

Vyhledávání na internetu

ve volně dostupných zdrojích

Internet obsahuje zhruba 5 mil. TB dat, Google (vyhledávač s největším indexem) zprostředkovává jen 0,004% celkové datové velikosti internetu. Jak se v takovém množství dat zorientovat a jak najít relevantní informaci? Jaký je rozdíl mezi vyhledávacím strojem a předmětovým katalogem, k čemu slouží metavyhledávače a co si představit pod pojmem neviditelný web? Jakou roli ve světě informací na internetu hrají sociální média a co je to real time vyhledávání? Nejen to se dozvíte po absolvování následujícího kurzu.

„Každé 3 dny vzniká zhruba stejné množství dat jako od počátku civilizace do roku 2003.“ (Jeff Jonas)

Tip na skvělou infografiku *How The Internet Works* <http://www.makeuseof.com/tag/internet-works-infographic/>



JORGE CHAM ©THE STANFORD DAILY

[HTTP://PHD.STANFORD.EDU/](http://phd.stanford.edu/)

Obsah:

1. Historie internetu a jeho struktura
2. Vyhledávací služby na internetu
3. Neviditelný web
4. Real time vyhledávání
5. Sémantické vyhledávání
6. Použité zdroje

1. Historie internetu a jeho struktura



V roce 1991 byly spuštěny první webové stránky na adrese <http://info.cern.ch>. Jsou stále funkční. V [následující prezentaci](#) se dozvíte, jak vznikl internet, na jakém principu je postaven, jak je definován, co je to WWW, Hypertext či jazyk HTML, k čemu slouží protokol HTTP a z čeho se skládá URL.

Web 2.0 × Web 3.0

Web 2.0 je ustálené označení vývojové etapy internetu. O tomto fenoménu se hovoří od roku 2004, i když termín „Web 2.0“ vznikl už několik let předtím. V současné době se už ale hovoří o Webu 3.0, i když jeho definice ještě není zcela jasná. Mnozí tvrdí, že stojíme na pomezí těchto dvou etap, jiní jsou přesvědčeni, že Web 3.0 je stále ještě hudbou budoucnosti.

Podle časové osy Nova Spivacka se právě teď nacházíme ve fázi Webu 3.0, která je označována též jako „sémantický web“. Dle jeho předpovědí se kolem roku 2020 web dostane do fáze 4.0. Na Spivackovu časovou osu se můžete podívat [ZDE](#). Jeho blog je dostupný na adrese <http://www.novaspivack.com/>.

Etapy vývoje webu (doplňkové odkazy směřují převážně na Wikipedii):

PC era: 1980-1990 (WINDOWS, MAC OS, [SGML](#), [BBS](#))

Web 1.0: 1990-2000 ([HTTP](#), [HTML](#), [JAVA](#), [XML](#), [FLASH](#))

Web 2.0: 2000-2010 ([OPEN ID](#), [ATOM](#), [RSS](#), [AJAX](#), [OWL](#), [SWRL](#))

Web 3.0: 2010-2020 ([SEMANTICKÝ WEB](#), [CLOUD COMPUTING](#))

Web 4.0: 2020-2030 (UMĚLÁ INTELIGENCE)

WEB 2.0	WEB 3.0
<ul style="list-style-type: none">▸ uživatelé se podílejí na obsahu či tvorbě různých webových projektů▸ komunitní servery, systémy pro sdílení, blogy, RSS agregátory▸ tagování obsahu (folksonomie)▸ ajaxové technologie, sdílené API rozhraní a „mashup projekty“▸ aplikace pro online tvorbu a správu dokumentů▸ linkovací systémy, slovníkové projekty (Wikipedie)	<ul style="list-style-type: none">▸ relativně malé aplikace, které mohou běžet na libovolném zařízení (PC, mobilní telefon, tablet...), rychlé, snadno modifikovatelné▸ sdílené aplikace, personalizace služeb▸ cloud computing▸ rozostření hranic profesionál / poloprofesionál / amatér (kdokoliv je schopný vytvořit program)▸ mikroformáty pro sémantický web▸ dotazování v přirozeném jazyce▸ narůst videoobsahu, 3D technologie

2. Vyhledávací služby na internetu

Z hlediska dostupnosti zdrojů rozlišujeme:

1. **informační zdroje dostupné přímo** (informace veřejného charakteru, poskytovány zpravidla bezplatně, vyhledatelné skrze různé vyhledávací služby)
2. **informační zdroje dostupné nepřímo** (např. profesionální a komerční databáze)

Vyhledávací služby dělíme na 2 základní typy:

1. **vyhledávací stroje** (indexují slova a termíny ve WWW dokumentech) – indexují dokumenty pomocí automatických programů, které jednotlivé stránky ukládají do své databáze, při vyhledávání dochází k tzv. „fulltextovému vyhledávání“, neboť stroje prohlíží stránku kompletně
2. **předmětové katalogy** (klasifikují dokumenty nebo celé servery podle předem dané předmětové klasifikace) – katalogy bývají budovány ručně a stránky jsou posuzovány jako celek, na rozdíl od vyhledávacího stroje lze v katalogu nejen vyhledávat, ale také prohlížet stránky zařazené v hierarchii

Mnohé služby nabízí jak katalogový přístup, tak i vyhledávání ve fulltextu. Jedna varianta ale bývá vždy preferována.

Příkladem je Seznam.cz, který je především předmětovým katalogem (dostupné i speciální databáze: Zboží.cz, Firmy.cz), ale nabízí i fulltextové vyhledávání.

Vyhledávací stroje

- **System, který na základě zadaného klíčového slova hledá v databázi nebo indexu a uživateli pak nabídne výsledek hledání.**
- Stěžejní pro použití vyhledávacího stroje je **klíčové slovo = vyhledávací termín = vyhledávací výraz**. (Správně zvolený vyhledávací výraz je základ úspěchu.)
- Z jednoho nebo více vyhledávacích výrazů se skládá **dotaz** (úplný vyhledávací požadavek, vyhledávací výrazy se spojují operátory).
- Dokument, který vyhovuje zadanému dotazu se jmenuje **hit**.

Kdy použijeme vyhledávací stroj?

1. Předmět zájmu je úzký (konkrétní) nebo obsahuje neobvyklé termíny
2. Hledáme zvláštní sídlo (konkrétní webovou stránku)
3. Chceme prohledávat maximální možnou část webu (Pozor! Včetně nerelevantních dokumentů.)
4. Chceme, aby byl výsledek vyhledávání co nejobsáhlejší

5. Chceme hledat určité typy dokumentů, souborů, jazyků, chceme nastavit upřesňující filtry
6. Chceme využít výhody těchto strojů jako je pojmové shlukování, řazení dokumentů dle popularity

Nevýhody vyhledávacích strojů

- Databáze vyhledávacích strojů vytváří roboti (=automaticky), proto je indexováno i mnoho stránek nevalné kvality.
- Roboti nejsou schopni vyhledat a indexovat vše (problém neviditelného webu), nečitelné jsou pro ně např. dynamicky se měnící stránky, obrázkové mapy apod.
- Některé stránky jsou chráněny heslem.
- Mnoho vyhledávacích strojů má omezení na počet indexovaných stránek z určité domény.



„Když něco vyhledáváte na Googlu, tak se Google ptá Chucka Norrise.“

„Google nehledá pouze Chucka Norrise. Každý totiž ví, že Chucka Norrise není třeba hledat, Chuck Norris si najde vás!“

Jak fungují vyhledávací stroje?

Znáte pojem „**Google bombs**“? Jedná se o jistý způsob zneužití vlastností vyhledávače Google, po kterém je výsledkem vyhledání nějaké fráze webová stránka, která danou frází neobsahuje, ale útočník na ni tak odkazuje, obvykle s humorným nebo politicko-kritickým úmyslem. Příklady různých „Google bomb“ najdete třeba na [Wikipedii](#).

Jak fungují vyhledávací stroje? – [Prezentace v Prezi](#)

Příklady vyhledávacích strojů

Zahraniční vyhledávače:



[Google](#) je nejpoužívanější světový internetový vyhledávač, který nabízí i české rozhraní. Jedná se o vyhledávač s největším počtem indexovaných stránek. Jeho popularita je tak vysoká, že se ustálilo slovo googlovat pro vyjádření vyhledávání na internetu a roku 2006 se toto sloveso objevilo v Oxford English Dictionary.



[Bing](#) je druhým největším vyhledávačem na světě a provozuje ho společnost Microsoft. Spuštěn byl v roce 2009 jako náhrada za Live Search.



[Blekkko](#) se snaží prorazit se zcela novým konceptem vyhledávání a coby „názorový vyhledávač“ prohledává web

jen z požadovaného úhlu pohledu. Vyhledávací dotaz tak vypadá např. takto: „politics /conservative“ (hledané slovo je politics, úhel pohledu = slashtag je conservative).

České vyhledávače:



[Morfeo](#) je textový vyhledávač založený na Open Source technologii Sherlock Holmes (Sherlock Holmes Search Engine), provozovaný na portálu Centrum.cz. Umožňuje hledání slov odvozených a synonym hledaných výrazů, opravuje překlepy. Prohledává pouze stránky v doméně .cz.



[Jyxo](#) nabízí české skloňování a časování, kontrolu pravopisu a třídění výsledků vyhledávání do skupin. Pro hodnocení stránek používá tzv. JyxoRank. [1.cz](#) používá vyhledávací technologii Morfeo založenou na systému Sherlock Holmes.

Metodika vytváření a upravování webových stránek takovým způsobem, aby jejich forma a obsah byly vhodné pro automatizované zpracování v internetových vyhledávačích, se nazývá **SEO**.

Předmětové katalogy

- Pokrývají menší část webového prostoru.
- Výhodou je **přesná hierarchická struktura a dělení na kategorie a podkategorie**.
- Předpokladem vyhledávání v katalogu je, že uživatel ví, co hledá a do které předmětové kategorie téma patří.
- Katalogy plní jednak tvůrci webových stránek a jednak informační pracovníci.
- Umožňuje postupovat kategoriemi a podkategoriemi od obecných věcí k těm specifickým.

Kdy použijeme předmětový katalog?

1. V případě, že naše téma je široké
2. Nemáme konkrétní představu o hledané informaci a chceme se nechat navést nabízenými podkategoriemi
3. Chceme získat seznam webových sídel
4. Hledáme informace o firmě, vyhledáváme ve výrobcích, hledáme nejnovější zprávy
5. Chceme se vyhnout nekvalitním stránkám, které indexují roboti vyhledávacích strojů

Výhody a nevýhody předmětových katalogů

- Největší výhodou je záruka kvality, která je dána zejména tím, že jsou katalogy vytvářeny „ručně“. Na druhou stranu mají omezený rozsah.
- Většina katalogů navíc informační zdroje hodnotí a anotuje.
- Používají se různé struktury kategorií a jiná schémata
- Procházení podkategorií je časově náročné
- Problém s platností odkazů, méně častá aktualizace katalogu
- Subjektivita při hodnocení zdrojů a jejich zařazování do kategorií
- Obecné popisy vznikající na základě zevrubného prozkoumání informačního zdroje

Příklady předmětových katalogů

Zahraniční předmětové katalogy:



[Yahoo!](#) byl původně pouze katalog s hierarchicky seřazenými odkazy, postupně se však stal neúspěšnějším a nejnavštěvovanějším internetovým portálem nabízející stovky nejrůznějších služeb. Ve snaze konkurovat Googlu se později vytvořil i fulltextový vyhledávač. Velmi populární službou je [Yahoo! answers](#).



[Google Directory](#) umožňuje vyhledávat stránky podle kategorií. V tomto projektu Google spolupracuje s katalogem Open Directory Project, který je, podobně jako Wikipedie, administrován komunitou dobrovolníků z řad uživatelů internetu.



[DMOZ](#) je mezinárodní katalog, na kterém pracují dobrovolníci z řad převážně profesionálů v dané oblasti, což zajišťuje velmi vysokou úroveň zatříděných odkazů. DMOZ je neziskový projekt. Tím se odlišuje od většiny katalogů a staví ho do pozice významného zdroje informací, který je neúplatný. Dostupná je také sekce [World: Česky](#).



[Galaxy](#) je jeden z nejstarších předmětových katalogů na internetu. Byl vyvinut firmou LOGIKA Corporation.

České předmětové katalogy:



[Seznam](#) je první český vyhledávač, byl založen v roce 1996. Kromě katalogu a doplňkových služeb nabízí také fulltextové vyhledávání. Je orientovaný převážně na český a slovenský internet. Vyhledávač podporuje také skloňování, časování a

technologii sitemap. Vedle HTML formátu indexuje i XML, PDF a DOC.



[Centrum](#) je další významný český katalog. V roce 2008 společnost Centrum.cz Holding BV koupilo portál Atlas.cz. Velmi populární je zejména jeho zpravodajský server [Aktuálně.cz](#). [Atlas](#) nabídl jako první fulltextové vyhledávání v ČR. Centrum i Atlas se nyní více zaměřují na zajímavé služby.

Metavyhledávací stroje

- Umožňují vyhledávání ve více než jednom vyhledávacím stroji nebo katalogu.
- Při vyhledávání kombinují výsledky vyhledávání a odstraňují duplicitní záznamy.
- Mohou být také seznamem vyhledávacích strojů, do kterých lze vstoupit z jednoho místa.
- Problém: každý vyhledávač má jiná pravidla vyhledávání, takže metavyhledávače umožňují efektivně vyhledávat pouze pomocí klíčových slov.
- Za zmínku stojí také tzv. „**vyhledávací centrály**“, které do jednoho místa sdružují vyhledávání ve vámi vybraných vyhledávacích

Výhody	Nevýhody
<ul style="list-style-type: none">▫ Při vyhledávání používáte pouze jedinou stránku▫ Práce pouze s jedním rozhraním▫ Není nutné sledovat novinky mezi vyhledávači a nové vyhledávací stroje▫ Dotaz zadáváte více vyhledávačům pouze jednou▫ Vyšší úplnost vyhledávání	<ul style="list-style-type: none">▫ Odtržení od rozhraní jednotlivých vyhledávačů▫ Přicházíme o informace reklamního charakteru▫ Někdy dochází k nízké výkonnosti metavyhledávačů▫ Metavyhledávače limitují počet záznamů z jednoho vyhledávacího stroje

Jak fungují metavyhledávače?



ROZESÍLACÍ MECHANISMUS – Algoritmus (dotazovací přístup), který určuje výběr vyhledávacích systémů, kterým bude dotaz zaslán.



AGENT ROZHRANÍ – Program, který ví, jak se navazuje spojení s vyhledávacím systémem. Umí konvertovat dotaz do formátu různých vyhledávacích systémů a interpretuje získané výsledky.



ZOBRAZOVACÍ MECHANISMUS – Odstraňuje duplicitní výsledky, řadí je dle míry relevance, provádí konverzi z interního formátu do vhodné podoby a výsledky zašle k zobrazení klientovi.

Příklady metavyhledávacích strojů



[Ask](#), dříve známý jako Ask Jeeves, je jeden z nejrozšířenějších metavyhledávačů. Základní myšlenkou Ask bylo poskytnout uživatelům odpovědi na jejich každodenní otázky položené v přirozeném jazyce. K dispozici je nyní nová beta verze, kde odpovědi na otázky nebudou jen strojově dohledávány, ale svou úlohu zde sehraje také komunita.



[Dogpile](#) je metavyhledávací stroj, který sbírá výsledky z Googlu, Yahoo!, Bing, Ask, Answers a z dalších populárních vyhledávačů včetně těch, které jsou zaměřeny na audio a video obsah.



[Metacrawler](#) sbírá výsledky z Googlu, Yahoo!, Bing, Ask, About, MIVA, LookSmart a dalších vyhledávačů.



[Search](#) nabízí kromě vyhledávání pomocí klíčových slov ve full textu také vlastní [předmětový katalog](#). Kritéria vyhledávání a seznam vyhledávacích strojů, ze kterých chcete zaslat dotaz, lze samozřejmě upravit v nastavení. V nabídce je také tzv. WebFerret software, vyhledávací systém pro Windows (volně ke stažení).



[Infodump](#) je metavyhledávač, který kromě tradičního fulltextového vyhledávání nabízí také vyhledávání obrázků, videí, novinek nebo práce. Napojený je také na Google Maps, Yahoo Maps, Mapquest a Live Maps.



[WebCrawler](#) nabízí nejrelevantnější výsledky z vyhledávačů Google, Yahoo!, Bing Search, Ask.com, About.com, MIVA, LookSmart atd. WebCrawler také poskytuje uživatelům možnost vyhledávat obrázky, audio, video, novinky, žluté stránky a bílé stránky. Patří k předním světovým metavyhledávačům.



Zajímavým metavyhledávačem je [PolyCola](#). Známy je zejména díky svému bývalému názvu **GahooYoogle.com**. Přimo na titulní straně PolyColy najdete možnost výběru kombinace vyhledávačů, přičemž přednastavenými vyhledávači jsou samozřejmě Yahoo! a Google.

Portály – brány do světa internetu?

Webové sídlo, které poskytuje širokou škálu služeb a informací, často s možností jejich přizpůsobení uživateli podle osobních potřeb a zájmů. Nabídka zahrnuje především vyhledávání webových informačních zdrojů a dokumentů, freemailovou službu, denní zpravodajství, mapy, vyhledávání kontaktních informací na osoby i firmy, zábavu (hry, horoskopy, vtipy apod.), elektronické obchody, bezplatný prostor pro publikování webových stránek apod. Naprostá většina nabízených služeb je pro koncového uživatele bezplatná díky ziskům z reklamy. Široce (obecně) zaměřené portály se označují jako horizontální portály. ([TDKIV](#))

Postupem času se začaly rovněž objevovat a rozvíjet **cíleněji profilované oborové portály**, někdy též zvané **vortály** (vertikální portály), které již neměly ambice pokrýt svým záběrem celou škálu lidských činností, ale zaměřily se jen na určitou tematickou oblast. Jejich návštěvnost je řádově nižší než u velkých portálů a rovněž nabídka služeb je chudší. Ovšem pro osoby, které daná problematika zajímá, představují oborové portály často velice cenný zdroj informací. Typickým případem jsou portály specializované na nákup zboží, nemovitostí nebo například na vyhledávání nabídek práce. Tyto oborové portály bývají často začleněny do struktury služeb výše zmiňovaných portálů – vyhledávačů.

Příklady portálů



[Yahoo!](#) byl původně pouze katalog s hierarchicky seřazenými odkazy, postupně se však stal nejúspěšnějším a nejnavštěvovanějším internetovým portálem nabízející stovky nejrůznějších služeb. Ve snaze konkurovat Googlu se později vytvořil i fulltextový vyhledávač. Velmi populární službou je [Yahoo! answers](#).



[Seznam](#) je první český vyhledávač, byl založen v roce 1996. Kromě katalogu a doplňkových služeb nabízí také fulltextové vyhledávání. Je orientovaný převážně na český a slovenský internet. Seznam je nejnavštěvovanější český portál.

Vyhledávání obrázků, hudby a videa

Jistě nemá valný smysl zdůrazňovat, že nejjednodušší vyhledávání obrázků realizujete např. přes [Google Images](#) a spousta videí je na [YouTube](#) nebo [Vimeu](#). O dostupnosti hudby ani nemluvě. V této kapitole se zaměříme spíše na díla (obrazová, auditivní i audiovizuální), která jsou vystavena pod licencí **Creative Commons** a lze je často opětovně používat a šířit, dokonce někdy i upravovat a užívat komerčně.



Licence Creative Commons fungují na jednoduchém principu: autor jejich prostřednictvím plošně uzavírá se všemi potencionálními uživateli díla smlouvu, na základě které jim poskytuje některá svá práva k dílu a jiná si vyhrazuje. Creative Commons nejsou popřením klasického pojetí copyrightu, ale jeho nadstavbou. Vycházejí z autorského zákona, který je upravuje jako licenční smlouvy (§ 46 -55 AutZ). Obliba licencí Creative Commons vychází především z jejich mezinárodní srozumitelnosti.

Licenční podmínky, neboli práva a povinnosti uživatele k dílu, jsou web: <http://www.creativecommons.cz/>



Díla pod licencí CC můžete hledat přes rozhraní na webu search.creativecommons.org nebo i např. přes pokročilé vyhledávání obrázků na Googlu.

Tagy

V prostředí internetu se často setkáte s tzv. **tagy**. Označujete jimi svá videa na YouTube, fotky na Flickru, prezentace na Slideshare, odkazy na Delicious apod. **Jedná se o slova nebo slovní spojení, která volně tvoří uživatel a která slouží k identifikaci obsahu.** Použití tagů a jejich množství závisí výhradně na uživateli.

Tags

oliver • žabák • knihovna utb • u13 •
univerzitní centrum • zlín • exkurze

ADD

[Choose from your tags](#)

Separate each tag with a space: *cameraphone urban moblog*. Or to join 2 words together in one tag, use double quotes: *"daily commute"*.

Ukázka tagů na Flickr.com

3. Neviditelný web



Neviditelný web (často označovaný jako „invisible web“ nebo „deep web“) obsahuje dokumenty na internetu, které jsou obtížně vyhledatelné. Opakem je viditelný web (též „visible web“ nebo „surface web“), který označuje běžně dostupné stránky. Přesnou velikost viditelného a neviditelného webu nelze určit. Odhaduje se, že neviditelný web je až 550x větší než web viditelný. Např. Google indexuje pouze 0,004% webu.

Důvody neviditelnosti

- odpojené stránky – robot nemůže sledovat spojení na stránku
- stránka obsahuje hlavně obrázky, video, audio – nedostatek textu, aby robot porozuměl obsahu
- stránka obsahuje hlavně soubory pdf, postscript, flash apod. – stránka se většinou neindexuje z ekonomických důvodů
- obsah v relačních databázích – roboty nedokážou vyplnit požadovaná pole v interaktivních formulářích
- obsah se mění v reálném čase – obrovské množství dat, které se neustále mění
- dynamicky vytvářený obsah – personalizovaný obsah není zajímavý pro většinu uživatelů
- soukromé weby a placený či heslem zabezpečený přístup

Vyhledávače pro neviditelný web



[CompletePlanet](#) umožňuje vyhledávání ve více než 100 000 databázích a specializovaných vyhledávacích nástrojích, u záznamů je uvedena míra relevance, popularita, spojení a kategorie, ve které se nachází v adresáři.



[Turbo 10](#) zpřístupňuje informace z více než 800 databází. Lze přizpůsobovat výsledky vyhledávání, zobrazují se se screen shotem.



[Scirus](#) obsahuje 410 mil. dokumentů odborné povahy. Najdete skrze něj preprinty, patenty, prohledává institucionální repozitáře.

4. Real time vyhledávání

Real time vyhledávání je dnes velice aktuální téma. Téměř každý velký vyhledávač implementuje nebo se chystá přidat výsledky real time vyhledávání, nebo-li posty, které jsou staré několik minut nebo sekund. Nejrozsáhlejším zdrojem takových informací je bezesporu Twitter a Facebook. Na Twitteru nám práci usnadňují tzv. [hashtagy](#), fotky lze vyhledávat např. na [TwitPic](#) a existuje také vyhledávač českého a slovesného twitteru [Klábosení.cz](#). Vyhledávání v reálném čase dnes nabízí také [Google](#). O čem lidé mluví právě teď, co je nejdiskutovanější a co se děje? Díky real time vyhledávání už máme možnost to zjistit. Ostatně Twitter je mnohdy rychlejší než CNN.

Real Time Search (YouTube video): <http://youtu.be/WRkYmx4A9Do>

5. Sémantické vyhledávání



„Se sémantickým webem je to podobné jako s Yettim či Járou Cimrmanem – víme toho o nich mnohé, ale ve skutečnosti je nikdo neviděl.“
(Michal Černý)

„Find What I Mean, Not What I Type.“ (Chris Sherman)

Co lze říct o sémantickém vyhledávání

- ▷ záležitost webu 3.0
- ▷ počítače mají větší pochopení pro vazbu mezi informacemi, nejde jen o vyhledávání odkazů podle zadaných klíčových slov
- ▷ dokumenty obsahují metainformace, se kterými umí vyhledávače pracovat
- ▷ vyhledávání zaměřené na získání odpovědí, ne výsledků vyhledávání
- ▷ sémantický vyhledávač neodkazuje na pravděpodobně relevantní dokument, ale přímo odpovídá na zadanou otázku
- ▷ používá sémantickou síť k mapování významů a vztahů slov, budování přesných sémantických sítí je časově a finančně náročné
- ▷ počítá s pokládáním dotazu v přirozeném jazyce
- ▷ se sémantickým webem lze snadno manipulovat, zkreslovat skutečnost, nebo dokonce lhát: „Téměř učebnicovým příkladem toho, že ne vždy se vyplatí věřit znalostním webům, je případ týkající se Wikipedie. Student Shane Fitzgerald na ni přiřkl (záměrně zcela nepravdivě) výrok „Někdo by mohl říci, že můj život byl jedním dlouhým soundtrackem“ francouzskému skladateli jménem Maurice Jarre. V jeho nekrologích u celé řady novin či webů se tento výrok objevoval a u některých článků dokonce tvořil nadpis. Na mystifikaci musel až po nějaké době upozornit sám student.“ (Lupa.cz)



Jedním z nejznámějších vyhledávačů, kteří používají sémantické vyhledávání je [WolframAlpha](#). Umí docela dobře počítat a derivovat, nicméně s otázkami se na něj musí chytře. Chvilí trvá, než se s ním uživatel seznámí a je nutné mu pokládat dotaz správně. Je vždy lepší postavit dotaz bez zbytečných slov. Velkou výhodou je i vizualizace výsledků. Vyzkoušejte např. jednoduchý dotaz: „Where Barack Obama was born?“ a uvidíte, co vše se dozvíte. Nebo zkuste počítat: $(A^{\sin(e^x)})$.

6. Použité zdroje



1. SPIVACK, Nova. How The WebOS Evolves?. *Nova Spivack Typepad Blog* [online]. February 09, 2007 [cit. 2011-01-18]. Dostupné z: http://novaspivack.typepad.com/nova_spivacks_weblog/2007/02/steps_towards_a.html.
2. TĚŠÍK, Miroslav. Web 2.0, 3.0, 4.0 ... In: *Symbio: tisková agentura* [online]. 12. 11. 2007 [cit. 2011-01-19]. Dostupné z: <http://www.symbio.cz/clanky/web-2-0-3-0-4-0.html>.
3. Slovník internetových výrazů: Web 2.0. In: *Symbio: tisková agentura* [online]. ©1999-2011 [cit. 2011-01-19]. Dostupné z: <http://www.symbio.cz/slovník/web-2-0.html>.
4. Slovník internetových výrazů: AJAX. In: *Symbio: tisková agentura* [online]. © 1999-2011 [cit. 2011-01-19]. Dostupné z: <http://www.symbio.cz/slovník/ajax.html>.
5. Slovník internetových výrazů: Folksonomie. In: *Symbio: tisková agentura* [online]. ©1999-2011 [cit. 2011-01-19]. Dostupné z: <http://www.symbio.cz/slovník/folksonomie.html>.
6. *Microformats.cz* [online]. copyright neuveden [cit. 2011-01-19]. Dostupné z: <http://microformats.cz/>.
7. *Cloud computing* [online]. © 2008-2010 [cit. 2011-01-19]. Dostupné z: <http://www.cloudcomputing.cz/>.
8. VÁCLAVEK, Petr. Přichází Web 3.0, těšte se a bojte se. In: *Lemonade* [online]. 18. únor 2010 [cit. 2011-01-19]. Dostupné z: <http://www.lemonadeprague.com/2010/02/p%C5%99ich%C3%A1z%C3%AD-web-3-0-t%C4%9B%C5%A1te-se-a-bojte-se.html>.
9. HLAVENKA, Jiří. *Vyhledávání na Internetu*. 2. vydání. Brno: Computer Press, 2004. ISBN 80-722-6759-0.
10. ČERVÁNKOVÁ, Hana. Nový vyhledávač Blekko je tu. Změní pravidla hry?. In: *Tyinternety.cz* [online]. 02.11.2010 [cit. 2011-01-19]. Dostupné z: <http://www.tyinternety.cz/sluzby/novy-vyhledavac-blekko-je-tu-zmeni-pravidla-hry-1991>.
11. Search Engine Optimization. In: *Wikipedia: the free encyclopedia* [online]. St. Petersburg (Florida): Wikipedia Foundation, 17. 8. 2005, last modified on 3. 10. 2010 [cit. 2011-01-19]. Dostupné z: http://cs.wikipedia.org/wiki/Search_Engine_Optimization.
12. KAPOUN, Jan. Historie Yahoo!. In: *Business World* [online]. 25.03.10 [cit. 2011-01-20]. Dostupné z: <http://businessworld.cz/ostatni/historie-yahoo-5842>.
13. Google: Google Directory. In: *Wikipedia: the free encyclopedia* [online]. St. Petersburg (Florida): Wikipedia Foundation, 17.9.2004, last modified on 16.1.2011 [cit. 2011-01-20]. Dostupné z: http://cs.wikipedia.org/wiki/Google#Google_Directory.
14. HELLER, Petr. *PetrHeller.info: Vítejte v moderním světě informací* [online]. 13. srpna 2005 [cit. 2011-01-20]. DMOZ – fenomén mezi katalogy. Dostupné z: <http://webzine.petrheller.info/?article=dmoz-fenomen-mezi-katalogy>.
15. MACICH, Jiří. Vyhledávač Ask.com hlásí návrat ke kořenům. In: *Lupa.cz* [online]. 29. 7. 2010 [cit. 2011-01-27]. Dostupné z:

<http://www.lupa.cz/zpravicky/vyhledavac-ask-com-hlasi-navrat-ke-korenum/>.

16. Internetový portál. In: *Wikipedia: the free encyclopedia* [online]. St. Petersburg (Florida): Wikipedia Foundation, 2008-01-15, last modified on 2011-01-08 [cit. 2011-03-07]. Dostupné z: http://cs.wikipedia.org/wiki/Internetov%C3%BD_port%C3%A1l.
17. ŠUBRT, Filip a Martin KRČÁL. *Vyhledávání na internetu* [online]. 2010-04-19 [cit. 2011-03-28]. Dostupné z: https://is.muni.cz/auth/el/1421/jaro2010/KPI11/um/12708882/opora_vyhledavani_na_internetu.pdf?fakulta=1421;obdobi=4705;kod=KPI11.
18. BOLDIŠ, Petr. Invisible web. *Bibliografické citace dokumentů podle ISO 690 a ISO 690-2* [online]. 2001-2006 [cit. 2011-03-28]. Dostupné z: <http://www.boldis.cz/index.php?iweb>.
19. Neviditelný web. *Infogram: portál pro podporu informační gramotnosti* [online]. ©2011 [cit. 2011-05-16]. Dostupné z: <http://www.infogram.cz/article.do?articleId=1765>.
20. ČERNÝ, Michal. Mýtus sémantického webu nebo znalostní vyhledávač Wolfram Alpha?. In: *Lupa.cz* [online] 18. 5. 2009 6:25 [cit. 2011-05-17]. Dostupné z: <http://www.lupa.cz/clanky/mytus-semantickeho-webu/>.
21. IKAROS, redakce. Jak lépe vyhledávat. *Ikaros* [online]. Roč. 14, č. 5/2 (2010) [cit. 2011-05-17]. Dostupné z: <http://www.ikaros.cz/najdi-to-co-mam-na-mysli-nikoliv-to-co-jsem-zadal-chris-sherman>.

Seznam obrázků



1. CHAM, Jorge. Addicted to the web. *Pilled Higher & Deeper: a grad student comic strip* [online]. 2010-05-08 [cit. 2011-01-06]. Obrázek ve formátu GIF. Dostupné z: <http://www.phdcomics.com/comics/archive.php?comid=185>.
2. SIMIEZZZ. *Iconscollection – Mail* [online]. 2007-09-02. Attribution-NoDerivs 2.0 Generic (CC BY-ND 2.0) [cit. 2011-01-27]. Obrázek ve formátu JPEG. Dostupné z: <http://www.flickr.com/photos/simiezzz/1301796545/>.
3. SIMIEZZZ. *Iconscollection – Mail* [online]. 2007-08-11. Attribution-NoDerivs 2.0 Generic (CC BY-ND 2.0) [cit. 2011-01-27]. Obrázek ve formátu JPEG. Dostupné z: <http://www.flickr.com/photos/simiezzz/1081911996/>.
4. SIMIEZZZ. *Iconscollection – Wait a minute* [online]. 2007-08-11. Attribution-NoDerivs 2.0 Generic (CC BY-ND 2.0) [cit. 2011-01-27]. Obrázek ve formátu JPEG. Dostupné z: <http://www.flickr.com/photos/simiezzz/1081057895/>.



Textová opora pro kurz v rámci portálu IVA – informační výchova na UTB ve Zlíně. <http://www.iva.knihovna.utb.cz>