

Univerzita Karlova v Praze

Filozofická fakulta

Ústav informačních studií a knihovnictví

Diplomová práce

Bc. Barbora Bjačková

Vyhledávání informací na internetu a jeho trendy a směry
Internet searching trends

Praha 2013

Vedoucí práce: Doc. Richard Papík, Ph.D

Prohlášení:

Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

V Praze, dne 7. srpna 2013

.....
Barbora Bjačková

Identifikační záznam

BJAČKOVÁ, Barbora. *Vyhledávání informací na internetu a jeho trendy a směry = Internet searching trends*. Praha, 2013-08-06. 110 s., 3 s. příl. Diplomová práce (Mgr.). Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví. Vedoucí diplomové práce Richard Papík.

Abstrakt (česky)

Vyhledávání na internetu se od jeho počátků hodně změnilo a změnilo i způsob vyhledávání informací. Nejprve byly vytvořeny první vyhledávací nástroje v síťovém prostředí, avšak větší rozvoj nástrojů pro vyhledávání na internetu přišel až se vznikem webu. Jedny z prvních nástrojů pro vyhledávání na internetu byly předmětové katalogy webových zdrojů, jako je katalog Yahoo! nebo katalog Open Directory Project. V současnosti jsou však převážně používány internetové vyhledávače. Kromě všeobecně zaměřených populárních vyhledávačů, jako je Google, Yahoo! nebo Bing, existují také vyhledávače specializované nebo vyhledávače specifické svým zaměřením či funkcemi, jako je DuckDuckGo se zaměřením na soukromí, Yandex či Seznam.cz zaměřené na určitý region či výpočetní vyhledávač WolframAlpha. Technologickým trendem v oblasti vyhledávání je vyhledávání multimediálních informací a vyhledávání přizpůsobené pro mobilní zařízení. K aktuálním trendům ve vyhledávání patří také personalizace, lokalizace a sociální vyhledávání. Dlouhodobým trendem je pak sémantické vyhledávání.

Klíčová slova (česky)

Internetové vyhledávače, vyhledávání informací, www služby, internet, Google

Abstrakt (anglicky)

Internet search has changed significantly since its beginning and it has also changed the way of information retrieval. Firstly, network search tools were created. However, greater development of internet search tools came after the creation of the Web. One of the first internet search tools were the web directories, such as Yahoo! or content directory Open Directory Project. Nowadays, web search engines are the most commonly used. Apart from general web search engines, there are also specialized or web search engines for particular aim or function, such as DuckDuckGo aimed at privacy, Yandex or Seznam.cz aimed at specific region or computational search engine WolframAlpha. Multimedia search and search adapted for mobile devices is technology trend in the field of internet search. Personalization, localization and social search belong among the contemporary trends. Semantic search is another long-lasting trend.

Klíčová slova (anglicky)

Web search engines, information retrieval, web services, internet, Google

Obsah

Předmluva	7
Úvod	9
1 Historie vyhledávání na internetu	11
1.1 První vyhledávací nástroje v síťovém prostředí	11
1.1.1 Archie	11
1.1.2 Usenet	11
1.1.3 Gopher	12
1.1.4 WAIS	12
1.2 Vznik webu a prvních prohlížečů	13
1.2.1 Hypertext	13
1.2.2 Vznik WWW	14
1.2.3 První webové prohlížeče	15
1.3 Počátky vyhledávání v prostředí WWW	16
1.3.1 What's new	16
1.3.2 W3Catalog	16
1.3.3 Wanderer	17
1.3.4 WebCrawler	17
1.3.5 Lycos	17
1.3.6 AltaVista	18
2 Současné internetové vyhledávače	19
2.1 Současný web a vyhledávání	19
2.1.1 Velikost webu a internetové vyhledávače	19
2.1.2 Možnosti vyhledávání na webu	20
2.1.3 Problémy vyhledávání na webu	20
2.1.4 Odraz terminologie spojené s vyhledáváním informací na internetu ve slovníku TDKIV	21
2.2 Předmětové katalogy webových stránek	23
2.2.1 Počátky předmětových katalogů	23
2.2.2 Principy pracování katalogů	24
2.2.3 Výhody a nevýhody předmětových katalogů a jejich využití v současnosti	24
2.2.4 Příklady katalogů webových stránek	26
Yahoo! Directory	26
Open Directory Project	27
Seznam odkazy	29
Firmy.cz	30
2.3 Internetové vyhledávače	31
2.3.1 Terminologie a technologie internetových vyhledávačů	31
2.3.2 Příklady konkrétních vyhledávačů	33

Google	35
Bing	45
Yahoo!	48
DuckDuckGo	50
WolframAlpha	53
Yandex	55
Seznam.cz	58
2.3.3 Porovnání vybraných internetových vyhledávačů	63
2.3.4 Metavyhledávače	66
2.3.5 Vyhledávání v hlubokém webu	67
Complete Planet	68
Scirus	68
2.3.6 Webové archivy	70
3 Trendy a směry vyhledávání na internetu	72
3.1 Vyhledávání multimediálních informací	72
3.1.1 Vyhledávání obrazových dokumentů	73
3.1.2 Vyhledávání zvukových dokumentů	76
3.1.3 Vyhledávání audiovizuálních dokumentů	77
3.2 Personalizace, lokalizace a socializace vyhledávání	79
3.2.1 Personalizované vyhledávání	79
3.2.2 Lokalizované vyhledávání	80
3.2.3 Sociální vyhledávání	81
3.2.4 Kritika a obavy spojené s personalizací vyhledávání	84
3.3 Mobilní vyhledávání	85
3.4 Sémantické vyhledávání	87
4 Závěr	90
Seznam použité literatury	93
Seznam obrázků	110

Příloha: Rešerše na téma vyhledávání informací na internetu
definována.

Chyba! Záložka není

Předmluva

V mé diplomové práci se zabývám vyhledáváním na internetu, především jeho současnými nástroji pro vyhledávání a aktuálními trendy. V textu jsem podrobně popsala a porovнала konkrétní internetové vyhledávače.

Jedním z nejdůležitějších důvodů pro výběr tohoto tématu byl můj zájem o vyhledávání na internetu. Tento zájem také souvisel s mou profesí, v době, kdy jsem si téma vybírala, jsem pracovala jako operátorka na informační telefonní lince Magistrátu hlavního města Prahy, kde bylo třeba volajícím rychle a efektivně vyhledat odpověď na jejich otázky. V současnosti pracuji v Národní knihovně na projektu *Webarchiv*¹, kde se s vyhledáváním na internetu setkáváme také velmi často, zejména při výběru elektronických zdrojů k archivaci.

Téma vyhledávání na internetu je velice rozsáhlé, proto nebylo možné obsáhnout vše, co k tomuto tématu patří. Zaměřila jsem se především na analýzu současných nejvýznamnějších a také určitým způsobem zajímavých vyhledávačů a nastínění aktuálních trendů.

Problémem se kterým jsem se při práci nejčastěji setkávala, souvisel s rychlým vývojem technologií v této oblasti. Mnoho nových funkcí, aplikací či technologických novinek bylo uvedených v průběhu psaní této práce. Nejnovější technologie jsem se pokusila zařadit především do poslední kapitoly zabývající se trendy ve vyhledávání na internetu.

V rámci této práce jsem si vytvořila uživatelské účty nebo se přihlásila do různých aplikací a sdílela s nimi svá data, abych je mohla důkladně vyzkoušet a ohodnotit.

Mezi použitou literaturou převažují elektronické zdroje, zejména z důvodu zaměření práce a také z důvodu aktuálnějšího publikování informací o novinkách ve vývoji vyhledávání na internetu.

Snahou této práce bylo vytvoření základního analytického přehledu významných a uživatelsky zajímavých vyhledávačů a popis a srovnání jejich uživatelských aspektů. Práce zaznamenává vývoj vyhledávání na internetu od jeho počátků až k nastínění jeho trendů a směrů. Stěžejní částí diplomové práce je analýza významných a progresivních internetových vyhledávačů. Vyhledávače jsou popsány ze všech významných uživatelských hledisek a jsou analyzovány jejich silné a slabé stránky. Součástí popisu

¹ Digitální archiv českých webových zdrojů, který slouží k jejich dlouhodobému uchování.

je také shrnutí podstatných rysů a uživatelské doporučení vyhledávačů. Další kapitola se zabývá nejnovějšími technikami vyhledávačů a jsou zde představeny nejvýznamnější aktuální trendy. V závěru práce jsou shrnuty nejdůležitější poznatky z oblasti vyhledávání na internetu, které jsou porovnány s minulostí, a je nastíněn možný budoucí vývoj.

Jako součást diplomové práce byla vypracována také rešerše zaměřená na vyhledávání na internetu. Tato rešerše obsahuje záznamy literatury z českého souborného katalogu, která se zabývá vyhledáváním na internetu.

Chtěla bych poděkovat vedoucímu diplomové práce doc. PhDr. Richardu Papíkovi, Ph. D. za konzultace, cenné rady a motivaci a také bych chtěla poděkovat své rodině a všem blízkým, kteří mě při psaní podporovali.

Úvod

Internet je fenomén, který naprosto zásadním způsobem změnil náš svět a ovlivnil téměř veškeré sféry lidské činnosti i života jednotlivců, od získávání informací, přes nákup zboží po způsob mezilidské komunikace. Slovníková definice charakterizuje síť internet jako "*celosvětovou počítačovou síť pracující na základě protokolů TCP/IP, které umožňují komunikaci mezi veřejnými a soukromými sítěmi, na různých typech komunikačních médií (telefonní linky, optické kabely, kabelová televize, satelit apod.) a různých technických platformách (PC, Macintosh, pracovní stanice aj.). Tato komplexní globální síť sestává z tisíců dalších nezávislých sítí provozovaných vládními agenturami, výchovně vzdělávacími a výzkumnými institucemi a soukromými obchodními společnostmi. K základním službám internetu patří elektronická pošta, World Wide Web, FTP, diskusní skupiny, elektronické konference aj. [SKLENÁK 2003a]*". Díky tomu, že se internet stále rozvíjí a zasahuje do stále více oblastí lidského života, je třeba, aby se nástroje umožňující nám orientaci v tomto virtuálním prostoru, ke kterým patří i nástroje pro vyhledávání informací, také stále přizpůsobovaly a vylepšovaly svou činnost.

Internet a jeho obsah také sehrává podstatnou roli v oblasti lidských práv. Jelikož je internet prostředí decentralizované, na rozdíl například od televize, je tak svobodným prostorem pro šíření nejrůznějších názorů a myšlenek. Jako decentralizovaný systém ho není tak snadné využít k propagandě a proto je v nedemokratických režimech přístup k síti internet omezován. Tímto však není konstatováno, že k manipulaci uživatele na internetu nedochází. Z těchto důvodů také dochází k různým snahám internet regulovat a kontrolovat, například aktuálním trendem rozšiřujícím se zejména díky sociálním sítím se rozšiřuje sbírání dat o uživateli. Navíc v dnešní době lidé využívají internet v opravdu velké míře a nechávají za sebou *digitální stopu*², která může být zneužita.

Síť internet se však neskládá pouze z toho, co je přístupné běžnými internetovými vyhledávači. Jeho součástí je také oblast nazývaná jako *hluboký web*, která je běžnými vyhledávači nedostupná, a také oblast zvaná *darknet*, která obsahuje neveřejné informace. Na rozdíl od hlubokého webu, jehož nedostupnost je způsobená především technicky, využívá *darknet* technologií anonymizace a sdílení souborů pouze

² Vzniká v důsledku našich běžných aktivit na internetu, ke kterým patří vlastní webové stránky a blogy, diskusní příspěvky, aktivita na sociálních sítích, registrace na stránkách atd.

s ověřenými uživateli. Příkladem je anonymizační síť Tor, která poskytuje takzvané skryté služby. Její součástí jsou i vyhledávače, které jsou zatím ještě na nízké úrovni. Z tohoto důvodu nebudou služby sítě *darknet* v práci dále rozebírány.

V počátcích internetu neexistovalo žádné specializované vyhledávání a malé množství uživatelů, které v té době mělo k internetu přístup, si odkazy na zdroje mezi sebou předávalo. Tato situace však nebyla dlouho udržitelná, počátkem 90. let vzniklo několik nástrojů pro vyhledávání na internetu, ovšem obrat nastal se vznikem webu. Internetové vyhledávače a portály obsahující katalogy odkazů nás provází již od vzniku webu a významně změnily způsob získávání informací.

V počátcích katalogů webových zdrojů bylo nutné se k informacím dostat, buď zdlouhavou navigací ve struktuře hierarchicky uspořádaných témat, nebo při vyhledávání pomocí internetových vyhledávačů vytvořit správný rešeršní dotaz. Díky obrovskému rozvoji informačních technologií dnes máme přístup k velkému množství informací s minimem úsilí. S masovým rozšířením výpočetní techniky a internetu se tak širokým vrstvám lidí dostaly do rukou informace a nástroje, které byly dříve vyhrazeny především odborníkům. S tím také souvisí vývoj nástrojů pro vyhledávání na internetu, které jsou dnes na takové úrovni, že uživatelé zvládnou většinu svých běžných informačních požadavků, jako je například nalezení dopravního spojení či kuchařského receptu, uspokojit sami bez pomoci informačního profesionála.

Tato snadnost, dostupnost a přístup k množství informací spolu se stále rostoucí velikostí internetu s sebou však nese určité rizika. Přesto, že se dnes internetové vyhledávače snaží uživatelům vyhledání požadované informace co nejvíce usnadnit, nastává problém posouzení objektivity a věrohodnosti informací.

Téma vyhledávání na internetu dnes zasahuje do mnoha různých oblastí. Jednou z významných oblastí, která úzce souvisí s vyhledáváním na internetu je technika nazývaná jako optimalizace pro vyhledávače (*Search Engine Optimization - SEO*), která se zabývá metodami, jak zlepšit pozici webových stránek ve výsledcích vyhledávačů. Díky svému komerčnímu využití je problematika SEO detailně rozpracována ve velkém množství literatury, a proto není předmětem této práce.

1 Historie vyhledávání na internetu

1.1 První vyhledávací nástroje v síťovém prostředí

V době před vznikem World Wide Web (WWW) bylo v síti internet sdílení dat realizováno především prostřednictvím e-mailu a služby pro přenos souborů, tzv. *File transfer protocol* (FTP). Počátky FTP, jednoho z nejstarších protokolů založeném na protokolu TCP/IP, spadají do 70. let, kdy byla vydána první specifikace pro FTP [RFC 114, 1971], k jeho velkému rozšíření došlo především v 80. letech a služba FTP se používá i vyvíjí dodnes, například pro distribuci software a dat.

FTP pracuje tak, že pokud chce někdo sdílet určitý soubor, nahraje tento soubor pomocí aplikace *FTP klient* na *FTP server*. Uživatel, který chce soubor stáhnout, použije také *FTP klienta*, kterým se připojí na tento *FTP server* a soubor stáhne k sobě.

S rostoucím množstvím souborů bylo těžké se v souborech orientovat, tak si uživatelé začali vytvářet seznamy souborů. Seznamy byly umístěny přímo na *FTP serverech* a informace o novinkách, přírůstcích či změnách si lidé posílali pomocí e-mailů či e-mailových konferencí. Po čase však počet *FTP serverů* narostl do takového množství, kdy bylo problematické stahovat všechny seznamy souborů a orientovat se v nich, proto bylo třeba vytvořit vyhledávací nástroje.

1.1.1 Archie

První vyhledávací nástroj pro prohledávání FTP serverů *Archie* vytvořili studenti McGillovy univerzity (*McGill University*) v roce 1990. Program *Archie* pracoval tak, že se pravidelně připojoval na *FTP servery* a indexoval na nich uložené soubory do svojí databáze. *Archie* indexoval pouze názvy souborů, takže bylo možné vyhledávat podle názvů souborů, nikoliv podle obsahu. Souborem mohl být nejen textový dokument, ale i obrázek, software či jiný typ dokumentu. Jako výsledek vyhledávání *Archie* vrátil uživateli cestu ke konkrétnímu adresáři, ve kterém byl uložený soubor s hledaným názvem [SEYMOUR, 2011, s. 49].

1.1.2 Usenet

Usenet je celosvětově rozšířený systém elektronických diskusních skupin. Vznikl na počátku 80. let a patří tak k jednomu z nejstarších způsobů síťové

komunikace. *Usenet* pracuje distribuovaně, na proměnné množství serverů ukládá, odesílá a přijímá příspěvky. Uživatelé si pak příspěvky vyhledávají sami na příslušném serveru. *Usenet* má hierarchickou strukturu, člení se na tématické skupiny, v nichž jsou diskusní vlákna obsahující jednotlivé příspěvky. Například skupina zabývající se počítači se nazývala *comp.** [USENET, 2001]. Již v 80. letech se správci diskusních skupin *Usenetu* rozhodli příspěvky na *Usenetu* archivovat. V roce 1995 vznikla velká webová databáze archivovaných příspěvků *Usenetu*, ve které se dalo vyhledávat. Tuto databázi v roce 2001 koupila firma Google a začlenila do své služby *Google Groups* [SHERMAN, 2001]. Tato služba dokáže vyhledávat v archivu Usenetových příspěvků do roku 1981. Dnes už systém *Usenet* z velké části nahradily e-mailové konference, diskusní fóra, blogy, online chat i sociální sítě.

1.1.3 Gopher

V roce 1991 byl vytvořen na Univerzitě v Minnesotě (*University of Minnesota*) protokol pro šíření, vyhledávání a získávání dokumentů nazvaný *Gopher*. Tento systém uživateli poskytoval hierarchicky strukturované menu, jimiž byl uživatel navigován buď ke konkrétnímu dokumentu, nebo byl odkázán na další *Gopher server* [GOPHER, 2012]. *Gopher* pracoval v textovém rozhraní a jeho ovládání bylo pro uživatele snadné, což také vedlo k jeho velkému rozšíření.

Pro ulehčení vyhledávání v *Gopher serverech* vznikly o rok později programy *Veronica* a *Jughead*. Vyhledávač *Veronica* dokázal vyhledávat podle klíčových slov v průběžně aktualizované databázi kopií strukturovaných menu *Gopher serverů* [MAKULOVÁ, 2002, s. 49]. Jako výstup vyhledávacího dotazu byla pak vytvořena nabídka výsledků hledání (*on-the-spot menu*), která měla strukturu upraveného *Gopher menu*. Vyhledávač *Jughead* se od *Veronicy* lišil v tom, že byl určený na rychlé prohledávání malého počtu serverů díky tomu, že si celou databázi držel v operační paměti [SEYMOUR, 2011, s. 49].

1.1.4 WAIS

Mezi další vyhledávací nástroje před vznikem webu patří *Wide Area Information Server*, zvaný *WAIS*. Tento systém vznikl na počátku 90. let a používal síťový protokol využívaný v knihovnách *Z39.50* pro komunikaci klientů se serverem. Tento systém pracoval distribuovaně, různé báze dat se nacházely na různých severech a dokázal

vyhledávat i další typy dokumentů jako jsou obrázky či zvuk [MAKULOVÁ, 2002, s. 50].

1.2 Vznik webu a prvních prohlížečů

1.2.1 Hypertext

Norma ČSN ISO definuje hypertext jako ukládání dokumentů ze souvisejících archivů s propojením selekčními údaji, umožňující bezprostřední přesuny v rámci dokumentu nebo do dokumentu jiného [ČSN ISO 5127:2003]. Hypertext je také definován jako nelineární uspořádání textu, jehož části jsou navzájem propojeny hypertextovými vazbami [SKLENÁK, 2003b].

K předchůdcům dnešního webu patří myšlenka Vanevara Bushe *Memex* publikovaná v roce 1945 v článku *As we may think* [1945]. V tomto článku popsal vizi stroje, ve kterém jsou uloženy informace a je možné je vyhledat. Přelomové bylo především to, že tento stroj měl pracovat na podobném principu jako lidský mozek, vytvářet asociace a tím propojovat informace mezi sebou.

Termín hypertext pak poprvé použil v první polovině 60. let Ted Nelson ve svém projektu uživatelsky jednoduché počítačové sítě *Xanadu*. Slovo hypertext je kombinací přípony hyper (nad, za) z řečtiny se slovem text [WHITEHEAD, J, 2000, s. 8]. V projektu *Xanadu* měl hypertext fungovat pro propojování souborů nebo různých verzí dokumentů. Tento projekt však nebyl nikdy úspěšně implementován [PROJECT XANADU, 2001].

Úspěšnějším pokusem o implementaci hypertextu se stal systém *oN-Line System* (NLIS), který v 60. letech vytvořil Douglas Engelbart. Engelbart pracoval na projektu *Augment*, kdy se snažil pomocí nových počítačových technologií zvýšit lidskou produktivitu a schopnosti (*Augmenting Human Intellect*). Jednou z částí tohoto projektu, který byl také využíván pro správu interních dokumentů a materiálů se stal kolaborativní systém NLIS, který tyto dokumenty hypertextově propojoval. V rámci projektu vzniklo několik technických novinek jako například správa verzí dokumentu, interaktivní multimedia (*hypermédia*) nebo také počítačová myš [GRIFFIN, 2000]. Nevýhodou systému bylo to, že byl uživatelsky náročný. Tento systém se neujal.

1.2.2 Vznik WWW

Síť World Wide Web (WWW) vytvořil absolvent Oxfordu Tim Berners-Lee na počátku 90. let 20. století v Evropské organizaci pro jaderný výzkum (CERN), která byla v té době jako jedna z mála organizací připojena k mezinárodní počítačové síti Internet.

Berners-Lee vytvořil nejprve hypertextový program *ENQUIRE* pro zaznamenání informací o lidech, programech či hardware v rámci CERNu a vztazích mezi nimi. Protože si Tim Berners-Lee uvědomil, že chce vytvořit hypertextový systém, který je ale dostupný všem i mimo CERN, začal pracovat na konceptu sítě WWW [BERNERS-LEE, 1994].

Tim Berners-Lee začal spolupracovat s programátorem Robertem Cailliu na projektu nazvaném *WorldWideWeb*, zkráceně W3. Cílem bylo vytvoření sítě hypertextově propojených dokumentů umístěných na různých serverech [BERNERS-LEE, 1990]. Berners-Lee využil hypertext tak, že jednotlivé dokumenty (webové stránky) napsané ve značkovacím jazyku *Hypertext Markup Language* (HTML) umožnil navzájem provázat pomocí odkazů.

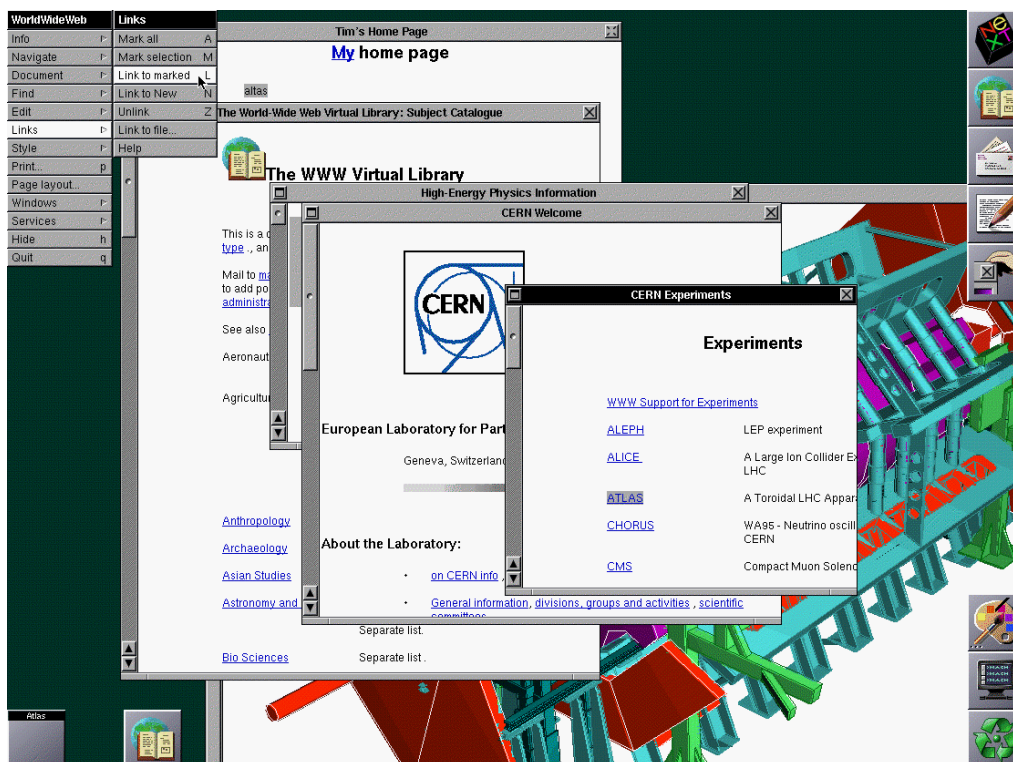
První webový server³ běžel na počítači *NeXT* a na něm také Berners-Lee vytvořil v prosinci roku 1990 první webový prohlížeč (*browser*) s názvem *WorldWideWeb*. Prohlížeč byl však brzy přejmenován na *Nexus*, protože původní název byl matoucí, jelikož původním názvem je označována i celá síť World Wide Web. Tento prohlížeč byl zároveň jednoduchým WYSIWYG (*what you see is what you get*)⁴ editorem webových stránek [BERNERS-LEE, 2013].

V srpnu roku 1991 zveřejnil Tim Berners-Lee základní informace o projektu World Wide Web v diskusní skupině *Usenetu alt.hypertext*. Od této doby se Web začíná stávat veřejným.

V roce 1994 Berners-Lee založil mezinárodní konsorcium *World Wide Web Consortium* (W3C), jehož cílem je vyvíjet spolu s veřejností otevřené standardy pro Web za účelem jeho dlouhodobého rozvoje [WORLD WIDE WEB CONSORTIUM, 1995].

³ První webový server byl nazvaný nxoc01.cern.ch, později přejmenovaný na info.cern.ch, a první webová stránka byla na adrese <http://nxoc01.cern.ch/hypertext/WWW/TheProject.html>.

⁴ Způsob editace dokumentu, kdy je verze zobrazená na obrazovce vzhledově shodná s výslednou verzí dokumentu.



Obrázek č. 1: První webový prohlížeč WorldWideWeb Tima Bernerse Lee [BERNERS-LEE, 2013]

1.2.3 První webové prohlížeče

Důležitým mezníkem v historii Webu se stalo vytvoření prvního grafického prohlížeče s názvem *Mosaic*, díky němuž došlo k velkému rozšíření Webu. Prohlížeč *Mosaic* vytvořil americký programátor Marc Andreessen v Národním centru pro superpočítačové aplikace (National Center for Supercomputing Applications, NCSA) v roce 1993 pro unixový systém *X Window*, poté vznikla verze pro Apple Macintosh a Microsoft Windows [VETTER, 1994]. *Mosaic* se stal brzy velice oblíbeným⁵ především pro své intuitivní grafické rozhraní, podporu multimédií, licenci umožňující využití pro nekomerční účely zdarma a výhodou byl také tým programátorů, který se o prohlížeč průběžně staral. K zániku tohoto prohlížeče přispěl vznik dalších prohlížečů, včetně jeho pokračovatele, prohlížeče *Netscape*.

Prohlížeč *Mosaic* tak dal vzniknout dvěma hlavním větvím prohlížečů, *Netscape* (později *Mozilla Firefox*), který vytvořila firma Marca Andreessena, a *Microsoft Internet Explorer*, který využil zdrojový kód prohlížeče *Mosaic*, jehož licenci prodala firma Spyglass Microsoftu [MOSAIC, 1996].

⁵ V lednu roku 1994 podle průzkumu Georgijského technického institutu (Georgia Institute of Technology) využívalo prohlížeč *Mosaic* 97 % dotázaných uživatelů [GEORGIA INSTITUTE OF TECHNOLOGY, 1994].

Koncem 90. let došlo k situaci, kdy trh ovládaly dva prohlížeče (*Netscape* a *Internet Explorer*), které soupeřily o dominanci na trhu, tomuto období se říká „válka prohlížečů“ (*browser wars*) [BROWSER WARS, 2013]. V tomto období byly do prohlížečů přidávány nové rozšiřující vlastnosti jako nové verze HTML, JavaScript, CSS a zásuvné moduly (plugíny) jako je Java nebo Flash. Rozšiřovala se také spolupráce s ostatními aplikacemi, například integrace schránky pro kopírování a vkládání a podpora gest myši (*pointing device gesture*). Toto období rozšiřování bylo na druhou stranu i obdobím nekompatibility stránek s různými prohlížeči.

V současnosti (2012) se stal nejpoužívanějším prohlížečem *Chrome* od firmy Google (29,4 %), druhým je *Internet Explorer* (27,8 %) a třetím *Mozilla Firefox* (20,1 %) [W3COUNTER, 2012]. Nově také přibývá na popularitě mobilních prohlížečů, kde dominuje především *Safari*, poté prohlížeč pro operační systém *Android* a *Opera Mini* [NETMARKETSHARE, 2013].

1.3 Počátky vyhledávání v prostředí WWW

1.3.1 What's new

Po velkém nárůstu webových stránek⁶ na počátku 90. let byli uživatelé informováni o vzniku nových webových stránek prostřednictvím oznámení typu *What's New Web Pages*. Takovou službou byla například výběrová *Netscape What's New*. Registry webových stránek vycházely i v tištěné podobě, v časopisech nebo jako publikace, například *New Rider's Official Internet and World Wide Web Yellow Pages* [MAKULOVÁ, 2002, s. 52].

1.3.2 W3Catalog

Jedním z prvních vyhledávacích nástrojů v prostředí WWW se stal *W3Catalog*, který vznikl v roce 1993 na Ženevské univerzitě (*Université de Genève*) a fungoval do roku 1996. *W3Catalog* využil již existující kvalitní manuálně vybrané seznamy webových stránek, ve kterých však nebylo možné vyhledávat. Vyhledávač vytvořil obrazy webových zdrojů a přeformátoval je do jednoho prohledávatelného seznamu [NIERSTRASZ, 1996].

⁶ Statistika nárůstu webových stránek získaná pomocí programu Wanderer na stránkách Massachusettského technického institutu <http://www.mit.edu/people/mkgray/net/web-growth-summary.html>.

1.3.3 Wanderer

V roce 1993 vytvořil student Massachusettského technického institutu (*Massachusetts Institute of Technology*, MIT) Matthew Gray vyhledávač *The World Wide Web Wanderer*, který měl sloužit k změření velikosti Webu. Tento vyhledávač byl prvním webovým robotem, který indexoval webové stránky a vytvářel index zvaný *Wandex* [SEYMOUR, 2011, s. 49-50].

1.3.4 WebCrawler

První vyhledávač, který dokázal vyhledávat full-textově, byl spuštěn v roce 1994. *WebCrawler* byl vytvořen studentem Washingtonské univerzity v Seattlu (*University of Washington*) Brianem Pinkertonem a zpočátku pracoval jako desktopová aplikace⁷ [SEYMOUR, 2011, s. 50]. *WebCrawler* indexoval obsah celé stránky a poté vypočítal poměr počtu slov na stránce k počtu vyhledávaných slov a výsledky pak řadil podle relevance. Jiné vyhledávače do této doby dokázaly indexovat pouze název stránky, její URL adresu a nejvíce 100 slov [MAKULOVÁ, 2002, s. 54]. *WebCrawler* v současnosti funguje jako metavyhledávač a vlastní ho firma InfoSpace (dnes pod názvem Blucora), která vlastní i další metavyhledávače jako je *Dogpile* či *MetaCrawler* [WEBCRAWLER, 2013].

1.3.5 Lycos

Vyhledávač *Lycos* vznikl v roce 1994 na Carnegie-Mellonově univerzitě (*Carnegie Mellon University*). Jeho název je odvozen od jména dravého pavouka Lycosidae (vlčí pavouk). Tento vyhledávač se vyznačoval zejména velkou databází indexovaných stránek. V době vzniku vyhledávače obsahovala jeho databáze 54 tisíc záznamů stránek, o rok později 1,5 milionů a v roce 1996 obsahovala již 60 milionů záznamů, což bylo více, než měly jiné vyhledávače [SHERMAN, 2002]. Na výstupu vyhledávání pomocí systému *Lycos* byly jednotlivé stránky uvedeny včetně abstraktu vytvořeného automatickou analýzou textu. *Lycos* také rozlišoval typy dokumentů, takže bylo možné hledat například pouze obrázky [MAKULOVÁ, 2002, s. 55]. *Lycos* dodnes změnil řadu majitelů, fungoval jako portál a začleňoval další funkce jako například vlastní e-mailovou schránku, nahrávání videí, online hry, video-chat a další.

⁷ Aplikace primárně pracuje s lokálními daty a používá lokálně instalované nástroje.

V současnosti jej vlastní indická společnost Ybrant Digital a zaměřuje se na služby multimediálního charakteru [LYCOS, 2013].

1.3.6 AltaVista

Vyhledávač *AltaVista* byl vytvořen v roce 1995 Louisem Monierem a Michaelem Burrowsem a měl ukázat potenciál v té době nejvýkonnějšího serveru společnosti Digital Equipment Corporation [ALTAVISTA, 2001]. *AltaVista* byl nejrychlejším vyhledávačem ve své době, zvládl přes 13 milionů dotazů za den a stal se také velmi populárním [SILVERSTEIN, 1999]. Tento vyhledávač byl nejen velmi rychlý, ale přinesl i novinky ve způsobu vyhledávání. Uživatel mohl své dotazy formulovat v přirozeném jazyce. Mimo klasických Booleovských operátorů také nabízel využití proximitního operátoru NEAR, kde mohly být vyhledávané termíny od sebe vzdálené až 10 slov [RALPH, 1996]. Výhodou vyhledávače byla také indexace souborů ve formátu PDF a stránek v různých jazycích. *AltaVista* také jako první spustila online překladač *Babelfish*, který dokázal překládat nejen slova, ale celé věty [SULLIVAN, 1997]. V konkurenci vyhledávače *Google* však *AltaVista* neobstál a v roce 2003 jej koupila firma Overture Services, kterou brzy převzala firma Yahoo!. *AltaVista* dnes stále existuje (na doméně <http://www.altavista.com/>), ale pracuje na vyhledávací technologii vyhledávače *Yahoo!* (dnes již *Bing*) [HASALÍK, 2007].



Obrázek č. 2: Vyhledávač AltaVista v roce 1999 (dostupné prostřednictvím služby Internet Archive)

2 Současné internetové vyhledávače

2.1 Současný web a vyhledávání

2.1.1 Velikost webu a internetové vyhledávače

Současné vyhledávače mají náročný úkol, vyhledat, projít a indexovat co největší počet webových stránek, protože počet webových stránek stále narůstá. Vznik vyhledávačů je spojen se snahou zmapovat web. V polovině 90. let byla změřena velikost webu pomocí vyhledávače s názvem *Wanderer*, který procházel webové stránky a zjišťoval jejich množství [SEYMOUR, 2011, s. 49-50].

Dnes se určuje velikost webu podle indexů nejpoužívanějších vyhledávačů. Ovšem takto zjištěná velikost webu není přesná, jelikož se omezuje pouze na povrchový web, který je běžně dostupný vyhledávači a nezohledňuje tak velikost hlubokého webu, kam vyhledávače přístup nemají. Odhaduje se, že velikost hlubokého webu je až 500 krát větší než je velikost webu povrchového [BERGMAN, 2001].

V roce 2008 oznámila společnost Google na svém oficiálním blogu, že jejich systémy zpracovávající odkazy na webu detekovaly již bilión jedinečných URL adres [ALPERT, HAJAJ, 2008]. Tento počet se ovšem nedá brát jako reálný počet webových stránek jelikož obsahuje duplicity i spam a vyhledávač *Google* ani neindexuje všechny tyto URL adresy. Odhadovaná skutečná velikost indexu Google byla v roce 2008 40 miliard webových stránek [ARRINGTON, 2008].

Velikostí webu se zabývá služba s názvem *WorldWideWebSize.com* (na adrese <http://www.worldwidewebsize.com/>), která zobrazuje odhad počtu webových stránek pomocí jazykové analýzy a statistických metod získaných vyhledávači *Google*, *Bing* a *Yahoo!*. Podle aktuálního měření (ze dne 21. 3. 2013) je minimální počet indexovaných webových stránek 14,67 miliard [WORLDWIDEWEBSIZE.COM, 2013].

Využitím a přínosem webu a především jeho socio-ekonomickými dopady na obyvatele konkrétních států se zabývá nadace *World Wide Web Foundation* v projektu nazvaném *The Web Index* (na adrese <http://thewebindex.org/>) [WORLD WIDE WEB FOUNDATION, 2012].

2.1.2 Možnosti vyhledávání na webu

Vyhledávat informace v prostředí sítě World Wide Web je možné několika způsoby. Nejsnazším způsobem vyhledání určité informace je zadání URL adresy do prohlížeče, to je ovšem podmíněno uživatelskou znalostí správné URL adresy. Tento způsob vyhledávání lze uplatnit pouze v případech, že uživatel vyhledává konkrétní stránky, ne už však informace či objekty. Dalším způsobem vyhledávání je navigace pomocí katalogů webových stránek, zejména prostřednictvím webového portálu, které kromě katalogu nabízí i další služby. Tyto katalogy obsahují obvykle předmětově a hierarchicky uspořádané odkazy na webové zdroje. Nejčastějším způsobem vyhledávání informací na internetu je v současnosti využití internetových vyhledávačů. Proces vyhledávání informací internetovým vyhledávačem se skládá z formulace dotazu, výběru z výsledků nabízených vyhledávačem, navigace pomocí odkazů na další stránky neboli "surfování po webu" a případná úprava vyhledávacího dotazu [LEVENE, 2010, s. 26-27].

V roce 2002 Andrei Broder klasifikoval uživatelské dotazy při vyhledávání na internetu podle uživatelského záměru do tří kategorií [2002, s. 3]:

1. navigační (*navigational*) – dotaz s cílem naleznout konkrétní stránku
2. informační (*informational*) – dotaz s cílem nelézt informace o nějakém tématu
3. transakční (*transactional*) – cílem je naleznout stránku, se kterou chce uživatel dále komunikovat (např. stáhnout film, objednat zboží aj.)

Podle Broderova průzkumu jsou nejčastější dotazy informační a nejméně časté dotazy navigační. Informační dotazy se nejvíce podobají klasickému vyhledávání informací například v akademických databázích.

2.1.3 Problémy vyhledávání na webu

Přesto, že vývoj nástrojů pro vyhledávání v internetu značně pokročil a vyhledávače dnes mají mnohem více funkcí a dovedností než tomu bylo v jejich počátcích, některé problémy spojené s vyhledáváním na internetu stále přetrvávají. Web je dynamický a stále se mění, velké množství stránek každý den vznikne, ale i zanikne. Vyhledávače proto musí svůj index často aktualizovat, aby jejich výsledky odpovídaly aktuální situaci.

Dalším problémem je různá kvalita obsahu. Díky masovému rozšíření internetu a stále větší části obsahu tvořeného přímo uživateli (sociální sítě, komentáře ke článkům, recenze a hodnocení produktů atd.) není snadné vybrat a posoudit, který zdroj

je skutečně relevantní. S tématem kvality obsahu je i silně spojena jeho subjektivita a to nejen ve výsledcích vyhledávání, ale i při formulaci dotazů.

Internetové vyhledávače si dnes kladou za cíl indexovat co největší množství obsahu na internetu. Díky tomu na dotaz uživatele nabízejí takové množství výsledků, ve kterém se uživatel není schopen orientovat, proto obvykle vybírá pouze z první či několika málo dalších stran nejrelevantnějších výsledků. Problémem kromě tohoto informačního přetížení může být také řazení výsledků podle individuálních informací o konkrétním uživateli, což je dnes běžná praxe velkých internetových vyhledávačů. Uživatel tak mezi těmito personalizovanými výsledky nemusí získat pro něj relevantní zdroj.

Po dlouhou dobu bylo problémem i vyhledávání netextových informací, ovšem v současnosti je tomuto tématu věnováno více pozornosti a již se dají i některé typy netextových dokumentů vyhledávat pomocí jejich obsahu.

2.1.4 Odraz terminologie spojené s vyhledáváním informací na internetu ve slovníku TDKIV

Česká terminologická databáze knihovnictví a informační vědy (TDKIV), kterou vytváří a spravuje *Národní knihovna ČR*, je osvědčeným zdrojem informací z této oblasti [KTD, 2003]. Ačkoliv je zaměřena především na oblast knihovnictví a informační vědy, zahrnuje i termíny spojené s vyhledáváním informací na internetu.

Vyhledávání informací na internetu

Databáze TDKIV nevymezuje pojem *vyhledávání informací na internetu*, obsahuje pouze obecnější termín *vyhledávání informací*. Ten definuje jako „*činnost, jejímž cílem je identifikace relevantních dokumentů nebo informací v informačních zdrojích (např. plnotextové nebo bibliografické databáze). Vyhledávání informací probíhá obvykle na základě konkrétního požadavku uživatele za pomoci dotazovacích a selekčních jazyků [ŠVEJDA, 2003].*“ Pro porovnání, anglická verze internetové encyklopedie Wikipedia obsahuje termín *Internet search*, který ovšem přesměruje na heslo *Web search engine*. To však není úplně přesné, jelikož internetové vyhledávače jsou pouze jedním z nástrojů vyhledávání na internetu.

Nástroje vyhledávání informací na internetu

Mezi nástroje sloužící k vyhledávání informací na internetu patří předmětové katalogy, internetové vyhledávače a metavyhledávače. TDKIV definuje obecný termín *vyhledávací nástroj (search tool)* jako „*obecné označení pro služby umožňující vyhledávání informací v prostředí internetu [SKLENÁK, 2003c].*“

Pro pojem předmětového katalogu webových stránek je v TDKIV používán termín *katalogový vyhledávací nástroj*, který je definován jako „*předmětově uspořádaný adresář webových zdrojů, který je založen na intelektuálním sběru informací o zdrojích dostupných na WWW. Jeho základní funkcí je vyhledávání informačních zdrojů pomocí hierarchicky uspořádaných tematických kategorií [SKLENÁK, 2003d].*“

Databáze TDKIV pro vymezení pojmu internetového vyhledávače používá termín *vyhledávací stroj* (indexační program) a jeho definice se týká především jeho činnosti a architektury: „*Jeden ze základních typů vyhledávacích nástrojů na internetu. Databáze vyhledávacího stroje je budována na základě automatizovaného sběru dat. Vyhledávací stroj tvoří čtyři základní funkční části: roboty (jejich hlavním úkolem je sběr informací na WWW), indexační program (zpracovávající informace získané roboty), vyhledávací program (vyhledávací algoritmus a související programy) a grafické rozhraní (sbírá dotazy od uživatele, předává je vyhledávacímu stroji a zobrazuje výsledky hledání uživateli) [SKLENÁK, 2003e].*“ Z jednotlivých prvků architektury internetového vyhledávače obsahuje slovník TDKIV vlastní heslo pouze pro termín *robot* internetového vyhledávače. *Indexovaný soubor* je zde vysvětlen obecněji a není přímo vztažen k internetovému vyhledávači, stejně jako *uživatelské rozhraní*.

Metavyhledávač, v TDKIV nazvaný jako *metavyhledávací stroj* je definován jako „*druh vyhledávací služby v prostředí internetu, která umožňuje uživateli na základě jednoho dotazu paralelní prohledávání databází několika vyhledávacích služeb [SKLENÁK, 2003f].*“

2.2 Předmětové katalogy webových stránek

Pro systémy anglicky nazývané jako *directories*, případně *web directories* se v české či slovenské literatuře vyskytuje pojmenování předmětové katalogy (Sklenák) nebo předmětové adresáře (Makulová), případně katalogové vyhledávací nástroje (TDKIV). Pro tyto systémy bude dále v textu používán název předmětový katalog, jelikož je rozšířenější.

Předmětový katalog je služba, která uživateli nabízí seznam odkazů, které jsou hierarchicky uspořádány v tématických kategoriích a podkategoriích. Charakteristikou předmětového katalogu je, že není vytvářen zcela automaticky, ale je zde zahrnuta i práce člověka. Cílem předmětových katalogů je snaha o strukturované uspořádání webových zdrojů a odstranění nevýhod, které přináší automatická indexace pomocí robotů [SKLENÁK, 2001, s. 252].

Faktorem přispívajícím k vzniku předmětových katalogů na počátku 90. let byla úroveň tehdejších vyhledávačů, které ještě nedokázaly prohledávat plné texty, ale hledaly především v názvech a klíčových slovech. Uživatel tak musel vědět, co přesně chce vyhledat. Naproti tomu katalogy webových stránek uživateli umožňovaly vybrat si z nabízených stránek zařazených do předmětové kategorie dle jeho zájmu.

2.2.1 Počátky předmětových katalogů

Úplně prvním předmětovým katalogem se stala *The WWW virtual library* (VLIB), kterou založil Tim Berners-Lee v roce 1991. Tento katalog vytvářejí dobrovolníci z celého světa, kteří přidávají a zařazují odkazy na internetové zdroje z témat, ve kterých vynikají [THE WWW VIRTUAL LIBRARY, 2008]. Dalším významným katalogem webových zdrojů je služba *EINet Galaxy*, která vznikla v roce 1994. Tato služba obsahuje indexovanou databázi, roztříděnou do kategorií a podkategorií podle věcných témat. Vyhledávat je možné buď procházením jednotlivých kategorií, nebo hledáním pomocí klíčových slov. Vyhledávat je možné jak v názvech kategorií a podkategorií, v názvech zdrojů i v samotných URL adresách [EINET, 1994].

Předmětové katalogy byly často součástí webových portálů (např. Yahoo! nebo Seznam.cz). Tyto portály před rozšířením vyhledávačů fungovaly jako jakási „brána do světa Internetu“, pro navigaci v internetu používaly katalog s tématicky uspořádanými kategoriemi, později začaly začleňovat i další služby jako například zpravodajství, e-mailové schránky či zábavu (hry, chat atd.).

V minulosti vznikaly také speciální akademické předmětové katalogy, které mohly být zaměřené všeobecně vědecky nebo na konkrétní téma. Na rozdíl od komerčních katalogů, které měly za cíl co nejvyšší návštěvnost, se zaměřovaly na vytváření souboru kvalitních ověřených zdrojů s cílem usnadnit získání akademických informací. Příkladem takového katalogu je služba *Academic Info* (v současnosti stále existuje na adrese <http://www.academicinfo.net/>) zaměřená na vzdělávání. Katalog obsahoval zdroje určené především pro studenty středních škol a vyšších. Byl rozčleněn podle vědních oborů. Jako příklad akademického katalogu s širším záběrem může být katalog *Librarian's Internet Index* (dnes na adrese <http://www.ipl.org/>) s kvalitními zdroji vybranými knihovníky veřejných knihoven, sloužící především pro jejich uživatele [PANDIA, 2001].

2.2.2 Principy pracování katalogů

Z počátku se o katalogy webových stránek starali především správci těchto služeb. Tito správci sledovali situaci a novinky v prostoru webu a na základě tohoto pozorování věcně spravovali katalog. To znamenalo především vyhledání a zařazení zdrojů do tematických kategorií. Později začaly katalogy umožňovat vkládání návrhů na zařazení zdroje uživatelům. Odkaz na webový zdroj je v katalogu obvykle opatřen názvem stránky, URL adresou, názvem kategorie, kam je zařazen a krátkým popisem obsahu. V současnosti také katalogy přidávají k odkazu náhled stránky [SKLENÁK, 2001, s. 253-4]. Mnoho stránek tematicky nenáleží pouze do jedné kategorie, u některých katalogů jako je například *EINet* může být jeden odkaz zařazen ve více kategoriích.

Vyhledávání v předmětových katalozích bylo realizováno v počátcích především volným prohlížením katalogu (*browse*), ale postupem času se katalogy zvětšovaly tak, že nebylo možné si v nich přehledně prohlížet zdroje. Do katalogů poté začalo být integrováno jednoduché vyhledávání, nejčastěji v kategoriích a podkategoriích nebo v názvech zdrojů.

2.2.3 Výhody a nevýhody předmětových katalogů a jejich využití v současnosti

Dnes jsou vyhledávače na vysoké úrovni, dokáží vyhledat mnoho relevantních informací, proto již není vysoká potřeba předmětových katalogů. V minulosti byly však katalogy vhodné zejména na určité typy vyhledávání. Předmětový katalog je vhodné využít, pokud je oblast našeho zájmu široká a máme dostatek času pro procházení

hierarchickou strukturou katalogu, pokud na výstupu požadujeme pouze seznam několika zdrojů k určitému tématu a nechceme procházet i nerelevantní zdroje, které nám poskytne vyhledávač [MAKULOVÁ, 2003, s. 380-381].

Výraznou výhodou katalogů, kterou zatím vyhledávače nemají, je jejich odborníky vytvářený kvalitní obsah. Ten je však podmíněn častým monitoringem katalogu i webu a také dostatečným množstvím editorů, kteří se věnují tématům svého odborného zájmu. Při každodenním výrazném nárůstu webových stránek je těžké udržovat předmětový katalog webových stránek aktualizovaný zejména v katalogích, které pokrývají mnoho témat s globálním záběrem.

V současnosti se klasické předmětové katalogy používají mnohem méně a ustupují tak vyhledávačům. Často fungují ještě v portálech jen jako doplněk vyhledávačů. To je dáno nevýhodami katalogů, především časem věnovaným vyhledávání a výrazným nepoměrem mezi počtem vyhledávači indexovaných stránek a počtem stránek zařazených v katalogích a jejich aktualizaci. Příkladem může být služba *Adresář Google (Google Directory)*, která v roce 2011 skončila. Společnost Google, která tuto službu spustila v roce 2000 jako konkurenci známému katalogu *Yahoo! Directory*, definitivně přestoupila pouze k fulltextovému vyhledávání s vysvětlením, že společnost Google věří, že vyhledávač je nejrychlejší cestou, jak najít požadované informace [YOUNG, 2011]. Dnes se předmětové katalogy používají především v komerční oblasti jako katalogy firem či zboží.

Výhody:

- vhodné pro úzce tematicky vymezené oblasti
- lidské zpracování, záruka kvality (stránky jsou vybírány odborníkem)
- vhodné pro firmy, výrobky
- hierarchická struktura, anotace

Nevýhody:

- malé pokrytí webu
- náročné na vytváření a správu (web se dynamicky rozvíjí – neplatné odkazy, zaniklé stránky)
- subjektivita (ve výběru zdrojů a jejich zařazení nebo popisu, různé tematické třídění v různých katalogích – neexistuje jednotné)
- uživatelsky i časově náročnější

2.2.4 Příklady katalogů webových stránek

Katalogy webových stránek – celosvětový záběr:

Yahoo! Directory



Jedním z nejznámějších předmětových katalogů webových stránek je *Yahoo! Directory* (na stránce <http://dir.yahoo.com/>), který vytvořili v roce 1994 doktorandi Stanfordovy univerzity (*Stanford University*) Jerry Yang a David Filo, v této době ještě pod názvem *Yahoo!*⁸ [PETERKA, 1996a]. Katalog *Yahoo! Directory* se později stal součástí portálu nazvaného celkově *Yahoo!*, který obsahuje mnoho služeb jako je e-mailová schránka, zpravodajství, finanční informace, hry a mnoho dalších. Součástí portálu je od roku 2003 také vlastní vyhledávač *Yahoo! Search*.

Yahoo! Directory obsahuje 14 předmětových kategorií, které jsou dále hierarchicky členěny. Některé podkategorie mohou být zařazeny ve více vyšších kategoriích, například podkategorie Fotografie (*Photography*) může být zařazena jak v kategorii Umění (*Art*) nebo Obchod a ekonomie (*Business and Economy*), ty jsou potom značeny znakem zavináče @. Záznam odkazu webové stránky obsahuje název stránky, stručný popis a URL adresu stránky.

Katalog je vytvářen editory manuálně, ale navrhnout přidání odkazu webové stránky může každý uživatel. Uživatel nejprve vybere kategorii, a pomocí tlačítka „Navrhnout stránku“ (*Suggest a site*) otevře formulář. Poslat návrh je však možné až po přihlášení uživatele pomocí *Yahoo! ID*. Návrh je pak posouzen a případně schválen editorem katalogu.

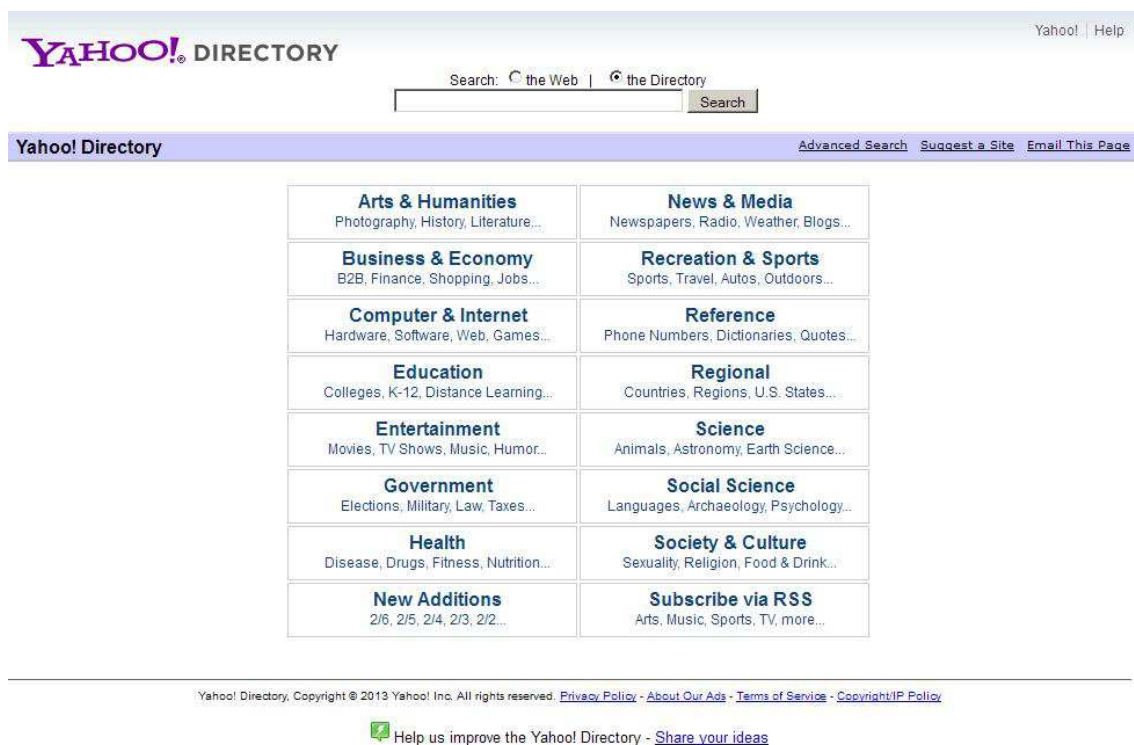
Uživatel může katalog volně prohlížet nebo v něm pomocí klíčových slov vyhledávat. Je také možné si prohlédnout nově přidané odkazy webových stránek podle data jejich přidání do katalogu nebo se přihlásit k odběru novinek pomocí služby RSS⁹ ve vybrané kategorii katalogu. Do katalogu jsou zařazeny jak nekomerční, tak i komerční stránky a v jakémkoliv jazyce. Pro odkazy z různých zemí je v katalogu speciální kategorie dle regionu (*Regional*).

Katalog *Yahoo! Directory* je komerční službou, pro uživatele je zdarma a je financována převážně z reklamy. Katalog také nabízí zpoplatněné nadstandardní služby

⁸ Akronym pro „*Yet Another Hierarchical Official Oracle*“ (volně přeloženo „ještě další hierarchický neformální/dotěrný Oracle“).

⁹ Formát pro publikování seznamu odkazů na aktualizované WWW stránky za účelem rychlého informování uživatele o novinkách a změnách (TDKIV).

jako například garance schválení či odmítnutí navržené webové stránky do 7 dnů [YAHOO!, 2013a].



Obrázek č. 3: Hlavní tematické kategorie katalogu Yahoo! Directory

Open Directory Project



Open Directory Project (ODP), známý také pod zkratkou *Dmoz*¹⁰ (na adrese <http://www.dmoz.org/>) je jedním z největších mezinárodních předmětových katalogů webových stránek. *ODP* vznikl ve Spojených státech v roce 1998. Dnes katalog vlastní firma Netscape, ale o jeho správu se stará velká skupina dobrovolníků z různých zemí. *ODP* je nekomerční katalog, proto také neobsahuje žádnou reklamu. Účel katalogu je především výběrový, *ODP* nechce pokrýt celý web, ale snaží se vytvořit objektivní nezávislý katalog kvalitních zdrojů ze všech oblastí lidského poznání a zájmů.

Impulsem pro vznik *ODP* se stal stav katalogu *Yahoo! Directory*. Tvůrci katalogu *ODP* Rich Skrenta a Bob Truel viděli problém v přílišné komercializaci společnosti Yahoo! [SHERMAN, 2000 s. 44]. Firma Yahoo! se začala koncem 90. let zaměřovat na další služby než byl katalog, zatímco se web dynamicky proměňoval a správce katalogu *Yahoo! Directory* začalo být na kvalitní pracovní katalogu příliš málo. To v důsledku vedlo k tomu, že v katalogu narůstalo množství nefunkčních odkazů, nerelevantních informací a bylo obtížné nechat zdroj zařadit do katalogu. Oba

¹⁰ Zkratka odvozená z původní URL adresy directory.mozilla.org.

tvůrci zaměstnaní ve firmě Sun Microsystems zaměřili svou pozornost na koncept software s otevřeným zdrojovým kódem (*open source software*)¹¹ a rozhodli se vytvořit předmětový katalog webových stránek na tomto principu [SHERMAN, 2000 s. 44].

Hlavním rozdílem mezi *ODP* a jinými předmětovými katalogy je, že *ODP* je zcela vytvářen a udržován dobrovolníky. V současnosti (leden 2013) na projektu *ODP* spolupracuje necelých 100 tisíc editorů a katalog obsahuje přes 5 milionů odkazů na webové stránky zařazených ve více než 1 milionu kategorií [NETSCAPE, 2013]. Zdroje do katalogu přidávají především editoři katalogu, ale stránky mohou navrhovat i uživatelé. Uživatel, který se chce stát editorem katalogu *ODP* si podá žádost, která musí být schválena současným editorem. Uživatel si vybírá tématickou kategorii, o kterou se zajímá nebo jí rozumí a do ní pak jako schválený editor může navrhovat webové stránky [NETSCAPE, 2012a]. Editoři mohou navrhovat stránky ve všech světových jazycích, katalog totiž obsahuje i kategorie podle jazyka¹². Editoři ovšem nemusí pouze navrhovat a zařazovat webové stránky do katalogu, ale podílejí se také na vývoji či radí novým editorům. Celý katalog *ODP* obsahuje 16 hlavních kategorií. Tvůrci katalogu nejprve uvažovali nad knihovnickým Deweyho desetinným tříděním, ale nakonec se rozhodli kategorie *ODP* postavit na konceptu struktury systému *Usenet* [SULLIVAN, 1998]. Kvalita katalogu je zajišťována především ve spolupráci a kontrole editorů navzájem. Pro komunikaci editorů slouží diskusní fórum nebo e-mailová komunikace [NETSCAPE, 2012b]. Pro aktuálnost odkazů byl vytvořen speciální nástroj, který pravidelně prochází odkazy v katalogu a kontroluje jejich funkčnost.

Vzhledem k tomu, že je *ODP* softwarem s otevřeným zdrojovým kódem, může si data ze systému kdokoliv převzít. Toho také využily vyhledávače jako například *Lycos*, který dával odkazům z *ODP* ve svém indexu význačné postavení, *HotBot*, *AltaVista*, *Dogpile* a další. Díky tomu, že se při použití převzatých dat z *ODP* zavazovali tito příjemci zveřejnit informaci o původu dat, se stal katalog *ODP* populárnějším a získal mnoho nových editorů [SHERMAN, 2000 s. 46]. Data z *ODP* se také stala základem dnes již zaniklého předmětového katalogu *Google Directory*.

¹¹ Software, který má dostupný zdrojový kód a jehož licenční ujednání poskytuje uživateli právo využívat program pro libovolné účely, právo studovat funkci programu a měnit ji dle svého uvážení, právo dále šířit kopie původního programu nebo pozměněné verze programu [HAVLOVÁ, 2003].

¹² Kategorie českých stránek katalogu *ODP* obsahuje 24 146 odkazů webových stránek ke dni 31. 1. 2013 [NETSCAPE, 2013].

dmz open directory project In partnership with AOL Search.

projekt | domov | sitemap | aktualizovat | spravovat | nahlast | zneuziti | pomoc

Hledat v celém adresáři ODP

Top: World: Český: Volný čas: Cestování: Cestopisy (146) [Informace](#)

- [Kolem světa \(11\)](#)

- [Afrika \(4\)](#)
- [Jižní Amerika \(1\)](#)
- [Asie \(27\)](#)
- [Oceánie \(4\)](#)
- [Evropa \(77\)](#)
- [Severní Amerika \(6\)](#)

Další informace:

- [World: Český: Volný čas: Cestování: Doprava: Autostop a spolujízda \(9\)](#)

Tato kategorie v jiných jazycích:

anglicky (933)	dánsky (29)	English (0)
esperanto (7)	finský (12)	francouzsky (849)
holandsky (164)	italsky (129)	japonsky (35)
Lituvu (13)	norsky (4)	německy (2,242)
polsky (91)	rumunsky (14)	rusky (99)
činsky (15)	čínsky (zj.) (35)	španělsky (43)
švédsky (9)		

- [Asmat - Cestovatelé](#) - Cestopisy, fotografie, zkušenosti a informace. Možnost publikování vlastních zážitků.
- [Austrálie - Singapur](#) - Fotografie a deník z cest po těchto dvou zemích.
- [Cestování net-mag.cz](#) - Na skok po světě, fotografie, rady, rekordy a zajímavosti.
- [Cestovatel](#) - Časopis pro cestovatele a dobrodruhy.
- [Cestování do Iránu](#) - Zápisky z cesty přes Rumunsko, Bulharsko, Turecko, do Iránu a zpět.

Obrázek č. 4: Kategorie „Cestopisy“ v českém jazyce katalogu ODP

Katalogy webových stránek – regionální záběr:

Seznam odkazy

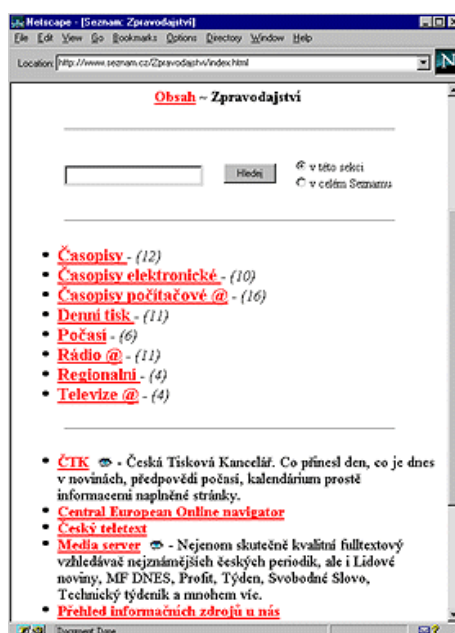


Na regionální úrovni existuje portál Seznam.cz (<http://www.seznam.cz/>) jako jeden z nejznámějších českých portálů, jehož součástí je tematicky uspořádaný katalog webových stránek. Portál Seznam.cz vznikl v roce 1996 a v této době fungoval především jako katalog webových stránek a obsahoval kolem tisíce odkazů na webové stránky [SEZNAM.CZ, 1996a]. V počátcích byla katalogu Seznam.cz vytýkána podobnost se světovým katalogem webových stránek Yahoo! [PETERKA, 1996b]. Vzhledem k velkému rozsahu katalogu bylo možné také vyhledávat v názvech a popisech zařazených webových stránek a názvech kategorií katalogu.

Katalog portálu Seznam.cz v roce 1996 obsahoval 12 hlavních tematických kategorií, později (v roce 2005) bylo změněno uspořádání katalogu na abecední seznam většího množství kategorií a dnes katalog služby *Seznam odkazy* obsahuje 14 hlavních tematických kategorií, které se dělí do dalších podkategorií, nejvíce do 3 úrovní. Jak již z názvu vyplývá, Seznam.cz nazývá jednotlivé záznamy webových stránek ve svém katalogu *odkazy*. *Odkaz* je zde stručnou informací o webové stránce a obsahuje název zdroje, krátký popis a URL adresu. V současnosti je součástí záznamu také náhled webové stránky [SEZNAM.CZ, 1996b]. Výběr webových stránek k zařazení do katalogu provádějí buď administrátoři katalogu, nebo sami uživatelé. Stránka zařazená do katalogu musí být dle pravidel v českém nebo slovenském jazyce a aspoň z části dokončená. Měla by být také nekomerční a informačně hodnotná [SEZNAM.CZ,

1996c]. Do katalogu může uživatel navrhnout webovou stránku prostřednictvím registračního formuláře. Uživatel může sám navrhnout, do jaké kategorie chce odkaz zařadit. Uživatelské návrhy webových stránek jsou před zveřejněním schvalovány administrátory služby [SEZNAM.CZ, 1996d].

Od roku 2005, kdy Seznam.cz spustil vlastní vyhledávač, se firma soustředí především na fulltextové vyhledávání a zájem o katalog začíná klesat [SEZNAM.CZ, 1996e]. V současnosti Seznam.cz katalog webových stránek pod názvem *Seznam odkazy* (na adrese <http://odkazy.seznam.cz/>) stále provozuje, ovšem už se téměř nevěnuje jejímu rozvoji. Na katalog nevede z úvodní stránky portálu žádný odkaz.



Obrázek č. 5: Kategorie „Zpravodajství“ v katalogu Seznam.cz v roce 1996 (dostupné prostřednictvím služby Internet Archive)

Firmy.cz



Další službou předmětového katalogu webových stránek jsou *Firmy.cz*, jsou založeny na komerční bázi. Služba *Firmy.cz* byla založena v roce 2001 [SEZNAM.CZ, 1996f]. *Firmy.cz* je rozsáhlou databází českých firem, která kombinuje fulltextové vyhledávání, katalog předmětových kategorií s výběrem regionu. Firma Seznam.cz se této službě dnes věnuje více než službě *Seznam odkazy*, v katalogu je zařazeno více než 500 tisíc firemních zápisů a katalog obsahuje přes 3600 tematických kategorií [SEZNAM.CZ, 1996g]. V tomto katalogu je možné vyhledávat procházením nebo pomocí vyhledávače.

Firmy.cz nabízí přidání záznamu webové stránky firmy zdarma, ale je možné si objednat další služby mimo reklamy, které jsou zpoplatněné. Je to například přidání loga firmy, přední místo ve výpisu výsledků vyhledávání, umístění do více kategorií katalogu nebo odstranění nejbližších podobných firem [SEZNAM.CZ, 1996h].

2.3 Internetové vyhledávače

2.3.1 Terminologie a technologie internetových vyhledávačů

V české literatuře není názvosloví zcela jednoznačné, internetové vyhledávače jsou nazývány také jako vyhledávací stroje (Sklenák, TDKIV), v anglické literatuře pak jako *search engine*, případně *web search engine*.

Vyhledávač se skládá ze 3 hlavních částí:

- robot
- index
- vlastní vyhledávací stroj

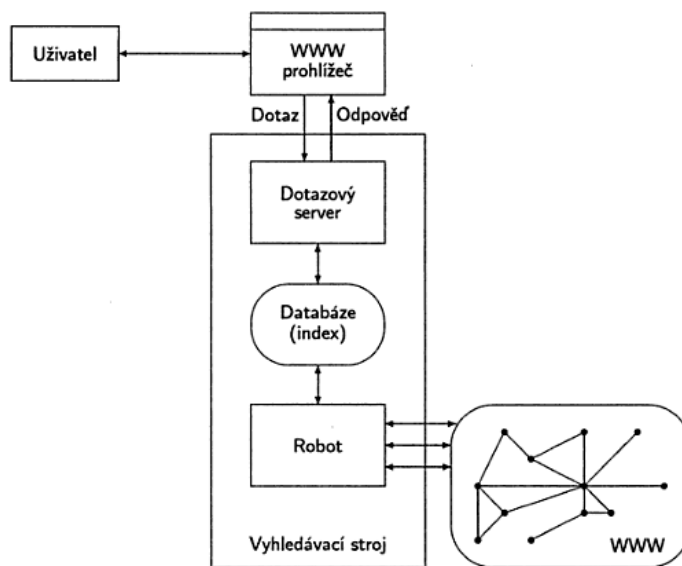
Někdy může být bráno jako další část vyhledávače i uživatelské rozhraní [MAKULOVÁ, 2002, s. 102].

Pro pojem *robot* se v angličtině vyskytuje více výrazů - *bot*, *crawler*, *spider*, *indexer*, *worm*, *wanderer*. První *roboty* byly vytvořeny za účelem zjištění velikosti webu (*The World Wide Web Wanderer*), v současnosti se využívají pro získávání informací o webových stránkách. *Roboty* jsou automatizované počítačové programy, které pravidelně prochází web a aktualizují svůj index tak, aby zjistily změny na webových stránkách a také získaly stránky nové. *Robot* nachází nové webové stránky prostřednictvím sledování hypertextových odkazů, které vychází z jím již indexovaných stránek. Dnes již většina vyhledávačů prohledává plný text dokumentu a některé i netextové prvky webových stránek.

Informace, které *robot* získá, se ukládají do databáze zvané *index*. Účelem *indexu* je organizovat data získaná o webových stránkách tak, aby se daly rychle a efektivně vyhledávat. *Index* vyhledávače obsahuje seznam klíčových slov a k nim přiřazené webové stránky, které dané klíčové slovo obsahují.

Pojmem *vyhledávací stroj* se označuje buď internetový vyhledávač jako celek, nebo pouze jeho součást, která provádí na základě uživatelských dotazů vyhledávání

v indexu. Uživatel komunikuje s vyhledávačem pomocí uživatelského rozhraní svého webového prohlížeče.



Obrázek č. 6: Architektura vyhledávacích strojů [SKLENÁK, 2001, s. 261]

Vyhledávač pracuje tak, že nejprve pomocí *roboty* prochází webové stránky. Roboti nedokáží indexovat celý web, obrovská část webu, kam se roboti nedostanou je nazývána jako hluboký web (*deep web*). Mohou to být stránky, které jsou pro procházení *roboty* zakázány (například pomocí souboru robots.txt) nebo je pro roboty technicky nemožné je indexovat (např. jsou dostupné pouze po zadání hesla). Pomocí *roboty* vyhledávač analyzuje, jakým způsobem má být obsah webové stránky indexován, například jiné indexování pro metadata nebo nadpis webové stránky a jiné pro vlastní text [SKLENÁK, 2001, s. 268].

Vyhledávač ukládá data do *indexu*, což je invertovaný soubor. Invertovaný soubor obsahuje seřazený seznam termů, které se v dokumentech vyskytují a k nim seznam zdrojových dokumentů. Ještě však před tím než jsou data o webové stránce uložena v *indexu*, jsou tato data zpracována. Většina vyhledávačů dnes indexuje plné texty webových stránek a v těchto textech se vyskytují synonyma, různé tvary slov, diakritika, slova nevýznamová a další. Pro vytvoření efektivního vyhledávacího indexu, který bude reprezentovat obsah dokumentu, je třeba tyto jevy odstranit.

Některé vyhledávače využívají tzv. *stemming*, což je automatické převedení slova na jeho základní tvar, kterou vyhledávače využívají při indexování i při zpracování dotazu [SKLENÁK, 2003g]. Vyhledávač pak dokáže vyhledávat různé tvary slova a slova příbuzná (například na dotaz na klíčové slovo „moudrý“ vyhledá slovo

„moudrá“ i slovo „moudrost“). Díky využití *stemmingu* může být index vyhledávače menší a na počet výsledků na zadaný dotaz vyšší. Nevýhodou *stemmingu* je však ta, že vrací současně i mnoho nerelevantních výsledků, například u ustálených slovních spojení. Technologie *stemmingu* je uplatnitelnější pro vyhledávání zejména v morfologicky složitějších jazycích než je anglický [MANNING, 2008, s. 34]. Vyhledávač Google začal částečně používat technologii *stemming* v roce 2003 [LEVENE, 2010, s. 95-96].

Problematickým prvkem jsou tzv. *stopslova* nebo také nevýznamová slova (*stop words*), která se v textu objevují často, ale nesou jen malý význam (např. spojky, předložky, členy). Dříve byla tato slova při indexaci ignorována, ale trendem současných vyhledávačů je při indexaci seznamy stop slov neignorovat, především kvůli možnosti vyhledání specifických frází (např. „to be or not to be“) [MANNING, 2008, s. 27]. Ovšem u běžných vyhledávacích dotazů, kde *stopslovo* není součástí fráze, vyhledávače obvykle *stopslovo* ignorují¹³ [LEVENE, 2010, s. 95].

2.3.2 Příklady konkrétních vyhledávačů

Podle společnosti *Alexa* patří mezi deset nejnavštěvovanějších webových stránek čtyři internetové vyhledávače. Nejnavštěvovanější stránkou je vyhledávač *Google*, internetový portál *Yahoo!* se umístil jako čtvrtý, pátý je čínský vyhledávač *Baidu.com* a sedmou nenavštěvovanější stránkou je vyhledávač *Windows Live* (dnes pod názvem *Bing*) od firmy Microsoft. V České republice je mezi deseti nejnavštěvovanějšími stránkami vyhledávač *Google Česká republika*, *Google* a české portály *Seznam.cz* a *Centrum.cz* [ALEXA INTERNET, 2013].

Společnost *WebCertain*, která se zabývá marketingem v oblasti vyhledávání v různých národních jazycích, vydává každoročně zprávu z průzkumu v oblasti jednotlivých národních trhů internetových vyhledávačů a sociálních sítí. V průzkumu jsou zveřejněna specifika těchto lokálních trhů. Z posledního průzkumu za rok 2012 bylo zjištěno, že společnost *Google* si na většině světových trhů upevňuje svou pozici, ovšem na největším internetovém trhu – v Číně zaznamenal výrazný pokles. V reportu byla zjištěna také skutečnost, že v zemích, kde má vyšší podíl na trhu domácí prohlížeč, je větší pravděpodobnost, že v oblasti sociálních sítí nebude mít největší podíl sociální síť *Facebook*. Ve většině světa má *Google* obrovskou převahu nad jinými vyhledávači. Jiné vyhledávače než je *Google* převažují především v asijských zemích. Jak již bylo

¹³ Při zadání dotazu „be cool“ Google vrátí výsledky, kde se vyskytuje celá fráze, naproti tomu při zadání dotazu „be a math student“ vrátí výsledky s výskytem výrazů „math“ a „student“.

zmíněno, v Číně je nejpopulárnějším internetovým vyhledávačem služba *Baidu* (s podílem na trhu 75-80 %), druhým nejpopulárnějším je také místní vyhledávač s názvem *Qihoo 360*, který je produktem firmy vlastníci stejnojmenný webový prohlížeč. Vyhledávač *Google* je s 6 % podílu na čínském trhu až čtvrtým vyhledávačem po dalších čínském vyhledávači *Sogou*. Další asijskou zemí, ve které vede domácí vyhledávač je Jižní Korea s vyhledávačem *Naver* používaným 70 % domácích uživatelů. V Rusku je před vyhledávačem *Google* nejpopulárnějším internetovým vyhledávačem ruský *Yandex* (60 % trhu) a v Japonsku vede před Googlem vyhledávač *Yahoo! Japan*. Do průzkumu je zahrnuta i Česká republika se svým poměrně specifickým trhem. V České republice do roku 2010 dominoval český vyhledávač v rámci portálu *Seznam.cz*, ale v roce 2010 se podílem na trhu vyhledávač *Google* vyrovnal vyhledávači *Seznam.cz*. V průzkumu za rok 2012 vyhledávač *Google* získal s 55 % větší podíl na českém trhu. *Google* toto vítězství podpořil předcházející reklamou v českých médiích [WEBCERTAIN, 2012].



Obrázek č. 7: Grafické znázornění populárních vyhledávačů a sociálních sítí podle regionu v průzkumu společnosti WebCertain [WebCertain, 2012]

Google



Žádný z internetových vyhledávačů nedosáhl takové popularity jako vyhledávač *Google* (<http://www.google.com/>). Firma Google je v současnosti určitým fenoménem, sloveso „to google“¹⁴ (v češtině: „googlit“, „vygooglovat“ něco) se používá jako synonymum pro vyhledání určité informace na internetu. Tento neologismus již byl zařazen do anglického slovníku a v roce 2002 byl vyhlášen Americkou společností pro nářečí (*American Dialect Society*) nejužitečnějším slovem roku [AMERICAN DIALECT SOCIETY, 2003].

Po vyhledávači *Google* byl také nazván psychologický jev, známý jako *Google Effect*. Experiment zaměřený na adaptaci lidské paměti na nové informační a komunikační technologie provedený na Kolumbijské univerzitě (*Columbia University*) ukázal, že lidé častěji spoléhají na to, že informace, které potřebují, jsou uloženy online. Experiment ověřil, že lidé zapomínají rychleji informace, které jsou snadno dostupné na internetu a ty, o kterých si myslí, že dostupné nejsou, si pamatují lépe a také si lépe pamatují, kde je informace uložena, než informaci samotnou. Internetové vyhledávače či online databáze tak začínáme využívat jako externí zdroj naší paměti [SPARROW, LIU, 2011].

Zakladatelé společnosti Google, doktorandi na Standfordské univerzitě (*Stanford University*) Larry Page a Sergey Brin, začali v polovině 90. let 20. století pracovat na projektu internetového vyhledávače. Cílem tohoto projektu bylo vytvořit prototyp vyhledávače, který bude indexovat a vyhledávat velké množství dokumentů. Důvodem bylo, že v této době uživatelé vyhledávali především pomocí ručně vytvářených předmětových katalogů webových stránek, jako byl například katalog *Yahoo!*, a internetové vyhledávače nebyly při vyhledávání příliš relevantní [BRIN, 1998]. Tento vyhledávač byl původně pojmenován *BackRub*, ovšem jeho název byl v roce 1997 změněn na *Google* [GOOGLE, 2001a]. Jméno vyhledávače je odvozeno od matematického termínu *googol*, který znamená číslo 10^{100} , což má symbolizovat určitou snahu uspořádat obrovské množství informací [SMITH, 2010a, s. 67].

Jeden z důvodů, které přispěly k úspěchu vyhledávače *Google*, bylo vytvoření algoritmu pro řazení výsledků vyhledávání založené na hodnocení stránky nazvané *PageRank*. Tento algoritmus vypočítává relevanci stránky podle relevance stránek, které

¹⁴ Kniha s tímto slovem v názvu *Googled: The End of the World As We Know It*, kterou vydal v roce 2009 spisovatel a novinář Ken Auletta, se zabývá úspěchem a vlivem firmy Google.

na ni odkazují¹⁵. Algoritmus PageRank se stal předmětem zájmu také vydavatelů webových stránek a specializovaných firem zabývajících se oblastí SEO, kteří se snaží tento algoritmus využít, aby své stránky nebo stránky svých klientů posunuli na vyšší místo v řazení výsledků.

Jedním z prvních úspěchů bylo zařazení *Googlu* časopisem PC Magazine do seznamu 100 nejlepších webů pro rok 1998 [GOOGLE, 2001a].

Google začal velmi brzy zpřístupňovat svůj vyhledávač v dalších světových jazycích kromě angličtiny, první jazykové verze pro několik světových jazyků byly spuštěny v roce 2000 a do pár let bylo přidáno další množství jazykových verzí, včetně české. V současnosti je vyhledávání dostupné ve více než 130 jazycích a Google vlastní více než 180 internetových domén [GOOGLE, 2001b].

Základní stránka vyhledávače *Google* je ve velmi jednoduchém designu a obsahuje pouze jedno vyhledávací pole (formulářové rozhraní pro zadání dotazu) a kromě tlačítka pro vyhledávání i tlačítko s názvem "Zkusím štěstí", které umožňuje přechod na první vyhledanou stránku ve výsledcích. Google v současnosti neumožňuje personalizaci vzhledu svého vyhledávače¹⁶, ovšem používá různá grafická a interaktivní loga k příležitosti významných dnů či výročí nazývaná jako *Google Doodles*. Pro přizpůsobení vzhledu a přidání různých aplikací provozuje společnost službu *iGoogle*, která ovšem bude na podzim roku 2013 ukončena [GOOGLE, 2013a].

Google provozuje kromě jednoduchého vyhledávacího rozhraní i rozhraní rozšířeného vyhledávání, kde je možné sestavit vyhledávací dotaz s pomocí funkcí Booleovských operátorů, filtrování výsledků dle jazyka dokumentu, místa vydání, času poslední aktualizace, webu nebo domény, typu souboru, místa výskytu výrazu na stránce (v textu stránky, adrese URL atd.), také dle licence k volnému užívání díla a je možné zapnout filtr bezpečného vyhledávání. Přímo z úvodní stránky je možné vybrat nastavení vyhledávání. V tomto nastavení lze nadefinovat filtr bezpečného vyhledávání, zobrazení dynamického vyhledávání (zobrazuje výsledky vyhledávání již při zadávání vyhledávacího dotazu), počet výsledků zobrazených na stránce a jejich zobrazování v novém okně, historii vyhledávání, preferované jazyky a také je možné zadat informace o geografické poloze.

¹⁵ Více informací a vysvětlení výpočtu PageRanku k dispozici v článku I. Rogerse The Google PageRank Algorithm and How It Works na adrese <http://www.cs.princeton.edu/~chazelle/courses/BIB/pagerank.htm>.

¹⁶ V minulosti bylo možné přidat libovolnou fotografii jako tapetu na stránku vyhledávače, tato funkce však již byla zrušena.

Filtrovat výsledky vyhledávání je možné podle mnohem většího počtu filtrů, než je základně nastaveno v rozšířeném vyhledávání. *Google* podporuje filtry podle typu obsahu dokumentu (pouze obrázky, knihy, blogy, diskuse, recepty atd.) nebo podle dalších kritérií (výsledky v okolí, ještě nenavštívené stránky, filtrování obrázků dle barvy, velikosti atd.) [GOOGLE, 2013b].

Vyhledávač *Google* pracuje s množstvím operátorů. Při vyhledávání je tak možné využít funkcí Booleovských operátorů (OR pro vyhledání alespoň jednoho z výrazů, - pro vyloučení výrazu, operátor AND není třeba zadávat), synonym (~), rozsahu čísel (číslo..číslo), vyhledat termín v přesném znění ("dotaz"), použít zástupný znak (*), omezit vyhledávání podle domény či stránky (site:). Je také možné vyhledat stránky, které odkazují na určitou adresu (link:) nebo související stránky (related:) [GOOGLE, 2013c]. *Google* nabízí množství dalších funkcí vyhledávání (*Google Features*), pomocí kterých lze převádět jednotky, zjistit předpověď počasí nebo výsledek sportovního zápasu a další¹⁷.

Při vyhledávání *Google* standardně používá funkci automatického doplňování. Tato funkce již při zapisování dotazu uživateli nabízí podobné dotazy na základě častých vyhledávání nebo také na základě vlastní vyhledávací historie [GOOGLE, 2013d]. V případě překlepu *Google* nabízí výsledky správného zápisu dotazu, ale je možné zvolit i vyhledávání původně zadaného termínu.

Výsledky vyhledávání se ve vyhledávači *Google* zobrazují v jednom panelu, v druhém panelu na pravé straně se zobrazují reklamní výsledky, případně i tzv. *Diagram znalostí*, což je souhrn základních a souvisejících informací o určitém objektu. Pod vyhledávací lištou jsou filtry vyhledávání, které umožňují výsledky omezit pouze na webové výsledky, obrázky, mapy, nákupy, videa, zprávy a další. Zde je také možnost použití vyhledávacích nástrojů, pomocí nichž lze výsledky filtrovat na základě data, aktuální polohy uživatele, typu výsledků či vybrat pouze výsledky ve vybraném jazyce. Na další výsledky vyhledávání lze přecházet listováním stránek, na rozdíl od vyhledávání obrázků, kde se další výsledky automaticky načítají.

Výsledky vyhledávání jsou řazeny podle relevance na základě hodnoty PageRanku, ale jsou také personalizovány podle uživatele. Do výsledků je tak promítnuta aktuální poloha uživatele (na základě zadání polohy uživatelem nebo na základě adresy IP a dalších faktorů), kdy vyhledávač nabízí výsledky z aktuálního místa

¹⁷ Celý výčet funkcí vyhledávání je k dispozici na adrese <http://www.google.com/intl/cs/help/features.html>.

uživatele, a historie jeho vyhledávání. Personalizaci na základě historie vyhledávání nebo ukládání této historie je možné deaktivovat.

Samotný výsledek vyhledávání obsahuje název stránky, URL adresu a úryvek (fragment) textu ze stránky. U některých více členitých stránek *Google* zobrazuje osnovu této stránky, lze tak rovnou přejít na určitou část stránky (např. do sekce kontakty). S výsledky je možné dále pracovat pomocí rozbalovací nabídky označené šipkou vedle URL adresy stránky. Lze zobrazit poslední indexovanou verzi stránky (Archiv), najít související stránky (Podobné) nebo stránku sdílet na sociální síti *Google+* (Sdílet). Pokud vyhledávač najde více výsledků ze stejné webové stránky, zobrazuje je s odsazením [GOOGLE, 2013e].

Centrum podpory *Google* poskytuje nápovědu k řadě služeb firmy *Google*. Nápověda je poskytována v mnoha světových jazycích včetně českého, ovšem počet služeb, pro které je nápověda dostupná v určitém jazyce se pro různé jazyky liší. *Google* také provozuje stránku s přehledem přerušení a výpadků služeb a stránku s informacemi a tipy, jak bezpečně používat internet, včetně vysvětlení základních pojmů spojených s internetem¹⁸. Nejvíce informací o vyhledávání je k dispozici na stránkách *Google: Vše o vyhledávání* (na adrese <http://www.google.com/insidesearch/>), které obsahují nejen nápovědu a tipy a triky, ale i vysvětlující informace o tom, jak pracuje vyhledávač, časovou osu vývoje s milníky ve vývoji vyhledávání *Google*, funkce vyhledávání a videa příběhů lidí, ve kterých se vyskytuje vyhledávání pomocí *Google* a další zajímavosti [GOOGLE, 2001c]. Odkaz na nápovědu pro vyhledávání se také zobrazuje v dolním panelu výsledků vyhledávání a zde je také možné odeslat zpětnou vazbu pomocí formuláře.

Služby Google

Vyhledávač *Google* je nejznámějším a nejpoužívanějším produktem firmy *Google*, avšak firma vyvíjí obrovské množství webových, softwarových aplikací a služeb a také hardware. Mezi oblastí, ve kterých firma *Google* vyvíjí produkty, patří vyhledávací nástroje, nástroje určené pro komunikaci, reklamu, vývoj aplikací, statistické nástroje či mapy, aplikace pro mobilní telefony a mnoho dalších. Mnoho z těchto služeb je vyvíjeno vlastním týmem nebo byly zakoupeny firmou *Google* už jako hotové produkty. Vzhledem k takové různorodosti jsou tyto služby integrovány v

¹⁸ Stránka výpadků služeb na adrese <http://www.google.com/appsstatus#> a stránka s radami pro bezpečné chování na internetu v českém jazyce na adrese <http://www.google.cz/intl/cs/goodtoknow/>.

mnoha případech pouze společným účtem (*Google Account*) a některé funkce určitých služeb a aplikací se také mohou částečně překrývat.

Firma Google vyvíjí a provozuje řadu služeb určených k vyhledávání. Mimo standardní službu pro vyhledávání na své domovské stránce existuje řada služeb, které vyhledávají konkrétní typy dokumentů, v různých platformách a další různé typy. V obrazových dokumentech lze vyhledávat pomocí služby *Google obrázky*, v audiovizuálních dokumentech pomocí služby *Google videa*. Google provozuje také služby pro vyhledávání v blozích (*Google blogy*), ve zprávách (*Google zprávy*), produktech (*Google Shopping*), vědeckých zdrojích (*Google Scholar*), službu pro vyhledávání a informace z finanční oblasti (*Google Finance*) či službu pro vyhledávání v patentových informacích (*Google Patents*), která dokáže vyhledávat v dokumentech amerického a od roku 2012 i evropského patentového úřadu [ORWANT, 2012]. V současnosti se stále populárnější službou stává vyhledávání pro mobilní telefony (*Google Search App*), která obsahuje i funkci hlasového vyhledávání. K dalším službám zaměřeným na vyhledávání od firmy Google patří například *Google upozornění*, které slouží k zasílání zpráv o nových výsledcích vyhledávání na nadefinovaný dotaz, *Google vlastní vyhledávač*, vhodný zejména k nadefinování vyhledávače například pro určitou webovou stránku.

Blogger

Tento oblíbený blogovací systém nebyl vyvinut firmou Google, ale společností Pyra Labs, kterou Google v roce 2003 koupila [BLOGGER, 2001].

Disk Google

Služba *Disk Google*, původně pod názvem *Dokumenty Google*, je cloudové úložiště¹⁹ elektronických dokumentů. Tato služba nabízí ukládání, sdílení a úpravu dokumentů v různých formátech a poskytuje uživatelům zdarma 15 GB prostoru [GOOGLE, 2001d].

Gmail

Jednou z nejpoužívanějších služeb firmy Google je bezplatná e-mailová služba *Gmail*. S touto e-mailovou službou je propojena i aplikace pro okamžité zasílání zpráv (*instant messaging*) s názvem *Gtalk*, která nabízí i možnost videochatu. S e-mailovou

¹⁹ Cloudové úložiště je služba, která umožňuje uložit data na servery poskytovatele.

schránkou je také možné propojit další aplikace každodenního použití jako je například *Google Kalendář* nebo *Google Kontakty*.

Google knihy

V roce 2004 společnost Google se v rámci svého projektu (původně nazvaného jako *Google Print Library Project*) rozhodla digitalizovat knihy z významných univerzitních a veřejných knihoven s cílem umožnit uživatelům vyhledávat nejen na webu, ale i v textech knih [GOOGLE, 2001e]. Během realizace tohoto projektu se firma potýkala s mnoha autorskoprávními spory a v současnosti zpřístupňuje podle autorských práv buď celý text knihy, omezený náhled nebo pouze krátký útržek textu. V textech knih je možné fulltextově vyhledávat.

Google Chrome

Vlastní webový prohlížeč firmy Google byl vytvořen v roce 2008 [GOOGLE, 2001a] a v současnosti je celosvětově nejpoužívanějším internetovým prohlížečem [W3COUNTER, 2013]. K popularitě tohoto prohlížeče přispěla především jeho rychlost, jednoduché ovládání, pokročilé funkce a také účinná reklama. Od roku 2012 existuje i mobilní verze tohoto prohlížeče pro operační systém Android [GOOGLE, 2001a].

Google+

Po úspěchu sociální sítě *Facebook* se Google rozhodl vytvořit konkurenční sociální síť s názvem *Google+*. Google se do této doby pokusil o vytvoření několika projektů charakteru sociální sítě (*Orkut*, *Google Wave*, *Google Buzz*), ovšem ne příliš úspěšně²⁰. Služba *Google+* byla spuštěna v létě roku 2011 a díky velké propagaci na počátku byl zájem o tuto sociální síť veliký. Přes kvalitní zpracování a snahu o vyšší ochranu soukromí se síti *Google+* nepodařilo získat více uživatelů než má síť *Facebook* a po rychlém nárůstu zájmu uživatelů krátce po jejím vydání tento zájem začal opadávat a podle výzkumů jsou uživatelé na této síti výrazně méně aktivní než na konkurenčním *Facebooku* [DOČEKAL, 2013].

²⁰ Firma Google oznámila v roce 2011 zrušení služby *Google Buzz* [HOROWITZ, 2011].

Google Mapy

V roce 2005 firma Google spustila službu *Google mapy* [GOOGLE, 2001a], která v současnosti poskytuje prohlížení panoramatických snímků z mnoha míst světa, včetně světových památek UNESCO či expozic galerií a muzeí (*Google Street View*). Dalšími souvisejícími mapovými službami firmy Google jsou například nástroje pro zobrazování fotografií vesmíru (*Google Sky*) či Měsíce (*Google Moon*) nebo plánovač cesty veřejnou dopravou (*Google veřejná doprava*).

Překladač

Služba Google překladač využívá pokročilých technologií pro překládání celých textů z velkého množství jazyků. Takovému překladači se říká strojový nebo také automatický překlad. *Google překladač* při práci vyhledává podobné texty mezi množstvím uložených textů a určuje ten nejvhodnější překlad [GOOGLE, 2011a]. Tento překladač také využívá metody strojového učení, a pokud uživatel není s překladem spokojen, může navrhnout lepší verzi překladu. V současnosti *Google překladač* podporuje 71 jazyků [GOOGLE, 2001f]. Překladač je možné nainstalovat do webových prohlížečů, kde umožňuje překládat celé webové stránky. V prohlížeči *Google Chrome* je překladač již integrován.

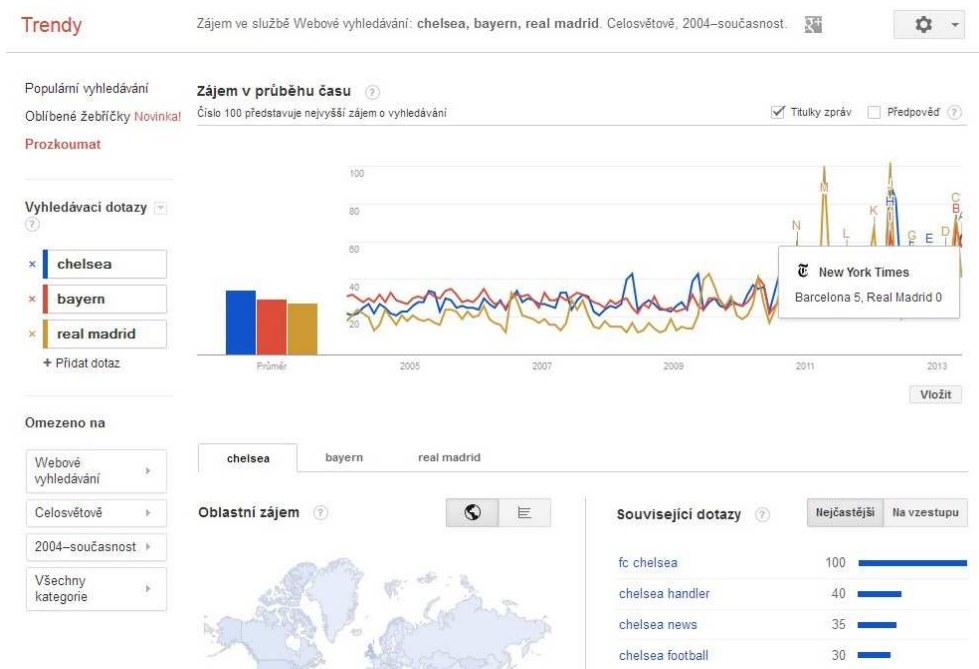
Reklamní služby

Jelikož získává firma Google největší zisk z reklamy, produkuje také množství reklamních aplikací a služeb. Takovou službou je například *AdWords*, která umožňuje firmám zaplatit si reklamu ve výsledcích vyhledávání pomocí *Google*. Podobnou službou pro mobilní telefony je *AdMob*.

Statistické služby

Jako vyhledávač s dominantním postavením na trhu, firma Google také vytváří služby statistického charakteru. Jednou ze statistických služeb vyhledávače *Google* je *Google Zeitgeist*, která ukazuje výběr ze statistik, jaké dotazy uživatelé nejvíce v určitém roce vyhledávali. V přehledu jsou dva typy žebříčků, nejhledanější dotazy (*most searched*) za daný rok celkově a trendy (*trending searches*), specifické nejvyhledávanější termíny určitého roku, které jsou odlišné od termínů předchozího roku. Tyto žebříčky jsou také rozděleny do tematických kategorií, například sport, události, videa a další [GOOGLE, 2001g]. Pokročilejším statistickým nástrojem

vyhledávače *Google* je služba *Google Trendy*, která zobrazuje aktuálně nejčastěji vyhledávané termíny ve vybraných zemích, případně vybrané kategorii nebo dokáže na zadaný termín zobrazit vývoj popularity ve vyhledávání tohoto termínu podle na časové ose, na mapě podle států a časté související dotazy. Dotazy je také možné omezit na základě výběru časového období, země, tematické kategorie a typu vyhledávání (v obrázcích, zprávách, aj.). U některých termínů lze na časové ose zobrazit i předpověď vývoje popularity a při výrazných nárůstech popularity je v některých případech zobrazen důvod²¹ ve formě titulku zpráv. Je také možné zadat více termínů pro porovnání vývoje jejich popularity [GOOGLE, 2013f].



Obrázek č. 8: Ukázka služby Google trendy

Více než sto firem a jejich produktů firma Google koupila. Prvním nákupem v roce 2001 se stala služba diskusních skupin *Usenet*, kterou Google začlenila do své služby *Google Groups*. Mezi další významné firmy, které Google koupila, jsou například tyto společnosti z nejrůznějších oblastí počítačových aplikací [LIST OF MERGERS AND ACQUISITIONS BY GOOGLE, 2013]:

- AdMob – v roce 2009 koupila Google firmu zaměřenou na reklamu v mobilních aplikacích

²¹ Například vysoký nárůst popularity termínu "meteor" v únoru 2013 z důvodu dopadu meteoritu v Rusku.

- Android – do roku 2005, kdy Android koupila firma Google, nepříliš známá firma zaměřená na software pro mobilní telefony, později se mobilní operační systém Android stal nejpoužívanější mobilní platformou na světě
- DoubleClick – firmu zaměřenou na reklamní aplikace koupila firma Google v roce 2007
- ITA Software – firmu zabývající se softwarem v oblasti letecké dopravy koupila firma Google v roce 2010, aby ji začlenila do své služby *Google Flights* a dostala se tak do vedení v oblasti trhu letecké dopravy
- Motorola – firmu Motorola Mobility, která vyráběla mobilní telefony, koupila firma Google v roce 2011 za 12,5 miliard jako svůj nejdražší nákup
- Picasa – nástroj pro správu a prohlížení fotografií byl zakoupen firmou Google v roce 2004
- Youtube – oblíbená služba pro sdílení videa, kterou firma Google koupila v roce 2006

Společnost Google se angažuje v mnoha projektech a spolupracuje s množstvím institucí a organizací. V oblasti kultury například spolupracuje s více než stovkou galerií, muzeí a dalších institucí zabývajících se výtvarným uměním na projektu *Art Project*, který obsahuje virtuální prohlídky galerií a muzeí či digitalizovaná umělecká díla ve vysokém rozlišení [GOOGLE, 2013g]. V rámci služby *Google Books* navázala spolupráci se světovými knihovnami, včetně Národní knihovny ČR, a zahájila digitalizaci historických dokumentů z fondů knihoven. Google se také angažuje v oblasti ekologie, prostřednictvím své dceřiné společnosti *Google Energy* investuje do výstavby zdrojů obnovitelné energie. Pomocí grantů, iniciativ a programů (jako je například (*Google Grants*, *Google.org* aj.) společnost Google podporuje neziskové organizace, vzdělávání, kulturu a další a to i na národní úrovni (například soutěž hudebních talentů *Český Youtube Fest*).

Jako největší a nejpoužívanější internetový vyhledávač, čelí společnost Google mnohým kritickým reakcím. Jednou z nich je obvinění společnosti z monopolu. Z tohoto důvodu byla společnost vyšetřována v minulosti již několikrát, Google prošetřovala z důvodu zneužívání vedoucí pozice na trhu například Evropská komise [STERLING, 2010]. Google se také v minulosti potýkal s porušováním autorských práv, například ve spojení se svou službou *Google Books*.

Mnoho kontroverzních reakcí přineslo působení firmy Google v Číně. Firma Google vstoupila na čínský trh v roce 2006 a z důvodu podmínek tamního politického režimu byly ve vyhledávači *Google* některé výsledky cenzurovány²². Po neúspěšných vyjednáváních společnosti Google s čínskou vládou ohledně zmírnění filtrace jejich obsahu a po hackerských útocích na své služby z Číny, se Google v roce 2010 z čínského trhu stáhla [GOOGLE CHINA, 2001].

Kontroverzní službou od firmy Google je *Street View*, která vyvolala po celém světě otázky týkající se ochrany soukromí. V mnoha zemích, zejména evropských, čelila firma kritice, kvůli níž firma Google nechala v této službě rozmazat tváře zachycených lidí.

Jedním z nejvýznamnějších témat, díky kterým společnost Google čelí kritice, je ochrana osobních údajů a soukromí. Ke kritice přispěla také změna politiky ochrany soukromí firmy Google v roce 2012 [WHITTEN, 2012]. Jako poskytovatel mnoha různorodých služeb, získává společnost Google velké množství uživatelských údajů. Google mnohé z těchto údajů shromažďuje a využívá nejen pro své služby, ale také k cílené reklamě a údaje mohou být poskytnuty i třetím stranám například za účelem dodržování právních předpisů nebo po udělení souhlasu uživatelem. Kritizována je také spolupráce Google s vládami světových zemí, zejména odstraňování informací na žádost vlády či vlastníků autorských práv či vydání osobních údajů uživatelů na žádost vlády²³.

Cílem společnosti Google bylo už od jejího počátku vytvořit vyhledávač, který by dokázal vyhledávat vše, co je na internetu. Úspěchem firmy Google je především to, že dokázala vytvořit takovou komplexní nabídku služeb, že uživatelé pro svou veškerou činnost na internetu mohou používat jejich služby. Společnost Google je také známá tím, že pro ni pracují nejlepší odborníci a také svým komfortním pracovním prostředím. Obrovskou výhodou společnosti Google je její globální zaměření, její základní služby jako vyhledávání jsou dostupné v množství jazyků, novinky jsou postupně lokalizovány do dalších jazyků kromě anglického na rozdíl od jiných významných vyhledávačů jako je například *Bing* zaměřený primárně na Spojené státy čínský *Baidu* či *Yandex* zaměřený na Rusko a okolní státy. To z firmy Google tvoří skutečně globálního poskytovatele služeb a přispívá k jeho výsadnímu postavení na trhu.

²² Prostředek cenzury internetu v Číně bývá také hovorově nazýván jako "velký čínský firewall" (*The great firewall of China*).

²³ Více informací a statistiku počtů žádostí o odstranění obsahu či vydání osobních údajů obsahuje dokument Google Transparency Report na adrese <http://www.google.com/transparencyreport/>.

Bing



Vyhledávač *Bing* (na adrese <http://www.bing.com/>), jehož předchůdci byly vyhledávače *MSN Search*, *Windows Live Search* a *Live Search*, provozuje firma Microsoft.

Svůj první vyhledávač *MSN Search*, který používal výsledky jiných vyhledávačů, především *LookSmart* a *Inktomi*, firma Microsoft spustila v roce 1998 [SULLIVAN, 2003]. Další vyhledávač *Windows Live Search* existoval pod tímto názvem od roku 2006, ale po krátké době byl vyhledávač přejmenován na *Live Search*, aby se odlišil od souboru služeb a software od firmy Microsoft [FOLEY, 2007]. Nový vyhledávač *Bing* byl spuštěn v roce 2009.

Vzhledem k tomu, že vyhledávač *Bing* dnes patří k jednomu z nejvýznamnějších vyhledávačů [DAVIES, 2012], jeho hlavním konkurentem je vyhledávač *Google*. Bing v roce 2012 spustil svou marketingovou kampaň nazvanou *Bing It On*, ve které na speciální webové stránce (na adrese <http://www.bingiton.com>) umožňuje uživateli zadat vyhledávací dotaz, ke kterému zobrazí výsledky vyhledávání z vyhledávačů *Bing* i *Google*, které ovšem nejsou označené jejich logem, takže uživatel neví, který výsledek patří kterému vyhledávači. Uživatel si pak v pěti kolech zvolí výsledky, které mu více vyhovují a tento web mu poté zobrazí, kdy preferoval *Bing* a kdy *Google*. V této kampani *Bing* odkazuje na nezávislou studii, podle které více uživatelů preferovalo výsledky vyhledávače *Bing* nad výsledky od *Google* (v poměru 2:1)²⁴ [MICROSOFT, 2013a]. K této kampani firma Microsoft spustila i televizní reklamu.

Jednou z vlastností, ve které se vyhledávač *Bing* liší od dalších vyhledávačů, je vzhled jeho domovské stránky. Na úvodní stránce je jako tapeta použita fotografie nebo video, nejčastěji s přírodním tématem, která je pro každý den jiná. Na domovské stránce je možné zobrazit základní informace o fotografii a prolistovat několik předchozích tapet. V americkém prostředí lze obrázek stáhnout do svého počítače a tato tapeta vyhledávače obsahuje také body (*hotspot*), které po najetí myši zobrazí další informace vztahující se k tématu obrázku, které slouží také jako odkaz pro další vyhledávání těchto informací v *Bingu*.

²⁴ Více informací na oficiálním blogu vyhledávače *Bing* na adrese http://bingcommunity.search.live.net/blogs/site_blogs/b/thedetails/archive/2012/09/06/bingchallenge.aspx



Obrázek č. 9: Ukázka informací s odkazem pro vyhledávání na tapetě domovské stránky vyhledávače Bing

Nastavení vyhledávače lze personalizovat. V rozhraní pro nastavení je možné vybrat ze tří možností filtrace obsahu pro dospělé od přísného po absenci filtrování, vybrat si jazyk pro zobrazení vyhledávače, změnit počet zobrazených výsledků na stránce či omezit vyhledávání na stránky ve vybraných jazycích či vybrané geografické oblasti.

Vyhledávat je možné primárně pomocí jednoduchého rozhraní s jedním vyhledávacím polem (formulářovým rozhraním pro zadání dotazu), ale lze využít i rozšířené vyhledávání (*advanced search*). Vyhledávač pracuje s Booleovskými operátory OR a NOT, operátor AND není třeba zadávat, jelikož ho Bing používá standardně [MICROSOFT, 2013b]. Pomocí pokročilého vyhledávání lze vyhledávací dotaz upřesnit přidáním dalších klíčových slov nebo omezit na základě domény či stránky, regionu a jazyka. Nově je také umožněno omezit výsledky vyhledávání podle času jejich zveřejnění.

Výsledky vyhledávání se při použití služby *Bing Social* a při vyhledávání určitých entit zobrazují na stránce ve třech sloupcích. První sloupec zobrazuje klasické výsledky (náhledy vyhledaných webových stránek, obrázků, videí atd.), sloupec uprostřed stránky obsahuje stručné a strukturované informace tzv. *Snapshots*²⁵, což je služba podobná *Diagramu znalostí* od Google [ČÍŽEK, 2013], a související vyhledávané termíny. Poslední sloupec pak představuje výsledky ze sociálních sítí (*Bing Social Sidebar*).

Vyhledávač *Bing* nabízí velké množství dalších služeb, pro Českou republiku je však mimo standardního webového vyhledávání dostupné pouze vyhledávání obrázků, videí, zpráv a překladač. Vyhledávání obrázků (*Bing Images*) umožňuje omezit obrázky podle mnoha kritérií jako je převažující barva obrázku, velikost, typ a tvar nebo tvář na

²⁵ Například při vyhledávání určitého státu zobrazí v tomto stručném souhrnu vlajku, hl. město, počet obyvatel aj., při vyhledávání významné osobnosti datum narození, vzdělání, jeho díla a další.

obrázku a vyhledávač také nabízí související témata obrázků. Pomocí služby vyhledávání videí (*Bing videos*) lze vyhledávat nebo prohlížet videa z mnoha různých zdrojů (internetové, ale i televizní kanály jako je *Animal Planet*), ovšem část těchto kanálů není dostupná v České republice. Americká verze vyhledávače *Bing* nabízí mnoho dalších služeb, jako jsou mapy (*Bing Maps*), které podobně jako služba *Google Maps* umožňuje i panoramatický pohled *Streetside*, související služba *Bing Local*, pomocí které lze vyhledávat místa v okolí (např. obchody, restaurace, knihovny) včetně podrobných informací jako je adresa a umístění na mapě, otevírací doba a také uživatelské recenze. Dalšími službami dostupnými v americké verzi vyhledávače *Bing* jsou *Bing Travel*, která umožňuje vyhledávat letenky nebo hotely, vyhledávání zpráv (*Bing News*), slovníkových definic (*Bing Dictionary*), zboží (*Bing Shopping*) a další.

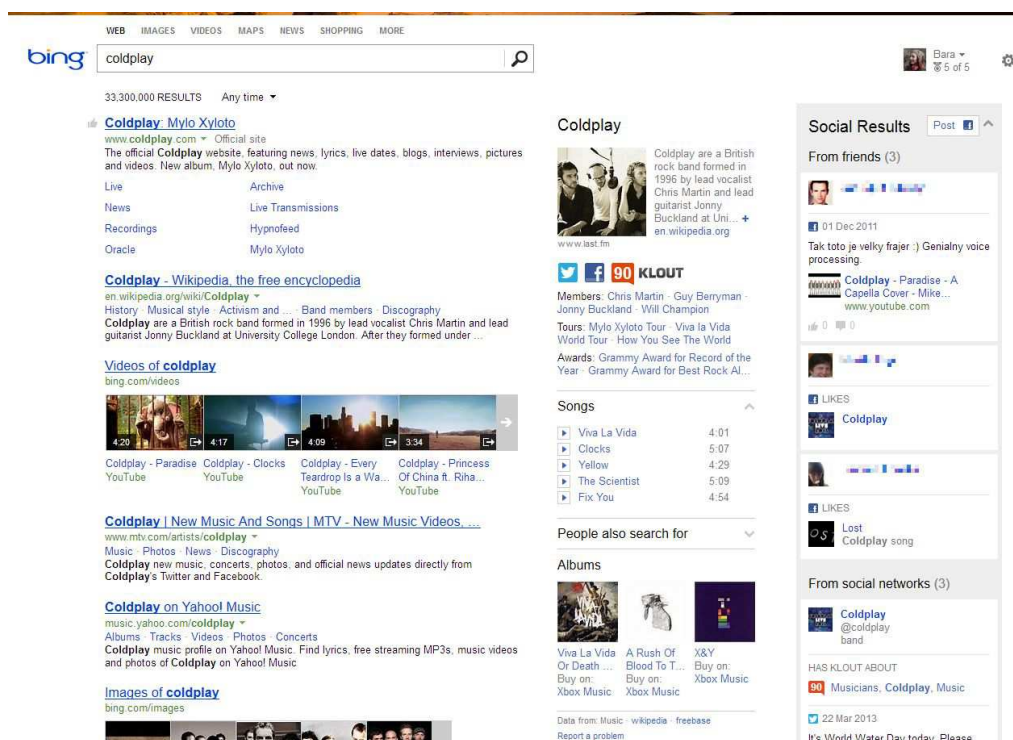
Díky smlouvám se společnostmi *Facebook* a *Twitter* provozuje *Bing* také služby související s obsahem z těchto sociálních sítí, ke kterým patří například služba vyhledávání informací ze sociálních sítí s názvem *Social search* a také služba pro prohlížení fotografií ze sociální sítě *Facebook Friend's Photos* [CARNEY, 2013].

Vyhledávač *Bing* se také vyznačuje poměrně rozsáhlým systémem nápovědy, ve které lze vyhledávat pomocí klíčových slov nebo procházet jednotlivá témata. Podobnou službou je i *Bing Explore*, která uživatelům představuje užitečné tipy a možnosti vyhledávání pomocí vyhledávače *Bing* jako například postup, jak si pomocí vyhledávače najít populární videa, naplánovat a zajistit výlet nebo program na páteční večer [MICROSOFT, 2013c]. Pro podrobnější dotazy nebo nahlášení problému je možné kontaktovat pomocí webového formuláře technickou podporu vyhledávače.

V americkém prostředí vyhledávač *Bing* nabízí svým uživatelům službu nazvanou *Bing Rewards*, díky které mohou uživatelé sbírat body (credits), které jsou směnitelné za ceny. Body jsou získávány vyhledáváním a prací se službami vyhledávače *Bing* a cenami jsou například kupony na elektronické služby.

Firma *Microsoft* jako jedna z nejvýznamnějších společností zabývajících se informačními technologiemi reaguje na nejnovější trendy a její internetový vyhledávač *Bing* je aktuálně alespoň v Americe největším konkurentem vyhledávače *Google*. Silnými stránkami tohoto vyhledávače je jeho vizuální prezentace (výsledků i domovské stránky), vyhledávání videí a obrázků, podpora uživatelské komunity sbíráním bodů s výhrami nebo uživatelskými tipy a triky pro vyhledávání a především v současnosti jeho pokročilé vyhledávání informací ze sociálních sítí.

Nevýhoda vyhledávače Bing spočívá především v jeho silném zaměření na americký, případně britský trh. Narozdíl od vyhledávače *Google* se *Bing* nesnaží všechny své služby zpřístupňovat i pro další země a jeho nabídka tak zůstává například v České republice výrazně chudší.



Obrázek č. 10: Ukázka výsledků vyhledávání pomocí Bing včetně služby Bing Social

Yahoo!



Webový portálů *Yahoo!* (na adrese <http://www.yahoo.com/>, vyhledávač na vlastní adrese <http://search.yahoo.com/>) je jednou z nejnavštěvovanějších webových stránek, zejména ve Spojených státech [ALEXA INTERNET, 2013]. Tento portál vznikl původně jako předmětový katalog webových stránek, v současnosti portál *Yahoo!* poskytuje mnoho služeb včetně vlastního vyhledávače s názvem *Yahoo! Search*.

Zpočátku využíval vyhledávač *Yahoo!* pro svou činnost výsledků jiných vyhledávačů, nejprve *Inktomi*, který později firma *Yahoo!* koupila, poté využíval služeb vyhledávače *Google*, přestože firma vlastnila další vyhledávače jako je například *AltaVista*, vlastní vyhledávání spustila firma *Yahoo!* v roce 2003 [YAHOO! SEARCH, 2001].

V roce 2009 uzavřela firma *Microsoft* se společností *Yahoo!* dohodu, že vyhledávání portálu *Yahoo!* bude realizovat vyhledávač od firmy *Microsoft Bing* a

Yahoo! bude dodávat pouze vlastní reklamní odkazy. Firma *Yahoo!* se chce zaměřovat především na své populární mediální služby jako je například *Yahoo! Sports* [JOHNSON, 2009]. Od roku 2010 tak vyhledávání na portálu *Yahoo!* pracuje na technologii vyhledávače *Bing* [ALLEN, 2010].

Portál *Yahoo!* provozuje velké množství služeb, k nejpobulárnějším patří například především služby mediálního charakteru jako jsou všeobecné zpravodajské informace (*Yahoo! News*), finanční informace (*Yahoo! Finance*), zprávy z oblasti sportu (*Yahoo! Sports*), bulváru (*omg!*) a další. *Yahoo!* také uživatelům umožňuje prostřednictvím služby *My Yahoo!* vytvořit vlastní personalizovaný vzhled a obsah portálu, je možné nastavit barevné schéma i strukturu stránky a přidat aplikace a služby dle svého výběru.

Vyhledávat pomocí *Yahoo!* je možné nejen ve webech, obrázcích či videích, ale také v blozích, ve zboží, místa pomocí lokálního vyhledávání a v dalších službách portálu *Yahoo!* jako jsou například recepty, zprávy z oblasti sportu, filmu nebo financí. Vyhledávač *Yahoo!* má základní i pokročilé vyhledávací rozhraní. Pomocí pokročilého vyhledávacího rozhraní je možné zadávat různé kombinace klíčových slov (vyhledávač pracuje s Booleovskými operátory i s uvozovkami pro vyhledávání přesné fráze), vyhledávat v určité stránce nebo doméně, vybrat formát vyhledávaného zdroje, aplikovat filtr bezpečného vyhledávání, omezit výsledky na základě země nebo jazyka a vybrat počet zobrazovaných výsledků.

Výsledky vyhledávání jsou zobrazeny v jednom sloupci, v levém panelu jsou pak související vyhledávací dotazy a možnost omezení výsledků podle času vydání. Zajímavostí je možnost shlukování výsledků do kategorií, ve vyhledávání v blozích je možné filtrovat výsledky podle jejich tématu nebo zdroje. Ve vyhledávání receptů je možné zaškrtnout použité ingredience a filtrovat například podle typu jídla nebo uživatelských hodnocení.

Portál *Yahoo!* získal svou popularitu již v druhé polovině 90. let, nejprve jako předmětově řazený katalog webových stránek, poté začal přidávat další služby a dnes jako velký portál s vyhledávačem a množstvím dalších služeb. Donedávna patřil vyhledávač *Yahoo!* vedle vyhledávačů *Google* a *Bing* k nejpobulárnějším vyhledávačům, ovšem v současnosti se však společnost *Yahoo!* rozhodla věnovat především svým úspěšným službám a vyhledávání na webu je realizováno prostřednictvím technologie vyhledávače *Bing*.

Poměrně nový internetový vyhledávač s názvem *DuckDuckGo* (na adrese <https://duckduckgo.com/>) je specifický především svým zaměřením na soukromí uživatelů a na relevantní výsledky a také na spolupráci se svými uživateli.

Vyhledávač *DuckDuckGo* vytvořil v roce 2008 Gabriel Weinberg [DUCKDUCKGO, 2013a]. *DuckDuckGo* je pouze vyhledávačem, neposkytuje další služby jako například e-mailovou schránku. Vyhledávač pracuje tak, že integruje výsledky z více než 50 různých vyhledávacích nástrojů a zdrojů včetně vlastního vyhledávacího robota nazvaného *DuckDuckBot*, ale i známých vyhledávačů jako je *Yahoo!*, *Bing*, *Yandex* či *WolframAlpha* a dalších zdrojů jako například *Wikipedia*. Rozdílem mezi výsledky těchto vyhledávačů a *DuckDuckGo* je specifická vrstva, která zpracovává výsledky tak, že změní relevanci dle vlastních kritérií, odstraní spam, vytvoří sumarizaci informací [DUCKDUCKGO, 2013b].

Podobně jako má vyhledávač *Google* grafiky s logem *Google* nazývané jako *Google Doodles* pro významné dny, *DuckDuckGo* vytváří také speciální loga pro příležitost některých svátků či významných dnů. Různé parametry prohlížeče jako je vzhled, nastavení soukromí nebo jazyk, je možné přizpůsobit. Vyhledávač také nabízí tzv. *vychytávky* (*goodies*) z různých tematických oblastí, obsahují zajímavé možnosti vyhledávače jako například výpočty rovnic, konverze měn, vygenerování náhodného hesla, nabídku alternativ k určitému software atd.

DuckDuckGo používá poměrně pokročilý systém syntaxe pro zapisování vyhledávacích dotazů. Nejvýraznější vlastností zápisu dotazu v *DuckDuckGo* je ta, že dokáže vyhledávat informace pouze v určitých stránkách pomocí znaku „!“ spojeného s termínem. Tuto vlastnost nazývá jako *!Bang Tags*. Pro vyhledávání klíčového slova například pouze na sociální síti *Facebook* stačí zapsat klíčové slovo a *!facebook*. Pro některé stránky je zavedena i zkratka, například pro vyhledávání ve Wikipedii stačí zapsat *!w*, pro internetový ochod *Amazon* například *!a*. Je možné hledat i typ obsahu, například *!image* pro obrázky nebo *!map* pro vyhledávání v mapách. Uživatelé také mohou pomocí formuláře na stránkách vyhledávače navrhnout vlastní *!Bang Tag* či úpravu stávajícího [DUCKDUCKGO, 2008a].

K dalším prvkům syntaxe vyhledávače *DuckDuckGo* patří například vlastnost nazvaná jako *I'm Feeling Ducky*, tzn. přejítí na první odkaz výsledků vyhledávání (klávesy „\“, „!“ , případně *!ducky*), využití Booleovských operátorů *AND*, *OR*, -

(pracuje jako operátor NOT), závorek a uvozovek (pro přesný termín). Při zápisu dotazu je také možné omezit vyhledávání na určitou doménu pomocí zápisu *site:* nebo na určitý stát *region:* či zkráceně *r:*. Pomocí zápisu *sort:date* je možné výsledky seřadit podle data vydání, pomocí *filetype:* či zkráceně *f:* lze omezit vyhledávání jen na určitý formát dokumentů [DUCKDUCKGO, 2013c].

V zobrazení výsledků se na prvním místě u některých dotazů objeví tzv. okamžité odpovědi (*instant answers*), což je tabulka se stručnou informací či definicí vyhledávaného termínu, případně výpočet příkladu, skupiny, do kterých termín patří a další. V případě, že vyhledávaný termín obsahuje více významů, nabídne vyhledávač různé významy tohoto slova k výběru pro další vyhledávání. Pokud uživatel vyhledává osobu, korporaci či jinou entitu, která vydává oficiální webové stránky, tak se odkaz na tyto stránky zobrazí jako první (pod rychlými odpověďmi) s označením, že se jedná o oficiální stránku. Další výsledky se na stránce načítají tím, jak uživatel postupuje stránkou dolů, není tak třeba klikat na další stránky. Pro přehlednost výsledků také slouží ikony loga webové stránky v náhledu výsledku, díky nimž je snadné rozlišit známé zdroje.

Nejdůležitějším rysem, kterým se *DuckDuckGo* vyznačuje je jeho ochrana soukromí. Jiné vyhledávače jako například *Google* ukládají historii navštívených stránek a vyhledávaných klíčových slov spojené s konkrétním uživatelem a na základě těchto informací pak každému uživateli zobrazují jiné – personalizované výsledky. Potenciálním rizikem takového shromažďování informací je jejich zveřejnění nebo zneužití další stranou. Informace o navštívených stránkách také mohou sloužit k lépe cílené reklamě.

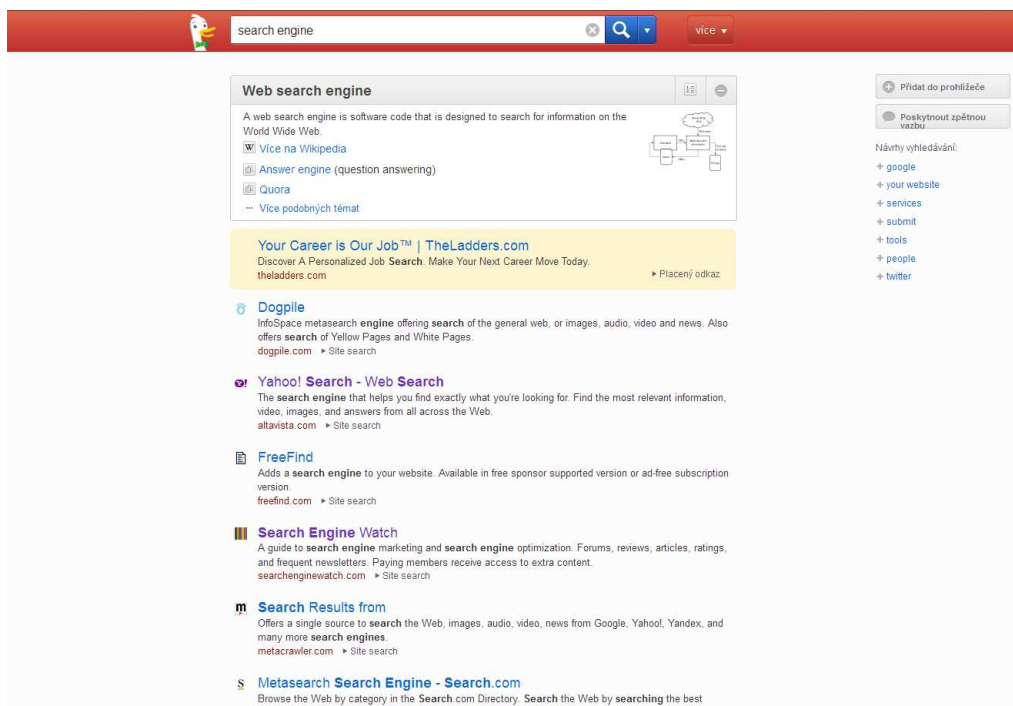
Při vyhledávání ovšem dochází i jevu nazývanému jako *information leakage*, který se dá přeložit jako únik informací a znamená to, že pokud uživatel vyhledá informaci a přejde na požadovanou webovou stránku, vyhledávač odešle této stránce informace o uživateli, například z jakého zdroje nebo z jaké IP adresy uživatel na stránku přišel. Tomu je možné zabránit použitím proxy serveru nebo zabezpečeného protokolu HTTPS.

Z těchto důvodů vyhledávač *DuckDuckGo* neukládá osobní informace o uživateli a běžně nepoužívá cookies, standardně využívá zabezpečení před únikem informací, provozuje zabezpečenou verzi vyhledávání pomocí protokolu HTTPS a také neobsahuje spam [DUCKDUCKGO, 2008b].

Vyhledávač *DuckDuckGo* nemůže v současnosti firmě Google konkurovat velikostí svého indexu ani rozsáhlou nabídkou služeb jako je vlastní prohlížeč, sociální síť, mapy atd. Nevýhodou tohoto vyhledávače je také menší rozsah lokalizace do různých jazyků, ovšem na této vlastnosti je průběžně pracováno a dobrovolníci mohou pomoci s překladem vyhledávače do svého mateřského jazyka. S tím souvisí i menší relevance výsledků při vyhledávání v jiných jazycích než je angličtina.

K největší výhodě vyhledávače *DuckDuckGo* patří především jeho politika týkající se soukromí vyhledávání. Další výhodou je také rychlé a efektivní vyhledávání pomocí možnosti okamžitého získání výsledků z vyhledávání na konkrétní stránce (*!Bang Tags*). K efektivitě přispívá také jednoduchý design webu a přehlednost zobrazení výsledků nebo také možnost získání okamžité odpovědi. Zajímavá je i možnost přidání zpětné vazby k výsledkům vyhledávání. Vyhledávač je zaměřen na uživatele, na platformě *DuckDuckHack* mohou uživatelé navrhovat a vyvíjet zásuvné moduly (*pluginy*) pro vyhledávač, uživatelé mohou také navrhovat svoje nápady a diskutovat s ostatními na diskusním fóru.

Vyhledávač *DuckDuckGo* je vhodný pro uživatele, pro které je důležitá otázka zabezpečení soukromí na internetu a také pro pokročilejší uživatele, kteří chtějí získat objektivně relevantní výsledky dotazu a vyhledáváním se nechtějí zdržovat, jelikož mohou používat klávesové zkratky. Vyhledávač *DuckDuckGo* využijí také uživatelé se znalostmi z oblasti informačních technologií, kteří se mohou podílet na tvorbě a úpravách vyhledávače. Uživatelé bez znalosti anglického jazyka budou zřejmě preferovat jiné vyhledávače, které jsou více přizpůsobeny jejich mateřskému jazyku.



Obrázek č. 11: Ukázka výsledků vyhledávání ve vyhledávači DuckDuckGo

WolframAlpha



Internetový vyhledávač *WolframAlpha* je nazývaný někdy také jako *answer engine* (nástroj na odpovídání dotazů) a na vlastní webové stránce (na adrese <http://www.wolframalpha.com/>) je představen jako *computational search engine* (výpočetní vyhledávač). Tento vyhledávač byl spuštěn v roce 2009 firmou Wolfram Research. Název vyhledávače je odvozen ze jména jeho zakladatele, britského matematika a fyzika Stephena Wolframa [WOLFRAM ALPHA, 2001].

WolframAlpha pracuje na jiném principu než klasické fulltextové internetové vyhledávače jako je například *Google* či *Yahoo!*, *WolframAlpha* uživatelské dotazy vypočítává. Jeho technologie je tvořená čtyřmi úzce propojenými celky – analýzou přirozeného jazyka, vybranými a průběžně spravovanými daty, množstvím algoritmů a rovnic a vizuální prezentací výsledků ve formě tabulek, grafů aj.

WolframAlpha dokáže zpracovávat dotazy z mnoha oborů a díky obrovskému množství algoritmů umí data vypočítat, porovnat, vytvořit graf a mnoho dalšího. Mezi některé příklady jeho funkcí patří například převody jednotek, výpočet kalorií v potravinách, vygenerování nákresu aktuální podoby hvězdné oblohy na vybraném místě, přehled četnosti písmen v určitém jazyce, zařazení živočicha či rostliny do taxonomie, vytvoření grafu ceny vybrané komodity v určité oblasti, porovnání nákladů na bydlení ve vybraných městech a mnoho dalších [WOLFRAM ALPHA, 2013a].

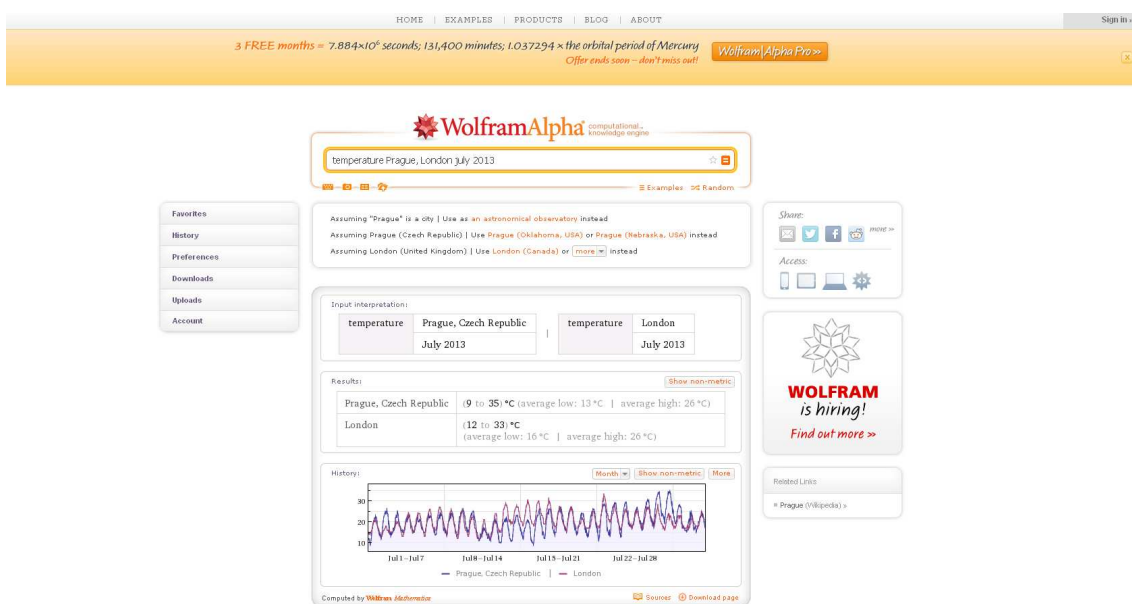
Uživatelské rozhraní vyhledávače *WolframAlpha* je jednoduché s jedním vyhledávacím polem (formulářovým rozhraním pro zadání dotazu), do kterého je možné zadávat klíčová slova, ale i celé fráze či věty v přirozeném jazyce (aktuálně lze pouze v angličtině). V placené verzi *WolframAlpha PRO* je také možné vyhledávat pomocí obrázku, soubor lze nahrát přímo z počítače nebo zadat URL adresu obrázku, případně pomocí přímého vložení dat či nahrání souboru v různém typu formátu (textové, audio, 3D modely, XML, formáty pro vědecké, medicínské, chemické informace a další).

Dotazy při vyhledávání ve *WolframAlpha* je možné zapsat jak pomocí klíčových slov, tak i celou větou. Vyhledávač pak větu rozdělí na slova či fráze, kterým rozumí a vyhledá vztahy mezi nimi. Ty pak při vypočítávání dotazu spojí.

Při příliš obecném či nejednoznačném dotazu nabídne vyhledávač několik významů pro upřesnění, při výpočtech, které potřebují více číselných informací, vyhledávač nabídne tabulku pro zadání doplňujících čísel. Po výpočtu dotazu je možné pomocí tlačítka *source* na konci stránky zobrazit zdroje takto vypočítaných dat. Mnoho informací získaných jako výsledek dotazu slouží i jako odkaz k získání podrobnějších informací o dané problematice. Vyhledávač nabízí v rámci stránky s výsledky vyhledávání i související dotazy.

WolframAlpha nabízí uživatelům možnost vytvořit si vlastní účet. Po zaregistrování je možné ukládat a spravovat oblíbené dotazy, mít přístup k historii svého vyhledávání, uložit si nastavení možností vyhledávače, nebo v rámci placené služby *WolframAlpha PRO* si stáhnout výsledky vyhledávání. Novinkou *WolframAlpha* je analýza účtu na sociální síti *Facebook*. Tato služba byla spuštěna na podzim roku 2012. Pro využití této služby je nutné si vytvořit účet na webu *WolframAlpha* a povolit tak přístup vyhledávače k datům z *Facebooku* a je dostupná po zadání klíčového slova *facebook analysis* do vyhledávacího pole. Analýza zahrnuje mimo základní statistické informace o uživateli analýzu jeho činnosti na *Facebooku* (počet nahraných fotografií, poměr mezi typy příspěvků, které uživatel přidal, průměrná denní aktivita podle dnů v týdnu a hodin atd.), interakce s přáteli (příspěvek s nejvíce komentáři, neoblíbenější fotografie atd.), analýzu přátel (poměr žen a mužů, nejstarší a nejmladší přátelé, statistika přátel dle místa bydliště atd.) nebo grafické znázornění sítě přátel a vztahů mezi nimi. Část těchto dat zároveň také funguje jako odkaz pro další vyhledávání, takže kliknutí například na datum narození některého z přátel přesměruje na informace o tomto datu, například jaké osobnosti se tento den narodily nebo v jaké fázi se zrovna nacházel měsíc.

Vyhledávač *WolframAlpha* není možné zcela porovnávat s jinými internetovými vyhledávači z důvodu jeho jiné technologie zpracování vyhledávání. *WolframAlpha* je vhodný pro rychlé vyhledávání strukturovaných odpovědí a dat, odpadá tak vlastní vyhledávání odpovědí z výsledků, které je někdy pracné a časově náročné. Zajímavá je také schopnost sémantické analýzy dotazů položených v přirozené řeči, ovšem aktuálně je k dispozici pouze anglická jazyková verze. Vyhledávač je tematicky zaměřený především na oblast přírodních věd a techniky, některá témata z oblasti humanitních nebo sociálních věd jako je například literatura nebo umění není zatím příliš zpracované. *WolframAlpha* není vhodný pro vyhledávání rozsáhlejších informací a souvislostí jako například internetová encyklopedie *Wikipedia*.



Obrázek č. 12: Ukázka výsledku výpočtu ve službě WolframAlpha

Regionální internetové vyhledávače

Yandex



Nejpopulárnější ruský internetový vyhledávač *Yandex* (na adrese <http://www.yandex.ru/>) s podílem 60 % veškerého vyhledávání v Rusku a s přibližně 50 miliony unikátních návštěvníků²⁶ [YANDEX, 1997a] je podle společnosti Comscore aktuálně čtvrtým nejpoužívanějším internetovým vyhledávačem [BONFILS, 2013]. *Yandex* má kromě ruské i další verze pro Ukrajinu, Bělorusko, Kazachstán a Turecko a na adrese <http://www.yandex.com/> zpřístupňuje i verzi v anglickém jazyce.

²⁶ Data za leden 2013.

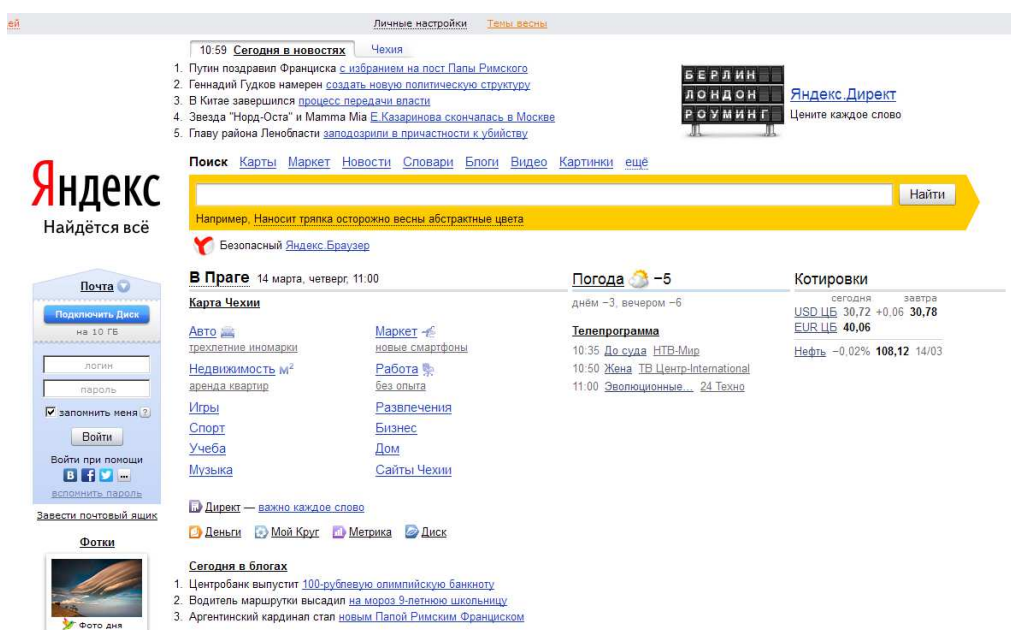
Yandex je portálem, který zahrnuje fulltextový vyhledávač, předmětový katalog webových stránek a další služby jako je e-mailová schránka, překladač, aktuální kurzy měn nebo předpověď počasí. Předmětový katalog portálu *Yandex* je rozdělen do 16 kategorií. V katalogu je možné také vyhledávat a lze to i pouze v jedné kategorii. Výsledky z katalogu je možné omezit podle regionu. Vyhledávač automaticky z IP adresy zjistí lokaci uživatele a na jejím základě mu nabízí výsledky a služby. Pomocí vyhledávače *Yandex* lze vyhledávat v mnoha typech informací jako v obrázcích, mapách, videích, zprávách, blozích, encyklopedii, ve službě *Market*, která slouží k vyhledávání a získání informací o zboží a dalších. Firma Yandex v roce 2009 vytvořila metodu strojového učení nazvanou *MatrixNet*, kterou uplatňuje při řazení výsledků (*ranking*). Algoritmus řazení výsledků *Yandexu* počítá s množstvím faktorů jako je oblíbenost webu, informační hodnota, region, citovanost a další [YANDEX, 1997b].

Dotaz je možné formulovat pomocí klíčových slov, ale i volně tvořenou větou, vyhledávač pracuje i s logickými operátory, ale má pro ně vlastní symboly (znak + pro vyhledání všech takto spojených slov, | pro funkci disjunkce a - pracuje jako operátor NOT). Pro vyhledání přesného tvaru slova či fráze je možné použít uvozovky „“, spojení klíčových slov znakem & vyhledá dokumenty, kde se tato slova nacházejí v jedné větě. Výsledky je také možné omezit dle jazyku dokumentu, umístění na určité doméně či typu dokumentu [YANDEX, 1997c]. Pro složitější dotazy nabízí *Yandex* rozhraní pokročilého vyhledávání, kde je možné vybrat typ či jazyk dokumentu z nabídky, omezit výběr časově nebo zvolit počet výsledků, který se bude zobrazovat [YANDEX, 1997d]. Výsledky jsou podobně jako u vyhledávače *DuckDuckGo* zobrazovány s ikonou loga vyhledané webové stránky. U výsledků v jiném než ruském jazyce vyhledávač umožňuje překlad stránky do vybraného jazyka pomocí svého překladače. *Yandex* také umožňuje uživatelům vytvořit vlastní účet nebo se přihlásit pomocí již vytvořeného účtu ze sociálních sítí do služby *My Findings*, pomocí které lze ukládat a spravovat historii i výsledky hledání [YANDEX 1997e]. *Yandex* návrhy klíčových slov i výsledky vyhledávání standardně personalizuje na základě historie vyhledávání, tuto vlastnost je však možné v nastavení vyhledávání zakázat [YANDEX 1997f].

Yandex je moderní vyhledávač a firma Yandex následuje aktuální trendy v oblasti internetového trhu, například vydala vlastní internetový prohlížeč *Yandex.Browser*, vlastní službu pro stahování a nákup mobilních aplikací *Yandex.Store*

nebo cloudové úložiště dokumentů s názvem *Yandex.Disk* podobně službě *Google Drive*. Yandex také nově umožňuje vyhledat i data z ruských, ale i globálních sociálních sítí jako je *Twitter* a také začal indexovat obrovské množství zahraničního webu [VYLEŤAL, 2012].

Tyto novinky, sofistikované vyhledávání, zpracování ruského jazyka a nabídka velkého množství služeb tak z *Yandexu* tvoří dominantní vyhledávač na ruském²⁷ a okolním trhu, který má potenciál expandovat a začít vytvářet konkurenci velkým globálním vyhledávačům.



Obrázek č. 13: Ukázka domovské stránky portálu Yandex

České internetové vyhledávače

K významným českým internetovým vyhledávačům patřily kromě vyhledávače *Seznam.cz*, který je populární dodnes, vyhledávače portálů *Atlas.cz*, *Centrum.cz* a *Jyxo*. Portály *Atlas.cz* a *Centrum.cz* provozovaly vlastní fulltextové vyhledávání. S rozvojem vyhledávače *Seznam.cz* a stoupající oblibou světového vyhledávače *Google* nebylo pro oba portály ekonomické vyvíjet vlastní vyhledávače a tak portál *Atlas.cz* začal pro vyhledávání na svých stránkách využívat v roce 2007 služeb firmy *Seznam.cz* a vyhledávání na portálu *Centrum.cz* je dnes realizováno prostřednictvím *Google* [SEZNAM.CZ, 1996i]. Dnes oba bývalé konkurenty vlastní jedna firma, ale ani tak nedokáže konkurovat návštěvnosti portálu *Seznam.cz* [DOČEKAL, 2008]. Vyhledávač

²⁷ Ruský internetový prostor bývá také označován jako *Runet* [RUNET, 2001].

Jyxo byl zajímavým projektem, zejména díky svému zpracování českého jazyka, k jeho vlastnostem patřilo ohýbání slov, navrhování opravy překlepů nebo chybně zadaných klíčových slov nebo rozčlenění výsledků do skupin [JYXO, 2013]. Vyhledávač *Jyxo* již ukončil svou činnost.

Seznam.cz



Česká firma Seznam.cz začala původně s předmětovým katalogem webových stránek a v současnosti provozuje stejnojmenný portál s fulltextovým vyhledávačem (na adrese <http://www.seznam.cz/>, vyhledávač je dostupný i na vlastní poddoméně <http://search.seznam.cz/>). Portál *Seznam.cz* je jedním z několika světových vyhledávačů, které mohou na regionální úrovni konkurovat vyhledávači *Google*. V současnosti je poměr firem *Google* a *Seznam.cz* na českém trhu poměrně vyrovnaný, avšak po dlouhou dobu zde dominoval portál *Seznam.cz*.

Vlastní fulltextový vyhledávač firma Seznam.cz spustila v roce 2005, do té doby využívala vyhledávačů jiných firem, např. *Jyxo* či *Google* [SEZNAM.CZ, 1996e]. Tento vyhledávač však vyhledával pouze v českých webových stránkách, pro vyhledávání v zahraničním webu využíval služeb firmy *Google*. V roce 2009 změnila firma Seznam.cz poskytovatele vyhledávání v zahraničních webových stránkách a zvolila tak nově vyhledávač *Bing* (původně *Live Search*) od firmy *Microsoft* [SEZNAM.CZ, 1996j]. Nově, od roku 2012, začal Seznam.cz indexovat i některé cizojazyčné stránky (především v anglickém, německém či slovenském jazyce) a jeho index tak vzrostl z 400 na 700 milionů dokumentů [SEZNAM.CZ, 2012], zahraniční výsledky jsou však standardně dodávány především službou vyhledávače *Bing*.

Popularita portálu *Seznam.cz* spočívá také v tom, že poskytuje množství dalších služeb. *Seznam.cz* v současnosti provozuje kolem třiceti služeb, k nejoblíbenějším patří e-mailová schránka, komunitní server *Lidé.cz*, vyhledávání v mapách *Mapy.cz*, které nabízí mimo jiné i turistické mapy, vyznačené cyklotrasy či historické mapy, zpravodajský server *Novinky.cz*, katalog komerčních společností *Firmy.cz* a další. Jednou z nejnovějších služeb portálu *Seznam.cz* je služba k přehrávání videoklipů *Mixér*.

Seznam.cz nabízí uživateli jednoduché nebo pokročilé vyhledávací rozhraní. V pokročilem vyhledávání je možné zvolit vyhledávání přesných frází nebo vyjmout konkrétní slova z vyhledávání, hledat přednostně v titulcích, adresách či v textech, omezit hledání na určitou doménu či doménu vyjmout z vyhledávání a omezit typ

souborů. *Seznam.cz* nevyhledává pouze webové stránky, ale umožňuje využít i mnoho dalších funkcí. Je možné zvolit vyhledávání firem a společností v katalogu *Firmy.cz*, v mapách pomocí služby *Mapy.cz*, využít slovník, vyhledávat v nabídkách zboží pomocí firmy *Zboží.cz*, vyhledat pouze encyklopedické informace. *Seznam.cz* nabízí také vyhledávání obrázků, kde lze obrázky filtrovat dle velikosti, orientace, typu obrázku či jeho barvy a také lze zapnout filtrování obrázků nevhodných pro děti. Pro vyhledávání obrázků *Seznam.cz* využívá databáze vyhledávače obrázků *Picsearch*. K vyhledávání obrázků *Seznam.cz* využívá vyhledávací technologie ruského vyhledávače *Yandex* a videa je možné řadit od nejnovějších nebo filtrovat pouze videa delší než 15 minut a opět lze zapnout bezpečné filtrování.

Vyhledávač *Seznam.cz* nabízí uživatelům také podpůrné funkce při hledání - opravy překlepů a navigační tip sloužící k upozornění uživatele, pokud zadal do vyhledávání celou URL adresu, zda chce navštívit přímo tento zadaný web.

Při zadávání dotazu je aktivována funkce automatické nabídky navrhovaných výrazů, tzv. našeptávač. Vyhledávač standardně pracuje s Booleovskými operátory AND a OR, které doplňuje automaticky a není třeba je zadávat. Operátory je možné využívat v pokročilém vyhledávání, pokročilé vyhledávání pracuje s nepovinným operátorem AND (zadáním znaku +), operátorem NOT (zadáním znaku -) a také s uvozovkami pro vyhledání přesné fráze, čárkou mezi dvěma slovy pro libovolnou vzdálenost obou slov od sebe. Další operátory suplují filtry pokročilého vyhledávání, např. operátor *site:* pro vyhledávání pouze v určité doméně či webu. Navíc je zde operátor *lang:* pro omezení jazyka a operátor *host:* stejně jako *site:* pro vyhledávání v doméně, ale ne již v poddoméně [SEZNAM.CZ, 1996k].

K dispozici jsou ve vyhledávání uživatelům i miniaplikace jako je kalkulačka, převod jednotek, měn, římských a arabských číslic, výpočet BMI a morseovka [SEZNAM.CZ, 1996l]. K jejich aktivaci stačí zadat vyhledávací dotaz ve správném tvaru.

Výsledky vyhledávání se zobrazují v širokém panelu, druhý menší panel slouží pro reklamy z vlastního reklamního systému firmy *Seznam.cz Sklik*, které se ovšem objevují i na začátku výsledků vyhledávání. Zobrazuje se maximálně 10 výsledků na stránku, na další je nutné listovat. Mezi výsledky jsou zařazeny i tzv. upoutávky dodané z dalších služeb společnosti *Seznam.cz* jako jsou například *Zboží.cz*, *Firmy.cz*, *Novinky.cz* a další. Na konci stránky s výsledky jsou zobrazeny související dotazy.

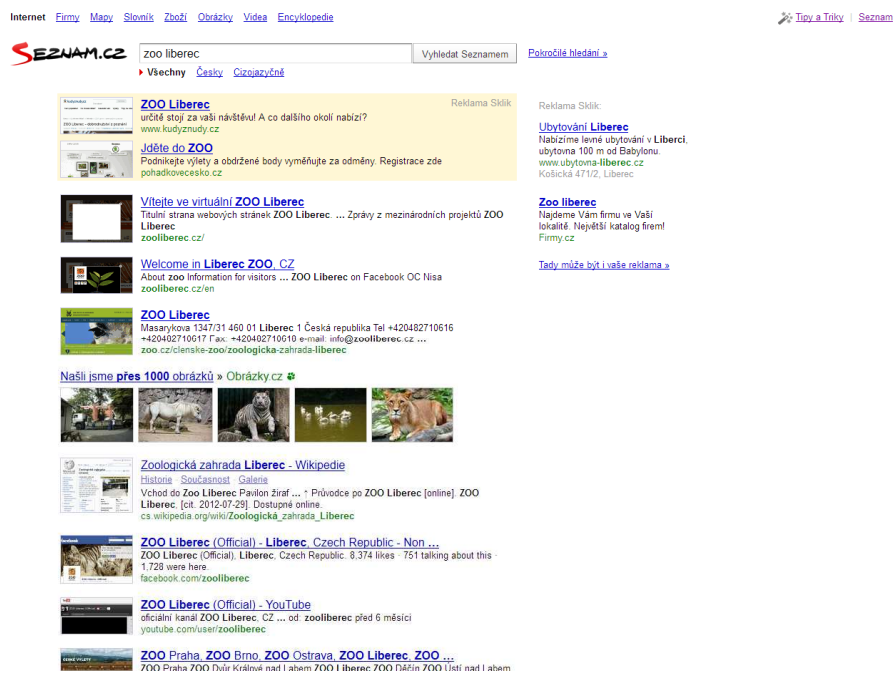
Výsledky řadí vyhledávač *Seznam.cz* podle vlastního neveřejného algoritmu s názvem *S-Rank*.

Jednotlivé výsledky vyhledávání obsahují kromě standardních údajů - názvu stránky, úryvku textu (*snippet*), URL adresy i automaticky generovaný náhled stránky, datum publikování, pokud je stránka členěna do různých tematických celků tak i osnova stránky (např. Wikipedia) a pokud je stránka zařazena v katalogu *Firmy.cz*, tak region a možnost zobrazení na mapě. Při běžném vyhledávání *Seznam.cz* vrací výsledky v různých jazycích dohromady, po zobrazení výsledků je možné filtrovat výsledky nebo vyhledávat pouze česky nebo cizojazyčně.

Portál *Seznam.cz* disponuje propracovanou nabídkou nápovědy (na adrese <http://napoveda.seznam.cz/>), která se vztahuje nejen k vyhledávání, ale i k dalším službám firmy *Seznam.cz*. Nápověda pro vyhledávání je podrobná, doplněná náhledy stránky a obsahuje informace pro běžné uživatele, ale i pro pokročilejší, například týkající se SEO (*Search Engine Optimization*) či administrace. *Seznam.cz* také provozuje speciální stránku věnovanou tipům a trikům pro vyhledávání (na adrese <http://www.chytrevyhledavani.cz/>), kde jsou k dispozici video-návody ke speciálním funkcím vyhledávání, například návod, jak zjistit aktuální kurzy měn. *Seznam.cz* nabízí i on-line podporu formou chatu nebo diskusní fórum, ve kterém si uživatelé mohou poradit navzájem.

Charakteristikou vyhledávače *Seznam.cz* je jeho orientace na český trh a poskytování množství dalších služeb. Jeho předností je kvalitní vyhledávání v českém webu, což z něj dělá silného konkurenta vyhledávači *Google* na lokální úrovni, ovšem jeho zaměřenost na český, částečně i slovenský trh limituje další mezinárodní expanzi.

Vyhledávač *Seznam.cz* je vhodný především pro uživatele, kteří neovládají cizí jazyk nebo chtějí vyhledávat české stránky a je také díky svému jednoduchému ovládní a podrobné nápovědě vhodnou volbou pro méně pokročilé uživatele. Pro vyhledávání v cizích jazycích je nutné se spokojit s výsledky dodávanými americkým vyhledávačem *Bing*. Popularita *Seznam.cz* také spočívá v tom, že je to portál a mnoho uživatelů tak oceňuje dostupnost mnoha služeb z jednoho místa.



Obrázek č. 14: Ukázka výsledků vyhledávání pomocí služby Seznam.cz

Specializované vyhledávače

Kromě všeobecných internetových vyhledávačů existuje i velké množství vyhledávačů specializovaných. Tyto vyhledávače jsou odlišné, protože se zaměřují na specifickou kategorii z hlediska obsahu, typu zdroje nebo způsobu vyhledávání.

Specializované vyhledávače jsou často zaměřeny na určité téma z hlediska obsahu. Takovým tématem jsou i informace o **pracovních nabídkách**. Mnoho vyhledávačů pracovních příležitostí (*job search engine*) pracuje jako metavyhledávače. Jedním z nejpopulárnějších vyhledávačů v oblasti pracovních nabídek je *Monster.com* (na adrese <http://www.monster.com/>), který koupil dříve také populární vyhledávač *Yahoo! HotJobs* [SCHONFELD, 2010]. Pomocí tohoto vyhledávače lze pracovní nabídky prohlížet nebo je vyhledávat v jednoduchém rozhraní pomocí názvu zaměstnání, firmy a regionu či v pokročilem vybrat další možnosti. *Monster.com* nabízí i související služby, uživatelé si mohou vytvořit svůj účet, poté konkrétní dotaz uložit a nechat si pravidelně zasílat výsledky vyhledávání. Zajímavou možností vyhledávače *Monster.com* je vložení životopisu do svého profilu [MONSTER, 2013]. Mezi další oblíbené vyhledávače patří například metavyhledávač *Indeed.com* nebo služba oblíbená především ve Spojených státech *CareerBuilder*. V českém prostředí patří k oblíbeným vyhledávačům pracovních příležitostí například služby *Jobs.cz*, *Prace.cz*, *Profesia*, *Sprace.cz*, provozovaná firmou Seznam nebo *Careerjet*.

Vyhledávače, jako jsou *Bing News*, *Google News*, *Yahoo News*, *MagPortal*, *Newslookup* a další, se zaměřují na **vyhledávání informací ze zpravodajství** (*news search engine*). Tyto vyhledávače agregují informace ze zpravodajských serverů, jako jsou noviny, časopisy, případně blogy. Vyhledávače, jako je například *Bing News*, se zaměřují na výběr z důvěryhodných a ověřených zdrojů. Vyhledávač článků *Topix* pracuje na podobném principu jako katalog webových stránek *Open directory project*, kde obsah tvoří skupina dobrovolníků. Díky zprávám, které navrhují a přidávají uživatelé, se *Topix* profiluje především jako vyhledávač zaměřený na lokální americké zdroje [TOPIX, 2001]. Vyhledávače jako *Google News* nebo *Yahoo News* mají mnoho regionálních jazykových verzí. Vyhledávače zpravodajských informací obvykle umožňují vyhledávat zprávy na základě jejich regionu, tematické kategorie, případně omezit dle data vydání nebo typu zdroje zpráv (*Newslookup*). Jedním z českých vyhledávačů zpravodajských informací je dnes už neaktualizovaný vyhledávač *Jyx* *Články*, jinak jsou zde zastoupené spíše agregátory zpravodajských informací, ke kterým patří například služba *Právědnes.cz* nebo *Chytré noviny*.

Zvláštním typem nástrojů pro vyhledávání informací jsou takzvané *question-and-answer* stránky (otázky a odpovědi), dále jako **Q&A**. Tyto nástroje se nedají nazývat internetovými vyhledávači, jelikož princip jejich fungování je odlišný. Q&A nástroje pracují s obsahem dodaným uživateli a s přirozeným jazykem. Příkladem takové služby je například *Yahoo! Answers*, která umožňuje uživatelům pokládat otázky a odpovídat na ně. Uživatelé mohou také odpovědi hodnotit a doporučit. Novějším nástrojem je služba *Quora* (na adrese <https://www.quora.com/>), která propojuje systém Q&A se sociálními sítěmi. Používat tuto službu je možné pouze po přihlášení, *Quora* nabízí přihlášení pomocí účtů ze sociálních sítí či e-mailového účtu. *Quora* pracuje s údaji ze zadaného účtu sociální sítě, takže zobrazuje například otázky, na které odpověděli nebo je položili přátelé ze sociální sítě. V tématech otázek je možné vyhledávat, třídít odpovědi podle nejlépe hodnocených, sledovat vybrané téma a další. *Quora* se snaží odpovědi třídít, vybírat relevantní, dobře hodnocené odpovědi z ověřených zdrojů [QUORA, 2013]. Tyto vyhledávací nástroje jsou vhodné v případě, že vyhledávané téma je natolik specifické, že není snadné ho nalézt pomocí klasického internetového vyhledávače.

2.3.3 Porovnání vybraných internetových vyhledávačů

Pro porovnání internetových vyhledávačů byly vybrány dva populární všeobecné vyhledávače *Google* a *Bing*. Portál *Yahoo!* byl z porovnání vynechán, jelikož dnes používá pro vyhledávání na internetu technologie společnosti Microsoft, která provozuje vyhledávač *Bing*. Dále byl zařazen alternativní vyhledávač *DuckDuckGo*, zaměřený na soukromí a ochranu osobních dat při vyhledávání, a dva regionální vyhledávače, *Yandex* jako jeden z nejvíce používaných regionálních vyhledávačů a český vyhledávač *Seznam.cz*. Do porovnání nebyl zahrnut výše popsáný vyhledávač *WolframAlpha*, jelikož pracuje na jiném principu než klasické internetové vyhledávače. Vyhledávače jsou porovnány z hlediska možností vyhledávání, jejich systému nápovědy a jsou zdůrazněny jejich hlavní rysy. Tabulka slouží k základnímu porovnání vybraných aspektů internetových vyhledávačů, jejich důkladnější analýzu obsahuje kapitola výše.

Internetové vyhledávače vybraných vyhledávačů pracují na podobné technologii, ovšem vyhledávače *DuckDuckGo* a *Seznam.cz* pracují i s výsledky jiných vyhledávačů. Z tabulky vyplývá, že všechny porovnávané vyhledávače umožňují uživatelům využít jak jednoduchého, tak i pokročilého vyhledávání a podporují Booleovské nebo i další operátory a samozřejmě je také filtrování výsledků dle různých kritérií. Největší vyhledávače *Google*, *Bing* a *Yandex* také výsledky vyhledávání personalizují. Unikátní technologií, kterou využívá pouze vyhledávač *Google*, je dynamické vyhledávání, které umožňuje vidět výsledky vyhledávání již v průběhu zadávání vyhledávacího dotazu. Technologii sémantického vyhledávání využívají pouze vyhledávače *Google* a *Bing*. Specifické funkce vyhledávání má také vyhledávač *DuckDuckGo*.

Uživatelská nápověda je dostupná u všech vyhledávačů. Vyhledávače také často obsahují stránky s informacemi a tipy pro vyhledávání. Většina vyhledávačů pak umožňuje odeslat dotaz či zpětnou vazbu pomocí formuláře, vyhledávač *Seznam.cz* umožňuje získat informace prostřednictvím on-line chatu s technickou podporou či diskusního fóra.

Aspekt vzhledu a prezentace výsledků vyhledávání není v tabulce rozepsán, protože v této oblasti se vyhledávače příliš neliší. Běžný záznam výsledků vyhledávání obsahuje název webové stránky, URL adresu a úryvek textu nazývaný také jako *snippet*. Záznamy výsledků některých členitějších webových stránek obsahují i osnovu, jejíž položky přímo odkazují na danou část webu. Záznamy výsledků vyhledávačů

DuckDuckGo a *Yandex* navíc obsahují také ikonu s logem dané webové stránky a záznamy výsledků vyhledávače *Seznam.cz* obsahují přímo náhledy webových stránek.

Hlavním rysem vyhledávače *Google* je jeho dominantní postavení a jeho mezinárodní zaměření, naopak pro vyhledávače *Bing*, *Yandex* a *Seznam.cz* je podstatným aspektem jejich regionální zaměření a znalost určitého regionu, například zpracování cyrilice v případě *Yandexu*. *DuckDuckGo* jako specificky zaměřený vyhledávač je významný svým důrazem na soukromí.

vyhledávač	možnosti vyhledávání	nápověda	hlavní rys
<i>Google</i>	<ul style="list-style-type: none"> - jednoduché i pokročilé rozhraní - podporuje Booleovské operátory a filtry - podpora automatického doplňování - dynamické vyhledávání - sémantický prvek vyhledávání (Knowledge Graph) - personalizace vyhledávání 	<ul style="list-style-type: none"> - podrobný návod s možností navigace a prohlížení témat nebo vyhledávání (v mnoha jazycích včetně češtiny) - rozsáhlá stránka informací o vyhledávání, včetně nápovědy, tipů, vysvětlení základů technologie vyhledávání - interaktivní formuláře nápovědy s možností výběru odpovědí - přehled výpadků a přerušení služeb - tipy pro bezpečné chování na internetu 	<ul style="list-style-type: none"> - dominantní postavení mezi vyhledávači - množství dalších služeb
<i>Bing</i>	<ul style="list-style-type: none"> - jednoduché i pokročilé rozhraní - podporuje Booleovské operátory a filtry - podpora automatického doplňování - sémantický prvek vyhledávání (Snapshots, není dostupný pro českou verzi) - personalizace vyhledávání 	<ul style="list-style-type: none"> - nápověda s možností prohlížení i vyhledávání (pouze v anglickém jazyce) - stránka s užitečnými návody a tipy jak vyhledávat - možnost kontaktovat podporu pomocí formuláře 	<ul style="list-style-type: none"> - zaměření na americký trh - vyhledávání v sociálních sítích (Bing Social) - multimediální prezentace titulní stránky

<i>DuckDuckGo</i>	<ul style="list-style-type: none"> - podporuje Booleovské operátory a filtry - tabulka se souhrnnými informacemi - zkratky pro vyhledávání - nepoužívá automatické doplňování 	<ul style="list-style-type: none"> - nápověda s možností prohlížení i vyhledávání (pouze v anglickém jazyce) - rozsáhlé diskusní fórum - možnost zapojení uživatelů do vývoje vyhledávače - možnost odeslat zpětnou vazbu pomocí formuláře 	<ul style="list-style-type: none"> - soukromí - syntax dotazů (zkratky pro vyhledávání) - personalizace vzhledu - vyhledává pouze webové stránky (ne obrázky, videa atd.)
<i>Yandex</i>	<ul style="list-style-type: none"> - jednoduché i pokročilé rozhraní - podporuje Booleovské operátory a filtry - podpora automatického doplňování - personalizace vyhledávání 	<ul style="list-style-type: none"> - podrobná nápověda pro uživatele (v ruském jazyce, stručná verze i v angličtině) - nápověda obsahuje náhledy webu nebo zadání dotazu - nápověda i pro webmastery - možnost odeslat zpětnou vazbu pomocí formuláře 	<ul style="list-style-type: none"> - regionální (orientace na Rusko a okolí) - zpracování azbuky - portál, mnoho služeb - vlastní prohlížeč a obchod s mobilními aplikacemi - vyhledává v sociálních sítích (Twitter, VKontakte aj.)
<i>Seznam.cz</i>	<ul style="list-style-type: none"> - jednoduché i pokročilé rozhraní - podporuje Booleovské operátory a filtry - podpora automatického doplňování 	<ul style="list-style-type: none"> - zpracovaný systém nápovědy (pouze v českém jazyce) - obsahuje náhledy webu - informace i pro webmastery - video-návody pro tipy a triky ve vyhledávání - helpdesk pomocí online chatu - diskusní fórum 	<ul style="list-style-type: none"> - regionální (české prostředí) - portál, mnoho služeb - pro vyhledávání obrázků a videí využívá cizí technologie

2.3.4 Metavyhledávače

Metavyhledávače jsou nazývány také jako metavyhledávací stroje, v anglickém jazyce jako *metasearch engines*, případně *metacrawlers*. Snahou metavyhledávačů je pokusit se pomocí kombinací více indexů pokrýt větší rozsah webu z důvodu malého překrývání indexů jednotlivých vyhledávačů. Nízké překrytí indexů vyhledávačů dokládá například studie pro metavyhledávač *Dogpile* vytvořená ve spolupráci s Queenslandskou technologickou univerzitou (*Queensland University of Technology*) a Pensylvánskou státní univerzitou (*Pennsylvania State University*), ve kterém bylo zkoumáno překrytí indexů vyhledávačů *Google*, *Yahoo*, *Live Search* a *Ask.com*²⁸ [DOGPILE, 2007]. Současné metavyhledávače je možné rozdělit do dvou skupin podle svého zaměření - všeobecně zaměřené metavyhledávače a metavyhledávače specializované na konkrétní oblast (např. pracovní nabídky).

Tyto vyhledávací nástroje pracují tak, že kombinují výsledky vyhledávání několika různých vyhledávačů. Uživatel tak může vyhledávat pomocí jednoho rozhraní v několika indexech různých vyhledávačů. Rozdílem mezi klasickým fulltextovým vyhledávačem a metavyhledávačem je z technického hlediska absence vlastního indexu a nutnost deduplikace vyhledaných záznamů. Uživatelskou výhodou metavyhledávačů je možnost prohledávat pomocí několika vyhledávačů z jednoho uživatelského rozhraní. Dnes metavyhledávače nevyhledávají pouze webové stránky, ale také obrázky, videa, zprávy a další typy obsahu.

První metavyhledávače začaly vznikat v polovině 90. let a jedním z prvních byla služba *MetaCrawler*, která vznikla v roce 1994. V současnosti (na adrese <http://www.metacrawler.com/>) jej vlastní firma Infospace (dnes pod názvem Blucora) a *MetaCrawler* shromažďuje výsledky vyhledávání ze služeb *Google*, *Yahoo* a *Yandex*. Firma Infospace vlastní mimo jiné i jeden z prvních fulltextových vyhledávačů *WebCrawler*, který dnes už pracuje jako metavyhledávač využívající výsledků *Google* a *Yahoo* (<http://www.webcrawler.com/>) a také populární metavyhledávač *Dogpile* (<http://www.dogpile.com/>). Některé metavyhledávače jako například služba *iBoogie* (na adrese <http://iboogie.com>) pracují s analýzou přirozeného jazyka a umí výsledky shlukovat do skupin podle jejich tématu (*clustering*).

²⁸ V této studii bylo zjištěno, že v roce 2007 bylo mezi těmito největšími vyhledávači průměrně 88 % unikátních výsledků (nalezených pouze jedním z vyhledávačů), přibližně 9 % výsledků bylo nalezeno současně dvěma vyhledávači, 2,2 % třemi a pouze 0,6 % všemi čtyřmi vyhledávači současně.

V dnešní době nejsou metavyhledávače již příliš používané. Důvodem je, že dnes již velké všeobecně zaměřené vyhledávače, zejména *Google*, dokážou najít velké množství relevantních dokumentů a není tak třeba kombinovat výsledky z více vyhledávačů. Kromě toho, fulltextové vyhledávače dnes poskytují množství dalších funkcí a služeb vyhledávání, že metavyhledávače jsou proti nim poměrně zastaralé. Metavyhledávače se dnes uplatňují spíše v určitých tematicky vymezených oblastech jako specializované metavyhledávače.

2.3.5 Vyhledávání v hlubokém webu

Pojmem hluboký web je označována část webu, jejíž informační zdroje jsou běžnými vyhledávacími nástroji nedostupné. Pro tento pojem se v češtině objevují výrazy *hluboký web* nebo také *neviditelný web*. V anglické literatuře je pak názvů hlubokého webu více - *deep web*, *invisible web*, *hidden web* nebo *undernet*. Jedním z prvních, kdo se o *hluboký web* začal zajímat, byl Michael Bergman, zakladatel společnosti BrightPlanet, ovšem pojem *neviditelný web* definovala již v roce 1994 Jill Ellsworth [BERGMAN, 2001, s. 2]. Oblasti webu, která je volně přístupná a prohledávatelná běžnými internetovými vyhledávacími nástroji se říká povrchový neboli viditelný web (*surface web*, *visible web*).

K důvodům, proč většina internetových vyhledávačů nemůže indexovat hluboký web, patří to, že často obsahuje databáze, které dynamicky generují obsah na základě požadavků uživatele (např. online katalogy knihoven, kalkulátory), přístup je možný pouze po zadání přihlašovacích údajů či hesla (např. komerční databáze), obsahuje různé formáty zdrojů (např. PostScript), které vyhledávač nedokáže zpracovat, neodkazují na ně žádné jiné webové stránky nebo využívají jiný protokol než http (např. Gopher, FTP). Velmi častým obsahem *hlubokého webu* jsou komerční databáze a databázová centra [PAPÍK, 2011, s. 68].

Vzhledem k tomu, že hluboký web není pro tradiční vyhledávače běžně dostupný, existují určité specializované nástroje určené k vyhledávání v oblasti hlubokého webu. Získávání informací z hlubokého webu se od práce robotů internetových vyhledávačů liší, nástroje pro prohledávání hlubokého webu například používají protokol pro sklizení metadat OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*), který je také často využíván v digitálních knihovnách.

Nástroje pro vyhledávání v hlubokém webu

Existují různé typy nástrojů, které umí vyhledávat v obsahu hlubokého webu, ovšem v současné době již umí indexovat hluboký web i významné internetové vyhledávače, ale indexují pouze jeho malou část. Z volně dostupných zdrojů, které indexují i hluboký web, jsou to služby typu předmětových katalogů jako je například katalog obrovského množství zdrojů z hlubokého webu *CompletePlanet* nebo dnes již zaniklý katalog vědecky zaměřených zdrojů *Intute* či již zmiňovaný nejstarší katalog *The WWW Virtual Library*. Existují také speciální vyhledávače, které indexují i hluboký web, některé z nich jsou tematicky zaměřené na oblast vědeckých informací jako je služba *Scirus*, *Infomine* nebo portál *Science.gov*. Nástroji pro přístup k hlubokému webu jsou také některé metavyhledávače jako například služba *SurfWax* nebo metavyhledávač s funkcí shlukování výsledků *Clusty*. Některé informace z hlubokého webu indexují i speciální služby všeobecně zaměřených populárních vyhledávačů. Na oblast vědy a výzkumu je zaměřena služba *Google Scholar* od firmy Google a služba firmy Microsoft nazvaná *Microsoft Academic Search*.

Complete Planet

Služba *CompletePlanet* (na adrese <http://aip.completeplanet.com/>) od firmy BrightPlanet je nástrojem pro vyhledávání v hlubokém webu. V *CompletePlanet* je možné buď vyhledávat pomocí zadání klíčových slov, nebo procházením hierarchicky a tematicky uspořádaného katalogu. Databáze *CompletePlanet* obsahuje přes 70 tisíc záznamů zdrojů či databází hlubokého webu [BRIGHTPLANET, 2000]. Nevýhodou tohoto nástroje je, že již není aktualizován (zhruba od roku 2004) a tak je také mnoho odkazů nefunkčních.

Scirus

Scirus (na adrese <http://www.scirus.com/>) je vyhledávací nástroj od firmy Elsevier, který se zaměřuje se na oblast vědeckých informací. Služba *Scirus* vyhledává ve stránkách univerzit, vědců, vědeckých organizací či společností nebo státních institucí zabývajících se výzkumem a také v odborných časopisech, zdrojích patentových informací, repozitářích a vědeckých databázích jako je například *MEDLINE*, *arXiv.org* či *RePEc*. Stránky, které *Scirus* indexuje, zahrnují různé typy dokumentů (patentové informace, preprinty, technické zprávy aj.), weby s omezeným

přístupem a stránky, které běžné fulltextové vyhledávače neumí indexovat [ELSEVIER, 2013].

Způsob práce tohoto tematicky zaměřeného vyhledávače se od běžných fulltextových vyhledávačů liší především ve způsobu získávání dat. Zdroje jsou pečlivě vybírány například pomocí automatické analýzy oblíbených stránek v určité vědní oblasti nebo mohou být nové zdroje manuálně navrženy správci a uživateli služby *Scirus*. Mimo sklizení dat pomocí protokolu OAI získává *Scirus* data do svého indexu pomocí takzvaného *focused crawling* (zaměřeného procházení webu), kdy roboty procházejí web podle určených pravidel (např. podle vlastního seznamu zdrojů služby *Scirus - seed list*) [ELSEVIER, 2004, s. 7-9].

Vyhledávat je možné pomocí jednoduchého nebo pokročilého rozhraní, v pokročilém rozhraní je možné použít Booleovské operátory mezi klíčovými slovy, omezit dotaz podle data, formátu dokumentu, typu zdroje či širšího tematického zaměření. Výsledky dotazu lze řadit dle relevance nebo data a dotaz je možné ještě upřesnit některým dalším ze systémem navržených klíčových slov.

Vyhledávač *Scirus*, který svým vzhledem a způsobem vyhledávání připomíná spíše komerční specializované databáze, je jedním z nejkvalitnějších volně dostupných vyhledávacích nástrojů v oblasti vědeckých informací a který dokáže prohledávat i hluboký web.

2.3.6 Webové archivy

Vzhledem k dynamické povaze webu, kdy každý den narůstá počet webových stránek a další obrovské množství stránek zaniká, mění svou podobu nebo adresu, jednou z možností, jak vyhledávat informace na internetu jsou digitální archivy webových zdrojů (webové archivy). Činností těchto archivů je výběr webových zdrojů a jejich archivace za účelem dlouhodobé ochrany a zpřístupnění. Webové archivy přispívají k zachování kulturního dědictví určitého regionu v době, kdy množství informací vzniká přímo v elektronické podobě (*born-digital*). V těchto archivech je tak možné nalézt již zaniklé webové zdroje nebo informace, které na příslušných stránkách již nejsou. Webové archivy obvykle obsahují kolekci webových zdrojů, které jsou v pravidelných intervalech archivovány. Výběr zdrojů k archivaci je prováděn buď manuálně pracovníky knihoven (*Webarchiv*) nebo automaticky (*Internet Archive*).

Archivací webových zdrojů se zabývají komerční firmy, jako je například organizace *Hanzo Archives*, a také státní instituce, zejména národní knihovny. Organizace a instituce, které se archivací webových zdrojů zabývají, sdružuje mezinárodní konsorcium IIPC (*International Internet Preservation Consortium*). Jedním ze zakladatelů konsorcia IIPC a v současnosti jedna z největších organizací, která archivuje webové zdroje a která je také členem Americké knihovní asociace (*American Library Association*), je nezisková organizace *Internet Archive* ([na adrese http://archive.org/](http://archive.org/)). Společnost *Internet Archive* byla založena v roce 1996 [INTERNET ARCHIVE, 2001a] a dnes obsahuje a zpřístupňuje více než 10 PB dat, včetně textových, audiovizuálních či softwarových materiálů [ONGE, 2012]. Kolekce zdrojů *Internet Archive* má nadnárodní záběr a archivované verze webových stránek jsou k dispozici online a zdarma, prostřednictvím programu Wayback Machine.

Důležitým tématem, které se týká vzniku i existence webových archivů, jsou autorská práva. V některých státech jako je například Francie, Dánsko či Nový Zéland je uzákoněn povinný výtisk i na elektronické publikace, prakticky to znamená, že v těchto zemích mohou archivovat webové zdroje bez vyjádření souhlasu vydavatele. V dalších státech, které vytváří webový archiv, však taková legislativa chybí a archivace a zpřístupnění jsou prováděny na základě dohody s vydavateli stránek. U organizace *Internet Archive* je situace složitější, knihy, audiovizuální, zvukové materiály nebo software zpřístupňuje převážně na základě veřejné licence (*public domain*) nebo jsou to volná díla, ale webové stránky archivuje *Internet Archive* jako nekomerční společnost

bez svolení vydavatelů. Dodatečně je však možné na žádost stránku vyřadit z archivu nebo archivovanou verzi znepřístupnit [INTERNET ARCHIVE, 2001b].

Národní knihovny archivují webové zdroje především na regionální úrovni. Jejich cílem je v ideálním případě archivovat celou národní doménu, to ovšem z důvodů omezeného úložiště není možné a cílem je tak především snaha získat reprezentativní vzorek webových zdrojů určitého státu, obvykle z jazykového a tematického hlediska. V České republice je takovýmto projektem *Webarchiv* Národní knihovny ČR, který archivované zdroje, vystavené pod licencí Creative Commons nebo s jejichž vydavateli má uzavřenou smlouvu, zpřístupňuje volně na svých webových stránkách (<http://www.webarchiv.cz/>). Jednou ročně je také prováděna celoplošná sklizeň české domény. Zdroje bez licence nebo smlouvy jsou dostupné pouze v budově Národní knihovny, legálně v rámci knihovní licence²⁹. Na stránkách *Webarchivu* je možné vyhledávat podle URL adresy stránky i podle klíčových slov a také lze prohlížet zdroje podle jejich oboru. Záznamy archivovaných stránek jsou také obsaženy v knihovním katalogu [KRATOCHVÍLOVÁ, 2012].

²⁹ Autorský zákon č. 121/2000 Sb., § 37.

3 Trendy a směry vyhledávání na internetu

3.1 Vyhledávání multimediálních informací

Dnešní internet disponuje mnohem větším multimediálním obsahem než dříve, což je způsobeno, jak rozvojem technologií pro přenos dat, tak i prudkým rozvojem zařízení pro zaznamenání multimediálního obsahu.

Lze předpokládat, že zájem o vyhledávání multimédií stále poroste. Podle průzkumu z roku 2012 patří činnost spojená s multimédií (přehrávání videí, poslech hudby, stahování atd.) k nejpoblárnějším informačním aktivitám na internetu mezi mladými lidmi ve věku 18-24 let [ST. JEAN, 2012]. Lidé obvykle vyhledávají multimediální informace z důvodu informační potřeby, pro zábavu, ke stažení a pro názornost [CHUNG, 2012]. Ke specifikům vyhledávání multimediálních informací, odlišným od vyhledávání textu, patří například velikost a kvalita souboru (např. vysoké rozlišení obrázku), rozměry, délka a čas, velké množství různých formátů a především obsah. Relevance a řazení výsledků u vyhledávání multimédií jsou také složitější než u textových dokumentů, jelikož není snadné vytvořit nějaké univerzální pravidlo pro relevanci multimediálních dokumentů. Oblast multimédií je velmi subjektivní a například jedna fotografie se dá popsat mnoha způsoby.

Multimediální dokumenty jako je obraz, audio či video, je možné vyhledávat dvěma typy způsobů. Jedná se o vyhledávání na základě metadat (např. textový popis obsahu obrázku či jméno interpreta skladby) a o vyhledávání na základě obsahu těchto dokumentů.

Multimediálních informace dnes standardně vyhledávají i největší všeobecně zaměřené internetové vyhledávače jako *Google*, *Bing*, *Yandex* a další, ale existují i specializované vyhledávače zaměřené obvykle na jeden typ multimediálního obsahu.

S oblastí využívání multimediálních dokumentů na internetu je také spojeno autorské právo. Pro sdílení na internetu je právně vhodné využívat volných děl. Pomocí rozcestníku pro vyhledávání na stránkách společnosti Creative Commons (na adrese <http://search.creativecommons.org/>) je možné vyhledávat různé typy multimediálních dokumentů, které jsou pod licenci *Creative Commons*. V tomto rozcestníku lze vybrat způsob dalšího využití díla (pro komerční účely nebo pro úpravu díla) a službu, ve které bude vyhledávat (např. Youtube, Flickr, Wikimedia Commons).

3.1.1 Vyhledávání obrazových dokumentů

Od vzniku digitálních fotoaparátů se začalo rozvíjet stále snadnější zachycení světa pomocí fotografií a dnes, kdy je možné nahrávat fotografie okamžitě na internet pomocí mobilního telefonu, velkou část elektronického obsahu na internetu tvoří obrázky. Díky lidské potřebě informace třídit a vyhledávat vznikly nejprve metody vyhledávání obrázků na základě jejich textového popisu a poté i na základě jejich obsahu. Vyhledávání obrázků se v anglickém jazyce nazývá jako *image retrieval* nebo *image search*.

Při vyhledávání obrázku pomocí klíčového slova, vyhledávače obvykle vyhledávají v názvu souboru obrázku, jeho adrese, textu, který na obrázek odkazuje a v textu, který je připojen k obrázku [IMAGE RETRIEVAL, 2001]. Pro popis obrázků samotnými uživateli jsou používána metadata ve formě tagů. Uživatelé tak k obrázkům připojují vybraná nebo volně tvořená klíčová slova. Tento systém je využit například ve službách hostingu fotografií jako je *Flickr* nebo *Picasa*. Společnost Google v roce 2006 pro zlepšení vyhledávání obrázků spustila hru s názvem *Google Image Labeler*, ve které uživatelé popisovali klíčovými slovy náhodné obrázky a získávali body, pokud navrhli stejné slovo jako protihráč. Tato služba již byla ukončena [GOOGLE IMAGE LABELER, 2001].

Další metodou, jak vyhledávat obrazové dokumenty je analýza obsahu dokumentu jako jsou barvy, tvary nebo textura. Tento přístup je v anglickém jazyce nazýván jako *query by image content (QBIC)*, případně *content-based image retrieval (CBIR)* nebo také *content-based visual information retrieval (CBVIR)* [CONTENT-BASED IMAGE RETRIEVAL, 2001]. Počátky vývoje vyhledávání obrazových dokumentů podle jejich obsahu spadají do druhé poloviny 90. let [DATTA, 2008, s. 5].

Rozpoznávání obsahu obrázku se neuplatňuje pouze ve vyhledávání pro uživatelské potřeby, ale například i v oblasti autorských práv. Při vyhledávání pomocí obrázku, lze zjistit, kde na internetu byl tento obrázek použit. Pro vyhledávání obrázků k volnému použití naopak slouží vyhledávání dle metadat.

Google Obrázky

Služba pro vyhledávání obrazových informací *Google Obrázky* (na adrese <http://images.google.com/>) nejpoužívanějšího vyhledávače *Google* vznikla již v roce 2001 [GOOGLE, 2001a]. Do roku 2011 bylo možné vyhledávat obrázky pouze pomocí

klíčových slov, od roku 2011 bylo přidáno do této služby vyhledávání pomocí obsahu obrázku [GOOGLE IMAGES, 2001]. Při vyhledávání pomocí klíčových slov je možné vybrat z několika dalších možností vyhledávání a filtrů jako je velikost obrázku, barva, typ obrázku (fotografie, obličej, kresba atd.) a čas zveřejnění. Obrázky je také možné třídit podle tématu a zobrazit jejich velikost. Ve výsledcích vyhledávání se uživateli také nabízí související klíčová slova a samozřejmě je také možnost nastavení filtru bezpečného vyhledávání. V rozšířeném vyhledávání obrázků je možné použít ještě více možností pro vyhledávání jako například vyhledat obrázky určitého formátu, podle tvaru, poměru stran nebo vyhledávat v určité doméně nebo podle typu licence (např. k volnému užívání, úpravám či komerčnímu sdílení). Výsledky vyhledávání není od roku 2010 nutné listovat, další výsledky se načítají automaticky [SMITH, 2010b]. Jakýkoliv nalezený obrázek je možné použít pro další vyhledávání pomocí obsahu obrázku. Pro vyhledávání pomocí obsahu obrázků ovšem primárně slouží ikona fotoaparátu ve vyhledávacím poli. Vyhledávat je možné dvěma způsoby, buď zadat URL adresu obrázku, nebo nahrát obrázek přímo ze svého počítače. Tato služba vyhledá vizuálně podobné obrázky, ale také webové stránky, na kterých se tento obrázek vyskytuje. Vyhledávání podle obsahu je možné zkombinovat s vyhledáváním pomocí metadat, po zadání URL adresy obrázku či jeho nahrání lze k tomuto vyhledávání přidat klíčová slova.

Společnost Google také vytvořila vyhledávač obrázků *Image Swirl*, založený na rozpoznávání obsahu a který shlukoval nalezené obrázky do skupin na základě jejich podoby a souvislosti. Tyto skupiny byly mezi sebou propojené a hierarchicky uspořádané, takže byl uživatel pomocí tohoto systému navigován k požadovanému obrázku [JING, 2012]. Projekt *Image Swirl* byl již zrušen.

Trendem je také využití vyhledávání pomocí obrazových informací v mobilních telefonech. Na principu rozpoznávání obrazových informací je založena služba *Google Goggles*. Prostřednictvím této služby lze mobilním telefonem vyfotit jakýkoliv objekt, služba fotografií analyzuje a na základě rozpoznávaných informací nabídne relevantní výsledky. *Google Goggles* zatím rozpoznává především textové informace, jako jsou například knihy, loga značek a z obrazových informací například zeměpisná místa či významná umělecká díla [GOOGLE, 2011b]. Aplikace zatím není velmi rozšířená, jelikož rozpoznávání obrázků se často nepodaří a je vhodné zatím pouze pro určité typy objektů.

Mimo tyto významné vyhledávače se širokým záběrem vyhledávaných typů dokumentů existují také specializované vyhledávače, zaměřené pouze na vyhledávání obrázků.

K neznámějším vyhledávačům obrázků podle obsahu patří služba *TinEye*, vytvořená kanadskou firmou Idée. *TinEye* sebe na svých stránkách (<http://www.tineye.com/>) prezentuje jako obrácený vyhledávač obrázků (*reverse image search engine*). Jako jeden z prvních specializovaných vyhledávačů obrazových dokumentů, který vyhledává podle obsahu obrázku, dokáže *TinEye* nalézt stejný obrázek, který se vyskytuje na více webových stránkách. Pomocí technologie *fingerprint* vytvoří digitální šifru zadaného obrázku, který pak porovnává s jinými. Díky této technologii lze vyhledat stejné nebo mírně pozměněné obrázky (jiná velikost, ořez, editace), ale není možné vyhledávat obrázky tematicky podobné [IDÉE, 2013a].

Jednou z novinek, které tato firma vyvíjí, je API (aplikační rozhraní) *MulticolorEngine*, jehož demoverze je k vyzkoušení na webových stránkách vyhledávače *TinEye*. Tato služba vyhledá obrázky podle zadaných barev. Je možné vybrat až pět barev a zvolit z kolika procent dané barvy se má obrázek skládat [IDÉE, 2013b].

Obrázek č. 15: Ukázka služby Multicolor Search Lab pro vyhledání obrázků podle barev

Oblast vyhledávání obrazových informací ještě není tak daleko jako je vyhledávání textových dokumentů. Problémy spojené s vyhledáváním obrázků podle jejich popisu jako je například subjektivní či nepřesný popis částečně odstraňuje metoda vyhledávání podle obsahu obrázků, která ovšem zatím není na takové úrovni, aby zcela vyhledávání dle popisu předstihla či nahradila. Aktuálním trendem, který využívá například firma Google, je kombinace vyhledávání pomocí obrázku se zpřesněním vyhledávání pomocí klíčových slov. Díky velkému rozvoji tabletů a chytrých telefonů se trendem stává také vyhledávání pomocí fotografií zachycených těmito přístroji. Na rozdíl od textového vyhledávání není tak snadné posoudit relevanci výsledků obrazového vyhledávání, ačkoliv se již vědci tímto tématem zabývají například v projektu vytvoření algoritmu *VisualRank* pro řazení obrázků dle relevance [SULLIVAN, 2008]. Vyhledávání obrazových dokumentů tak skýtá potenciál pro další rozvoj.

3.1.2 Vyhledávání zvukových dokumentů

Podobně jako při vyhledávání obrazových dokumentů, zvukové dokumenty je možné vyhledávat buď podle jejich popisu (*metadata-based music information retrieval*), v tomto případě zejména názvu skladby, interpreta či žánru nebo přímo podle obsahu (*content-based music information retrieval*, případně *music search*). Vyhledávání zvukových dokumentů pomocí obsahu je dále ještě možné dělit na množství různých způsobů zadání obsahu jako je vyhledávání ukázkou (*query by example*), zahráním (*query by playing*), broukáním (*query by humming*), vyťukáváním (*query by tapping*), vyhledávání obrysem melodie (*query by contour*) nebo vyhledávání notovým zápisem (*query by musical notation*) [VOCŮ, 2012]. V oblasti vyhledávání zvukových dokumentů se více uplatňují specializované vyhledávače, na úkor velkých všeobecných vyhledávačů na rozdíl například od situace v oblasti obrazových dokumentů. Ve vyhledávání hudby se také stále častěji uplatňují mobilní aplikace.

Vyhledávání hudby pomocí metadat využívají internetové obchody s hudbou jako je například *iTunes* nebo *Amazon* a také hudební streamovací služby jako je *Deezer*, *Rdio* či *Spotify*.

Jednou z nejpoužívanějších služeb, které umí rozpoznat hudbu podle zvukové ukázky je aplikace *Shazam*³⁰ (<http://www.shazam.com/>). Tato služba pracuje tak, že do mobilu nebo jiného zařízení je prostřednictvím mikrofonu nahrána zvuková ukázka,

³⁰ Především pro mobilní aplikace, dnes dostupná i pro operační systém Windows 8.

služba vytvoří akustický otisk skladby (*acoustic fingerprint*) a porovná se svou databází [ACOUSTIC FINGERPRINT, 2001]. Výsledkem je pak název skladby a jméno interpreta s ukázkou skladby. S výsledkem je možné dále pracovat, zakoupit skladbu v elektronickém obchodě *Amazon*, poslechnout si skladbu pomocí služby *Rdio*, sdílet výsledek na sociálních sítích, vyhledat videa s touto skladbou na *Youtube*, podívat se na termíny koncertů daného interpreta a další. *Shazam* dokáže rozpoznat pouze zvuk v kvalitě studiové nahrávky, podobná služba *SoundHound* (dříve *Midomi*, na adrese <http://www.soundhound.com/>) vyhledává skladbu nejen na základě přesné ukázky skladby, ale i na základě její interpretace, například živé verze z koncertu, vlastního zpěvu či broukání dané skladby.



Obrázek č. 16: Ukázka zpracování a zobrazení výsledku vyhledávání v aplikaci Shazam

3.1.3 Vyhledávání audiovizuálních dokumentů

Se snadnou dostupností fotoaparátů, webových kamer, mobilních telefonů a dalších zařízení pro nahrávání videa se stále zvyšuje počet audiovizuálních dokumentů na internetu. Jen na populární službu pro sdílení videí *Youtube* je každou hodinu nahráno kolem sta hodin videa [YOUTUBE, 2013].

Audiovizuální dokumenty jsou vyhledávány podle metadat jako je název, informace o autorovi, popis videa a také datum nahrání videa, jeho délka, kvalita a dalších parametrů. Ve službách pro nahrávání a sdílení videí jsou také pro popis obsahu videa často využívány tagy zadané autorem videa. Na rozdíl od vyhledávání obrazových nebo zvukových dokumentů není zatím technologie v oblasti audiovizuálních dokumentů na takové úrovni, aby bylo možné bezproblémově vyhledávat videa na základě analýzy jejich obsahu. Jednou z technologií, která napomáhá částečné analýze

obsahu videa je rozpoznávání řeči (*speech recognition*), kterou využívá například služba *Youtube* pro automatické vytváření titulků k videím [SPEECH RECOGNITION, 2001].

Vzhledem k aktuálnímu stavu technologií, kdy vyhledávání videí spočívá především ve vyhledávání pomocí metadat, včetně tagů, které uživatelé zadávají subjektivně a často s cílem získat vyšší sledovanost videa přidávají i zcela nerelevantní tagy, není přesnost vyhledávání audiovizuálních dokumentů vysoká a často obsahuje mnoho nerelevantních výsledků.

Populární všeobecné vyhledávače mají obvykle vlastní vyhledávání audiovizuálních dokumentů. Nejznámější službou nahrávání, sdílení a vyhledávání videí je služba *Youtube*, kterou v současnosti vlastní firma Google. Vyhledávač *Google* má ale také vlastní vyhledávání pro videa (<http://www.google.com/videohp>), jehož výsledky je možné filtrovat podle délky videa, času zveřejnění či aktualizace, kvality videa, jeho zdroje a také je možné vybrat pouze videa s titulky. Součástí záznamu výsledků vyhledávání je v tomto případě i náhled videa. Propracované rozhraní vyhledávání videí má také vyhledávač *Bing*. *Bing* indexuje videa populárních sítí ke sdílení videí, ale v americké verzi zpřístupňuje také televizní pořady vybraných amerických televizí. Vyhledávat je buď možné pomocí klíčových slov, nebo prohlížením kategorií jako jsou například virální videa, zpravodajská videa, kde je možné vybrat i z několika zpravodajských agentur, trailery k filmům, televizní pořady a další. Výsledky vyhledávání je možné filtrovat podle délky videa, rozlišení a zdroje a řadit dle relevance nebo času. *Bing* také na rozdíl od vyhledávače *Google* nabízí související témata k vyhledávání.

Mimo všeobecných vyhledávačů je možné audiovizuální dokumenty vyhledávat i pomocí specializovaných vyhledávačů. Příkladem může být vyhledávač *Blinkx* (<http://www.blinkx.com/>), který umožňuje vyhledávání videí pomocí klíčových slov nebo prohlížením tematických kategorií. Ke službě je možné se přihlásit pomocí účtu ze sociální sítě, poté zobrazuje proud videí seřazený podle času na základě preferovaných tematických kategorií a umožňuje ukládání oblíbených videí, sdílení videí a prohlížení oblíbených videí přátel ze sociálních sítí [BLINXX, 2013].

Specializovaným vyhledávačem videí zaměřeným na oblast vědy je služba *Yovisto* (<http://www.yovisto.com/>). Pomocí této služby je možné vyhledávat vzdělávací videa, například z univerzitních přednášek. Vyhledávač spolupracuje s významnými

americkými univerzitami a také s populárním cyklem konferencí *TED*³¹ [YOVISTO.COM, 2006].

3.2 Personalizace, lokalizace a socializace vyhledávání

3.2.1 Personalizované vyhledávání

Před několika lety přinášely vyhledávače každému uživateli stejné výsledky, dnes však každý uživatel dostává výsledky upravené podle jeho předchozí aktivity na internetu. Ukládání informací o uživatelské aktivitě a snaha přizpůsobit výsledky konkrétnímu člověku probíhá již delší dobu, ovšem přelomem v oblasti personalizace vyhledávačů se stal rok 2007, kdy největší internetové vyhledávače koupily společnosti zaměřené na online reklamu [RÖHLE, 2007]. Personalizace na základě webových historie jedním ze způsobů, jak se vyhledávač pokouší nalézt jedinečné a relevantní výsledky pro daného uživatele. Podle průzkumu společnosti Yandex personalizované vyhledávání dokáže ušetřit uživateli až 14 % času při vyhledávání, což je dané také tím, že podle Yandexu je 30 % vyhledávacích dotazů uživatele opakovaných [THE RUNET, 2012]. Jedním z důležitých důvodů personalizace vyhledávání je však také marketing pomocí cílené reklamy. Vzhledem k tomu, že je většina internetových vyhledávačů dostupná pro uživatele zcela zdarma, jejich největší zisky plynou z reklamy³². Cílem nejen internetových vyhledávačů, ale všech, kdo získává příjmy pomocí reklamy na internetu, se tak stává udržet uživatele na svých webových stránkách či aplikacích co nejdéle.

Vyhledávač *Google* začal standardně personalizovat všechny výsledky vyhledávání od roku 2009 [SULLIVAN, 2009]. To znamená, že vyhledávač *Google* personalizuje výsledky přihlášených i nepřihlášených uživatelů a pokud uživatelé nechtějí dostávat personalizované výsledky, tak musí tuto možnost vypnout v nastavení účtu či počítače. Personalizace ovlivňuje pořadí a zobrazování výsledků vyhledávání na základě činnosti na internetu, ale i informací o poloze a interakcí na sociálních sítích. K metodám sbírání dat za účelem personalizace patří sestavení profilu uživatele z hlediska demografických informací, jeho zájmů, chování na internetu (zda je pokročilý uživatel, kolik času tráví na stránkách atd.). Příkladem může být průzkum Marry Elen Bates, ve kterém zjistila, že vyhledávač *Google* v rámci personalizovaného vyhledávání nabízí několika uživatelům různé a různě uspořádané výsledky při dotazu na aktuální

³¹ *TED (Technology, Entertainment, Design)* pořádá každoročně konference s přednáškami z různých oblastí vědy, techniky, kultury a dalších. Záznamy z přednášek jsou volně dostupné ke zhlédnutí na internetu.

³² Více finančních informací o hospodaření společnosti Google na adrese <http://investor.google.com/>.

zpravodajství na vybrané téma, upravené na základě jejich zájmů, zpravodajských serverů a novin, které preferují a dalších [BATES, 2011].

Vyhledávač *Google* ve službě nazvané *Dashboards* (na adrese <http://www.google.com/dashboard/>) umožňuje po přihlášení zobrazit a spravovat přehled osobních dat spojených s účtem a v aplikacích od firmy *Google*. V nastavení reklam vyhledávače *Google* je možné zjistit a upravit informace, které se služba na základě uživatelských předchozí aktivity v rámci účtu *Google* pokouší odhadnout (např. věk, pohlaví, zájmy) a na jejich základě uživateli nabízet cílené reklamy.

Personalizace vyhledávání je současným trendem a vyhledávače soustředí na tuto oblast mnoho pozornosti. *Google* není jediným vyhledávačem, který personalizuje výsledky uživatelů, *Bing* spustil v roce 2011 funkci personalizace nazvanou *Adaptive Search* [CROOK, 2011] a ruský vyhledávač *Yandex* oznámil novinku v personalizaci svých výsledků, nově se výsledky vyhledávání budou aktualizovat i v rámci aktuální vyhledávací relace, to znamená v řádu sekund [YANDEX, 2013]. Existují ovšem i vyhledávače, které personalizaci nepodporují a zaměřují se na ochranu soukromí, příkladem je vyhledávač *DuckDuckGo*, *Blekk* nebo vyhledávač *StartPage*.

3.2.2 Lokalizované vyhledávání

S personalizací vyhledávání úzce souvisí i lokalizace vyhledávání. Lokalizace je založena zejména na metodě geolokace. Ke zjištění polohy touto metodou jsou používány údaje z IP adresy, informace z webového prohlížeče jako například nastavení jazyka nebo časové pásmo. V případě mobilních zařízení údaje o poloze poskytuje přímo operační systém zařízení, který ho získává z *GSM*, *GPS* (*global positioning system*) a také *Wi-Fi sítě*. V oblasti webu se lokalizace využívá například pro cílenou reklamu, automatické nastavení jazyka, omezení dostupnosti služeb na základě země. Lokalizace je využívána také pro služby vyhledávání na internetu jako například takzvané hledání míst (*local search*) provozuje například firma *Google* původně jako *Google Places* a v současnosti integrované do sociální služby *Google+* jako *Google+ Local*, firma *Microsoft* jako službu *Bing Local* nebo vyhledávač *Yahoo!* službu *Yahoo! Local*. Rozdílem oproti klasickému vyhledávání jsou u hledání míst především geografické informace. Velký zájem na vyhledávání míst mají především komerční firmy, které se snaží prostřednictvím těchto služeb získat zákazníky.

Služba *Google+ Local* (<http://www.google.com/+/learnmore/local/>) umožňuje vyhledat konkrétní místa v okolí. Město, oblast nebo adresu zjistí automaticky nebo je

možné adresu změnit manuálně. Tímto způsobem je možné vyhledávat například restaurace, obchody, muzea, hotely a další. Záznam takového místa obsahuje informace o typu místa, jeho umístění včetně náhledu mapy s propojením na *Google mapy*, kontaktní informace a otevírací dobu, fotografie, recenze a doporučení podobných míst. Přidanou hodnotou oproti předchozí službě *Google Places* je zde propojení na sociální služby. Uživatelé přihlášení prostřednictvím služby *Google+* mohou přidávat recenze a hodnocení míst, fotografie a mohou si místa přidávat mezi oblíbené nebo zobrazit recenze či oblíbená místa svých přátel v rámci sociální sítě [GOOGLE, 2001h].

3.2.3 Sociální vyhledávání

V současnosti se už uživatelé příliš nepotýkají s problémem nalezení informací, naopak vyhledávače jim předkládají obrovské množství výsledků. Dnešním problémem uživatelů při vyhledávání je rozpoznání kvalitních a ověřených informací v množství výsledků, které jim nabízí vyhledávače. Ještě před vznikem webu, vyhledávání informací často obsahovalo i sociální kontext, což zahrnovalo komunikaci s osobou, která danému tématu rozumí [BURGHARDT, 2012, s. 20]. Moderní vyhledávače se snaží pomocí nejnovějších technologií integrovat do vyhledávání i dostupné sociální informace. Nejvíce osobních informací a sociálních vazeb je dnes obsaženo v sociálních sítích.

Během posledních několika let, zhruba od roku 2004, kdy vznikla největší sociální síť *Facebook*, došlo k obrovskému rozvoji sociálních sítí (*social networking service*). Sociální sítě jsou služby, které uživatelům umožňují vytvoření osobního profilu a komunikovat, sdílet informace a data prostřednictvím služby s dalšími uživateli. K nejpobulárnějším sociálním sítím na světě patří kromě sítě *Facebook*, síť umožňující sdílet krátké textové příspěvky *Twitter*, profesionální sociální síť *LinkedIn*, síť zaměřená na vytváření a sdílení kolekcí obrázků *Pinterest*, hudebně zaměřená síť *MySpace* nebo sociální síť firmy Google *Google+* [EBIZMBA, 2013]. Mezi populární sociální sítě patří i regionální sociální sítě, zejména v oblasti Ruska a Asie dosahují sítě jako je *Odnoklassniki*, *Qzone* nebo *Vkontakte* popularity jako největší globální sociální sítě. V České republice patří k oblíbeným například služby *Lidé* a *Spolužáci* od firmy Seznam.cz. Kromě všeobecně zaměřených sociálních sítích sloužících především ke komunikaci s ostatními, existují i sociální sítě specializované, k oblíbeným patří například služba *Instagram* zaměřená na sdílení fotografií, *Foursquare* umožňující sdílet prostřednictvím mobilního telefonu svou aktuální polohu. Zajímavou sociální sítí,

tematicky zaměřenou na knihy a čtení je služba *Goodreads* (<http://www.goodreads.com/>). Uživatelé této sítě mohou označovat a sdílet knihy, které přečetli, psát recenze a doporučovat knihy přátelům. Další zajímavou specializovanou sítí je služba *ResearchGate* (<http://www.researchgate.net/>) umožňující komunikaci a spolupráci vědců.

Vzhledem k velkému rozvoji sociálních sítí a přítomné personalizaci vyhledávání, mají internetové vyhledávače zájem o indexaci a vyhledávání v obsahu sociálních sítí. Internetové vyhledávače dnes indexují veřejné profily a některé další veřejné informace ze sociálních sítí, například komentáře přidané na veřejnou stránku.

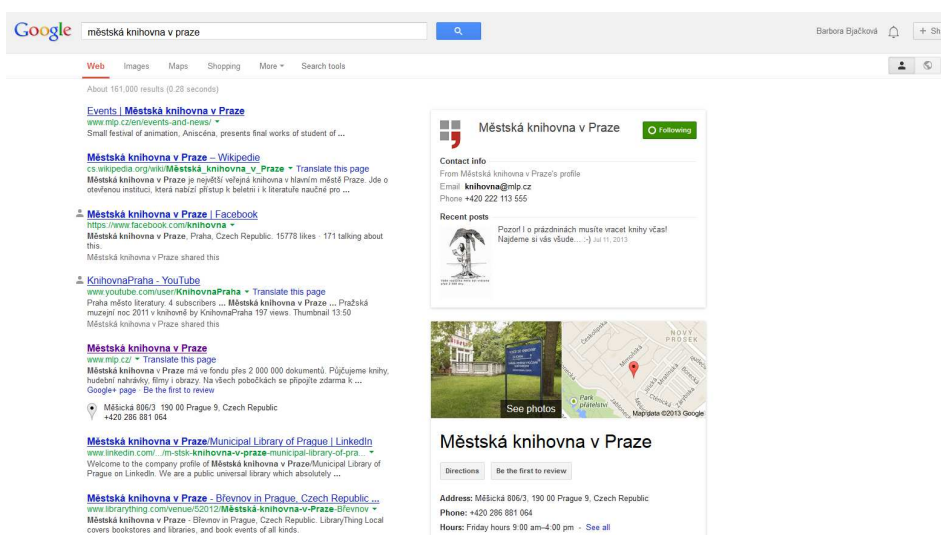
Jednu z nejpokročilejších integrací informací ze sociálních do svého vyhledávání sítí uvedl vyhledávač Bing. Služba nazvaná *Bing Social* (<http://www.bing.com/social>) umožňuje vyhledávat informace v nejnovějších veřejných příspěvcích sociálních sítí *Facebook* a *Twitter* a sledovat aktuální trendy. Bing také umožňuje prostřednictvím služby *Bing Social Sidebar* získávat informace ze sociálních sítí *Facebook*, *Twitter*, *Klout*, *Quora*, *LinkedIn* a dalších [MICROSOFT, 2013d]. Po přihlášení ze svého účtu na síti *Facebook* se ve výsledcích vyhledávání zobrazují i výsledky ze sociálních sítí (*Social Results*) v podobě panelu na pravé straně okna. V tomto panelu je zobrazena činnost (oblíbené stránky, sdílení statusu, videa či fotky atd.) facebookových přátel v souvislosti s vyhledávaným tématem a také další všeobecné informace ze sociálních sítí o daném tématu (např. poslední příspěvek vyhledávané osoby na *Twitteru*). Nově je možné prostřednictvím této služby sdílet výsledky vyhledávání nebo přidat komentář k příspěvku na síti *Facebook* přímo ze stránky vyhledávače. Další sociální službou vyhledávače Bing je *Friend's Photos*, která je zatím pouze ve verzi beta a umožňuje prohlížení fotoalb ze sociální sítě *Facebook* přímo na stránce vyhledávače s ohledem na nastavení jejich soukromí.

Služba zaměřená na sociální vyhledávání vyhledávače *Google* se nazývá *Search Plus Your World (SPYW)*. Pomocí této služby se ve výsledcích vyhledávání zobrazují i informace ze sociálních sítí, stránky a veřejné profily vyhledávaných osob či institucí na sociálních sítích, osobní informace uživatele nebo informace sdílené s uživatelem [SINGHAL, 2012a]. *Google* však ve svých výsledcích ze sociálních sítí preferuje vlastní sociální síť *Google+*, ovšem menší počet výsledků z jiných sociálních sítí je způsobeno i zamezováním přístupu ke svým datům od společnosti *Facebook* a *Twitter* [SULLIVAN, 2012]. Na rozdíl od vyhledávače *Bing*, *Google* tyto výsledky ze sociálních sítí začleňuje přímo mezi běžné výsledky vyhledávání, tyto výsledky jsou

pak odlišeny ikonou. Služba je stejně jako *Social Search* vyhledávače *Bing* dostupná pouze v anglické verzi.

Samotné vyhledávání v sociálních sítích není v současnosti na příliš vysoké úrovni, důkazem toho je i největší sociální síť *Facebook*. V sociální síti *Facebook* je možné vyhledávat pouze ve jménech lidí, názvech stránek, skupin a dalších entit. Pro výsledky vyhledávání osob nabízí *Facebook* filtry podle vzdělání (znamená školu, kterou navštěvují nebo absolvovali), místa bydliště nebo narození a pracoviště. Vyhledávat je také možné v textech nejnovějších příspěvků přátel nebo ve veřejných příspěvcích. Facebook však v současnosti vyvíjí nové vyhledávání *Graph Search* založené sociálních propojení a umožňující zadávat dotazy v přirozeném jazyce.

Díky obrovskému rozvoji sociálních sítí dnes společnosti největších sociálních sítí jako je *Facebook* a *Twitter* zaujímají na trhu stejně vysoké postavení jako největší internetový vyhledávač *Google*. Pro internetové vyhledávače tak spolupráce se sociálními sítěmi bude stále důležitější. V současnosti se v oblasti partnerství se sociálními sítěmi daří velmi dobře vyhledávači *Bing*, jelikož firma Microsoft uzavřela se společností *Twitter* dohodu o výhradním poskytování dat služby *Twitter* pro vyhledávač *Bing* a spolupracuje také se sítí *Facebook* [TAYLOR, 2010]. Pro vyhledávač *Google* by se *Facebook* mohl stát v oblasti vyhledávání velkou konkurencí, pokud se firmě *Facebook* podaří vytvořit kvalitní vyhledávač využívající osobních informací, které uživatelé sítě *Facebook* sdílí.



Obrázek č. 17: Ukázka vyhledávání pomocí vyhledávače Google se zapnutou službou Search Plus Your World

3.2.4 Kritika a obavy spojené s personalizací vyhledávání

S personalizací a lokalizací a socializací vyhledávání se pojí i obavy o zachování soukromí na internetu. Mezi rozšířené obavy spojené s personalizací patří únik či zneužití osobních dat nebo poskytnutí dat třetím stranám (například vládním organizacím), nezobrazení relevantních informací, pokud nekorespondují s osobním profilem uživatele uloženého v internetovém vyhledávači a použití osobních dat pro cílenou reklamu.

Podle průzkumu z roku 2012 než polovina respondentů (68 %) vnímá negativně využívání dat o předchozím vyhledávání pro úpravu jejich výsledků vyhledávání z důvodu omezení přístupu k požadovaným informacím [PURCELL, 2012, s. 20]. Množství respondentů (73 %) také vyjádřilo obavu z personalizace, protože to považují za zásah do soukromí [PURCELL, 2012, s. 21]. Přesto, že většina respondentů s personalizací nesouhlasí, pouze 38 % dotazovaných si uvědomuje, jak mohou sami omezit sběr svých osobních dat [PURCELL, 2012, s. 25]. Podobné výsledky přinesl i výzkum společnosti Google, podle kterého si 45 % uživatelů Google myslí, že by výsledky vyhledávání měly být pro každého stejné a 39 % respondentů personalizace nevdá, ale mají obavy o osobní data a pouze 7,5 % dotazovaných by bylo ochotno používat sociální síť *Google+* pokud by díky tomu mohli získávat více pro ně přizpůsobených výsledků vyhledávání [GOODWIN, 2012].

Dlouhodobým jevem je také omezování anonymity na internetu prostřednictvím vytváření verifikovaných osobních účtů a profilů a umožnění využívání mnoha služeb pouze po přihlášení jedním z takových účtů. Získaná data, tak mohou být jednoznačně spojena s konkrétní osobou.

Z důvodu personalizace došlo také ke změnám podmínek a zásad ochrany soukromí internetových vyhledávačů. Firma Google v roce 2012 sjednotila podmínky pro ochranu osobních dat u svých různých produktů a vytvořila tak nové zásady ochrany soukromí (dostupné na adrese <http://www.google.com/policies/privacy/>). Díky této změně může firma Google získaná data ze svých různých služeb (například vyhledávání, youtube atd.) přiřadit k jedinečnému účtu uživatele a používat je pro personalizaci ve svých službách. Do té doby byla data spojená s vyhledáváním uložena zvlášť a nebyla využívána v jiných službách [BLAKEMAN, 2012].

Problematické je také ukládání a další poskytování osobních údajů spojených s vyhledáváním. Firma Microsoft prohlašuje na svých stránkách, že data o vyhledávání pomocí svého vyhledávače *Bing* ukládá odděleně od identifikačních údajů přímo

spojených s uživatelem jako je jméno, e-mailová adresa nebo telefonní kontakt. V těchto záznamech o vyhledávání jsou také po 6 měsících vymazány propojení s IP adresou a s dalšími identifikátory jako jsou cookies po 18 měsících [MICROSOFT, 2013e].

Způsobů, jak omezit ukládání osobních údajů internetovými vyhledávacími a zabránit tak personalizaci, je více. K nejběžnějším patří vymazání historie vyhledávání a zákaz jejího ukládání. Tyto kroky umožňují vyhledávače obvykle v nastavení uživatelského účtu. Další možností je prohlížení stránek v anonymním režimu, což umožňují moderní internetové prohlížeče, nebo vyhledávání pomocí vyhledávačů, které neukládají osobní data. Pokročilejšími metodami je také použití proxy serveru nebo vymazání identifikátorů jako jsou cookies.

I přes to, že existuje množství lidí, kterým sbírání dat za účelem personalizace vadí, je nutné s personalizovaným vyhledáváním do budoucna počítat. Vyhledávače díky těmto údajům mohou efektivněji využívat reklamu a zvyšovat tím svoje zisky, proto není pravděpodobné, že by se tohoto sbírání dat vzdaly. Navíc mnoho uživatelů nepřestane těmto službám využívat i přesto, že personalizaci vnímá negativně.

3.3 Mobilní vyhledávání

Stále více uživatelů využívá pro práci a zábavu mobilní zařízení jako jsou chytré mobilní telefony (*smartphone*) a tablety. Podle průzkumu společnosti ComScore vlastní více než 129 milionů Američanů chytrý telefon [COMSCORE, 2013]. Také roste přístup z mobilních telefonů k internetu, aktuálně z celkové uživatelské činnosti na internetu je kolem 15 % realizováno přes mobilní telefony³³ oproti 8,5 % na začátku roku 2012 [STATCOUNTER, 1999]. Popularita mobilních zařízení roste zejména v Asii, v Číně počet přístupů k internetu z mobilního zařízení začíná předstihovat přístup ze stolního počítače [WEE, 2012] a v počtu uživatelů zařízení s mobilními operačními systémy Android a iOS letos již Čína, jako země s jedním z nejrychleji rostoucích trhů mobilních zařízení, předstihla Spojené Státy [FARAGO, 2013].

Mobilní zařízení tak do jisté míry mění podobu internetu, webové stránky jsou upravovány do podoby, která je lépe čitelná pro mobilní zařízení, vzniká velké množství aplikací určených pro mobilní telefony a pro mnoho aplikací využívaných již na stolních počítačích je vytvářena i mobilní verze. Tento trend má také vliv na

³³ Data k červnu 2013, nezahrnuje tablety.

vyhledávání. Díky mobilním zařízením je vyhledávání dostupné prakticky všude, kde je k dispozici mobilní síť. Internet však dnes proniká i do mnoha dalších oblastí, nové typy televizí nebo počítačů v automobilech jsou připojeny k internetu, platební terminály nebo veřejné informační terminály například s dopravními informacemi a další.

Vzhledem k omezenějším možnostem ovládání telefonů oproti počítačům bylo vyvinuto i hlasové ovládání. Při hledání prostřednictvím mobilního telefonu je pak možné využít hlasového vyhledávání (*voice search*). Hlasové vyhledávání pracuje s technologií rozpoznávání řeči a databází slov. V hlasovém vyhledávání od firmy Google jsou záznamy o vyhledávání ukládány a slovník je tak stále aktualizován. Vzhledem k velkému rozvoji mobilních zařízení skýtá trend hlasového vyhledávání velký potenciál. Podle společnosti Google bylo v roce 2010 pro 25 % vyhledávání v mobilních zařízeních s operačním systémem Android využito hlasového vstupu [KINCAID, 2010].

Jednou z neznámějších služeb, která využívá hlasové ovládání je aplikace *Siri*, osobní asistent a navigátor od firmy Apple, který byl spuštěn v roce 2011 [SIRI, 2001]. Aplikace *Siri* je ovládána hlasovými příkazy a kromě vyhledávání informací dokáže například odesílat a číst zprávy, zaznamenávat poznámky, nahrávat příspěvky na sociální sítě, získávat aktuální informace o počasí a další. Do aplikace je také integrováno vyhledávání pomocí výpočetního vyhledávače *WolframAlpha*. Konkurentem *Siri* je služba od firmy Google s názvem *Google Now*, což je také osobní asistent s hlasovým vyhledáváním, který ovšem stejně jako *Siri* není zatím dostupný v českém jazyce. Google však podporuje i hlasové vyhledávání jako aplikaci v rámci mobilní verze svého vyhledávače. Vlastní hlasové vyhledávání určené především pro mobilní telefony s operačním systémem *Windows Phone* vyvíjí také firma Microsoft pod názvem *Bing Voice Search*.

Velký rozvoj mobilních zařízení měl významný vliv na vývoj hledání míst (*local search*), které je popsáno výše.

Dalším stupněm ve vývoji stále mobilnějších zařízení jsou takzvané *wearable computers* (nositelné počítače). Jedno z nejnovějších zařízení tohoto typu vyvíjí firma Google pod názvem *Google Glass*. Pro běžné uživatele však zatím *Google Glass* k dispozici nejsou, současnou verzi *Explorer Edition* pro vývojáře bylo možné získat pouze za přísných podmínek. Brýle jsou vybaveny miniaturní obrazovkou s kamerou a modulem pro poslech zvuku a ovládají se hlasovými příkazy nebo touchpadem na straně brýlí. V současnosti jsou *Google Glass* stále ve vývoji a zatím umí fotografovat, nahrávat video, telefonovat a odesílat zprávy, vyhledávat pomocí *Google* a obsahuje

základní aplikace jako je kalendář, čas, navigace, překladač a další [GOOGLE, 2013h]. Pro *Google Glass* již existuje několik externích aplikací, které je možné do brýlí nainstalovat a nedávno byl vytvořen webový prohlížeč určený pro toto zařízení [WELCH, 2013].

Nositelné počítače jako je zařízení *Google Glass* jsou zatím ještě ve fázi vývoje, ovšem rostoucí požadavky na stále vyšší dostupnost informací a jejich vyhledávání z nich mohou vytvořit trend budoucnosti a změnit tak způsob získávání informací jako jej změnily mobilní telefony.

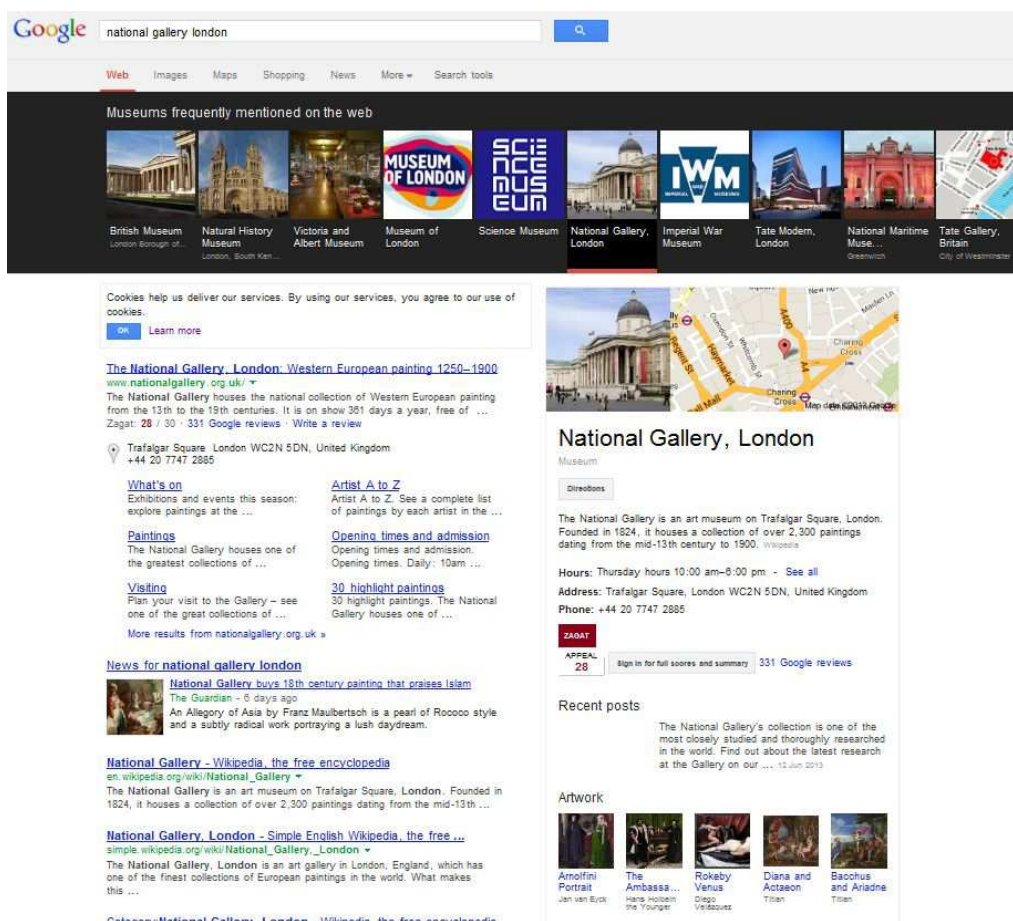
3.4 Sémantické vyhledávání

Sémantické vyhledávání je vyhledávání, které pracuje s významy termínů v kontextu dokumentů. Jeho charakteristikou je přímé odpovídání odpovědí na otázky místo zobrazení výsledků vyhledávání. Technologicky pracuje sémantické vyhledávání s analýzou přirozeného jazyka, sémantickou sítí objektů a vztahů a vazbami mezi informacemi.

Jedním z aktuálních projektů sémantického vyhledávání je služba firmy Google nazvaná *Knowledge Graph*, česky *Diagram znalostí*³⁴. *Diagram znalostí* pracuje na principu databáze, jejíž objekty jsou entity jako místa, osoby, umělecká díla a další a ty jsou mezi sebou propojeny vztahy. Informace tato databáze čerpá z veřejně dostupných encyklopedických zdrojů jako je *Wikipedia*, *Freebase* nebo *CIA World Factbook* a obsahuje přes 500 milionů objektů s více než 3,5 miliardami informací a vztahů [SINGHAL, 2012b]. Každý objekt má vlastní identifikátor a kategorii, do které patří, takže například pro vyhledávání klíčového slova "Waltari" nabídne k výběru informace o hudební skupině i finském spisovateli. Informace z *Diagramu znalostí* se ve výsledcích vyhledávání zobrazují jako tabulka se strukturovanými informacemi o vyhledávaném objektu (včetně obrázků, událostí, map, recenzí atd.), ale i se souvisejícími objekty (například knihy, které daná osoba napsala), případně objekty, které v této souvislosti lidé také často vyhledávají. Jednou z funkcí *Diagramu znalostí* je také tzv. karusel (*carousel*), což je skrolovací panel souvisejících objektů. Tento panel se zobrazuje při vyhledávání entity začleněné do nějaké skupiny, například při vyhledávání uměleckého díla, kdy služba zobrazí související díla od stejného autora, nebo je zobrazen při dotazu na výčet informací, například muzea v Londýně, herci v

³⁴ Česká verze této služby začala být uváděna pro vybrané uživatele v červnu 2013.

určitému filmu či renesanční malíři. *Diagram znalostí* je také propojen s lokálním vyhledáváním *Google+ Local* a zobrazuje i statistické údaje, jsou-li k dispozici. Výhodou této služby je okamžité získání informací bez nutnosti návštěvy další webové stránky, ale také vyhledávání souvislostí.



Obrázek č. 18: Ukázka vyhledávání s funkcí *Diagram znalostí*

Funkci podobnou *Diagramu znalostí* má i vyhledávač *Bing*. Tato služba vyhledávače *Bing* se nazývá *Snapshots* a je založená na technologii *Satori*, která také přidává objektům vzájemné vztahy [QUIAN, 2013]. Rozdílem oproti *Diagramu znalostí* je, že *Bing* do *Snapshots* pro osoby začleňuje i informace ze sociálních sítí, například odkaz na účet na sociální síti *Facebook*, *Twitter* a dalších. Oproti vyhledávači *Google* je služba *Snapshots* dostupná pouze v angličtině a není k dispozici v České republice.

Skutečnost, že se vyhledávače snaží pochopit význam vyhledávaných termínů, naznačují i další vyhledávače. *DuckDuckGo* umožňuje uživateli na dotaz s více významy výběrem tematické kategorie specifikovat požadovaný význam termínu a dále upraví vyhledávací dotaz, tak aby vyhledával daný význam. Jedním z nejpokročilejších vyhledávacích nástrojů, který pracuje s analýzou přirozeného jazyka je *WolframAlpha*,

který ovšem nepracuje jako klasický vyhledávač, ale dotazy vypočítává. Existují i experimentální vyhledávače zaměřené na sémantické vyhledávání, příkladem může být vyhledávač *Lexxe* (<http://www.lexxe.com/>). Vyhledávač používá technologii nazvanou sémantický klíč (*semantic key*), pomocí které lze vyhledávat významy. Sémantických klíčů je přes 500 a uživatelé mohou další sami navrhnout, příkladem sémantického klíče je cena, rychlost, datum a další [LEXXE, 2012]. Vyhledávací dotaz má pak strukturu sémantického klíče a klíčového slova, například dotaz ve tvaru "speed: cheetah" (rychlost: gepard). Mezi výsledky, které jsou zobrazeny jako v klasickém vyhledávači, jsou pak barevně vyznačeny odpovědi a je zobrazena také statistika nejčastějších odpovědí. Trendem se také stávají aplikace inteligentních osobních asistentů a navigátorů pro mobilní zařízení, které pracují se sémantickým zpracováním řeči, jako je například aplikace *Siri* od firmy Apple, aplikace *Evi* či *Assistant*.

Sémantické vyhledávání je odlišné v tom, že pracuje s pojmy a vztahy mezi nimi podobně jako přirozený jazyk, na rozdíl od klasického modelu vyhledávačů, které vyhledávají podle četnosti izolovaných termínů. Funkce sémantického vyhledávání je z uživatelského hlediska zatím poměrně omezená, ale současný vývoj ukazuje, že sémantické vyhledávání je významným trendem a můžeme očekávat, že v budoucnu vyhledávače mnohem lépe porozumí vyhledávacím dotazům v přirozeném jazyce a dokážou na ně odpovědět.

4 Závěr

Vyhledávání na internetu se od vzniku prvních nástrojů pro vyhledávání v něm, zejména později také od vzniku webu výrazně změnilo. Tento skutečný technologický vývoj je dnes možné porovnat s tím, jak se měnily představy a názory vědců, spisovatelů, filmařů a dalších umělců na technologické možnosti budoucnosti, které byly ztvárněny především v žánru *science fiction (sci-fi)*. V populárních filmech a seriálech *sci-fi* žánru se objevilo množství technologií, které byly později skutečně realizovány. Příkladem může být komunikační zařízení podobné mobilnímu telefonu zobrazené v 60. letech v seriálu *Star Trek* nebo univerzální překladač mluvené řeči, který se objevil v mnoha *sci-fi* filmových ztvárněních i literatuře, například postava zvaná Babylónská rybka (*Babel fish*) v knize *Stopařův průvodce po galaxii*, po které byl pojmenován jeden z prvních internetových překladačů. Ve svém díle 1984 načrtl spisovatel George Orwell antiutopistickou vizi světa budoucnosti. Mnoho z jeho představ technologií a možností sledování jsou dnes aktuálnější než kdy dříve.

Internet a jeho nástroje, včetně vyhledávání na internetu, jsou předmětem řešení řady etických problémů. Internet je poměrně svobodný prostor, což dokládají i jeho snahy o omezení v nedemokratických režimech, ovšem poslední zprávy ukazují, že se stává také nástrojem pro sledování lidí a získávání jejich osobních dat.

V současnosti se informace a jejich vyhledávání stávají stále dostupnějšími. Je to umožněno výkonnějšími počítači, které umožnily vzniku složitějšího softwaru s pokročilými funkcemi. K velkému rozvoji také došlo na straně internetových prohlížečů, od jednoduchého zobrazení textu a obrázků po současný interaktivní, multimediální obsah a pokročilý design stránek. Dochází ke stálému zlepšování práce internetových vyhledávačů, ale i vzniku dalších nových technologií jako jsou mobilní telefony, tablety a nově i nositelná zařízení jako jsou *Google Glass*.

Mění se také struktura informací na internetu, dříve byla většina informací tvořena vydavateli webových stránek, zatímco dnes se mění chování uživatelů od pasivních konzumentů k aktivním tvůrcům obsahu a tak stále přibývá obsahu tvořeného přímo uživateli. Velký vliv na toto chování mají především sociální sítě. Tato etapa obsahu webu, nazývaná také jako web 2.0 ovlivnila také vývoj internetových vyhledávačů. Díky nárůstu multimediálního obsahu je současným trendem vývoj vyhledávání multimediálních informací, jako jsou obrázky, zvuk či video. S vzrůstající

popularitou sociálních sítí je spojeno také úsilí internetových vyhledávačů o indexaci a umožnění vyhledávání těchto dat tvořených uživateli.

Situace na trhu s internetovými vyhledávači se také změnila, z množství konkurenčních vyhledávačů v minulosti si získala vedoucí místo na globálním trhu společnost Google, která se může dnes obávat spíše konkurence významných sociálních sítí než dalších internetových vyhledávačů. Velký vliv na internetové vyhledávače má také reklama.

Podle průzkumu z roku 2012 je spokojenost uživatelů s výsledky vyhledávání pomocí internetových vyhledávačů vysoká, 91 % respondentů se podaří najít vždy nebo téměř vždy informaci, kterou hledají a 73 % respondentů připadají informace získané pomocí internetového vyhledávače důvěryhodné [PURCELL, 2012, s. 3]. V oblasti uživatelských požadavků na nástroje pro vyhledávání informací na internetu tak již není v centru pozornosti relevance výsledků a nalezení požadované informace, uživatelé to totiž už dnes považují za samozřejmost.

V současnosti je ovšem velmi aktuální otázka soukromí a ochrany osobních údajů na internetu. Díky tomu, že společnosti provozující velké internetové vyhledávače zároveň nabízejí řadu dalších internetových služeb a uživatel tak často používá produkty a služby jedné společnosti téměř k celé své činnosti na internetu, získávají tyto společnosti množství osobních informací a dat, často dokonce demografický a zájmový profil uživatele. Získaná data jsou internetovými vyhledávači dnes již běžně využívána pro personalizaci a lokalizaci výsledků vyhledávání. Roste také tlak na snížení anonymity na internetu. Tyto jedinečné soubory dat se ovšem také stávají středem zájmů poskytovatelů cílené reklamy, ale i státních subjektů a mohou být i velkým rizikem pro uživatele při zneužití těchto dat.

Dalším problémem v současnosti není to, jak nalézt informaci, ale naopak přesycení informacemi, kdy je pro uživatele těžké se orientovat v množství informací na internetu a posoudit, které informace jsou důvěryhodné.

S rostoucí dostupností informačních technologií a internetu se vyhledávání stává stále běžnější formou získávání informací, uživatel začíná zapomínat různé informace, protože je může kdykoliv vyhledat.

Jedním z trendů, které se snaží vyhledávání usnadnit a přiblížit přirozenému jazyku je sémantické vyhledávání. Sémantickým vyhledáváním se zabývá výpočetní vyhledávač *WolframAlpha* a prvky sémantického vyhledávání, které by měly dotazu

uživatelé porozumět a přinést mu odpověď přímo a v kontextu, nově testují největší vyhledávače *Google* i *Bing*.

Dlouhodobým trendem je také integrace vyhledávání přímo do elektronických služeb, příkladem může být stránka pro nákup letenky, kde je třeba nejprve vyhledat požadovanou variantu a poté je umožněn nákup.

Pro pohled do budoucnosti internetových technologií a získávání informací se stačí podívat na asijské země. Prudký hospodářský a technologický rozvoj zemí jako je Čína, Indie, Jižní Korea či Indonésie umožňuje růst počtu internetových uživatelů či trhu s mobilními zařízeními. Číňané dnes tráví více času na internetu než Američané a také více využívají mobilní telefony [MEEKER, 2013].

Lze předpokládat, že v budoucnu se uplatní především nositelné počítače a připojení běžných předmětů k internetu. Očekávat lze také větší využití různých způsobů ovládání jako je ovládání hlasové či dotykové. Je tak možné, že se v budoucnosti změní způsob vyhledávání a místo dnešního vyhledávání pomocí klávesnice a internetového vyhledávače budou lidé využívat mobilní asistenty ovládané hlasově či dotykově, které budou rozumět přirozenému jazyku a rovnou předkládat vyhledané informace. Můžeme se domnívat, že v budoucnosti se získávání informací z internetu, ale i jejich sdílení, stane naprosto samozřejmou a přirozenou lidskou činností bez ohledu na to, kde se budeme nacházet.

Seznam použité literatury

- ACADEMICINFO, 2013. Palm Coast: Academic Directories, c2013 [cit. 2013-08-03]. Dostupné z: <http://www.academicinfo.net/>
- ACOUSTIC FINGERPRINT, 2001. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2013-08-04]. Anglická verze. Dostupné z: http://en.wikipedia.org/wiki/Acoustic_fingerprint
- ALEXA INTERNET, 2013. Top Sites: The top 500 sites on the web. *Alexa: The Web Information Company* [online]. Alexa Internet, [cit. 2013-04-27]. Dostupné z: <http://www.alexa.com/topsites/global>
- ALLEN, J., 2010. Yahoo! Organic Results Switch To Bing. *Search Engine Watch* [online]. 24. 8. 2010 [cit. 2013-05-03]. Dostupné z: <http://searchenginewatch.com/article/2050616/Yahoo-Organic-Results-Switch-To-Bing>
- ALPERT, J. a N. HAJAJ, 2008. We knew the web was big... In: *Google Official Blog* [online]. Google, 2008 [cit. 2013-07-29]. Dostupné z: <http://googleblog.blogspot.cz/2008/07/we-knew-web-was-big.html>
- ALTAVISTA, 2001. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001-, 24. 1. 2013 [cit. 2013-01-25]. Dostupné z: <http://en.wikipedia.org/wiki/AltaVista>
- AMERICAN DIALECT SOCIETY, 2013. 2002 Words of the Year. *American Dialect Society: ADS* [online]. American Dialect Society, 13. 1. 2013 [cit. 2013-07-30]. Dostupné z: http://www.americandialect.org/2002_words_of_the_year
- ARRINGTON, M, 2008. Google's Misleading Blog Post: The Size Of The Web And The Size Of Their Index Are Very Different. *TechCrunch* [online]. AOL, 2008 [cit. 2013-07-03]. Dostupné z: <http://techcrunch.com/2008/07/25/googles-misleading-blog-post-on-the-size-of-the-web/>
- BATES, M. E., 2011. Is Google really filtering my news?. In: *Librarian of Fortune: Mary Ellen Bates contributes white noise to the blogosphere* [online]. Mary Ellen Bates, 2011 [cit. 2013-07-07]. Dostupné z: http://www.librarianoffortune.com/librarian_of_fortune/2011/09/is-google-really-filtering-my-news.html
- BERGMAN, M. K., 2001. *The Deep Web: Surfacing Hidden Value* [online]. BrightPlanet, 2001 [cit. 2013-02-15]. Dostupné z: <http://brightplanet.com/wp-content/uploads/2012/03/12550176481-deepwebwhitepaper1.pdf>
- BERNERS-LEE, T, 1994. A Brief History of the Web. In: *W3C* [online]. W3C, 1994 [cit. 2013-01-27]. Dostupné z: <http://www.w3.org/DesignIssues/TimBook-old/History.html>
- BERNERS-LEE, T. a R. CAILLIAU, 1990. WorldWideWeb: Proposal for a HyperText Project. In: *W3C* [online]. W3C, 1990 [cit. 2013-01-27]. Dostupné z: <http://www.w3.org/Proposal.html>

BERNERS-LEE, T., 2013. The WorldWideWeb browser. In: *W3C* [online]. W3C, c2013 [cit. 2013-01-26]. Dostupné z: <http://www.w3.org/People/Berners-Lee/WorldWideWeb.html>

BLAKEMAN, K., 2012. Google personalisation: web history isn't the only problem. In: *Karen Blakeman's Blog: News and comments on search tools and electronic resources for business information* [online]. Karen Blakeman, 2012 [cit. 2013-07-12]. Dostupné z: <http://www.rba.co.uk/wordpress/2012/02/22/google-personalisation-web-history-isnt-the-only-problem/>

BLINXX, 2013. Blinxx, c2013 [cit. 2013-08-04]. Dostupné z: <http://www.blinkx.com/>

BLOGGER (SERVICE), 2001. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2013-08-03]. Anglická verze. Dostupné z: [http://en.wikipedia.org/wiki/Blogger_\(service\)](http://en.wikipedia.org/wiki/Blogger_(service))

BONFILS, M., 2013. Yandex Just Passed Bing to Become 4th Largest Global Search Engine. *Search Engine Watch* [online]. 2013 [cit. 2013-03-09] Dostupné z: <http://searchenginewatch.com/article/2242374/Yandex-Just-Passed-Bing-to-Become-4th-Largest-Global-Search-Engine>

BRIGHTPLANET, 2000. *CompletePlanet* [online]. BrightPlanet, c2000-2010 [cit. 2013-04-19]. Dostupné z: <http://aip.completeplanet.com/>

BRIN Sergey, Lawrence PAGE. *The Anatomy of a Large-Scale Hypertextual Web Search Engine*. 1998 [cit. 2013-05-05] Dostupné z: <http://infolab.stanford.edu/~backrub/google.html>

BRODER, A., 2002. A taxonomy of web search. *ACM SIGIR Forum* [online]. 2002-09-01, vol. 36, issue 2, s. 3-10 [cit. 2013-07-29]. DOI: 10.1145/792550.792552. Dostupné z: <http://portal.acm.org/citation.cfm?doid=792550.792552>

BROWSER WARS, 2001. In: *Wikipedia: the free encyclopedia* [online]. St. Petersburg (Florida): Wikimedia Foundation, 2001-, 18. 1. 2013 [cit. 2013-01-27]. Anglická verze. Dostupné z: http://en.wikipedia.org/wiki/Browser_wars

BURGHARDT, M., M. HECKNER a Ch. WOLFF, 2012. Chapter 2 The Many Ways of Searching the Web Together: A Comparison of Social Search Engines. LEWANDOWSKI, Dirk. *Web search engine research* [online]. Bingley: Emerald Group Publishing, 2012, s. 19 - 46 [cit. 2013-07-12]. Library and Information Science. ISBN 9781780526379.

BUSH, V., 1945. As we may think. *The Atlantic* [online]. 1945 [cit. 2013-01-13]. ISSN 1072-7825. Dostupné z: <http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>

CARNEY, M., 2013. Microsoft is emerging as the social search leader and Google should be worried. In: *PanDodaily* [online]. PandoDaily, 1. 7. 2013 [cit. 2013-08-04]. Dostupné z: <http://pandodaily.com/2013/07/01/microsoft-is-emerging-as-the-social-search-leader-and-google-should-be-worried/>

COMSCORE, 2013. ComScore Reports January 2013 U.S. Smartphone Subscriber Market Share. In: *ComScore* [online]. ComScore, 6. 3. 2013 [cit. 2013-08-06].

Dostupné z:

http://www.comscore.com/Insights/Press_Releases/2013/3/comScore_Reports_January_2013_U.S._Smartphone_Subscriber_Market_Share

CONTENT-BASED IMAGE RETRIEVAL, 2001. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2013-08-04].

Anglická verze. Dostupné z: http://en.wikipedia.org/wiki/Content-based_image_retrieval

CREATIVE COMMONS, 2013. *CC Search* [online]. Mountain View, CA: Creative Commons [cit. 2013-08-04]. Dostupné z: <http://search.creativecommons.org/>

CROOK, A., 2011. Adapting Search to You. In: *Bing Blogs: Search Blog* [online].

Microsoft, 2011 [cit. 2013-07-06]. Dostupné z:

http://www.bing.com/blogs/site_blogs/b/search/archive/2011/09/14/adapting-search-to-you.aspx

ČÍŽEK, J., 2013. Knowledge Graph vs. Bing Snapshots. Kdo je na tom lépe?. In:

Živě.cz [online]. Mladá fronta, 2. 7. 2013 [cit. 2013-07-30]. Dostupné z:

<http://www.zive.cz/clanky/knowledge-graph-vs-bing-snapshots-kdo-je-na-tom-lepe/sc-3-a-169535/default.aspx>

ČSN ISO 5127:2003. Praha: *Informace a dokumentace - slovník*. Praha: Český normalizační institut, 2003. 159, [1] s. Třídící znak 01 0162.

DATTA, Ritendra, Dhiraj JOSHI a James Z. WANG, 2008. Image Retrieval: Ideas, Influences, and Trends of the New Age. *ACM Computing Surveys*. 2008, roč. 40, č. 2, s. 1-60.

DAVIES, D., 2012. Major Search Engines and Directories. *Search Engine Watch*

[online]. 2012 [cit. 2013-03-28]. Dostupné z:

<http://searchenginewatch.com/article/2048976/Major-Search-Engines-and-Directories>

DOČEKAL, D., 2008. Centrum a Atlas mají společného vlastníka. In: *Lupa.cz: server o českém internetu* [online]. Lupa.cz, 27. 2. 2008 [cit. 2013-08-04]. Dostupné z:

<http://www.lupa.cz/clanky/centrum-a-atlas-maji-spolecneho-vlastnika/>

DOČEKAL, D., 2013. Google+ vs. Facebook? 6 minut vs. 6 hodin. In: *Lupa.cz: server o českém internetu* [online]. Lupa.cz, 11. 5. 2013 [cit. 2013-07-29]. ISSN 1213-0702.

Dostupné z: <http://www.lupa.cz/clanky/google-vs-facebook-6-minut-vs-6-hodin/>

DOGPILE, 2007. *Different Engines, Different Results : Web Searchers Not Always*

Finding What They're Looking for Online [online]. 2007 [cit. 2013-04-06]. Dostupný z:

<http://biblis2011.files.wordpress.com/2013/03/different-engines-searchs.pdf>

DOGPILE, 2013. Infospace, c2013 [cit. 2013-08-04]. Dostupné z:

<http://www.dogpile.com/>

DUCKDUCKGO, 2008. *DuckDuckGo* [online]. DuckDuckGo, [2008] [cit. 2013-03-29]. Dostupné z: <https://duckduckgo.com/>

DUCKDUCKGO, 2008a. !Bang. *DuckDuckGo* [online]. DuckDuckGo, [2008] [cit. 2013-03-29]. Dostupné z: <https://duckduckgo.com/bang.html>

- DUCKDUCKGO, 2008b. Privacy. *DuckDuckGo* [online]. DuckDuckGo, [2008] [cit. 2013-03-29]. Dostupné z: <https://duckduckgo.com/privacy>
- DUCKDUCKGO, 2013a. History. *Support Center* [online]. DuckDuckGo, 3. 6. 2013 [cit. 2013-03-29]. Dostupné z: <http://help.duckgo.com/customer/portal/articles/216406-history>
- DUCKDUCKGO, 2013b. Sources. *Support Center* [online]. 2013 [cit. 2013-03-29]. Dostupné z: <http://help.duckgo.com/customer/portal/articles/216399-sources>
- DUCKDUCKGO, 2013c. Syntax. *Support Center* [online]. 2013 [cit. 2013-03-29]. Dostupné z: <http://help.duckgo.com/customer/portal/articles/300304-syntax>
- EBIZMBA, 2013. Top 15 Most Popular Social Networking Sites: June 2013. In: *EBizMBA: The Business Knowledgebase* [online]. eBizMBA, 2013 [cit. 2013-07-07]. Dostupné z: <http://www.ebizmba.com/articles/social-networking-websites>
- EINET, 1994. *Einet.net* [online]. Chicago, Illinois: eiNET, 1994 [cit. 2013-01-14]. Dostupné z: <http://www.einet.net/>
- ELSEVIER, 2004. *Scirus White Pages: How Scirus Works* [online]. Amsterdam: Elsevier, 2004 [cit. 2013-04-20]. Dostupný z: http://www.scirus.com/press/pdf/WhitePaper_Scirus.pdf
- ELSEVIER, 2013. About Scirus. *Scirus: for scientific information only* [online]. Amsterdam: Elsevier, c2013 [cit. 2013-04-20]. Dostupné z: <http://www.scirus.com/srsapp/aboutus/>
- FARAGO, P., 2013. China Knocks Off U.S. to Become World's Top Smart Device Market. In: *Flurry Blog* [online]. San Francisco: Flurry, 2013 [cit. 2013-06-27]. Dostupné z: <http://blog.flurry.com/bid/94352/China-Knocks-Off-U-S-to-Become-World-s-Top-Smart-Device-Market>
- FOLEY, M. J., 2007. Microsoft severs Live Search from the rest of the Windows Live family. *ZDNet* [online]. 2007 [cit. 2013-03-17]. Dostupné z: <http://www.zdnet.com/blog/microsoft/microsoft-severs-live-search-from-the-rest-of-the-windows-live-family/339>
- GEORGIA INSTITUTE OF TECHNOLOGY, 1994. *GVU's First WWW User Survey Results* [online]. Atlanta: Georgia Institute of Technology, 1994 [cit. 2013-01-20]. Dostupné z: http://www.cc.gatech.edu/gvu/user_surveys/survey-01-1994/
- GOODREADS, 2013. *Goodreads* [online]. Goodreads, c2013 [cit. 2013-08-05]. Dostupné z: <http://www.goodreads.com/>
- GOODWIN, D., 2012. Google Users Dislike Personalized Search Results. *Search Engine Watch* [online]. 10. 2. 2012 [cit. 2013-06-20]. Dostupné z: <http://searchenginewatch.com/article/2145297/Google-Users-Dislike-Personalized-Search-Results-Survey>
- GOOGLE CHINA, 2001. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001-, 26. 7. 2013 [cit. 2013-08-02]. Anglická verze. Dostupné z: http://en.wikipedia.org/wiki/Google_China

GOOGLE IMAGE LABELER, 2001. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2013-08-04]. Anglická verze. Dostupné z: http://en.wikipedia.org/wiki/Google_Image_Labeler

GOOGLE IMAGES, 2001. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2013-08-04]. Dostupné z: http://en.wikipedia.org/wiki/Google_Images

GOOGLE, 2001. *Google* [online]. Google, [2001] [cit. 2013-08-03]. Dostupné z: <http://www.google.com/>

GOOGLE, 2001a. Podrobná historie společnosti. *Google* [online]. Mountain View, CA: Google, [2001] [cit. 2013-05-20]. Dostupné z: <http://www.google.com/about/company/history/>

GOOGLE, 2001b. Desatero, kterým se řídíme. *Google* [online]. Mountain View, CA: Google, [2001] [cit. 2013-05-20]. Dostupné z: <http://www.google.com/about/company/philosophy/>

GOOGLE, 2001c. *Vše o vyhledávání* [online]. Google, [2001] [cit. 2013-08-03]. Dostupné z: <http://www.google.com/insidesearch/>

GOOGLE, 2001d. *Google Disk* [online]. Google, [2001] [cit. 2013-08-03]. Dostupné z: <http://www.google.com/intl/cs/drive/about.html>

GOOGLE, 2001e. Google Books History. *Google Books* [online]. Google, [2001] [cit. 2013-08-03]. Dostupné z: <http://www.google.com/googlebooks/about/history.html>

GOOGLE, 2001f. O překladači Google. *Google Překladač* [online]. Mountain View, CA: Google [cit. 2013-07-30]. Dostupné z: http://translate.google.com/about/intl/cs_ALL/

GOOGLE, 2001g. Google Zeitgeist. *Google* [online]. Mountain View, CA: Google, 2001 [cit. 2013-05-17]. Dostupné z: <http://www.google.com/intl/en/zeitgeist/>

GOOGLE, 2001h. Místa: objevte a navštivte skvělé místní podniky. *Google+ Funkce* [online]. Google, [2001] [cit. 2013-08-04]. Dostupné z: <http://www.google.com/+learnmore/local/>

GOOGLE, 2011a. Překladač Google – často kladené dotazy. *Google* [online]. Mountain View, CA: Google, c2011 [cit. 2013-06-10]. Dostupné z: http://www.google.cz/intl/cs/help/faq_translation.html#statmt

GOOGLE, 2011b. Google Goggles. *Google Mobile* [online]. Google, c2011 [cit. 2013-08-04]. Dostupné z: <http://www.google.com/mobile/goggles/>

GOOGLE, 2013a. What's happening to iGoogle?. In: *Google Inside Search* [online]. Google, 28. 5. 2013 [cit. 2013-07-14]. Dostupné z: <https://support.google.com/websearch/answer/2664197?hl=en>

GOOGLE, 2013b. Vyhledávací nástroje a filtry. *Web Search: nápověda* [online]. Mountain View, CA: Google, c2013 [cit. 2013-05-20]. Dostupné z: https://support.google.com/websearch/answer/142143?hl=cs&ref_topic=3081620

- GOOGLE, 2013c. Vyhledávací operátory. *Web Search: nápověda* [online]. Mountain View, CA: Google, c2013 [cit. 2013-05-20]. Dostupné z https://support.google.com/websearch/answer/136861?hl=cs&ref_topic=3036305
- GOOGLE, 2013d. Automatické doplňování. *Web Search: nápověda* [online]. Mountain View, CA: Google, 19. 4. 2013 [cit. 2013-05-21]. Dostupné z https://support.google.com/websearch/answer/106230?hl=cs&ref_topic=3037079
- GOOGLE, 2013e. Úplný přehled stránky s výsledky. *Web Search: nápověda* [online]. Mountain View, CA: Google, 24. 4. 2013 [cit. 2013-05-21]. Dostupné z: https://support.google.com/websearch/answer/35891?hl=cs&ref_topic=3036305
- GOOGLE, 2013f. *Google Trendy* [online]. Google, c2013 [cit. 2013-08-02]. Dostupné z: <http://www.google.cz/trends/>
- GOOGLE, 2013g. *Art Project* [online]. Google, 2013 [cit. 2013-08-02]. Dostupné z: <http://www.google.com/culturalinstitute/project/art-project>
- GOOGLE, 2013h. *Google Glass* [online]. Google [cit. 2013-08-05]. Dostupné z: <http://www.google.com/glass/>
- GOPHER (PROTOCOL), 2001. In: *Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida): Wikimedia Foundation, 2001- , 17. 8. 2012 [cit. 2012-08-19]. Anglická verze. Dostupné z: http://en.wikipedia.org/wiki/Gopher_%28protocol%29
- GRIFFIN, Scott, 2000. Douglas Engelbart. *Internet Pioneers* [online]. Scott Griffin, [2000] [cit. 2013-01-27]. Dostupné z: <http://www.ibiblio.org/pioneers/engelbart.html>
- HASALÍK, Radim, 2007. Proč prohrála Altavista souboj s Googlem. *Lupa.cz* [online]. 21. 12. 2007 [cit. 2013-01-26]. ISSN 1213-0702. Dostupné z: <http://www.lupa.cz/clanky/proc-prohrala-altavista-souboj-s-googlem/>
- HAVLOVÁ, J., 2003. Software s otevřeným zdrojovým kódem. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR, 2003- [cit. 2013-08-03]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000013968&local_base=KTD
- HOROWITZ, B, 2011. A Fall Sweep. In: *Google Official Blog* [online]. Mountain View, CA: Google, 14. 10. 2011 [cit. 2013-05-25]. Dostupné z: <http://googleblog.blogspot.cz/2011/10/fall-sweep.html>
- CHUNG, EunKyung a JungWon YOON, 2012. Analysis of multimedia needs and searching features: An exploratory study. *Proceedings of the American Society for Information Science and Technology* [online]. 2012, vol. 49, issue 1, s. 1-5 [cit. 2013-07-14]. DOI: 10.1002/meet.14504901297. Dostupné z: <http://doi.wiley.com/10.1002/meet.14504901297>
- IBOOGIE, 2013. Austin: CyberTavernTV [cit. 2013-08-04]. Dostupné z: <http://iboogie.com/>
- IDÉE, 2013a. Frequently Asked Questions. *TinEye* [online]. Idée, c2013 [cit. 2013-08-04]. Dostupné z: <http://www.tineye.com/faq#how>

IDÉE, 2013b. MulticolorEngine. *TinEye* [online]. Idée, c2013 [cit. 2013-08-04]. Dostupné z: <http://services.tineye.com/MulticolorEngine>

IMAGE RETRIEVAL, 2001. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2013-08-04]. Anglická verze. Dostupné z: http://en.wikipedia.org/wiki/Image_retrieval

INTERNET ARCHIVE, 2001. *Internet Archive: universal access to all knowledge* [online]. San Francisco: Internet Archive, 2001 [cit. 2013-06-06]. Dostupné z: <http://archive.org/>

INTERNET ARCHIVE, 2001a. About the Internet Archive.. *Internet Archive* [online]. San Francisco: Internet Archive, 2001 [cit. 2013-06-06]. Dostupné z: <http://archive.org/about/>

INTERNET ARCHIVE, 2001b. Removing Documents From the Wayback Machine. *Internet Archive* [online]. San Francisco: Internet Archive, 2001 [cit. 2013-06-06]. Dostupné z: <http://archive.org/about/exclude.php>

IPL2, 1995. *Ipl2: information you can trust* [online]. Drexel University's College of Information Science & Technology, c1995- [cit. 2013-08-03]. Dostupné z: <http://www.ipl.org/>

JING, Yushi a kol., 2012. Google image swirl. *Proceedings of the 21st international conference companion on World Wide Web - WWW '12 Companion on World Wide Web* [online]. New York, New York, USA: ACM Press, 2012, s. 539-540 [cit. 2013-08-04]. DOI: 10.1145/2187980.2188116. ISBN 978-1-4503-1230-1. Dostupné z: <http://dl.acm.org/citation.cfm?doid=2187980.2188116>

JOHNSON, N., 2009. It's Official: Microsoft and Yahoo! Finally Strike Search Deal. *Search Engine Watch* [online]. 29. 7. 2009 [cit. 2013-05-03]. Dostupné z: <http://searchenginewatch.com/article/2052024/Its-Official-Microsoft-and-Yahoo-Finally-Strike-Search-Deal>

JYXO, 2013. Informace. *Jyxo* [online]. CET 21, Stránka ze dne 23. 1. 2012 dostupná ze služby Internet Archive. [cit. 2013-08-04]. Dostupné z: <http://jyxo.vybereme.cz/d/info>

KINCAID, J., 2010. Google: 25% Of Queries From Android 2.0 Devices Use Voice Search. *TechCrunch* [online]. AOL, 12. 8. 2010 [cit. 2013-07-03]. Dostupné z: <http://techcrunch.com/2010/08/12/googles-hugo-barra-25-of-android-queries-are-voice-based/>

KRATOCHVÍLOVÁ, Z., 2012. Jak nalézt zdroje uložené ve WebArchivu. In: *WebArchiv* [online]. Praha: Národní knihovna ČR, 2012 [cit. 2013-06-06]. Dostupné z: http://blog.webarchiv.cz/2012/03/jak-nalezet-zdroje-ulozene-ve-webarchivu_06.html

KTD, 2003. *KTD : Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online databáze]. Praha: Národní knihovna České republiky, 2003- [cit. 2013-02-23]. Dostupné z WWW: <http://aleph.nkp.cz/cze/ktd>

LEVENE, M, 2010. *An introduction to search engines and web navigation* [online]. Hoboken: Wiley, 2010 [cit. 2012-11-22]. ISBN 978-0-470-87423-3. Dostupné z: <http://onlinelibrary.wiley.com/book/10.1002/9780470874233>

LEXXE, 2012. Technology. *Lexxe beta* [online]. Lexxe, c2012 [cit. 2013-07-04]. Dostupné z: <http://www.lexxe.com/technology.html>

LIST OF MERGERS AND ACQUISITIONS BY GOOGLE, 2001. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001-, 29. 7. 2013 [cit. 2013-07-29]. Dostupné z: http://en.wikipedia.org/wiki/List_of_mergers_and_acquisitions_by_Google

LYCOS, 2013. About Lycos. *Lycos* [online]. Waltham: Lycos, c2013 [cit. 2013-01-26]. Dostupné z: <http://info.lycos.com/about/company-overview>

MAKULOVÁ, S., 2002. *Vyhľadavanie informácií v internete: problémy, východiská, postupy*. Bratislava: EL&T, 2002. 376 s. ISBN 80-88812-16-X.

MAKULOVÁ, S., 2003. Okienko do internetu. Kedy používame predmetové adresáre webu? *Knižnica*. 2003 [cit. 2013-01-14]. Roč. 4, č. 8. Dostupné z: http://www.snk.sk/swift_data/source/casopis_kniznica/2003/august/380.pdf

MANNING, Christopher D., Prabhakar RAGHAVAN a Hinrich SCHÜTZE, 2008. *Introduction to information retrieval*. New York: Cambridge University Press, 2008. xxi, 482 s. ISBN 978-0-521-86571-5.

MEEKER, M. a L. WU, 2013. Internet Trends D11 Conference. In: *KPCB* [online prezentace]. KPCB, 29. 5. 2013 [cit. 2013-08-06]. Dostupné z: <http://www.kpcb.com/insights/2013-internet-trends>

METACRAWLER, 2013. *Metacrawler: search the search engines* [online]. Infospace, c2013 [cit. 2013-08-04]. Dostupné z: <http://www.metacrawler.com/>

MICROSOFT, 2013. *Bing* [online]. Microsoft, c2013 [cit. 2013-08-04]. Dostupné z: <http://www.bing.com/>

MICROSOFT, 2013a. *Bing it on: Bing vs Google* [online]. Microsoft, c2013 [cit. 2013-08-04]. Dostupné z: <http://www.bingiton.com/>

MICROSOFT, 2013b. Search effectively. *Bing Help Home* [online]. Microsoft, c2013 [cit. 2013-04-02]. Dostupné z: <http://onlinehelp.microsoft.com/en-us/bing/ff524480.aspx>

MICROSOFT, 2013c. Bing Explore. *Bing* [online]. Microsoft, c2013 [cit. 2013-08-04]. Dostupné z: www.bing.com/explore/home

MICROSOFT, 2013d. Bing Explore Social. *Bing* [online]. Microsoft, c2013 Dostupné z: <http://www.bing.com/explore/social?sf=flt19:bing51socpreview>

MICROSOFT, 2013e. Bing Privacy Statement. *Microsoft* [online]. Microsoft, c2013, July 2013 [cit. 2013-08-05]. Dostupné z: <http://www.microsoft.com/privacystatement/en-us/bing/default.aspx#>

MONSTER, 2013. New York: Monster Worldwide, c2013 [cit. 2013-03-20]. Dostupné z: <http://www.monster.com/>

- MOSAIC, 1996. Mosaic -- The First Global Web Browser. In: *Livinginternet.com* [online]. William Stewart, c1996-2011 [cit. 2013-01-27]. Dostupné z: http://www.livinginternet.com/w/wi_mosaic.htm
- NÁRODNÍ KNIHOVNA ČR, 2001. *WebArchiv: archiv českého webu* [online]. Praha: Národní knihovna ČR, 2001, 6. 6. 2013 [cit. 2013-06-06]. Dostupné z: <http://www.webarchiv.cz/>
- NETMARKETSHARE, 2013. *Netmarketshare: Market Share Statistics for Internet Technologies* [online]. Aliso Viejo: Net Applications, [cit. 2013-01-26]. Dostupné z: <http://www.netmarketshare.com/>
- NETSCAPE, 2012a. Becoming an ODP Editor. *Open Directory Project: DMOZ* [online]. Netscape, c2012, 20. 1. 2013 [cit. 2013-02-07]. Dostupné z: <http://www.dmoz.org/docs/en/help/become.html>
- NETSCAPE, 2012b. Open Directory Editing Guidelines. *Open Directory Project: DMOZ* [online]. Netscape, c2012, 1. 9. 2012 [cit. 2013-02-07]. Dostupné z: <http://www.dmoz.org/docs/en/guidelines/>
- NETSCAPE, 2013. *Open Directory Project: DMOZ* [online]. Netscape, c2013 [cit. 2013-02-07]. Dostupné z: <http://www.dmoz.org/>
- NIERSTRASZ, Oscar, 1996. *W3 Catalog History* [online]. 1996 [cit. 2013-01-14]. Dostupné z: <http://scg.unibe.ch/archive/software/w3catalog/>
- ONGE, R. S., 2012. Internet Archive Reaches 10 Petabytes Of Saved Data. In: *TheTechJournal* [online]. Dhaka: TheTechJournal.com, 2012 [cit. 2013-06-06]. Dostupné z: <http://thetechjournal.com/internet/internet-archive-reaches-10-petabytes-of-saved-data.xhtml>
- ORWANT, J., 2012. Improving Google Patents with European Patent Office patents and the Prior Art Finder. In: *Research Blog: The latest news from Research at Google* [online]. Mountain View, CA: Google, 2012 [cit. 2013-05-24]. Dostupné z: <http://googleresearch.blogspot.de/2012/08/improving-google-patents-with-european.html>
- PANDIA, 2001. Alternative Academic Search Directories. *Pandia* [online]. Oslo: P&S Koch, 2001 [cit. 2013-02-07]. Dostupné z: <http://www.pandia.com/sw-2002/16-directories.html>
- PAPÍK, R., 2011. *Strategie vyhledávání informací a elektronické informační zdroje*. Praha: Velryba, 2011. 192 s. ISBN 978-80-85860-22-1.
- PETERKA, J., 1996a. Yahoo!. In: *EArchiv.cz: archiv článků a přednášek Jiřího Peterky* [online]. Jiří Peterka, 1996 [cit. 2013-02-07]. Dostupné z: <http://www.earchiv.cz/a96/a609k140.php3>
- PETERKA, J., 1996b. Seznam, nebo „České Yahoo“?. In: *EArchiv.cz: archiv článků a přednášek Jiřího Peterky* [online]. Jiří Peterka, 1996 [cit. 2013-02-07]. Dostupné z: <http://www.earchiv.cz/a96/a623k140.php3>

- PROJECT XANADU, 2001. In: *Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida): Wikimedia Foundation, 2001- , 19.12.2012 [cit. 2012-01-14]. Anglická verze. Dostupné z: http://en.wikipedia.org/wiki/Project_Xanadu
- PURCELL, K., J. BRENNER a J. RAINIE, 2012. *Search Engine Use 2012* [online]. Washington, D.C.: Pew Research Center, 2012. 42 s. Dostupné z: <http://pewinternet.org/Reports/2012/Search-Engine-Use-2012.aspx>
- QUIAN, R., 2013. Understand Your World with Bing. In: *Bing Blogs: Search Blog* [online]. Microsoft, 21. 3. 2013 [cit. 2013-07-04]. Dostupné z: http://www.bing.com/blogs/site_blogs/b/search/archive/2013/03/21/satorii.aspx
- QUORA, 2013. Quora [cit. 2013-08-04]. Dostupné z: <https://www.quora.com/>
- RALPH, Randy D., 1996. AltaVista Tutorial. In: *Netstrider* [online]. Clemmons: Netstrider, 1996 [cit. 2013-01-26]. Dostupné z: <http://www.netstrider.com/search/altavista.html>
- RESEARCHGATE.NET. *ResearchGate* [online]. Researchgate.net, c2013 [cit. 2013-08-05]. Dostupné z: <http://www.researchgate.net/>
- RFC 114, 1971. *A File Transfer Protocol* [online]. A. K. BHUSHAN. April 1971 [cit. 2012-08-19]. Dostupné z: <http://www.rfc-editor.org/rfc/rfc114.txt>
- RÖHLE, T., 2007. Desperately seeking the consumer: Personalized search engines and the commercial exploitation of user data. *First Monday* [online]. 2007, roč. 12, č. 9 [cit. 2013-07-06]. ISSN 1995-2013. Dostupné z: http://firstmonday.org/issues/issue12_9/rohle/index.html
- RUNET, 2001. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2013-08-04]. Dostupné z: <https://en.wikipedia.org/wiki/Runet>
- SEYMOUR, Tom, Dean FRANTSVOG a Kumar SATHEESH, 2011. History Of Search Engines. *International journal of management & Information Systems* [online]. 2011, vol. 15, no. 4 [cit. 2013-01-14]. ISSN 2157-9628. Dostupné z: <http://journals.cluteonline.com/index.php/IJMIS/article/view/5799>
- SEZNAM.CZ, 1996. *Seznam.cz* [online]. c1996-2013 [cit. 2013-08-03]. Dostupné z: <http://www.seznam.cz/>
- SEZNAM.CZ, 1996a. Rok 1996. In: *Seznam.cz: o nás* [online]. Praha: Seznam.cz, c1996-2013 [cit. 2013-01-30]. Dostupné z: <http://onas.seznam.cz/cz/o-firme/historie-firmy/1996/>
- SEZNAM.CZ, 1996b. Co je to odkaz. In: *Seznam.cz: nápověda* [online]. 1996. Praha: Seznam.cz, c1996-2013 [cit. 2013-01-30]. Dostupné z: <http://napoveda.seznam.cz/cz/odkazy/nejcastejsi-otazky-k-odkazum-1/co-je-to-odkaz/>
- SEZNAM.CZ, 1996c. Závazná pravidla pro odkazy. In: *Seznam.cz: nápověda* [online]. Praha: Seznam.cz, c1996-2013 [cit. 2013-01-30]. Dostupné z: <http://napoveda.seznam.cz/cz/odkazy/pravidla-katalogu-1/zavazna-pravidla-pro-odkazy/>

- SEZNAM.CZ, 1996d. Jak přidám odkaz? In: *Seznam.cz: nápověda* [online]. Praha: Seznam.cz, c1996-2013 [cit. 2013-01-30]. Dostupné z: <http://napoveda.seznam.cz/cz/odkazy/nejcastejsi-otazky-k-odkazum-1/jak-pridam-odkaz/>
- SEZNAM.CZ, 1996e. Rok 2005. In: *Seznam.cz: o nás* [online]. Praha: Seznam.cz, c1996-2013 [cit. 2013-01-30]. Dostupné z: <http://onas.seznam.cz/cz/o-firme/historie-firmy/2005/>
- SEZNAM.CZ, 1996f. Rok 2001. In: *Seznam.cz: o nás* [online]. Praha: Seznam.cz, c1996-2013 [cit. 2013-01-31]. Dostupné z: <http://onas.seznam.cz/cz/o-firme/historie-firmy/2001/>
- SEZNAM.CZ, 1996g. Firmy.cz. In: *Seznam.cz: o nás* [online]. Praha: Seznam.cz, c1996-2013 [cit. 2013-01-31]. Dostupné z: <http://onas.seznam.cz/cz/firmy-cz.html>
- SEZNAM.CZ, 1996h. Firemní zápisy. In: *Seznam.cz: nápověda* [online]. Praha: Seznam.cz, c1996-2013 [cit. 2013-01-30]. Dostupné z: <http://napoveda.seznam.cz/cz/firmy/placene-produkty/firemni-zapisy/>
- SEZNAM.CZ, 1996i. Rok 2007. In: *Seznam.cz: o nás* [online]. 1996. Praha: Seznam.cz, c1996-2013 [cit. 2013-01-25]. Dostupné z: <http://onas.seznam.cz/cz/o-firme/historie-firmy/2007/>
- SEZNAM.CZ, 1996j. Rok 2009. In: *Seznam.cz: o nás* [online]. Praha: Seznam.cz, c1996-2013 [cit. 2013-01-25]. Dostupné z: <http://onas.seznam.cz/cz/o-firme/historie-firmy/2009/>
- SEZNAM.CZ, 1996k. Pokročilé vyhledávání. In: *Seznam.cz: nápověda* [online]. 1996j. Praha: Seznam.cz, c1996-2013 [cit. 2013-04-27]. Dostupné z: <http://napoveda.seznam.cz/cz/fulltext-hledani-v-internetu/pokrocile-hledani/#operatory>
- SEZNAM.CZ, 1996l. Miniaplikace. In: *Seznam.cz: nápověda* [online]. Praha: Seznam.cz, c1996-2013 [cit. 2013-04-27]. Dostupné z: <http://napoveda.seznam.cz/cz/fulltext-hledani-v-internetu/hledani-fulltext-miniaplikace/>
- SEZNAM.CZ, 2012. Seznam testuje vyhľadavanie v cudzojazyčných výsledkoch. In *Seznam.cz: Blog fulltextového týmu* [online]. Praha: Seznam.cz, 2012 [cit. 2013-04-26]. Dostupné z: <http://fulltext.sblog.cz/2012/04/03/seznam-testuje-vyhľadavanie-v-cudzojazycných-vysledkoch/>
- SHAZAM ENTERTAINMENT, 2002. *Shazam* [online]. Shazam Entertainment, c2002-2013 [cit. 2013-08-04]. Dostupné z: <http://www.shazam.com/>
- SHERMAN, 2001. Google Restores Usenet Archive. *Search Engine Watch*. 8. 5. 2001 [cit. 2013-01-15]. Dostupné z: <http://searchenginewatch.com/article/2065667/Google-Restores-Usenet-Archive>
- SHERMAN, Ch., 2000. Humans do it better. *Online*. 2000, roč. 24, č. 4, s. 43-48. ISSN 01465422.

SHERMAN, Ch., 2002. Happy Birthday Lycos! *Search Engine Watch*. 2002 [cit. 2013-01-15]. Dostupné z: <http://searchenginewatch.com/article/2064656/Happy-Birthday-Lycos>

SCHONFELD, E., 2010. Monster Buys HotJobs from Yahoo for \$225 Million. *TechCrunch* [online]. AOL, 2010 [cit. 2013-03-20]. Dostupné z: <http://techcrunch.com/2010/02/03/yahoo-hotjobs-monster-225-million/>

SILVERSTEIN, Craig a kol., 1999. Analysis of a very large web search engine query log. *ACM SIGIR Forum* [online]. 1999-09-01, vol. 33, issue 1, s. 6-12 [cit. 2013-01-25]. DOI: 10.1145/331403.331405. Dostupné z: <http://portal.acm.org/citation.cfm?doid=331403.331405>

SINGHAL, A., 2012a. Search, plus Your World. In: *Google Official Blog* [online]. Google, 10. 1. 2012 [cit. 2013-08-05]. Dostupné z: <http://googleblog.blogspot.cz/2012/01/search-plus-your-world.html>

SINGHAL, A., 2012b. Introducing the Knowledge Graph: things, not strings. In: *Google Official Blog* [online]. Google, 16. 5. 2012. Dostupný z: <http://googleblog.blogspot.co.uk/2012/05/introducing-knowledge-graph-things-not.html>

SIRI, 2001. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2013-07-06]. Anglická verze. Dostupné z: <http://en.wikipedia.org/wiki/Siri>

SKLENÁK, Vilém a kol., 2001. *Data, informace, znalosti a Internet*. Praha: C.H. Beck, 2001. xvii, 507 s. ISBN 80-7179-409-0.

SKLENÁK, Vilém a Ludmila CELBOVÁ, 2003a. Internet. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR, 2003- [cit. 2013-08-01]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000536&local_base=KTD

SKLENÁK, Vilém, 2003b. Hypertext. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR, 2003- [cit. 2013-08-01]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000617&local_base=KTD

SKLENÁK, Vilém, 2003c. Vyhledávací nástroj. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR, 2003- [cit. 2013-02-24]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000677&local_base=KTD

SKLENÁK, Vilém, 2003d. Katalogový vyhledávací nástroj. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR, 2003- [cit. 2013-02-24]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000647&local_base=KTD

SKLENÁK, Vilém, 2003e. Vyhledávací stroj. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR, 2003- [cit. 2013-02-24]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000678&local_base=KTD

SKLENÁK, Vilém, 2003f. Metavyhledávací stroj. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR, 2003- [cit. 2013-02-24]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000628&local_base=KTD

SKLENÁK, Vilém, 2003g. Stemming. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR, 2003- [cit. 2013-07-29]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000664&local_base=KTD

SMITH, R., 2010a. GOOGLE MEANS EVERY. *Research Technology Management*. 2010, roč. 53, č. 1, s. 67-69.

SMITH, N., 2010b. Ooh! Ahh! Google Images presents a nicer way to surf the visual web. In: *Google Official Blog* [online]. Google, 20. 7. 2010 [cit. 2013-06-29]. Dostupné z: <http://googleblog.blogspot.cz/2010/07/ooh-ahh-google-images-presents-nicer.html>

SOUNDHOUND, 2013. *SoundHound: Instant Music Search and Discovery* [online]. SoundHound, c2013 [cit. 2013-08-04]. Dostupné z: <http://www.soundhound.com/>

SPARROW, B., J. LIU a D. M. WEGNER, 2011. Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips. *Science* [online]. 2011, vol. 333, issue 6043, s. 776-778 [cit. 2013-05-02]. DOI: 10.1126/science.1207745. Dostupné z: <http://www.sciencemag.org/content/333/6043/776>

SPEECH RECOGNITION, 2001. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2013-08-04]. Dostupné z: http://en.wikipedia.org/wiki/Speech_recognition

ST. JEAN, Beth a kol., 2012. An analysis of the information behaviors, goals, and intentions of frequent Internet users: Findings from online activity diaries. *First Monday* [online]. 2012, roč. 17, č. 2 [cit. 2013-07-12]. Dostupné z: <http://firstmonday.org/ojs/index.php/fm/article/view/3870/3143>

STATCOUNTER, 1999. Mobile vs. Desktop from Jan 2012 to June 2013. *StatCounter: GlobalStats* [online]. StatCounter, c1999-2013 [cit. 2013-08-06]. Dostupné z: http://gs.statcounter.com/#mobile_vs_desktop-ww-monthly-201201-201306

STERLING, G., 2010. EU Steps Up The Pressure: Pursuing Formal Anti-Trust Investigation Against Google. In: *Search Engine Land* [online]. Third Door Media, 30. 11. 2010 [cit. 2013-08-02]. Dostupné z: <http://searchengineland.com/eu-steps-up-the-pressure-pursuing-formal-anti-trust-investigation-against-google-57299>

SULLIVAN, D., 1997. AltaVista Debuts Translation Service. *Search Engine Watch*. 31. 12. 1997 [cit. 2013-01-25]. Dostupné z: <http://searchenginewatch.com/article/2064828/AltaVista-Debuts-Translation-Service>

SULLIVAN, D., 1998. NewHoo: Yahoo Built By The Masses. *Search Engine Watch* [online]. 30. 6. 1998 [cit. 2013-02-03]. Dostupné z: <http://searchenginewatch.com/article/2068086/NewHoo-Yahoo-Built-By-The-Masses>

SULLIVAN, D., 2003. Where Are They Now? Search Engines We've Known & Loved. *Search Engine Watch* [online]. 2003 [cit. 2013-03-17]. Dostupné z: <http://searchenginewatch.com/article/2064954/Where-Are-They-Now-Search-Engines-Weve-Known-Loved>

SULLIVAN, D., 2008. Google Paper: Better Image Search Though VisualRank / Image Rank. In: *Search Engine Land* [online]. Third Door Media, 28. 4. 2008 [cit. 2013-08-04]. Dostupné z: <http://searchengineland.com/google-paper-better-image-search-though-visualrank-image-rank-13872>

SULLIVAN, D., 2009. Google Now Personalizes Everyone's Search Results. *Search Engine Land* [online]. Third Door Media, 4. 12. 2009 [cit. 2013-07-06]. Dostupné z: <http://searchengineland.com/google-now-personalizes-everyones-search-results-31195>

SULLIVAN, D., 2012. Google's Results Get More Personal With "Search Plus Your World". In: *Search Engine Land* [online]. Third Door Media, 10. 1. 2012 [cit. 2013-08-05]. Dostupné z: <http://searchengineland.com/googles-results-get-more-personal-with-search-plus-your-world-107285>

ŠVEJDA, Jan, 2003. Vyhledávání informací. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR, 2003- [cit. 2013-02-24]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000001825&local_base=KTD

TAYLOR, B., 2010. Bringing Your Friends to Bing: Search Now More Social. In: *The Facebook Blog* [online]. Facebook, 13. 10. 2010 [cit. 2013-07-09]. Dostupné z: <https://blog.facebook.com/blog.php?post=437112312130>

THE RUNET, 2012. Yandex presents new search algorithm. In: *The Runet: From Russia with web* [online]. The Runet, 2012 [cit. 2013-07-07]. Dostupné z: <http://eng.therunet.com/en/news/549/>

THE WWW VIRTUAL LIBRARY, 1991. *The WWW Virtual Library* [online]. 1991 [cit. 2013-01-14]. Dostupné z: <http://vlib.org/>

TOPIX, 2001. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2013-08-05]. Dostupné z: <http://en.wikipedia.org/wiki/Topix.net>

USENET, 2001. In: *Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida): Wikimedia Foundation, 2001- , 11. 1. 2013 [cit. 2013-01-15]. Anglická verze. Dostupné z: <http://en.wikipedia.org/wiki/Usenet>

VETTER, R. J., C. SPELL a C. WARD, 1994. Mosaic and the World Wide Web. *Computer* [online]. 1994, vol. 27, issue 10, s. 49-57 [cit. 2013-08-02]. DOI: 10.1109/2.318591. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=318591>

VOCŮ, O., 2012. Vyhledávání hudbou a jeho vazby na portály se sdíleným videem. *Knihovna* [online]. 2012, roč. 23, č. 1, s. 63-83 [cit. 2013-06-19]. ISSN 1801-3252. Dostupné z: http://knihovna.nkp.cz/knihovna121/12_63.htm

- VYLEŤAL, M., 2012. Rusové dál vzdorují Googlu, tamní jednička Yandex zvažuje i expanzi do Evropy. *Lupa.cz* [online]. 16. 10. 2012 [cit. 2013-03-10]. ISSN 1213-0702. Dostupné z: <http://www.lupa.cz/clanky/rusove-dal-vzdoruji-googlu-tamni-jednicka-yandex-zvazuje-i-expanzi-do-evropy/>
- W3COUNTER, 2012. Global Web Stats: December 2012. *W3Counter* [online]. Lansdale: Awio Web Services LLC, 2012 [cit. 2013-01-26]. Dostupné z: <http://www.w3counter.com/globalstats.php?year=2012&month=12>
- W3COUNTER, 2013. Global Web Stats: April 2013. *W3Counter* [online]. Lansdale: Awio Web Services LLC, 2013 [cit. 2013-05-24]. Dostupné z: <http://www.w3counter.com/globalstats.php?year=2013&month=4>
- WEBCERTAIN, 2012. *The Webcertain Global Search & Social Report* [online]. Webcertain, 2012 [cit. 2013-02-21]. Dostupný z: <http://globalcentral.net/assets/cb757434/Search-Social-2012-Done.pdf>
- WEBCRAWLER, 2013. About WebCrawler. *WebCrawler* [online]. Infospace, c2013 [cit. 2013-01-15]. Dostupné z: <http://www.webcrawler.com/support/aboutus>
- WEE, W., 2013. China's Mobile Internet Users Surpass Desktop Users. In: *Tech in Asia* [online]. Singapore: Tech in Asia, c2013 [cit. 2013-06-27]. Dostupné z: <http://www.techinasia.com/china-mobile-internet-users-statistics-behavior/>
- WELCH, Ch., 2013. Google Glass gets full-on web browser, new hands-free functions in monthly update. In: *The Verge* [online]. Vox Media, 1. 7. 2013 [cit. 2013-07-06]. Dostupné z: <http://www.theverge.com/2013/7/1/4483772/google-glass-gets-web-browser-new-hands-free-functions-software-update>
- WHITEHEAD, J., 2000. As We Do Write : Hyper-terms for Hypertext. *ACM SIGWEB Newsletter* [online]. 2000-06-01, vol. 9, 2-3, s. 8-18 [cit. 2013-01-14]. Dostupné z: <http://portal.acm.org/citation.cfm?doid=763988.763990>
- WHITTEN, A., 2012 Updating our privacy policies and terms of service. In: *Google Official Blog* [online]. Google, 24. 1. 2012 [cit. 2013-08-04]. Dostupné z: <http://googleblog.blogspot.cz/2012/01/updating-our-privacy-policies-and-terms.html>
- WOLFRAM ALPHA, 2001. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001-, 27. 5. 2013 [cit. 2013-07-30]. Anglická verze. Dostupné z: http://en.wikipedia.org/wiki/Wolfram_Alpha
- WOLFRAM ALPHA, 2013. *WolframAlpha: computational knowledge engine* [online]. Wolfram Alpha - A Wolfram Research Company, c2013 [cit. 2013-08-04]. Dostupné z: <http://www.wolframalpha.com/>
- WOLFRAM ALPHA, 2013a. Examples by Topic. *WolframAlpha: computational knowledge engine* [online]. Wolfram Alpha - A Wolfram Research Company, c2013 [cit. 2013-08-04]. Dostupné z: <http://www.wolframalpha.com/examples/>
- WORLD WIDE WEB CONSORTIUM, 1995. A Little History of the World Wide Web. *W3C* [online]. World Wide Web Consortium, 1995, 2000 [cit. 2013-02-05]. Dostupné z: <http://www.w3.org/History.html>

WORLD WIDE WEB FOUNDATION, 2012. The Web Index [online]. World Wide Web Foundation, c2012 [cit. 2013-02-19]. Dostupné z: <http://thewebindex.org/>

WORLDWIDEWEBSIZE.COM, 2013. Maurice de Kunder, [2013] [cit. 2013-02-21]. Dostupné z <http://www.worldwidewebsize.com/>

YAHOO! SEARCH, 2001. Yahoo! Search. In: *Wikipedia: the free encyclopedia* [online]. St. Petersburg (Florida): Wikimedia Foundation, 2001-, 1. 3. 2013 [cit. 2013-04-28]. Anglická verze. Dostupné z: https://en.wikipedia.org/wiki/Yahoo!_Search

YAHOO!, 2013. *Altavista* [online]. Yahoo!, c2013 [cit. 2013-01-25]. Dostupné z: <http://www.altavista.com/>

YAHOO!, 2013. *Yahoo!* [online]. Yahoo!, c2013 [cit. 2013-08-04]. Dostupné z: <http://www.yahoo.com/>

YAHOO!, 2003a. *Yahoo! Directory* [online]. Yahoo!, c2003 [cit. 2013-02-07]. Dostupné z: <http://dir.yahoo.com/>

YANDEX, 1997. *Janděks* [online]. Yandex, c1997-2013 [cit. 2013-03-09]. Dostupné z: <http://www.yandex.ru/>

YANDEX, 1997a. About. *Yandex Company* [online]. Moskva: Yandex, c1997-2013 [cit. 2013-03-09]. Dostupné z: http://company.yandex.com/general_info/yandex_today.xml

YANDEX, 1997b. MatrixNet: New Level of Search Quality. *Yandex Company* [online]. Moskva: Yandex, c1997-2013 [cit. 2013-03-09]. Dostupné z: <http://company.yandex.com/technologies/matrixnet.xml>

YANDEX, 1997c. Bazovye vozmozhnosti. *Janděks : pomošč'* [online]. Yandex, c1997-2013 [cit. 2013-03-10]. Dostupné z: <http://help.yandex.ru/search/>

YANDEX, 1997d. Rasširenyj poisk. *Janděks* [online]. Yandex, c1997-2013 [cit. 2013-03-10]. Dostupné z: <http://yandex.ru/search/advanced>

YANDEX, 1997e. Moi nachodki → pomošč'. *Janděks : pomošč'* [online]. Yandex, c1997-2013 [cit. 2013-03-10]. Dostupné z: <http://help.yandex.ru/nahodki/>

YANDEX, 1997f. Personal'nye otvety. *Janděks : pomošč'* [online]. Yandex, c1997-2013 [cit. 2013-03-10]. Dostupné z: <http://help.yandex.ru/search/?id=1125352>

YANDEX, 2013. It May Get Really Personal – we have rolled out our second-generation personalised search program. In: *Yandex company* [online]. Yandex, 30. 5. 2013 [cit. 2013-08-04]. Dostupné z: http://company.yandex.com/press_center/blog/entry.xml?pid=20

YOUNG, R, 2011. Google Directory Has Been Shut Down. *Search Engine Watch*. 25. 11. 2011 [cit. 2013-01-20]. Dostupné z: <http://searchenginewatch.com/article/2096661/Google-Directory-Has-Been-Shut-Down>

YOUTUBE, 2013. Statistics. Youtube [online]. YouTube, c2013 [cit. 2013-08-04].
Dostupné z: <http://www.youtube.com/yt/press/statistics.html>

YOVISTO.COM, 2006. *Yovisto: academic video search* [online]. yovisto.com, c2006-
2013 [cit. 2013-08-04]. Dostupné z: <http://www.yovisto.com/>

Seznam obrázků

Obrázek č. 1: První webový prohlížeč WorldWideWeb Tima Bernerse Lee [BERNERS-LEE, 2013]

Obrázek č. 2: Vyhledávač AltaVista v roce 1999 (dostupné prostřednictvím služby Internet Archive)

Obrázek č. 3: Hlavní tematické kategorie katalogu Yahoo! Directory

Obrázek č. 4: Kategorie „Cestopisy“ v českém jazyce katalogu ODP

Obrázek č. 5: Kategorie „Zpravodajství“ v katalogu Seznam.cz v roce 1996 (dostupné prostřednictvím služby Internet Archive)

Obrázek č. 6: Architektura vyhledávacích strojů [SKLENÁK, 2001, s. 261]

Obrázek č. 7: Grafické znázornění populárních vyhledávačů a sociálních sítí podle regionu v průzkumu společnosti WebCertain [WebCertain, 2012]

Obrázek č. 8: Ukázka služby Google trendy

Obrázek č. 9: Ukázka informací s odkazem pro vyhledávání na tapetě domovské stránky vyhledávače Bing

Obrázek č. 10: Ukázka výsledků vyhledávání pomocí Bing včetně služby Bing Social

Obrázek č. 11: Ukázka výsledků vyhledávání ve vyhledávači DuckDuckGo

Obrázek č. 12: Ukázka výsledku výpočtu ve službě WolframAlpha

Obrázek č. 13: Ukázka domovské stránky portálu Yandex

Obrázek č. 14: Ukázka výsledků vyhledávání pomocí služby Seznam.cz

Obrázek č. 15: Ukázka služby Multicolor Search Lab pro vyhledání obrázků podle barev

Obrázek č. 16: Ukázka zpracování a zobrazení výsledku vyhledávání v aplikaci Shazam

Obrázek č. 17: Ukázka vyhledávání pomocí vyhledávače Google se zapnutou službou Search Plus Your World

Obrázek č. 18: Ukázka vyhledávání s funkcí Diagram znalostí