	10
DuckDuckGo	6
Lycos	0
Yandex	7,5

Tabulka 16: Dosažené body při vyhledávání

Vyhledávaný dotaz - PaedDr. Petr Pexa	
Vyhledávač	body
Google	10
Seznam	10
Jyxo	7
Bing	6
Morfeo	9,5
Ask	6
DuckDuckGo	7
Lycos	5
Yandex	3

Tabulka 14: Dosažené body při vyhledávání

Vyhledávaný dotaz - Total Commander	
Vyhledávač	body
Google	9,5
Seznam	8
Jyxo	2
Bing	10
Morfeo	1
Ask	9,5
DuckDuckGo	9,5
Lycos	9,5
Yandex	9,5

Tabulka 17: Dosažené body při vyhledávání

Vyhledávaný dotaz - Miroslav Pešťák	
Vyhledávač	body
Google	8
Seznam	7
Jyxo	2
Bing	2
Morfeo	7,5
Ask	7,5
DuckDuckGo	5
Lycos	5
Yandex	4

Tabulka 15: Dosažené body při vyhledávání

Vyhledávaný dotaz - Sčítání lidu 2011	
Vyhledávač	body
Google	9,5
Seznam	6
Jyxo	5
Bing	3,5
Morfeo	5
Ask	6
DuckDuckGo	2
Lycos	2
Yandex	3

Tabulka 18: Dosažené body při vyhledávání aktuálního téma

Vyhledávaný dotaz - Matějská pout'	
Vyhledávač	body
Google	9
Seznam	7,5
Jyxo	1
Bing	5,5
Morfeo	9
Ask	7
DuckDuckGo	5,5
Lycos	5,5
Yandex	3

Tabulka 19: Dosažené body při vyhledávání aktuálního téma

Vyhledávaný dotaz - Majáles 2011	
Vyhledávač	body
Google	10
Seznam	10
Jyxo	7
Bing	5
Morfeo	9
Ask	9,5
DuckDuckGo	6
Lycos	6,5
Yandex	1,5

Tabulka 20: Dosažené body při vyhledávání aktuálního téma

11.1.3 Vyhodnocení testování relevantnosti

Výsledná tabulka (Tabulka 21) je průměrem všech jednotlivých testovacích vyhledávání (jednotlivé testované fráze jsou v tabulkách 1 až 20), maximum tedy bylo 10 bodů a minimum 0. Nejlépe dopadl vyhledávač Google, který dosáhl průměru 9,55 bodů, druhý nejlépe hodnocený byl vyhledávač Ask, třetí Seznam a jako čtvrtý v pořadí skončil vyhledávač Bing od Microsoftu.

Testování odpovídá standardnímu vyhledávání českého uživatele, tedy při vyhledávání anglických frází může být relevantnost jednotlivých vyhledávačů jiná. Přesto ale fulltextový vyhledávač Ask v testu dokázal, že lokalizace vyhledávače do místních jazyků nemá vliv na výsledky vyhledávání a může nabídnout i kvalitní výsledky vyhledávání frází, které nejsou v anglickém jazyce.

Hodnocení relevantnosti - bodový průměr	
Vyhledávač	body
Google	9,55
Seznam	6,65
Jyxo	3,60
Bing	5,73
Morfeo	4,98
Ask	6,95
DuckDuckGo	4,68
Lycos	5,30
Yandex	4,28

Tabulka 21: Výsledná tabulka hodnocení relevantnosti fulltextových vyhledávačů

11.2 Test rychlosti indexování

Rychlost indexování je jedním z nejdůležitějších faktorů relevantnosti fulltextových vyhledávačů, zvláště pokud uživatel požaduje relevantní výsledky vyhledávání na aktuální témata.

Indexování jsem ověřoval příkazem `site: s` s parametrem url adresy webu nebo testovacího článku, pokud vyhledávač parametr `site` neznal, nahradil jsem ho názvem článku v uvozovkách nebo konkrétní url adresou. Kontrola probíhala ručně v intervalu 5 minut během první hodiny, poté v pravidelném intervalu 1 hodiny po dobu 24 hodin, následně jednou za 24 hodin. V případě aktuálního zpravodajského článku jsem jako nejdelší přijatelnou dobu k zaindexování zvolil 24 hodin, v případě článků na menších blozích jsem horní hranici zaindexování článku zvolil na 1 měsíc, stejně tak u nové webové stránky. Bodové hodnocení jsem vždy volil u konkrétního případu tak, že nejkratší doba měla plný počet bodů, tedy 10 a nejdelší doba zaindexování, nebo žádné zaindexování, obdrželo 0 bodů.

Rychlost indexování jsem otestoval na vybraných vyhledávačích v několika simulovaných případech.

- Sledování rychlosti zaindexování nového článku na velkém zpravodajském serveru `idnes.cz`. Zpravodajský server `idnes.cz` je největším zpravodajským serverem v Čechách a tak k němu také české fulltextové vyhledávače přistupují. Četnost procházení tohoto serveru je mnohonásobně větší než u běžných webových stránek. Nejlépe dopadl vyhledávací robot Google, který novou webovou stránku na serveru

idnes.cz zaindexoval během jedné minuty. Druhým v pořadí z testovaných vyhledávačů, který novou stránku zaindexoval, byl Seznam následně Bing. České vyhledávače Jyxo ani Morfeo nový článek na portálu iDnes nezaznamenaly během jednoho dne, což je podle mého názoru u aktuálního zpravodajství horní hranice. Konkrétní doba indexování jednotlivých vyhledávačů je v následující tabulce. (Tabulka 22)

Rychlost indexování nové stránky na portálu idnes.cz	
Vyhledávač	doba
Google	1 minuta
Seznam	9 minut
Jyxo	vice než 1 den ∞
Bing	10 minut
Morfeo	vice než 1 den ∞
Ask	30 minut
DuckDuckGo	5 hodin
Lycos	40 minut
Yandex	6 hoodin

Tabulka 22: Tabulka rychlosti indexování nového článku na serveru idnes.cz

- Druhý test spočíval ve sledování rychlosti indexování nového článku na blogu, který má více jak roční historii, více jak deseti zaindexovaných stránek v testovaných vyhledávačích. K testování jsem využil dva internetové blogy www.mirecekp.net a www.bedlive.info. V obou případech mnohonásobně předčil rychlost všech Google, vyhledávače Jyxo ani Morfeo opět novou stránku nestihly do minimální doby (1 měsíc) zaindexovat. Naopak překvapivě rychle, v porovnání s ostatními, zaindexoval nové články vyhledávač Ask, který přestože není lokalizovaný do češtiny, indexuje větší české webové stránky poměrně pravidelně.

Rychlost indexace nové stránky na blogu mirecekp.net	
Vyhledávač	doba
Google	2 minuty
Seznam	4 dny
Jyxo	více než 1 měsíc ∞
Bing	27 dní
Morfeo	více než 1 měsíc ∞
Ask	2 hodiny
DuckDuckGo	více než 1 měsíc ∞
Lycos	30 dní
Yandex	více než 1 měsíc ∞

Tabulka 23: Tabulka rychlosti indexování nové stránky na blogu mirecekp.net

Rychlost indexace nové stránky na blogu bedlive.info	
Vyhledávač	doba
Google	1 den
Seznam	30 dní
Jyxo	více než 1 měsíc ∞
Bing	více než 1 měsíc ∞
Morfeo	více než 1 měsíc ∞
Ask	28 dní
DuckDuckGo	více než 1 měsíc ∞
Lycos	více než 1 měsíc ∞
Yandex	více než 1 měsíc ∞

Tabulka 24: Tabulka rychlosti indexování nové stránky na blogu bedlive.info

- V poslední části testování rychlosti indexování jsem založil novou webovou stránku na nově pořízené doméně 2. řádu (959.cz), bez dřívějšího obsahu a bez zpětných odkazů. Nový web jsem jen přidal přes přidávací formuláře jednotlivých vyhledávačů do jejich indexu, spíše tedy do seznamu stránek, který má indexovací robot navštívit. Tato metoda se používá, právě u nových webových stránek, na které nevedou žádné zpětné odkazy a vyhledávače by tak novou webovou stránku nemohly najít přirozenou cestou. Přidávací formulář není k dispozici u vyhledávačů Ask, DuckDuckGo, Lycos, Morfeo. I přesto vyhledávače DuckDuckGo a Lycos novou webovou stránku zaregistrovaly, i když na ni žádný jiný web neodkazuje. Domnívám se, že toto zaindexování způsobilo mé časté testovací dotazování, zda se stránka v indexu nachází či nikoliv. Vyhledávač Ask ani po více jak dvou měsících stránku neindexuje, tím

potvrzuje své pravidlo napsané v manuálu, že indexuje pouze weby, na které vedou zpětné odkazy. [22] Vyhledávač Jyxo, přestože jsem novou stránku přidal pomocí přidávacího formuláře, stránku také po více než dvou měsících stále neindexuje. (Tabulka 25)

Rychlost indexace nového webu na nové doméně	
Vyhledávač	doba
Google	2 dny
Seznam	7 dní
Jyxo	více než 1 měsíc ∞
Bing	14 dní
Morfeo	více než 1 měsíc ∞
Ask	více než 1 měsíc ∞
DuckDuckGo	17 dní
Lycos	7 dní
Yandex	15 dní

Tabulka 25: Tabulka rychlosti indexování nového webu

11.2.1 Vyhodnocení rychlosti indexování

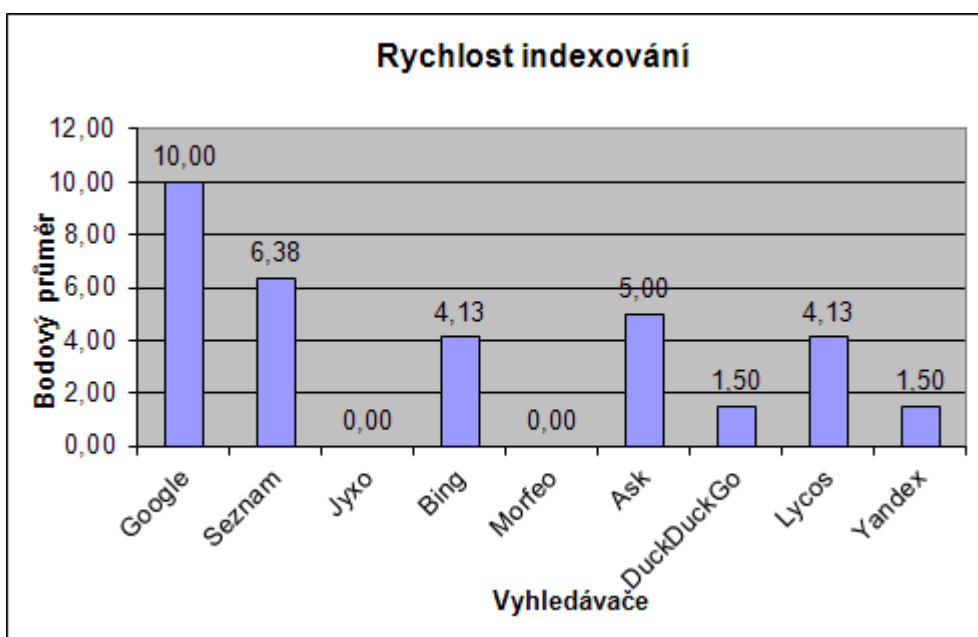
Z výsledné tabulky (Tabulka 26) vyplývá, že jednoznačně nejrychleji z testovaných fulltextových vyhledávačů indexuje obsah českého internetu vyhledávač Google, který ve všech testovaných případech zaindexoval stránku nejrychleji, z maximálního počtu 10 bodů získal všech 10. Druhý v pořadí se umístil český Seznam, který indexuje pouze obsah českého (a z části Slovenského) internetu, tedy mnohonásobně méně než globální Google, přesto nedosahuje potřebné rychlosti. Rychlost indexace se u Seznamu také odvíjí od kvality zdrojového serveru, a tak naměřených 9 minut do doby zaindexování nového článku na serveru idnes.cz je dostačující, u webových stránek menšího rozsahu je doba indexování znatelně pomalejší.

Jako třetí nejrychlejší vyhledávač v indexování skončil překvapivě vyhledávač Ask, který není lokalizovaný do češtiny, přesto mu české vyhledávání nedělá problémy. Až jako čtvrtý v pořadí skončil Bing, který je plně lokalizovaný do češtiny, přesto v některých případech indexování měl značné prodlevy, například 14 dní uplynulo než vyhledávač Bing navštívil nový web, na který byl indexovací robot upozorněn přes formulář k tomu určený. V případě indexování článků na velkém zpravodajském serveru potvrdil Bing třetím nejrychlejším časem (nový článek na

serveru idnes.cz zaindexoval za 10 minut), svou třetí příčku mezi fulltextovými vyhledávači na českém internetu za Seznamem a Googlem.

Rychlost indexování					
Vyhledávač	Nový web na nové doméně	Nová stránka na blogu 1	Nová stránka na blogu 2	Nová stránka na portálu idnes.cz	bodový průměr
Google	10	10	10	10	10,00
Seznam	7	8	1	9,5	6,38
Jyxo	0	0	0	0	0,00
Bing	5	2	0	9,5	4,13
Morfeo	0	0	0	0	0,00
Ask	0	9,5	2	8,5	5,00
DuckDuckGo	3	0	0	3	1,50
Lycos	7	1,5	0	8	4,13
Yandex	4	0	0	2	1,50

Tabulka 26: Výsledná tabulka rychlosti indexování fulltextových vyhledávačů



Graf 2: Výsledný graf rychlosti indexování fulltextových vyhledávačů

11.3 Celkové vyhodnocení testování fulltextových vyhledávačů

Do celkového hodnocení fulltextových vyhledávačů byly zahrnuty následující faktory:

- **relevantnost** – Více o posuzování relevantnosti v kapitole 11.1
- **rychlost indexování** – Více o hodnocení rychlosti indexování v kapitole 11.2
- **reklama ve výsledcích vyhledávání** – Reklama ve výsledcích vyhledávání je rušivým elementem, který je navíc v mnoha případech pro neznalé těžce odlišitelný od přirozených výsledků vyhledávání. Vyhledávače umisťují reklamu na první pozice výsledků vyhledávání a někdy také do pravého sloupce. Výjimkou z porovnávaných vyhledávačů je Ask, který umisťuje reklamní odkazy za 1. výsledek vyhledávání. V případě, že je uživatel spokojen s prvním výsledkem vyhledávání, je tak ušetřen prohlížení reklamy. Bodové hodnocení jsem zde volil následovně: 0 bodů pro vyhledávače, kteří reklamu mají, 10 bodů pro ty které reklamu v SERP nemají. Výjimkou je právě zmiňovaný Ask, kde jsem zvolil 4 body, jelikož největší počet návštěvníků skončí při vyhledávání právě na stránce, která se ve výsledcích zobrazila na prvním místě.
- **ovladatelnost a přehlednost** – Do této skupiny jsem zařadil několik faktorů, které mají vliv na přívětivost a použitelnost vyhledávače jako takového. V úvahu byly brány následující funkce a vlastnosti: intuitivní našeptávač, možnost procházet výsledky vyhledávání pouze klávesnicí, čitelnost a přehlednost samotných výsledků a také jejich grafické zobrazení, například font, velikost písma a kontrast barvy písma a barvy pozadí.
- **zobrazení náhledů stránek** – Náhled webové stránky ve výsledcích vyhledávání je podle mého názoru velice nápomocný, pokud uživatel vyhledává stránku, kterou již někdy navštívil, pamatuje si její vzhled a po krátkém prohlédnutí výsledků vyhledávání může odhalit tu stránku, kterou hledá. Náhledy jsou také dobré pro ty, kteří stránku ještě nikdy nenavštívili. Ve výsledcích vyhledávání je ale barevnost a rozložení

stránky zaujalo, v náhledu samotném je obrázek zboží nebo produktu, který vyhledávají. Bodové hodnocení v případě zobrazování náhledů mohla dosahovat pouze 0 nebo 10 bodů, 10 bodů v případě, že náhledy jsou k dispozici a 0 pokud náhledy webových stránek k dispozici nejsou.

- **doplňkové funkce** – Jako doplňkové a přídatné funkce nebo vlastnosti vyhledávače lze považovat přednostní vypsání odpovědi bez nutnosti navštívit vyhledanou stránku. Například při vyhledání informací „počasí Praha“ je před prvním výsledkem vyhledávání zobrazeno aktuální počasí a případně i předpověď počasí na delší období. Další užitečnou funkcí jsou například kalkulačka nebo převodníky měn. Podobných doplňkových funkcí nabízí některé vyhledávače desítky, jiné méně. Opět jsem využil bodování od 0 do 10. Množství funkcí, které jednotlivé vyhledávače nabízejí, jsem posuzoval dle jejich návodů a vlastních zkušeností.

Ve výsledné tabulce (Tabulka 27) je vidět konečné bodové hodnocení v jednotlivých kategoriích včetně celkových počtů bodů, kterých mohl vyhledávač dosáhnout maximálně 10. Plného počtu bodů nezískal žádný z testovaných vyhledávačů, nejvíce 9,22 bodů získal fulltextový vyhledávač Google a umístil se tak na prvním místě v tabulce.

Druhé místo, možná trochu překvapivě získal, v Čechách méně známý, fulltextový vyhledávač Ask, který získal 6,72 bodů. Druhé místo získal hlavně díky svým kvalitním výsledkům vyhledávání v českém internetu a také díky své přehlednosti a ovladatelnosti, i když není lokalizovaný do češtiny.

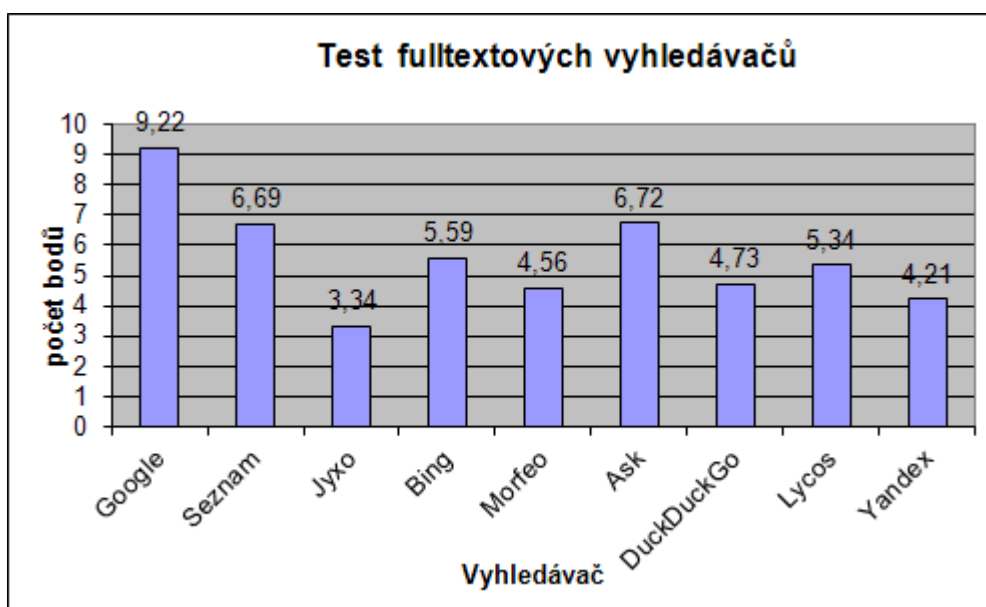
Třetí v hodnocení skončil Seznam s 6,69 body, čtvrtý pak Bing s 5,59 body. Další vyhledávače jsou pro českého uživatele, který vyhledává české výrazy spíše nepoužitelné, výjimkou může být vyhledávání v angličtině, kdy nám naopak předloží lepší výsledky vyhledávání, než v pořadí třetí, Seznam.cz, který je v Čechách nejpoužívanějším vyhledávačem.

Dříve známé a využívané, typicky české vyhledávače, Morfeo a Jyxo dopadly v testu velmi špatně. Jyxo se umístil na posledním 9. místě a Morfeo na místě 7. Z toho vyplývá mé doporučení omezit využívání těchto dvou českých vyhledávačů, jelikož nemají dostatek aktuálních a nových webů ve svém indexu a často jsou také výsledky nerelevantní. Naopak mohou doporučit využívání vyhledávače Google,

jelikož v testu prokázal, že disponuje adekvátními výsledky vyhledávání a nabízí řadu dalších funkcí, které mohou uživateli zrychlit vyhledávání informací na internetu. Také vyhledávač Ask mohou doporučit, ovšem k jeho plnému využití je potřeba minimální znalost anglického jazyka, protože jeho rozhraní ani příkazy funkcí nejsou lokalizované do češtiny.

procentuelní zastoupení v celkovém hodnocení	84%	6%	2%	3%	2%	3%	100%
Vyhledávač	relevanost vyhledávání	rychlost indexování	reklama ve výsledcích vyhledávání	ovladatelnost / přehlednost	náhled stránky	doplňkové funkce	Celkové hodnocení
Google	9,55	10	0	10	0	10	9,22
Seznam	6,65	6,38	0	9,5	10	8	6,69
Jyxo	3,60	0	10	4	0	0	3,34
Bing	5,73	4,13	10	8	0	3	5,59
Morfeo	4,98	0	10	6	0	0	4,56
Ask	6,95	5	3	9,5	0	8	6,72
DuckDuckGo	4,68	1,5	10	9	0	8	4,73
Lycos	5,30	4,13	10	8	10	0	5,34
Yandex	4,28	1,5	10	7	0	4	4,21

Tabulka 27: Výsledná tabulka testování fulltextových vyhledávačů

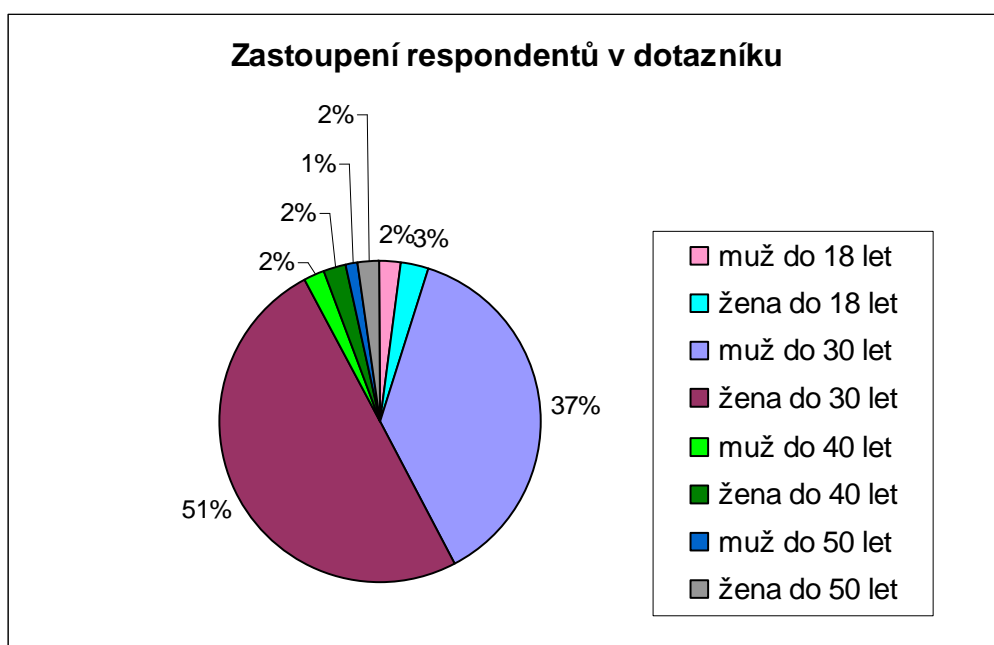


Graf 3: Výsledný graf testování fulltextových vyhledávačů

12 Zpracování výzkumu

Součástí mé bakalářské práce je dotazník („Dotazník internetového vyhledávání“) na téma internetové vyhledávání, kde jsem zkoumal u uživatelů znalosti fulltextových vyhledávačů, znalost rozšiřujících funkcí a vyhledávacích operátorů, jejich spokojenost s vyhledávačem, který používají primárně, a znalosti jiných vyhledávačů.

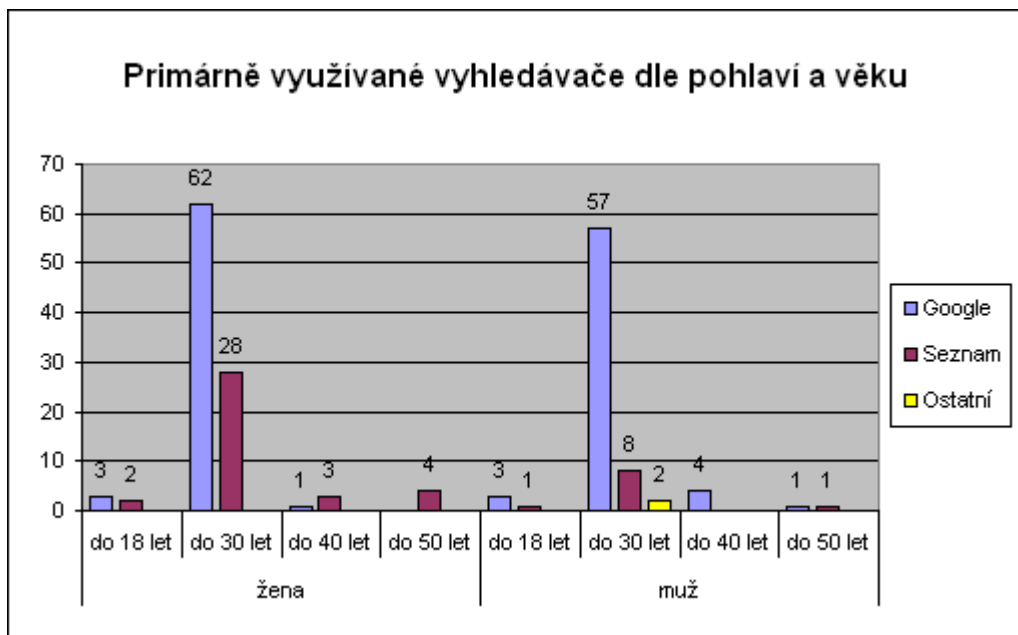
Dotazník vyplnilo 180 uživatelů internetu, 57% všech dotázaných byly ženy a 43% bylo mužů, z toho více jak polovina patří do skupiny ženy ve věku do 30 let, 37% pak muži také ve věku do 30 let. Kompletní zastoupení respondentů v dotazníku je na následujícím grafu. (Graf 4)



Graf 4: Celkové zastoupení respondentů v dotazníku

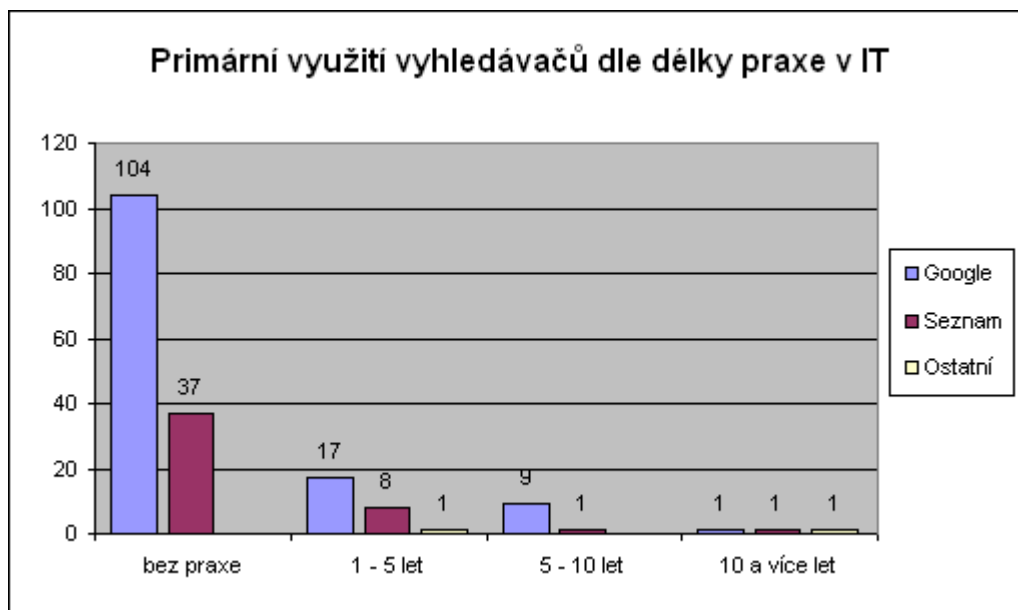
Domníval jsem se, že většina žen oproti mužům využívá vyhledávač Seznam a dále jsem se domníval, že ti, kteří mají větší praxi v IT, využívají více Google než jiné vyhledávače. První výše uvedenou hypotézu výsledky mého výzkumu nepotvrdily. Vyhledávač Seznam využívají jako primární vyhledávač spíše starší ročníky bez ohledu na pohlaví. Tento fakt si vysvětluji jako určitý zvyk uživatelů. Převážná většina starších uživatelů v době, kdy se internet v České republice začínal rozvíjet, používala Seznam jako výchozí stránkou svých prohlížečů. V té době nebyl Google ani jiný fulltextový vyhledávač v Čechách dostupný. Od té doby svůj zvyk

neměnili a využívají stále vyhledávač Seznam jako primární, přestože jsou si vědomi, že jiné vyhledávače existují (někdy je také využijí).



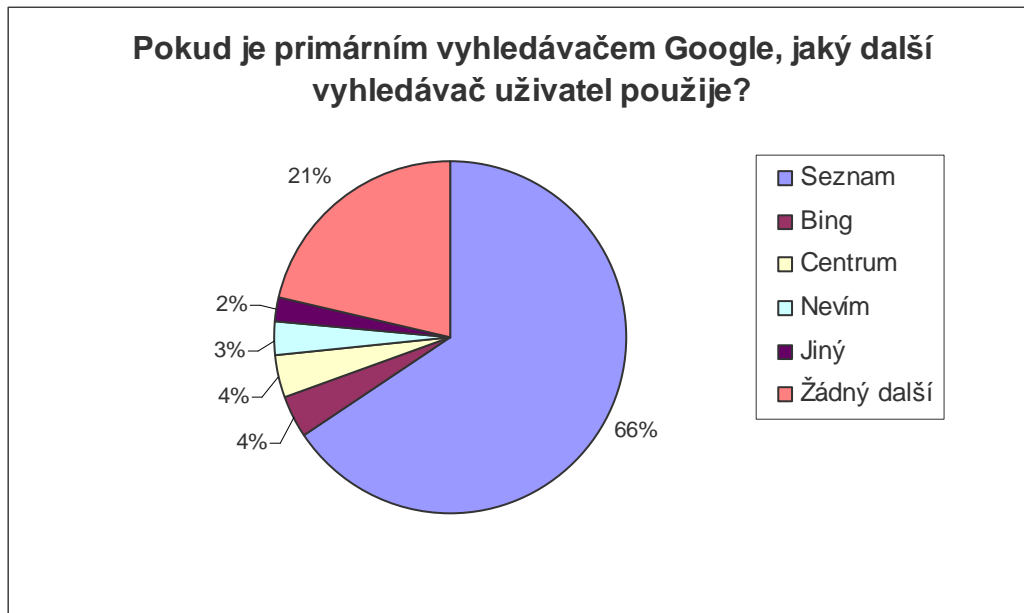
Graf 5: Primárně využívané vyhledávače dle pohlaví a věku

Druhou domněnku zjištěné výsledky potvrdily. Vyhledávač Google používá více uživatelů ve všech skupinách podle délky praxe v IT. Ve skupině respondentů, kteří délku své praxe v IT uvedli v rozmezí 5 – 10 let, využívá Google jako hlavní vyhledávač celých 90 % uživatelů, tento fakt potvrzuje již zmíněnou domněnku. Ovšem velké procentuální zastoupení vyhledávače Google bylo také ve skupině respondentů bez praxe v IT. V této skupině využívá primárně vyhledávač Google 74% dotázaných. (Graf 6)



Graf 6: Primární využití vyhledávačů dle délky praxe v IT

Z mého dotazníku také vyplývá, že většina uživatelů, kteří nejsou spokojeni s výsledkem vyhledávače Google při prvním pokusu vyhledávání, se uchyluje k vyhledávání převážně na Seznamu. Ale nezanedbatelná část uživatelů (21%) je s výsledky vyhledávání Google téměř vždy spokojena. (Graf 7)



Graf 7: Nejčastěji používaný vyhledávač, pokud primární vyhledávač Google nenajde to, co uživatel hledá

Uživatelé, kteří primárně používají fulltextový vyhledávač Seznam a nejsou spokojeni s výsledky vyhledávání, použijí jako alternativu vyhledávač Google a to v 96% všech případů. (Graf 8)



Graf 8: Nejčastěji používaný vyhledávač, pokud primární vyhledávač Seznam nenajde to, co uživatel hledá

Tyto ukazatele vypovídají o povědomí uživatelů. Uživatelé, kteří používají primárně Seznam a nejsou spokojeni s výsledkem vyhledávání, vědí, že existuje alternativní vyhledávač Google, o jiném z pravidla neuvažují nebo nevědí, že je k dispozici. Ovšem uživatelé, kteří nejsou spokojeni s výsledkem vyhledávání Google, jsou podle rozmanitosti odpovědí, zkušenější nebo informovanější.

Graf 8 také potvrzuje myšlenku Matěje Nováka, se kterou se také velmi ztotožňuji. Předpokládá, že podobně jako zanikají české vyhledávače Jyxo a Morfeo, skončí i Seznam a Google bude mít většinový podíl ve vyhledávání také v Čechách. [16] Uživatelé, kteří mají Seznam jako hlavní vyhledávač a často vyhledávají v alternativním vyhledávači Google, zde totiž najdou to, co hledají a mohou za určitou dobu zjistit, že častěji získávají relevantnější výsledky vyhledáváním právě na Google, a tak časem změni svůj primární vyhledávač.

12.1 Reklama ve vyhledávacích

Reklama je ve výsledcích vyhledávání rušivý element, který často dovede uživatele na stránky, které původně nehledal. Z tohoto důvodu jsem se respondentů, kteří primárně využívají Seznam nebo Google, dotazoval, zda jsou si vědomi přítomnosti reklamních odkazů na prvních pozicích ve výsledcích vyhledávání.

Téměř polovina dotázaných, kteří využívají primárně vyhledávač Seznam nevědí, že na prvních místech ve výsledcích vyhledávání jsou reklamní odkazy. (Graf 9, Obrázek 15)



Obrázek 15: Ukázka zobrazení výsledku vyhledávání Seznam.cz včetně reklamních odkazů na prvních dvou pozicích

Uživatelé vyhledávače Google jsou na tom s vědomím o reklamních odkazech ve výsledcích vyhledávání podstatně lépe. 70% dotázaných, kteří využívají primárně vyhledávač Google si je vědoma reklamních odkazů ve výsledcích vyhledávání. (Graf 9, Obrázek 16)

Hledat



Přibližný počet výsledků: 494 000 (0,17 s) Rozšířené vyhledávání


▶ **[Autobazar Č. Budějovice - Prověřené vozy, nízké ceny!](#)** Reklamy
 Záruka na mechanický stav vozu 12 měsíců.
www.aaaauto.cz/Autobazar



[Autobazary České Budějovice](#)
 Najděte si nejbližší **autobazar** Možnost volání zdarma!
www.zlatestranky.cz

[Autobazary české budějovice - Nákup a prodej silničních vozidel](#)
 Velký výběr nových i ojetých vozů
euautomobile.cz

[Místa pro dotaz **autobazar blízko lokality **České Budějovice**](#)**

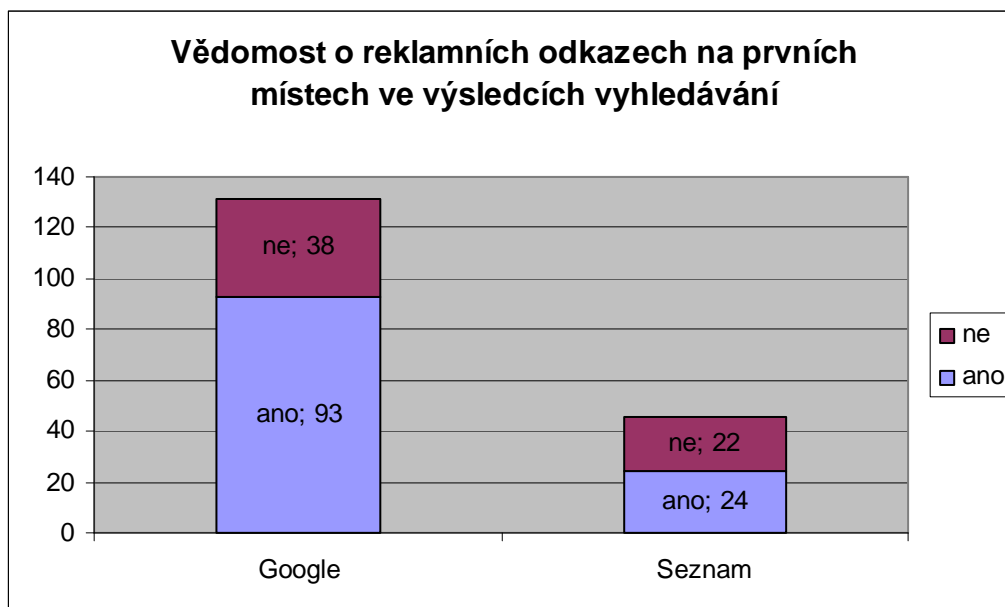
 **[Aston eurotrading s.r.o](#)**  - [Stránka v Adresáři míst](#)
www.astoncb.cz - J.Š.,Baara 77, České Budějovice - 608 966 569

 **[Autobazar DuoCar - HD SERVIS s.r.o.](#)**  - [Stránka v Adresáři míst](#)
www.sauto.cz - Okružní 1, České Budějovice - 604 724 566

 **[Jh Autocentrum - Josef Hron](#)**  - [Stránka v Adresáři míst](#)
www.jhauto.centrum.cz - Okružní 621/1A, České Budějovice 4 - 387 310 008

Obrázek 16: Ukázka zobrazení výsledku vyhledávání Google včetně reklamních odkazů na prvních 3 pozicích

Hlavním zájmem vyhledávačů by mělo být nabízet co nejlepší výsledky vyhledávání. Pro samotné vyhledávače je ovšem důležitý také zisk, který vyhledávače mají převážně z reklamních odkazů, přesněji tedy z prokliku na reklamní odkaz, a proto tyto reklamní odkazy umisťují ve výsledcích vyhledávání takovým způsobem, aby měly pokud možno, co největší počet prokliků. Přestože je reklama ve výsledcích vyhledávání odlišena od výsledků vyhledávání, oba zmiňované vyhledávače, tedy Google i Seznam, volí podbarvení reklamních odkazů takovou barvou, která je na dnešních standardních LCD monitorech jen těžko odlišitelná od okolního bílého pozadí, a tak si někteří uživatelé mohou reklamní odkazy od přirozených výsledků vyhledávání jednoduše splést.



Graf 9: Vědomost uživatelů o reklamních odkazech na prvních místech ve výsledcích vyhledávání vyhledávače Google a Seznam

Obdobným způsobem zobrazují kontextovou reklamu ve výsledcích vyhledávání také vyhledávače Ask, Lycos, Zoohoo a mnoho dalších. Výjimkou je například vyhledávač Bing, který kontextovou reklamu ve výsledcích vyhledávání nezobrazuje vůbec, stejně tak Jyxo, Morfeo, Yandex nebo DuckDuckGo.

13 Závěr

Hlavním cílem této bakalářské práce bylo seznámit běžného uživatele internetu s největšími internetovými vyhledávači a také s několika vyhledávači menšími, tzv. alternativními. Představované vyhledávače jsem podrobil testování tak, aby byly uživateli ukázány jejich četné výhody i nevýhody. Ze všech výsledků všech testovacích částí (test na rychlost, relevantnost a všeobecný test na přehlednost a ovladatelnost) jsem vybral nejúspěšnější vyhledávač, který mohu na základě jednotlivých kritérií a zjištěných výsledků doporučit jako nejlepší pro vyhledávání informací na českém internetu.

Součástí práce je také výzkum široké veřejnosti na téma internetové vyhledávání, z kterého jsem vycházel v praktické části. Snahou bylo hlouběji představit nejpoužívanější vyhledávače Seznam a Google, jejich skryté funkce, jejich výhody a nevýhody. Z výsledků výzkumu vznikl rozbor znalostí uživatelů internetu jednotlivých vyhledávačů a jejich schopností tyto vyhledávače správně využívat.

Díky praktické i teoretické části bakalářské práce jsem si prohloubil zkušenosti s vyhledávači, které jsou důležité pro úspěšnou optimalizaci webových stránek a jejich následnou propagaci v online marketingu. Výzkumná část práce mi jasně potvrdila i vyvrátila některé hypotézy týkající se korelace mezi typem vyhledávače a věkem, pohlavím nebo zkušenostmi uživatelů v oboru IT. Provedený výzkum a rozbor tak splnili stanovený cíl práce - představit a porovnat fulltextové vyhledávače na českém internetu.

Použité zdroje

- [1] *SEOMoz* [online]. 2010 [cit. 2011-01-15]. Meta Description SEO Best Practices. Dostupné z WWW: <<http://www.seomoz.org/learn-seo/on-site-topics>>.
- [2] *Nápověda Nástroje pro webmastery* [online]. 2010, updated 03/20/2011 [cit. 2011-01-15]. Pokyny pro webmastery. Dostupné z WWW: <<http://www.google.com/support/webmasters/bin/answer.py?hl=cs&answer=35769>>
- [3] *Kaviárové toasty.com* [online]. ? [cit. 2011-01-15]. Faktory ovlivňující hodnocení vyhledávačů. Dostupné z WWW: <<http://kaviarovetoasty.com/preklady/faktory-ovlivnujici-hodnoceni-vyhledavacu.php>>.
- [4] *Google Inc.* [online]. 1999 [cit. 2011-01-15]. The Company. Dostupné z WWW: <<http://web.archive.org/web/19990221202430/www.google.com/company.html>>.
- [5] *Google - Informace o korporaci* [online]. ? [cit. 2011-01-15]. Přehled technologie. Dostupné z WWW: <<http://www.google.com/corporate/tech.html>>.
- [6] JANOVSKEÝ, Dušan. *Yuhůův weblog o webu* [online]. 2008-10-12 [cit. 2011-01-15]. BadRank, seznamte se. Dostupné z WWW: <<http://weblog.jakpsatweb.cz/d/1225209900-badrank-seznamte-se.html>>.
- [7] *Financial Times* [online]. 2008-09-16 [cit. 2011-03-13]. Google still struggling to conquer outposts. Dostupné z WWW: <http://us.ft.com/ftgateway/superpage.ft?news_id=fto091620081227400726>.
- [8] *Seznam nápověda* [online]. ? [cit. 2011-03-13]. Algoritmus. Dostupné z WWW: <<http://napoveda.seznam.cz/cz/hledani-fulltext-algoritmus-vyhledavani-razeni-vysledku-faq-dotazy.html>>.
- [9] *Jyxo.cz* [online]. ? [cit. 2011-03-15]. Odpovědi na časté otázky - webmasteři. Dostupné z WWW: <<http://jyxo.vybereme.cz/d/faqw>>.
- [10] *SEO Book.com* [online]. 2008-07-27 [cit. 2011-03-20]. Microsoft Search BrowseRank Research Reviewed. Dostupné z WWW: <<http://www.seobook.com/microsoft-search-browserank-research-reviewed>>.

- [11] *JustIT.cz* [online]. 2010-11-11 [cit. 2011-03-20]. Ask.com, bývalý Ask Jeeves, ruší vlastní vyhledávání. Dostupné z WWW: <<http://www.justit.cz/wordpress/2010/11/11/ask-com-byvaly-ask-jeeves-rusi-vlastni-vyhledavani/>>.
- [12] Teoma. In *Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 2004-02-24, last modified on 2011-04-07 [cit. 2011-04-01]. Dostupné z WWW: <<http://en.wikipedia.org/wiki/Teoma>>.
- [13] Yahoo! Search. In *Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 2005-03-18, last modified on 2011-03-03 [cit. 2011-04-01]. Dostupné z WWW: <http://en.wikipedia.org/wiki/Yahoo!_Search>.
- [14] *TechRockets* [online]. 2010-07-14 [cit. 2011-04-01]. Yandex to increase its share in Runet. Dostupné z WWW: <<http://www.techrockets.com/2010/07/yandex-to-increase-its-share-in-runet/>>.
- [15] *Search Engine Land* [online]. 2010-05-19 [cit. 2011-04-01]. Russia's Yandex Search Engine Goes Global. Dostupné z WWW: <<http://searchengineland.com/russias-yandex-search-engine-goes-global-42381>>.
- [16] *Tyinternety.cz* [online]. 2011-03-16 [cit. 2011-04-15]. Ve sklepe s Matějem Novákem. Dostupné z WWW: <<http://www.tyinternety.cz/socialni-site/ve-sklepe-s-matejem-novakem1-3150>>.
- [17] *GoogleGuide* [online]. ?, Sunday May 9, 2010 [cit. 2011-01-01]. Google Search Operators. Dostupné z WWW: <http://www.googleguide.com/advanced_operators.html>.
- [18] *Bing Help* [online]. 2011 [cit. 2011-01-15]. How Bing delivers search results. Dostupné z WWW: <<http://onlinehelp.microsoft.com/en-us/bing/ff808447.aspx>>.
- [19] *MSDN* [online]. 2011 [cit. 2011-04-15]. Advanced Operator Reference. Dostupné z WWW: <<http://msdn.microsoft.com/en-us/library/ff795620.aspx>>.
- [20] MEYER, Carl D. *Google's PageRank and Beyond: The Science of Search Engine Rankings*. [s.l.] : Princeton University Press, 2006. 234 s. ISBN 0691122024.
- [21] Kubíček, Michal. *Velký průvodce SEO.: Jak dosáhnout nejlepších pozic ve vyhledávačích*. Computer Press, 2008. 320 s. ISBN 978-80-251-2195-5.

[22] *About Ask Jeeves* [online]. ? [cit. 2011-04-11]. Webmasters Information. Dostupné z WWW: <<http://sp.uk.ask.com/en/docs/about/webmasters.shtml>>.

[23] *About Ask Jeeves* [online]. ? [cit. 2011-04-11]. Webmasters Information. Dostupné z WWW: <<http://sp.uk.ask.com/en/docs/about/webmasters.shtml>>.

[24] *Seznam nápověda* [online]. ? [cit. 2011-04-18]. Pokročilé hledání. Dostupné z WWW: <<http://napoveda.seznam.cz/cz/fulltext-hledani-v-internetu/pokrocile-hledani/>>.

[25] *Jyxo* [online]. 200 [cit. 2011-04-18]. Informace. Dostupné z WWW: <<http://jyxo.vybereme.cz/d/info>>.

Přílohy

Dotazník internetové vyhledávání

V rámci bakalářské práce provádím výzkum o vědomostech uživatelů v oblasti internetového vyhledávání. Budu rád pokud mi věnujete chvilku času na vyplnění tohoto anonymního dotazníku.

1. Jste...
 - a. Muž
 - b. Žena
2. Do jaké věkové kategorie patříte?
 - a. do 18 let
 - b. do 30 let
 - c. do 40 let
 - d. do 50 let
 - e. do 60 let
 - f. nad 60 let
3. Jaké je vaše nejvyšší dosažené vzdělání?
 - a. ZŠ
 - b. SŠ nebo SOŠ
 - c. VOŠ
 - d. VŠ
4. Jak velké je město nebo obec ve které žijete?
 - a. do 1000 obyvatel
 - b. do 10 000 obyvatel
 - c. do 100 000 obyvatel
5. Jaké je vaše povolání?
 - a. Student
 - b. Student v oboru IT

- c. Zaměstnanec, podnikatel
 - d. Zaměstnanec nebo podnikatel v oboru IT
 - e. Nezaměstnaný
6. Délka praxe v oboru IT? Pokud jste na předchozí otázku odpověděli, že pracujete v oboru IT nebo studujete IT.
- a. Bez praxe
 - b. 1 – 5 let
 - c. 5 – 10 let
 - d. 10 a více let
7. Jaký internetový vyhledávač znáte a alespoň jednou jste ho využili?
- a. Google
 - b. Seznam
 - c. Jyxo
 - d. Morfeo
 - e. Bing, MSN, Live
 - f. Yahoo
 - g. Ask
 - h. Altavista
 - i. Lycos
 - j. DuckDuckGo
 - k. Zoohoo
 - l. Index
 - m. Jiné
8. Jaký internetový vyhledávač používáte jako hlavní?
- a. Google
 - b. Seznam
 - c. Bing, MSN, Live
 - d. Jyxo
 - e. Yahoo

- f. Vyhledávám pomocí svého internetového prohlížeče
 - g. Nevím
 - h. Jiné
9. Jaký internetový vyhledávač používáte jako druhý?
- a. Google
 - b. Seznam
 - c. Bing, MSN, Live
 - d. Jyxo
 - e. Yahoo
 - f. Nevím
 - g. Žádný jsme spokojen s výsledky vyhledávání mého hlavního vyhledávače.
 - h. Jiné
10. Jaký používáte alternativní internetový vyhledávač, pokud nejste spokojeni s výsledky vašeho primárního a druhého vyhledávače?
- a. Google
 - b. Seznam
 - c. Bing, MSN, Live
 - d. Jyxo
 - e. Yahoo
 - f. Žádný další
 - g. Jiné
11. Používáte operátory, jako např. uvozovky, pro upřesnění vyhledávání?
- a. Ano
 - b. Ne
12. Pokud ano jaké? (Jednotlivé operátory vepište níže.)
- _____

13. Jak byste hodnotily své zkušenosti s IT všeobecně? Na stupnici od 1 do 10.

1 = Začátečník

10 = Profesionál

14. Toto byla poslední otázka, děkuji za váš čas. Pokud máte nápady na rozšíření dotazníku, případně připomínky vyplňte je do kolonky níže.