



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Univerzita Karlova v Praze Ústav informační studií a knihovnictví

Modul č. 1

Informace a Internet

Milena Medková

Určeno pro projekt:

Název: Studium informační vědy a znalostního managementu v evropském kontextu

Reg. číslo: CZ.1.07/2.2.00/07.0284

Operační program Vzdělávání pro konkurenceschopnost

Oblast podpory. 2.2 Vysokoškolské vzdělávání

Realizace: Vysoká škola báňská - TU Ostrava, Ekonomická fakulta, pracoviště Obchodní akademie a VOŠ Valašské Meziříčí

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky

OBSAH

Obsah.....	2
Úvod.....	3
1. Stručná historie Internetu	3
1.1 Od ARPANETU k Internetu	6
1.2 Provoz Internetu	7
1.3 Internet v České republice.....	8
2. Protokoly Internetu.....	11
2.1 Vrstvový model TCP/IP	11
2.2 Dokumenty k Internetu.....	14
2.3 Adresy počítačů v Internetu	15
2.3.1 Číselné (IP) adresy	15
2.3.2 Doménové adresy	19
2.3.3 Pravidla pro tvorbu doménových jmen	20
3. Služby Internetu	21
3.1 Vzdálené přihlašování - Remote Login.....	21
3.2 Přenos souborů pomocí FTP	22
3.3 Elektronická pošta	23
3.4 Gopher.....	27
3.4 World Wide Web – WWW	28
3.6 Elektronické konference a diskusní skupiny	30
4. Úvod do strategie vyhledávání v Internetu	33
4.1 Jak vyhledávat - tipy pro snazší hledání.....	34
4.1.1 Hlavní kroky vyhledávací strategie.....	35
4.2 Vyhledávací dotazy	36
4.2.1 Jednoduché dotazy	36
4.2.2 Pokročilé dotazy	38
5. Klasifikace vyhledávacích služeb	41
5.1 Seznamy zdrojů (Resource List)	41
5.2 Předmětové katalogy (Webové portály).....	41
5.3 Vyhledávače (Search Engines).....	43
5.4 Metavyhledávače (Metasearch Engines).....	45
5.5 Další typy vyhledávacích nástrojů	46
5.6 Virtuální knihovny	47
5.7 Neviditelný web	47
5.8 Sociální sítě	48
6. Google.com	49
6.1 Chronologie klíčových událostí	49
6.2 Vyhledávací strategie Google.com	50

„Learning about the Internet is not something one can complete in an afternoon – learning never stops because the Internet keeps changing“

Douglas E. Comer

ÚVOD

Studijní materiál Internet a informace je úvodem do zpracování informací. Jeho cílem je podpora informační gramotnosti studentů především prvních ročníků vysokých škol. Internet se stává nedílnou součástí jejich přípravy na studium, je potřeba využívat nástroje Internetu efektivně a především kriticky. Vzhledem k tomu, že vše kolem Internetu se neustále mění a vyvíjí, nejde text do hloubky, ale drží se širšího tedy ne tolik proměnlivého rámce. Učební text byl vytvořen v rámci projektu Studium informační vědy a znalostního managementu v evropském kontextu, který byl financován z Operačního programu Vzdělávání pro konkurenceschopnost.

1. STRUČNÁ HISTORIE INTERNETU

Internet změnil svět výpočetní komunikační techniky jako nic předtím. Vynálezy telegrafu, telefonu, rádia a počítače připravily půdu pro integrační možnosti Internetu. Historie Internetu začíná před rokem 1962, v době, kdy počítače přestože jsou primitivní a s malou pamětí stojí stovky, tisíce dolarů. Firma RAND Corporation (<http://www.rand.org>) v této době řešila nemalý úkol - jak by jednotlivé orgány administrativy USA mohly úspěšně komunikovat i po jaderné válce. Bylo jisté, že řešení by mělo být založeno na existenci vhodné komunikační sítě. RAND Corporation přišla s řešením, které bylo v roce 1964 zveřejněno. Je založeno na následujících dvou principech:

- síť nebude mít žádnou centrální složku
- síť bude od začátku navrhována tak, aby fungovala, i když jsou některé její části v troskách

Konkrétní naplnění výše uvedených principů spočívalo v tom, že síť se navrhla takovým způsobem, aby všechny její uzly měly v zásadě rovnocenné postavení a předem počítaly s tím, že přenosy mezi jednotlivými uzly nejsou spolehlivé. Právě tento druhý požadavek stál u zrodu myšlenky, jenž spočívala v tom, že přenášená data se rozdělí na vhodně velké části (nazývané pakety), které se budou přenášet jako samostatné celky - tedy obdobně, jako například dopisy listovní poštou. Každý datový paket bude opatřen plnou adresou svého příjemce, cesta jeho přenosu bude vždy volena samostatně, nezávisle na cestě jiných paketů. Různé pakety se pak mohou ubírat různými cestami, takže když například jedna z možných cest k určenému cíli bude náhle zničena, další pakety budou přenášeny jinou cestou.

Světlo světa spatřila technika, označovaná příznačně jako přepojování paketů (Packet Switching), navíc ještě v takové variantě, které se dnes říká datagramová služba (Datagram Service). Obvykle je sice méně efektivní než takový způsob přenosu dat, který modeluje fungování telefonní sítě (a který je označován jako přepojování okruhů (Circuit Switching), ale na druhé straně zase přináší maximální možnou robustnost. [PETERKA, 1997].

Je potřeba podotknout, že na projektu přepojování paketů pracovala nejen RAND Corporation, ale současně i MIT (Massachusetts Institute of Technology) a Národní fyzikální laboratoř ve Velké Británii. Na výzkumu pracovali Paul Baran, Dobald Davies, Leopard Kleinrock aj. Přepínání paketů ve skutečnosti není žádný moderní vynález, ale jen vhodná aplikace dávnověké myšlenky dynamického využívání dopravních prostředků používaných poštovními i telegrafními systémy již po staletí. Zásilky či zprávy se obvykle převážejí či předávají najednou a teprve v okamžiku, kdy jejich množství naplní vůz nebo vytíží přenosovou trasu natolik, aby se doprava vyplatila.

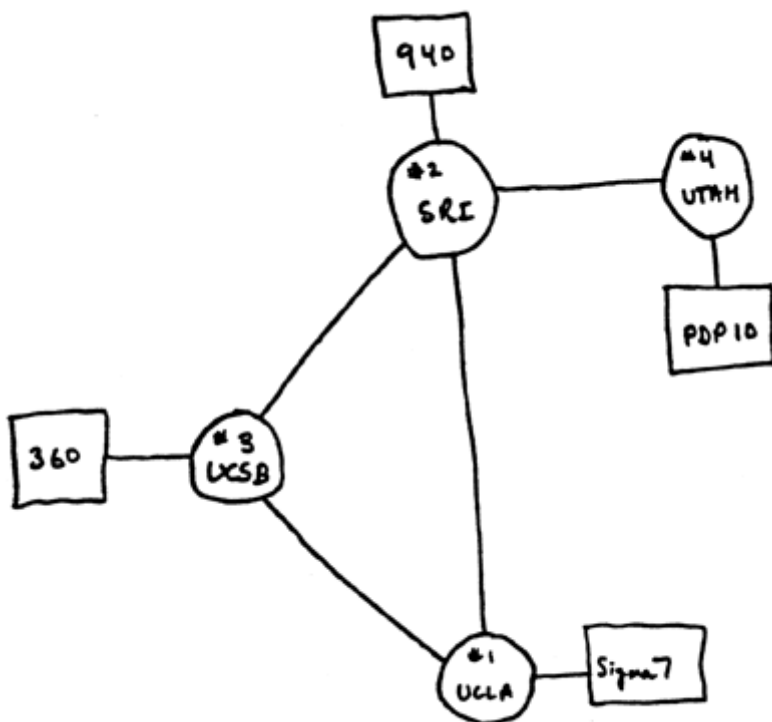
Myšlenka silně decentralizované, maximálně robustní a "výbuchuvzdorné" sítě, pracující na principu přepojování paketů, byla poprvé v praxi implementována ve Velké Británii. V roce 1968 vznikla experimentální síť v National Physical Laboratory (Národní laboratoře pro fyziku).

USA se k tomuto kroky odhodlaly v roce 1969. Podnět a finanční prostředky přišly z resortu obrany, konkrétně od grantové agentury ministerstva obrany USA, s názvem ARPA (Advanced Research Projects Agency)¹. Podle této grantové agentury byla experimentální síť také pojmenována jako ARPANET. První uzly sítě ARPANET byly umístěny na 4 amerických univerzitách - UCLA (University of California Los Angeles, Vint Cerf, Steve Crocker, Jon Postel), UCSB (University of California Santa Barbara, Glen Culle, Burton Fried), ve Stanfordu (Stanford Research Institute, SRI, Doug Engelbart) a na univerzitě v Utahu. Vlastní uzel přitom byl realizován univerzálním počítačem (konkrétně šlo o počítač Honeywell DDP516), který byl naprogramován tak, aby fungoval jako tzv. Interface Message Processor (IMP). Pro vzájemnou komunikaci používaly uzly IMP pevné okruhy s přenosovou rychlostí 50 kbps, a přenosový protokol NCP (Network Control Protocol). Ještě do konce roku 1969 byla vlastní síť uvedena do provozu, a její uživatelé tak mohli na dálku využívat možnosti a výpočetní kapacitu superpočítačů, které tato síť propojovala.

ARPANET však rychle rostl, a v roce 1971 již měl celkem 15 uzlů, přičemž tento počet neustále rostl: v roce 1972 měl ARPANET 37 uzlů, a v roce 1973 se k němu připojují také první zahraniční uzly, ve Velké Británii a v Norsku.

V říjnu 1972 demonstroval Kahn ARPANET velmi úspěšně veřejnosti na první Mezinárodní konferenci počítačové komunikace (International Conference on Computer Communication).

¹ ARPA se posléze přejmenovala na DARPA (Defense Advanced Research Project Agency), což mělo zdůraznit její zaměření na potřeby resortu obrany. <http://www.darpa.mil>



Obr. č. 1: Původní náčrt sítě Arpanet, 4-node ARPANET diagram

Zdroj: http://www.computerhistory.org/exhibits/internet_history/ [Internet history, 2006]

Skupina pedagogů a postgraduálních studentů z UCLA (University of California Los Angeles), zřídila na základě grantu od agentury ARPA laboratoř (Network Measurement Center), ve které testovali a měřili chování ARPANETu. Členy této skupiny byli takoví lidé, jako např. Vinton Cerf, Steven Crocker či Jon Postel. Výsledky svého šetření publikovali ve formě pracovních dokumentů, které nazvali "Request For Comment". Úplně první materiál RFC (Request For Comment) byl publikován 7. dubna 1969. Napsal ho Steven Crocker, a týkal se problematiky programového vybavení uzlových počítačů sítě.

<http://tools.ietf.org/html/rfc1>

Klíčovou osobností v dalším vývoji ARPANATu byl Vinton G. Cerf, který v roce 1972 po ukončení svých postgraduálních studií nastoupil na místo docenta na univerzitě ve Stanfordu. Krátce nato začal pořádat síťové semináře, v rámci kterých byly položeny základy pro novou generaci komunikačních protokolů pro síť ARPANET – protokolů rodiny TCP/IP.

Jejich vývoj pod vedením Vintona Cerfa ovšem určitou dobu trval. Byl financován z prostředků agentury ARPA (mezitím přejmenované na DARPA). První verze specifikací protokolu TCP byla v září roku 1973 prezentována na počítačové konferenci na University of Sussex. Publikována byla následujícího roku, v květnovém čísle časopisu IEEE Transactions on Communications (autory byli Vinton Cerf a Robert Kahn). První implementace protokolu TCP pak probíhaly souběžně na třech místech, kromě Stanfordu i u firmy BBN (Bolt, Beranek and Newman) a na University College v Londýně. Šlo tedy od začátku o mezinárodní záležitost.

V roce 1977 proběhly první praktické zkoušky TCP, které prokázaly životaschopnost nového protokolu. Celková koncepce přenosových protokolů nakonec vyústila v rozdělení původního protokolu TCP na dva samostatné protokoly - na protokol IP (Internet Protocol), který se staral zejména o vlastní přenos a o zakrytí všech specifík konkrétních přenosových cest, ale nikoli na spolehlivém principu (tj. když došlo k nějakému výpadku, ztrátě či poškození dat, sám se nestaral o nápravu), a na "nový" protokol UDP, který sám využíval přenosových služeb protokolu IP a navíc k nim přidával i zajištění spolehlivosti. Jako alternativa ke spolehlivému protokolu TCP pak byl vytvořen protokol UDP (User Datagram Protocol), který také využíval přenosových služeb protokolu IP, ale již k nim nepřidával žádné zajištění spolehlivosti (a je tedy určen pro takové aplikace, které před spolehlivostí dávají přednost rychlosti a pravidelnějšímu přísunu dat). [*TCP/IP Overview and History*]

1.1 Od ARPANETU k Internetu

Osudy sítě ARPANET, se odehrávaly plně v režii resortu obrany USA - který také vše financoval. V roce 1983 byly odděleny od původního ARPANETu ty jeho části, které měly něco společného s vojenstvím. Byla ustanovena samostatná síť MILNET. ARPANET tak získal mnohem civilnější náplň práce, ale přesto byl nadále financován z prostředků resortu obrany.

ARPANET ovšem nebyl zdaleka jedinou počítačovou sítí na světě či alespoň v USA. Svě počítačové sítě si budovaly i jiné resorty. Vzhledem ke kvalitám a veřejné dostupnosti protokolů TCP/IP byly tyto sítě stále častěji budovány na bázi těchto protokolů. Pro provozovatele sítí se ukázalo jako velmi výhodné, když je mohli propojit s ARPANETem. Ten se tak stále více stával spíše zárodečnou sítí, na kterou se postupně "nabalovaly" další a další sítě, až vznikl celý konglomerát vzájemně propojených sítí, kterému se začalo říkat Internet.

Souviselo to ostatně i s nástupem lokálních sítí, ke kterému dochází někdy po roce 1983, kdy plně vyžívá nejnámější a nejpoužívanější přenosová technologie pro sítě LAN - Ethernet. Také pro nesčetné množství těchto lokálních sítí bylo velmi výhodné využít protokoly TCP/IP a připojit se ke stále se rozšiřující soustavě vzájemně propojených sítí na bázi těchto protokolů, tj. k Internetu.

Zřejmě nejvýznamnější sítí, která se kdy "nabalila" na ARPANET a stala se tak součástí Internetu, byla síť NSFNET. Díky štědrému přísunu finančních prostředků mohla agentura NSF financovat doslova masové připojování dosud nepřipojených akademických institucí k NSFNETu a tím i k Internetu. NSFNET díky svým přenosovým schopnostem a taktéž na základě koncepčního záměru, postupně přejímal původní roli ARPANETu - tedy roli páteřní sítě, skrz kterou prochází největší část provozu v rámci celého Internetu

V březnu roku 1990 byl původní ARPANET v tichosti odstaven a zrušen. NSFNET úplně přebíral úlohu páteřní sítě celé obrovské soustavy sítí, která tvořila Internet. NSFNET byl původně koncipován jako výzkumná síť, ale s postupem času se stal především provozní sítí, zajišťující spíše rutinní přenosy velkých objemů dat. Nikoli ale přenosy, sloužící ryze komerčním účelům, kterým se NSFNET nadále brání.

V roce 1989 spolupracovali Tim Berners-Lee a Robert Caillau na propojeném informačním systému pro CERN (Evropská organizace pro jaderný výzkum, poblíž Ženevy ve Švýcarsku). V té době se pro tvorbu dokumentů obvykle používaly jazyky TeX, Postskript, a také SGML. Berners-Lee si uvědomoval, že potřebují něco jednoduššího a v roce 1990 byl tedy navržen jazyk HTML a protokol pro jeho přenos v počítačové síti – HTTP (*HyperText Transfer Protocol – přenosový protokol hypertextu*). Zároveň také Tim Berners-Lee napsal první webový dokument, který nazval WorldWideWeb.

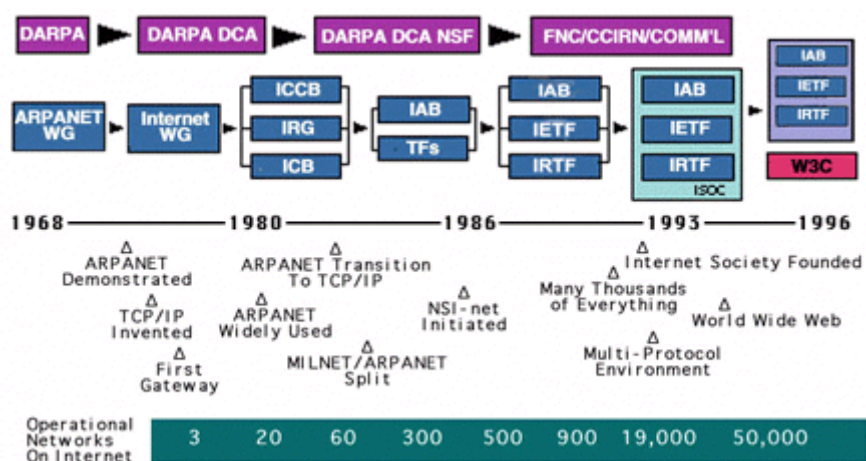
V roce 1991 zakládá NSFNET novou síť NREN, the National Research and Education Network

1.2 Provoz Internetu

Jedním z celkem přirozených trendů v rámci rozvoje Internetu bylo i to, že instituce, které se postaraly o vybudování významných sítí, záhy svěřily péči o jejich rutinní provoz specializovaným institucím. Agentura ARPA tak učinila již v červenci roku 1975, kdy svůj ARPANET předala jako plně funkční síť do správy organizaci DCA (Defense Communications Agency, později přejmenované na DISA, Defense Information Systems Agency). Řízením Internetu jako takového (přesněji provozováním střediska NCC, Network Control Center) pak ministerstvo obrany pověřilo firmu BBN (Bolt, Beranek and Newman), která stála již u zrodu samotného ARPANETu a "postavila" jeho první uzly. Tato firma v provozování řídicího střediska NCC pokračovala až do roku 1993 (později již na základě grantů, pocházejících od agentury NSF). Podobně se zachovala i agentura NSF, když po necelém roce od zprovoznění své páteřní sítě NSFNET předala její správu firmě Merit Network Inc. (kterou založilo konsorcium univerzit z Michiganu), a spolu s firmami IBM a MCI ji pověřila realizací potřebného "upgrade" páteřní sítě NSFNET - především jejím zrychlením.

Kromě vlastní sítě NSFNET a jejího provozu však agentura NSF financovala i celý propracovaný systém shromažďování a distribuce nejrůznějších informací o Internetu. Základem tohoto systému bylo informační středisko NIC (Network Information Center), které dlouhá léta fungovalo ve Stanfordu na tamním institutu SRI (Stanford Research Institute). Toto středisko například zajišťovalo všechny registrace nových sítí, připojovaných k Internetu, a přidělovalo také celosvětově unikátní síťové adresy (tzv. IP adresy). Kromě toho bylo středisko NIC ve Stanfordu také jedním z hlavních deponitářů dokumentů Internetu (mj. dokumentů RFC).

Od dubna 1993 však byla většina informačních a dalších služeb kolem Internetu svěřena konsorciu tří firem, které je zajišťují na komerčním základě: Network Solutions, Inc., AT&T a General Atomics. Toto konsorcium, kterému se říká InterNIC, je přitom zčásti financováno opět z prostředků agentury NSF. [A Brief History of the Internet]



Obr. č. 2: Časová osa

Zdroj: *A Brief History of the Internet*

1.3 Internet v České republice

Historie Internetu se v českých zemích začala psát v listopadových dnech roku 1989. V té době padly politické bariéry, které až do té doby z principiálních důvodů bránily našemu zapojení do významnějších celosvětových sítí - mimo jiné i do Internetu. S ústupem politických překážek však rázem nabyly na významu překážky jiné, tentokrát již technického charakteru - naše země, po dlouholetému "ne-rozvoji" telekomunikací, neměla vhodnou komunikační infrastrukturu pro větší rozvoj počítačových sítí. Prvními sítěmi, které se k nám po sametové revoluci dostaly, byly takové, které mají jen minimální nároky na potřebnou komunikační infrastrukturu. V březnu roku 1990 se do českých zemí dostává síť FIDO, na kterou se napojují především stanice BBS, Bulletin Board System.

V květnu roku 1990 se pak do Československa dostává i síť EUnet, propojující především Unixové počítače. V říjnu roku 1990 se k nám dostává evropská odnož sítí Bitnet, síť EARN (European Academic and Research Network), která již vyžaduje trvalé spojení po pevných okruzích. Prvním uzlem této sítě u nás, a současně i tzv. národním uzlem sítě EARN pro tehdejší Československo, se stal střediskový počítač IBM 4381 na Oblastním výpočetním centru (OVC) ČVUT Praha (nyní VC ČVUT). Síť EARN přitom poskytovala pouze služby dávkového charakteru (zejména elektronickou poštu a přenos souborů), takže vystačila i

s relativně pomalými pevnými okruhy. Uzel CSEARN začínal s linkou o přenosové rychlosti 9600 b/s, kterou byl připojen na rakouský národní uzel sítě EARN v Linzi.

První pokusy s připojením do Internetu se pak objevují až v listopadu roku 1991. Toto datum je také v různých oficiálních i neoficiálních statistikách uváděno jako datum připojení tehdejšího Československa k Internetu. Zpočátku šlo pouze o komutované napojení z Prahy (konkrétně z VC ČVUT na uzel Internetu v rakouském Linzi), později byla k připojení použita pevná linka do Linze, uměle "rozpůlená" tak, aby jedna její polovina přenášela provoz v rámci sítě EARN a druhá provoz Internetu. [Peterka, J.]

V listopadu 1991 se u nás do evropského Internetu připojily první univerzity prostřednictvím sítě CESNET, a tím prakticky začala historie Internetu v České republice.

Již v roce 1991 byl podáván návrh na vybudování celorepublikové páteřní sítě. Ta měla propojovat všechna tuzemská akademická centra a dále by pak měly Internet rozvádět metropolitní sítě. Na základě tohoto návrhu vnikly v Československu dva projekty na vybudování národních páteřních sítí, slovenské a české, přičemž propojení těchto dvou páteří bylo součástí projektu české strany (konkrétně šlo o spoj vedený z Prahy do Bratislavy). Český projekt dostal jméno FESNET (Federal Educational and Scientific NETwork).

13. února 1992 pak na ČVUT Praha dochází ke slavnostnímu aktu formálního připojení Československa k Internetu, mj. i za účasti představitelů agentury NSF.

V červnu 1992 české ministerstvo školství schválilo projekt a uvolnilo na něho 20 milionů korun. V průběhu roku 1992 se písmeno F v názvu FESNET změnilo na C a tím vlastně vznikl CESNET (Czech Educational and Scientific NETwork). Na Slovensku se souběžně začala budovat síť SANET (Slovak Academic NETwork).

Síť CESNET byla zpočátku realizována hvězdicovou topologií se středovými uzly v Brně a v Praze. Ty byly propojeny pevnou linkou o rychlosti 64 kbps. K těmto dvěma uzlům byla připojovány další uzly umístěné v našich akademických městech. Postupně byly připojeny České Budějovice, Hradec Králové, Liberec, Plzeň a další. V březnu roku 1993 byly uzly CESNETu rozmístěny už v 11 městech. Řada z těchto připojení byla realizována pomocí pronajaté pevné telefonní linky. A jaké byly tehdejší přenosové rychlosti? Kromě linky z Prahy do Brna (která měla již zmiňovaných 64 kbps), pracovaly všechny ostatní spoje na rychlosti 19,2 kbps. Postupem času byla většina uzlů připojena dvěma nezávislými spoji, kvůli zachování konektivity v případě výpadku některé z linek. Také se rozrostl počet linek směřujících za hranice naší země. Z původně jediného spoje vedoucího z Prahy do rakouského Lince byla vytvořena linka Praha-Vídeň. Později přibylo spojení Praha-Amsterdam realizované spojem o rychlosti 64kbps. (A rozpadem Československa vznikl další zahraniční spoj Praha-Banská Bystrica). Na přelomu let 1994 a 1995 byla komunikační infrastruktura CESNETu v podstatě dobudována a pozornost se přesunula především na zvyšování přenosových rychlostí a s ním související zlepšování spolehlivosti.

Literatura:

1. *A Brief History of the Internet* [online]. c2010 Internet Society. [cit. dne 2010-03-17]. Dostupné z: <http://www.isoc.org/internet/history/brief.shtml>.
2. *Internet history: A history of the Internet 1962 - 1992* [online]. c2006 Computer History Museum. [cit. dne 2010-03-17]. Dostupné z: http://www.computerhistory.org/internet_history/.
3. PETERKA, J. Internet. *Computerworld*, 1995, roč. 6, č. 4, s. 17 – 19. ISSN 1210-9924.
4. PETERKA, J. *Internet u nás* [online]. [cit. dne 2010-03-17]. Dostupné z: <http://www.earchiv.cz/a95/a504c504.php3>
5. *TCP/IP Overview and History: Modern TCP/IP Development and the Creation of TCP/IP Architecture* [online] c2001 - 2005 Charles M. Kozierek [cit. dne 2010-01-30]. Dostupné z: http://www.tcpipguide.com/free/t_TCPIPOverviewandHistory-2.htm.

2. PROTOKOLY INTERNETU

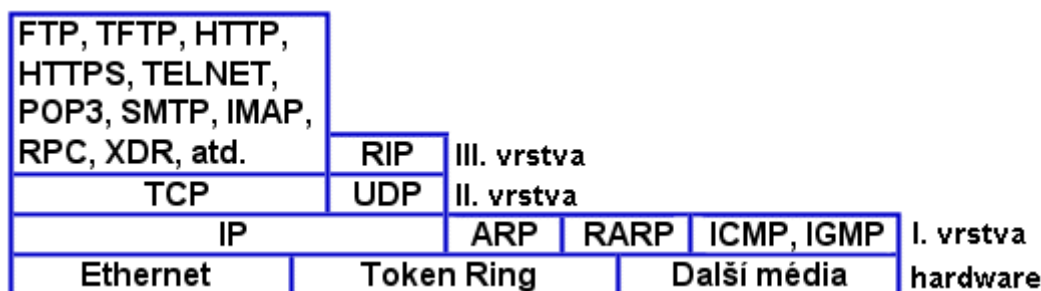
V souvislosti s Internetem a počítačovou sítí obecně se často setkáme s pojmem protokol. Protokol je soubor určitých pravidel (jinak řečeno norem nebo standardů), které musí dané zařízení nebo program splňovat, aby byl schopen komunikovat s jiným odpovídajícím zařízením nebo programem. Proces tvorby síťových protokolů je dlouhodobý a protokoly se neustále vyvíjí. Na Internetu se používá mix rodiny protokolů TCP/IP, protokolů ITU a ISO. Protokoly ITU a ISO se vyskytují zejména na spodních dvou vrstvách, fyzické a linkové.

TCP/IP (Transmission Control Protocol/Internet Protocol) označuje celou síťovou architekturu, nejen dva protokoly, které má ve svém názvu a které tvoří její jádro. Podle jednoho z tvůrců, Douga Comeru, by její správné a kompletní označení mělo být The TCP/IP Internet Protocol Suite. [PUŽMANOVÁ, 2009, s.83]

2.1 Vrstvový model TCP/IP

Na rozdíl od sedmivrstvého referenčního modelu OSI² protokolovou architekturu TCP/IP tvoří 4 vrstvy:

- Vrstva rozhraní sítě (network interface)
- Vrstva mezisíťová (internet layer)
- Transportní vrstva (transport layer)
- Aplikační vrstva (application layer)



Zdroj: *Počítačové sítě...*

² Referenční model OSI (Open Systems Interconnection) vytvořila Mezinárodní normalizační organizace a přijala ho jako mezinárodní normu IS 7498 v roce 1984. OSI se začala zabývat i Mezinárodní telekomunikační unie (ITU), která přijala a vydala kompletní text zmíněné mezinárodní normy jako své doporučení X.200. [PUŽMANOVÁ, 2009, s.42]

Hardware (vrstva síťového rozhraní)

Nejnižší vrstva architektury umožňuje přístup k fyzickému přenosovému médiu. Je přímo zodpovědná za přístup k síti, a je proto specifická pro každou síť podle její implementace. V současnosti se může využívat všech známých typů přenosových prostředí, lokálních sítí (Ethernet, Token Ring, FDDI) i rozlehlých sítí (X.25, ATM apod) pro podporu TCP/IP

I. vrstva (mezisíťová vrstva, vrstva Internetu)

Odpovídá svými funkcemi a službami vrstvě síťového referenčního modelu OSI. Její funkce zahrnují především síťovou (logickou) adresaci, směrování a předávání datagramů přes komunikační podsíť, včetně směrování. Další funkcí je provádění segmentací a znovusestavování datagramů do a z rámců specifikovaných protokolem nižší vrstvy.

IP - Internet Protocol

Nezákladnější protokol, neobsahuje potvrzování (počítač neví jestli data která vyslal, přijmul vzdálený počítač). Zabezpečuje správné doručování dat k jednotlivým počítačům v síti. Protokol IP provádí vysílání datagramů (TCP/IP používá místo pojmu paket pro síťovou datovou jednotku pojem datagram) na základě síťových adres obsažených v jejich záhlavích a poskytuje síťovou službu bez spojení. Data jsou od odesílatele k příjemci dopravována (směřována) přes směšovače. Na cestě od odesílatele k příjemci se může vyskytnout celá řada směšovačů. IP protokol je tedy protokol umožňující spojit jednotlivé lokální sítě do celosvětového Internetu. Protokol IP prošel několika verzemi své specifikace a momentálně se za „normu“ považuje verze 4 protokolu (RFC 791, STD 5). Verze IP protokolu jsou autoritativně zdokumentovány IANA³.

ARP - Address Resolution Protocol

Převádí 32 bitovou IP adresu na 48 bitovou MAC adresu.

RARP - Reverse Address Resolution Protocol

Naopak převádí MAC adresu na IP adresu. Tento protokol používají bezdiskové pracovní stanice, které neznají svojí IP adresu.

ICMP - Internet Control Message Protocol

Používá se k signalizaci chyb a různých nestandardních situací (ale pouze potřebám signalizace, ICMP sám nezajišťuje jejich nápravu).

IGMP - Internet Group Management Protocol

Podporující tzv. skupinové vysílání (multicasting).

II. vrstva (transportní vrstva)

Poskytuje mechanismus pro koncový přenos dat mezi dvěma stanicemi.

³ <http://www.iana.org/assignments/version-numbers>

TCP/UDP

Na jednom počítači lze provozovat několik programů, které poskytují své služby. Aby se rozlišilo na kterou službu program přistupuje, musí být nějak rozlišeny, a to takzvaným portem. Adresa portu, neboli IP adresa páru tvoří socket, to jest zdrojová adresa portu a zdrojová IP adresa tvoří zdrojový socket. Pro veřejně přístupné služby jako webové servery a poštovní servery je číslo portu důležité. Například, port 80 se připojí k webovému serverovému softwaru a port 25 k poštovnímu serverovému SMTP démonovi. Maximálně může být najednou spuštěno 65 tisíc portů (programů). Socket = IP adresa + port.

TCP - Transmission Control Protocol

Potvrzovaný protokol. Vytváří takzvané virtuální spojení. Toto spojení trvá po dobu než aplikace spojení ukončí.

UDP - User Datagram Protocol

Nepotvrzovaný protokol. Od IP se liší jen tím, že má navíc port. Lze tak poslat konkrétnímu programu dotaz. Moc se nepoužívá, spíše jen na služební komunikaci. Např. routery když každých 30 sec. hlásí, kdo je připojen.

III. vrstva (aplikační vrstva)

Obsahuje protokoly (aplikace), které se už přímo využívají ke komunikaci po síti.

FTP/TFTP - File Transfer Protocol/Trivial FTP

Slouží k přenosu souborů mezi počítači spojenými do sítě. TFTP je jednodušší varianta k FTP.

HTTP/HTTPS - Hyper Text Transfer Protocol

Slouží k přístupu na www stránky. HTTPS je zabezpečený (šifrovaný) přenos www stránek.

TELNET - Telecommunication Network

Vytváří terminálový provoz. Můžeme pracovat se vzdáleným počítačem stejně jako bychom seděli u terminálu bezprostředně k němu připojeném. Protože komunikace probíhá nešifrovaně představuje jeho používání bezpečnostní riziko. Náhradou za TELNET je SSH (Secure Shell) který komunikuje šifrovaně.

POP3 - Post Office Protocol

Slouží k přijímání elektronické pošty poštovním klientem.

SMTP - Simple Mail Transfer Protocol

Slouží k odesílání elektronické pošty poštovním klientem

RPC/XDR

Vzdálené volání procedur. Používá se při požadavku provést výpočet programu na jiném počítači než kde jsou uložena data.

2.2 Dokumenty k Internetu

Všechny informace o síti Internet a TCP/IP (konvencích a síťových protokolech) jsou od samotného počátku dostupné na síti a od roku 1969 se označují jako Request for Comments (RFC). RFC jsou volně dostupné prostřednictvím webu, elektronické pošty nebo FTP serveru. RFC jsou považovány za de facto normy, tj. žádnému z orgánů spravujících vývoj Internetu a jeho technických specifikací nebyl přiznán statut mezinárodní normalizační instituce. První dokument byl vytvořen v dubnu 1969.

Typy RFC [PUŽMANOVÁ, 2009, s. 31 - 34]

:

- Standards track – Proposed (první stupeň návrhu normy), Draft (druhý stupeň návrhu vyžadující dvě nezávislé implementace) Internet standard (STD, plná de facto norma)
- Best current practice (BCP)
- Informational (některé označené jako FYI)
- Experimental
- Historic

Vlastnosti RFC dokumentů [MUSIL, 2003]

- Nemění se, nestárnou, nejsou vydávány jejich aktualizace, opravy atd. Tato vlastnost je výhodná z hlediska distribuce a vedení archivů RFC.
- Jsou volně šiřitelné a bezplatné.
- Některé z nich jsou užitečné i pro laického uživatele, ty pak dostávají označení FYI a nějaké menší číslo.
- Všechny dokumenty RFC jsou číslovány (pořadovým číslem) a nikdy nejsou měněny (nahrazují se vyšším RFC). Pokud je potřeba nějak upravit jejich věcný obsah, je vydán nový dokument RFC, který „zneplatňuje“ příslušný předchozí dokument RFC⁴.

⁴ Bližší informace o procesu tvorby RFC jsou uvedeny v [RFC 2026](#)

Volně k dispozici na adrese:

<http://www.ietf.org/rfc.html>

Editor RFC:

<http://www.rfc-editor.org/>

2.3 Adresy počítačů v Internetu

Jsou dvojího typu:

1. adresy číselné (IP)
2. adresy doménové

2.3.1 Číselné (IP) adresy

Počítače a jiná zařízení, propojená v jedné společné síti, mají každý svou vlastní adresu. Ta je pro daný počítač jedinečná, nikde v Internetu se neopakuje. Každá adresa jednoznačně určuje počítač, kterému byla přidělena. Pakety putující Internetem s sebou nesou mj. cílovou adresu. Internet doručuje pakety správným směrem tak dlouho, až dorazí na adresu určený počítač. Každý paket též nese adresu odesílatele, takže komunikace může být obousměrná. Adresy, o kterých je zde řeč, jsou adresy protokolu IP (Internet Protocol), zkráceně IP adresy. IP adresa je tedy v informatice číslo, které jednoznačně identifikuje síťové rozhraní v počítačové síti, která používá IP (internetový protokol). V současné době je nejrozšířenější verze IPv4, která používá 32bitové adresy. Z důvodu nedostatku IP adres bude nahrazen protokolem IPv6, který používá 128bitové IP adresy.

IPv4

Protokol IP verze 4 používá IP adresu o délce čtyři bajty. IP adresa adresuje jednoznačné síťové rozhraní systému – **unicast**. Pokud má systém více síťových karet (více síťových rozhraní) a na všech je provozován protokol IP, pak má každé rozhraní svou IP adresu. Je možná i opačná varianta, kdy na jedné síťové kartě (fyzicky jednom síťovém rozhraní) je podporováno několik IP adres. První adresa se obvykle nazývá primární, další adresy pak sekundární nebo aliasy. Využití sekundárních IP adres je běžné např. pro WWW servery, kdy na jednom počítači běží WWW servery několika firem a každý se má tvářit jako samostatný WWW server.

V praxi se využívání sekundárních IP adres pro WWW servery považuje za plýtvání – používají se tzv. virtuální WWW servery, kdy mnoha WWW serverům stačí jedna společná IP adresa. Specifikace serveru se potom provádí na aplikační úrovni v protokolu HTTP (pomocí hlavičky host). [KABELOVÁ, A., DOSTÁLEK, L., 2008]

Adresa IPv4 se zapisuje notací, kde se jednotlivé bajty mezi sebou oddělují tečkou.

Rozeznáváme:

- Dvojkovou notaci, kde je každý ze čtyř bajtů adresy zapsán číslem ve dvojkové soustavě, např. 10101010.01010101.11111111.11111000.
- Desítkovou notaci, kde se čtyři osmiciferná dvojková čísla (z dvojkové notace) převedou do desítkové soustavy, tj. zde: 170.85.255.248
- Šestnáctkovou notaci, kde jsou jednotlivé bajty IP adresy vyjádřeny šestnáctkově (hexadecimálně), tj. zde: aa.55.ff.f8

Desítková notace se používá nejčastěji, a to pro přehlednost; každé z písmen a-d vyjadřuje desítkovým číslem 8 bitů adresy. Celkem tedy $8 \times 4 = 32$ bitů. Jednotlivé číslice a-d mohou nabývat hodnot 0 až 255. Limit 255 vyplývá z toho, že nejvyšší osmimístné dvojkové číslo (11111111) je právě 255 desítkově. Kombinací čtyř takových čísel získáme $256 \times 256 \times 256 \times 256 = 4\,294\,967\,296$ (4 miliardy) možných IP adres. IP adres na světě tedy může být 2^{32} , nepočítaje nepovolené hodnoty jako samé nulové bity (0.0.0.0) a samé jedničkové bity (255.255.255.255), stejně jako speciální vyhrazené adresní rozsahy pro zvláštní účely. IP adresa je tedy v podstatě libovolná kombinace třiceti dvou nul a jedniček. V počátcích Internetu se to zdálo být nepředstavitelné množství, ale dnes je masové nasazování privátních IP rozsahů důkazem jejich nedostatku.

IPv6

IPv6 vznikl v polovině 90. let 20. století, především jako reakce na rychlé vyčerpávání adres pro IPv4. Jeho autoři se rozhodli nabídnout kromě mnohem většího adresního prostoru i některé nové vlastnosti, reagující na vývoj počítačových sítí. Stanovili si následující hlavní cíle:

- dostatečně bohatý adresní prostor - pokud možno by už nikdy neměla nastat nouze o adresy
- podpora služeb se zaručenou kvalitou
- design odpovídající vysokorychlostním sítím
- bezpečnostní mechanismy přímo v IP
- podpora mobilních zařízení
- automatická konfigurace
- kooperace s IPv4 a co nejjednodušší přechod ze stávajícího protokolu na nový

[*Vlastnosti protokolů*, 2008]

Hlavní změnou oproti IPv4 je délka síťové adresy. Adresy IPv6 jsou 128 bitů dlouhé (jak je určeno RFC 4291). Zatímco IPv4 obsahuje zhruba 4 miliardy adres, IPv6 má dostatek prostoru pro $3,4 \times 10^{38}$ unikátních adres. Adresy IPv6 se typicky skládají ze dvou logických částí: 64bitová (pod)síťový prefix a 64bitové části hosta, buď automaticky vytvářené na základě MAC adresy rozhraní nebo přiřazené následně. Jelikož globálně unikátní MAC adresa umožňuje vystopovat uživatelské vybavení - a tedy uživatele - IPv6 adresy se mění s časem. IPv6 adresy s obvykle zapisují jako osm skupin čtyř hexadecimálních čísel. Například 2001:0db8:85a3:08d3:1319:8a2e:0370:7334 je platná adresa IPv6. Pokud je jedna nebo více ze čtyřčlenných skupin 0000, nuly mohou být vynechány a nahrazeny dvěma dvojtečkami (::). Např. 2001:0db8:0000:0000:0000:0000:1428:57ab lze nahradit 2001:0db8::1428:57ab.⁵

⁵ Více informací: RFC 4291: <http://tools.ietf.org/html/rfc4291>

Přidělování IP adres

Přidělování adres neboli adresních rozsahů probíhá hierarchicky. Kdesi úplně na vrcholu pyramidu sídlí ICANN (Internet Corporation for Assigned Names and Numbers), který zajišťuje koordinaci na nejvyšší úrovni.

Zeměkoule je rozdělena na regiony, které obhospodařují regionální registrátoři (Regional Internet Registry, RIR). Každý z nich disponuje určitým adresním prostorem pro síť ve svém regionu a eviduje přidělené rozsahy. Přestože spolu RIR navzájem komunikují, jedná se o nezávislé instituce s vlastní organizací a pravidly.

Dalším článkem hierarchie jsou lokální registrátoři (Local Internet Registry, LIR). Typickým příkladem LIR je poskytovatel internetového připojení, který chce přidělovat adresy svým zákazníkům. Pokud splní podmínky příslušného regionálního registrátora (což zpravidla znamená zavázat se k dodržování pravidel a zaplatit poplatek), dostane od něho přidělen velký blok adres.

Existují také firemní LIR, a to v případě, že daná firma je natolik velká, že potřebuje spoustu adres⁶.

Seznam LIR v České republice: <http://www.ripe.net/membership/indices/CZ.html>

Během uplynulých zhruba deseti let byl svět rozdělen na tři regiony. Obsluhovali je následující registrátoři:

RIPE NCC	(Réseaux IP Européens Network Coordination Centre) Evropa, Rusko, severní Afrika
ARIN	(American Registry for Internet Numbers) Amerika, jižní Afrika
APNIC	(Asia Pacific Network Information Centre) Asie, Austrálie a Oceánie

⁶ Seznam LIR v České republice: <http://www.ripe.net/membership/indices/CZ.html>

Síťaři některých zemí se odmítli spojit a usilovali o získání vlastního regionu. ICANN na tyto tlaky reagoval vydáním podmínek pro vytvoření nových RIR. Ve stručnosti obsahují následující požadavky:

- Uchazeč o RIRství musí pokrýt velkou geografickou oblast.
- Musí prokázat, že jej podporují LIRové v dané oblasti.
- Musí vypracovat otevřené a transparentní postupy pro správu svěřených zdrojů.
- Musí být neutrální, zejména vůči svým LIRům.
- Musí být technicky schopen vykonávat tuto úlohu.
- Jeho postupy musí odpovídat globálním pravidlům pro přidělování adres.
- Měl by předložit plán aktivit podporovaný svými svěřenci.
- Musí publikovat model svého financování (doporučenou formou je nezisková společnost).
- Musí vést evidenci spravovaných a přidělených zdrojů.
- Musí zajistit důvěrnost informací, které mu LIRové poskytují⁷.

Zajímavou částí dokumentu je i důrazné doporučení, že RIR by neměl mít nic společného se správou doménových jmen pro "normální" (dopředné) dotazy. RIR je z principu věci angažován v přidělování reverzních domén, které bezprostředně souvisejí s adresami. Běžné domény by však měly jít zcela mimo regionálního registrátora.

Zmíněný dokument vznikl na objednávku, protože tou dobou byl zjevný zájem o zřízení nových RIR ve dvou oblastech - Latinské Americe a Africe. V současné době proces jejich vzniku spěje ke konci a lze očekávat, že během letošního roku začnou oba noví regionální registrátoři pracovat. Jedná se o:

LACNIC	(Latin American and Caribbean Network Information Centre) Latinská Amerika, Karibik
AfriNIC	(African Network Information Centre) Afrika

Zdroj: *Změny v přidělování adres* [online]. Dostupné z: <http://www.lupa.cz/clanky/zmeny-v-pridelovani-adres/>

⁷ Více viz: <http://www.icann.org/en/aso/emerging-rirs-01oct00.htm>

2.3.2 Doménové adresy

Vzhledem k tomu, že číselné IP adresy jsou těžko zapamatovatelné, používá se místo IP adresy síťového rozhraní název síťového rozhraní. Pro každou IP adresu je zavedeno jméno síťového rozhraní (počítače), přesněji řečeno doménové jméno. Jedna IP adresa může mít přiřazeno žádné, jedno nebo několik doménových jmen. Převod číselných adres na adresy doménové zajišťují systémy Domain Name Service (DNS). Doménové jméno se skládá z řetězců vzájemně oddělených tečkou.

Domény 1. stupně (Top Level Domain)⁸

A. Generické TLD (gTLD) – mají tři a více znaků (com., org., gov., edu., mil., biz., name., museum., net. Atd)

B. Národní TLD (ccTLD) – jsou dvouznakové. Tyto znaky jsou totožné s identifikací země podle normy ISO-3166⁹.

⁸ O přidělování TLD se stará agentura ICANN (www.icann.org)

Pro ČR: agentura NIC.CZ (www.nic.cz)

⁹ http://cs.wikipedia.org/wiki/ISO_3166-1

2.3.3 Pravidla pro tvorbu doménových jmen

Doménová jména musí vyhovovat normám RFC 1034, 1035, 1122, 1123 a jakýmkoliv je nahrazujícím nebo doplňujícím normám. Jedná se zejména o tyto podmínky:

- doménové jméno může obsahovat pouze znaky [a-z,0-9,-];
- délka doménového jména je nejvýše 63 znaků;
- doménové jméno nesmí začínat, ani končit znakem „-“;
- doménové jméno nesmí obsahovat dva znaky „-“ za sebou¹⁰.

Nejdůležitější a zajímavé RFC dokumenty, vztahující se k DNS [ŠŤASTNÝ, 2007]:

[RFC 1034](#) – Domain Names - Concepts and Facilities

[RFC 1035](#) – Domain Names - Implementation and Specification

[RFC 1537](#) – Common DNS Data File Configuration Errors

[RFC 1886](#) – DNS Extensions to support IP version 6

[RFC 1912](#) – Common DNS Operational and Configuration Errors

[RFC 1996](#) – A Mechanism for Prompt Notification of Zone Changes (DNS NOTIFY)

[RFC 2181](#) – Clarifications to the DNS Specification

[RFC 2308](#) – Negative Caching of DNS Queries

[RFC 2870](#) – Root Name Server Operational Requirements

[RFC 3330](#) – Special-Use IPv4 Addresses

[RFC 3467](#) – Role of the Domain Name System

Litertura:

1. PUŽMANOVÁ, Rota. *TCP/IP v kostce*. 2. uprav. a rozšiř. vyd. České Budějovice: Koop, 2009. 619 s. ISBN 8978-80-7232-388-3.
2. DOSTÁLEK, Libor, KABELOVÁ, Alena. *Velký průvodce protokoly TCP/IP a systémem DNS*. 5. aktualit. vyd. Brno: Computer Press, 2008. 488 s. ISBN 978-80-251-2236-5.
3. MUSIL, Marek. *Historie internetu* [online]. Poslední modifikace 23. července 2003. [cit. Dne 1. března 2010]. Dostupné z: <http://ihistory.webzdarma.cz/chap/sites/RFC.php>
4. ŠŤASTNÝ, Petr. *Dokumenty RFC a jiné – DNS* [online]. Poslední změna 29. dubna 2007. [cit. Dne 1. března 2010]. Dostupné z: <http://www.pweb.cz/dns/rfc.html>
5. *Vlastnosti protokolů* [online] Poslední editace 6. srpna 2008. [cit. Dne 1. března 2010]. Dostupné z: https://www.ipv6.cz/Vlastnosti_protokolu

¹⁰ Více na: http://www.nic.cz/files/nic/doc/Pravidla_registrace_CZ_ccReg_20080930_fin.pdf

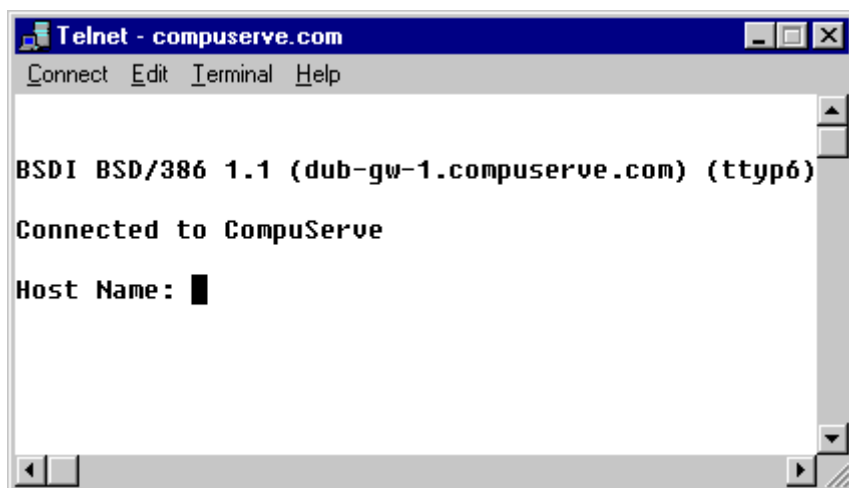
3. SLUŽBY INTERNETU

V současné době Internet nabízí nepřehledné množství služeb. Historicky nejstaršími službami byly elektronická pošta, přenos souborů a vzdálené přihlašování (tzv. remote login). V současnosti je maximálně využívanou službou WWW.

3. 1 Vzdálené přihlašování - Remote Login

Vzdálené přihlašování se realizuje protokolem Telnet, podle kterého se také obvykle jmenují utility (programy) které jej implementují. Uživatel si na svém počítači spustí program Telnet a přihlásí se ke vzdálenému počítači (zadá jméno svého uživatelského účtu na tomto počítači a heslo). Pak již může na tomto počítači pracovat stejně, jako kdyby seděl na skutečném terminálu ve fyzické blízkosti vzdáleného počítače.

Problémem vzdáleného přihlašování je ale jeho malá uživatelská přítulnost – pracuje pouze v textovém režimu.



Obr. č. 3 : Přihlášení k telnetu

Zdroj: *TechTrain* [online].

3.2 Přenos souborů pomocí FTP

FTP neboli File Transfer Protocol je služba, která umožňuje práci se vzdálenými soubory a to převážně jejich upload (kopírování na vzdálený server) nebo download (stažení), dále pak mazání, změnu práv, práci s adresáři atd. FTP lze využívat ve dvou případech:

- download z anonymních FTP serverů
- administrace webových stránek.

První návrh specifikace tohoto protokolu je již z roku 1971 (RFC 114), současná podoba je popsána ve standardu RFC 959.



FTP root at ftp.microsoft.com

```
This is FTP.MICROSOFT.COM
230-Please see the dirmap.txt file for
230-more information. An alternate
230-location for Windows NT Service
230-Packs is located at:
230-ftp://198.105.232.37/fixes/
```

```
03/13/98 10:09PM      Directory bussys
11/05/97 12:00AM      Directory deskapps
12/12/97 12:00AM      Directory developr
11/05/97 12:00AM           8,102 dirmap.htm
11/05/97 12:00AM          4,405 dirmap.txt
04/13/93 12:00AM           710 DISCLAIM.TXT
```

Obr. č. 4: Kořen protokolu FTP na <ftp://ftp.microsoft.com>
Zdroj: *TechTrain* [online].

Příklad: Spustit příkazový řádek ve Windows; např. Start > Spustit: command .

```
V příkazovém řádku zadat:
c:\>ftp ftp.linux.cz
Uživatel: anonymous
Heslo: vas@email.cz
ftp>help
ftp>ls
ftp>quit
```

3.3 Elektronická pošta

MF Dnes, 18.2. 1994, Washington: “ Americký president Bill Clinton a švédský premiér Carl Bild se stali prvními nejvyššími představiteli, kteří si vyměnili diplomatickou poštu pomocí světové elektronické sítě Internet “.

Elektronická pošta patří mezi nejstarší a nejoblíbenější služby všech počítačových sítí včetně Internetu. Její obliba a význam jsou tak velké, že dokonce mění zvyky a chování uživatelů.

Přednosti elektronické pošty

- rychlost
- její samotná elektronická forma – tím, že textové zprávy jsou elektronickými dokumenty, lze je snadno archivovat, zpracovávat, tisknout, vyhledávat v nich apod., množství úkonů lze zautomatizovat a zvýšit tak pohodlí a výkonnost uživatele – adresáře, šablony, e-podpis apod.
- možnost rozesílání zpráv mnoha příjemcům najednou – každá poštovní zpráva může mít až tři druhy adresátů najednou:
 - **hlavní příjemce** (v angličtině příjemce v kategorii To: - jemu je zpráva adresována
 - **příjemce kopie** (příjemce v kategorii Cc:, z anglického Carbon Copy, doslova průklep – tomuto příjemci odpovídá u úředních dopisů zaslání „na vědomí“). Důležité je, že příjemce kopie pozná její statut, neboli vidí že hlavním příjemcem je někdo jiný a on dostává pouze kopii na vědomí. Stejně tak hlavní příjemce se dozví, kdo dostal kopii zprávy.
 - **příjemce slepé kopie** (příjemce v kategorii Bcc., z anglického Blind Carbon Copy) – příjemce této slepé kopie vidí, komu je zpráva adresována (včetně kopií), vidí i její obsah, ale naopak to neplatí. Hlavní příjemce ani příjemce kopie se nedozví, kdo dostal slepou kopii. V žargonu existuje pro tento druh slepé kopie název “bonzácká kopie“

Pro tvorbu a správné doručení elektronického dopisu jsou zapotřebí dva typy programů:

1. Mail Transfer Agent (MTA) – tento program běží na serveru a má za úkol s ostatními MTA na jiných serverech zajišťovat doručení dopisu do poštovní schránky správného adresáta. Různí MTA se dorozumívají pomocí protokolu **SMTP** (Simple mail transfer protokol)

2. Mail User Agent (MUA) - s tímto programem pracuje bezprostředně uživatel. Tento program tedy vytváří uživatelské rozhraní pro práci s poštou. Pro přístup do poštovní schránky musí použít MUA speciální protokoly:

a) Post Office Protokol 3 (POP3) – umožňuje vybrání schránky, čili přesun dopisu ze vzdáleného počítače na lokální počítač. Protokol má některé slabiny – nelze dopisy nijak filtrovat – vždy se přenášejí všechny, nejde vybrat dopisy třeba jen podle předmětu, není možné přenášet jen části dopisů atd.

b) Internet Message Access Protokol 4 (IMAP4) – protokol IMAP se snaží odstranit nevýhody protokolu POP3 – např. umožňuje přenést jen záhlaví dopisu apod.

Email a Netiquett

Netiquett (z angl. slova Network Etiquett, síťová etiketa), tzn. netiket

7 pravidel, jak napsat e-mail podle netikety:

1. Adresování aneb kolika lidem píšete

Posíláme-li e-mail více příjemcům, je většinou diskrétnější uvést jejich seznam skrytý, tedy v políčku BCC, skrytá kopie (Blind Carbon Copy). Výjimka je, když chceme, aby lidé v seznamu o sobě věděli – například pořádáme sraz třídy. Jednotlivé adresy oddělujeme čárkou.

2. Předmět zprávy

Předmět zprávy má za úkol zaujmout (podobně jako titulek v novinách). Pokud dobře napíšeme předmět, nezapadne v množství přijatých zpráv a je jednodušší se zprávami manipulovat. **Předmět je vlastně stručný, výstižný text, který charakterizuje obsah dokumentu.**

Příklad: Šéfredaktorovi časopisu Zahrádkář posíláme článek o přemnožení slimáků na Valašsku. Špatně: Zpráva. Hlemýžď se přemnožil. Přemnožení škůdce. Předmět nejasný, nevíme zda je od rozzuřeného zemědělce, nebo od ochránců přírody. Lépe: Přemnožení hlemýžďů ve středních Čechách – článek k publikování..

Do předmětu zprávy je možné u obecných e-mailových adres napsat konkrétního příjemce. Pro paní Novákovou, předmět: učební plán VOŠ. Mimochodem pokud firma uvádí na svých www stránkách pouze jednu povšechnou adresu, svědčí to o jejím značně amatérském přístupu k zákazníkovi.

Špatné příklady předmětu: Zpráva, Problém, Prosba, Co mám dělat?

3. Priorita zprávy čili důležitost

Každé zprávě můžeme přiřadit tzv. prioritu, většinou volíme ze tří úrovní důležitosti: nízká – normální – vysoká. Jakou zvolíme prioritu nemá žádnou souvislost s rychlostí doručení zprávy, jen to jen upozornění pro příjemce.

4. Začínáte oslovením

I na začátek elektronické zprávy patří oslovení, nezapomínat na skloňování ...pane Nováku.

5. Vlastní text

Platí pravidlo „jedno téma jeden mail“, snažíme se být struční. V emailu není vhodné psát VERZÁLKAMI, takto psaný text je obecně chápán jako křičení. Pokud chceme nějaké slovo zdůraznit a z jakéhokoliv důvodu nemůžeme použít formátování textu (např. tučné), pomůžeme si hvězdičkou a podtržítkem. * na začátku a konci slova stimuluje tučné písmo, podtržítka kurzívu.

Opatrně s emotikony. Maximálně 2-3 v soukromé korespondenci, v obchodním styku na ně raději zapomeňme.

Používat háčky a čárky? V současné době většina uživatelů dokáže přijímat diakritiku¹¹.

6. Ani e-mailu nesmí chybět podpis

Na konec zprávy patří rozloučení a podpis. Podpis je možné předem vytvořit a nemusíme ho pokaždé psát, můžeme mít několik popisů pro různé formy písemností.

Oficiální podpis pod e-mailem obsahuje: jméno a příjmení, , pozice v zaměstnání, jméno firmy, (její webovou stránku), elektronická adresa a telefonní a faxový kontakt.

7. Opatrně s přílohami

U posílání příloh se tedy držíme několika zásad:

u příloh nad 200 KB předem upozorníme adresáta

nikdy neposíláme coby přílohu spustitelné soubory (přípona EXE, COM, BAT, VBS, WSH, SCR) a dokumenty obsahující makra (makro je uložena posloupnost činností, kterou lze vyvolat jediným příkazem)

¹¹Podrobnější informace na <http://www.cestina.cz>

Základní pravidla pro odeslání odpovědi

Jak rychle odpovědět? Samozřejmě záleží na prioritě zprávy, u důležité nejlépe ihned, jinak je běžná praxe ve firemním styku – **odpověď do druhého dne**.

Píšeme-li odpověď, poštovní klient vytvoří novou zprávu, do které umístí text, na který odpovídáme. Nemažte ho – příjemce lépe ví, na co odpovídáte. Můžeme smazat jen věty, které nesouvisí s jádrem problému.

Respektujte soukromí. Pokud vám někdo pošle osobní zprávu, neposílejte ji dále bez souhlasu odesílatele.

Neměňte obsah přeposílaného e-mailu bez svolení odesílatele – on je autor.

Z firemních e-mailových adres neposílejte řetězové dopisy

Pokud víte, že delší dobu budete mimo vaši e-mailovou schránku, je dobré se odhlásit z konferencí a ve schránce nastavit automatickou odpověď.

Pozor na spamming

Od 7. září 2004 začal platit nový Zákon o některých službách informační společnosti (č. 480/2004 Sb.), který problematiku spamu upravuje a vyžaduje prokazatelný souhlas příjemce zprávy. Dohledem nad dodržováním zákona byl pověřen Úřad pro ochranu osobních údajů. Novely zákona – v letech 2005, 2006, 2007.

Zákon byl vytvořen podle směrnice Evropského společenství č. 2000/31/ES. Spam definuje jako obchodní sdělení, což jsou všechny formy sdělení určeného k přímé či nepřímé podpoře zboží či služeb nebo image podniku fyzické či právnické osoby. Zákon řeší nejen internetový spam, ale také jiné formy elektronické komunikace (SMS, telemarketing).

Podle zákona se za obchodní sdělení nepovažují údaje umožňující přímý přístup k informacím o činnosti fyzické či právnické osoby nebo podniku, zejména doménové jméno nebo adresa elektronické pošty; za obchodní sdělení se dále nepovažují údaje týkající se zboží, služeb nebo image fyzické či právnické osoby nebo podniku, získané uživatelem nezávisle. [4]
Šíření obchodních sdělení upravuje § 7 tohoto zákona.

DESATERO PŘIKÁZÁNÍ POČÍTAČOVÉ ETIKY

- Nepoužiješ počítače ke škodě jiného.
- Nebudeš ničivě zasahovat do práce druhých lidí.
- Nebudeš slídit v souborech jiných lidí.
- Nepoužiješ počítače ke krádeži.
- Nepoužiješ počítače pro křivé svědectví.
- Nepoužiješ nebo nepořídíš kopii softwaru, který jsi nezaplatil(a).
- Nepoužiješ neoprávněně počítačového zdroje jiných lidí.
- Nepřivlastníš si intelektuální dílo jiného.
- Budeš přemýšlet o společenských následcích programu, který jsi stvořil(a).
- Budeš používat počítače ohleduplně a s úctou.¹²

3.4 Gopher

Služba Gopher byla vyvinuta v USA, na univerzitě v Minnesotě, o něco málo dříve než World Wide Web. Poté spolu obě služby soupeřily, s tím, že zpočátku měl navrch (co do počtu serverů) Gopher. Kolem roku 1995 bylo celosvětově v provozu přes 6000 serverů této služby. S postupem času se ale stále více prosazoval Web. Hlavním důvodem jeho prohry byl především neatraktivní „obal“. V současné době stále funguje několik Gopher serverům, které na Internetu udržuje pár nadšenců¹³.

¹² Převzato z The Computer Ethics Institute, <http://www.computerethicsinstitute.com/>

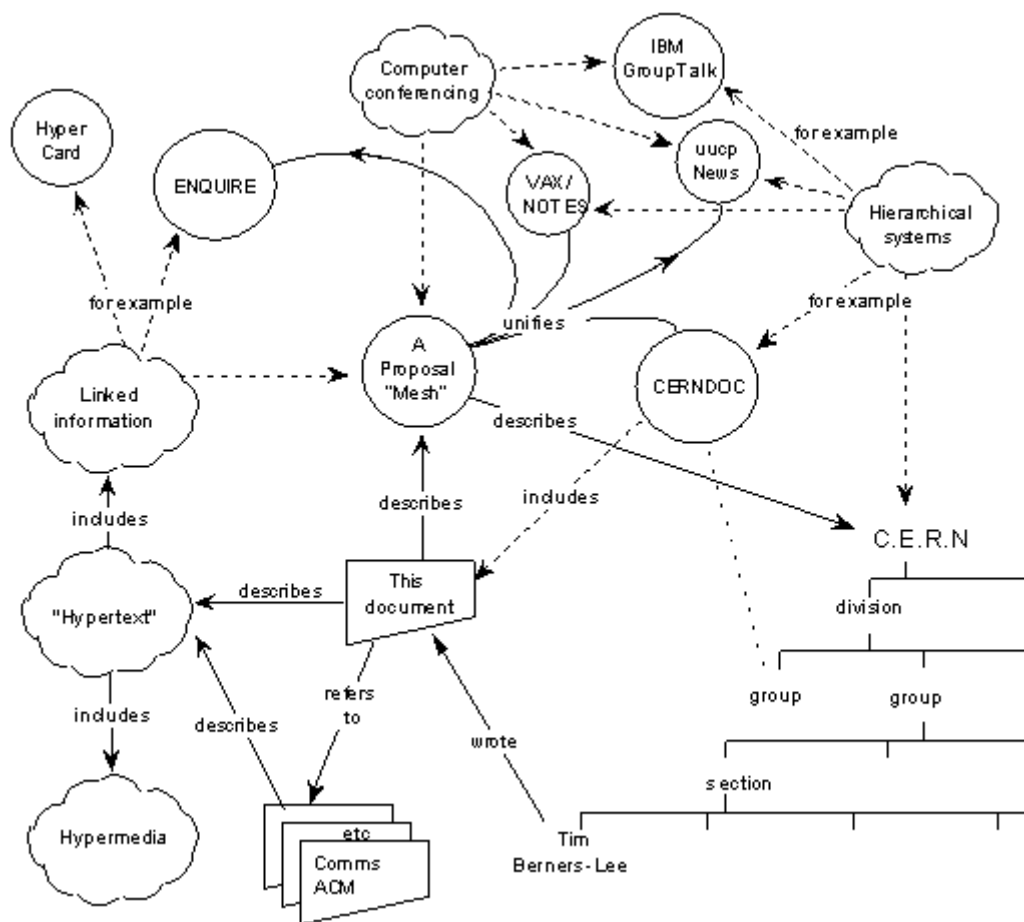
¹³ <http://galerie.earchiv.cz/videos/gophers.php3> - Videá o službě Gopher
<http://galerie.earchiv.cz/scshots/gopher/index.php3> - Služba Gopher - Screenshoty

3.4 World Wide Web – WWW

Dnes nejznámější a nejoblíbenější služba Internetu. Tato služba začala jako poměrně jednoduchá služba pro zpřístupnění textových informací. Dnes je spíše platformou pro poskytování celé řady dalších konkrétních služeb.

Stručná historie:

- 1945
 - Vannevar Bush – (1890 – 1974) publikuje článek „As We May Think“ v časopise The Atlantic Monthly [1945], zamýšlí se nad problémem narůstajícího množství dat, článek se zabývá využitím přístrojů pro komunikaci, zmiňuje se o zařízení Memex (MEMory a Extended, rozšířená paměť), fungujícího na principu hypertextu., předpověď možného propojení více smyslů při komunikaci s přístrojem – dnes multimédia
- 60. léta 20. st
 - Doug Engelbart - Stanford Research Institute, ONLine Systém (NLS) – první hypertextové zařízení
 - Ted Nelson – 1963 definoval pojem hypertext, 1965 pojem hypertext publikoval
- 1989
 - Tim Berners Lee, CERN – návrh webového dokumentu



Obr. č.5 : Information Management. A Proposal
Zdroj: BERNERS-LEE, 1989.

Web je informační systém založený na hypertextové technologii, která zajišťuje velmi progresivní a atraktivní komunikaci elektronických (digitálních) informačních zdrojů (textových, obrazových...) jejich koncovým uživatelům. Elektronické textové zdroje jsou tvořeny na bázi jazyka HTML a jsou fyzicky ukládány na počítačích označovaných www servery. Ukládání není živelné. Má pevná pravidla. Uživatel využívá tyto dokumenty v rámci komunikace www server/www klient prostřednictvím protokolů HTTP. Optimální využití všech možností www zajišťují jen klienti (PC, terminál) pracující v grafickém režimu (např. Netscape, Explorer apod.). Přístup k dokumentům se pak v rámci klientů realizuje jednoduše přes adresu dokumentu v prostoru www zapsanou ve formě URL. Adresa umožňuje jednoznačně identifikovat jakýkoliv objekt v Internetu.

WWW je postaven na architektuře klient/server. Hlavním úkolem klienta je komunikovat s uživatelem a vytvářet mu potřebné uživatelské rozhraní, včetně zobrazování obsahu jednotlivých webových stránek. Serverem je počítač, na kterém jsou umístěny jednotlivé webové stránky. Webový server sám nikomu tyto stránky nevnučuje - pouze pasivně čeká, až některý klient doslova řekne o nějakou konkrétní webovou stránku. K tomu dojde v okamžiku, kdy uživatel nějakého webového klienta (browseru) klikne na nějaký aktivní hypertextový odkaz, čímž dává najevo, že se chce přesunout na jinou webovou stránku. Aby toto vše mohlo fungovat, musí existovat vhodná konvence pro vzájemnou komunikaci mezi webovým serverem a klientem. tato konvence má podobu přenosového protokolu HTTP (HyperText Transfer Protokol), prostřednictvím něhož klient vznáší své požadavky na server a server zasílá požadované stránky. Další částí této konvence je jazyk HTML (HyperText Markup Language) - jde o jazyk ve kterém jsou psány samotné webové stránky tak, aby klient jejich obsahu rozuměl

3.6 Elektronické konference a diskusní skupiny

A. Elektronická konference

- pracuje s využitím elektronické pošty
- účastníci diskutují o předem smluvených tématech
- po přihlášení je elektronická adresa přidána do **seznamu účastníků konference**
- jakýkoliv **dopis** zaslaný na adresu konference je **automaticky rozeslán na adresy všech účastníků** dané konference

Podle způsobu rozesílání příspěvků lze konference rozdělit na:

nemoderované - každý příspěvek zaslaný do konference je automaticky rozeslán

moderované - každý příspěvek jde **přes moderátora konference**

Z hlediska členství lze konference rozdělit na:

otevřené

uzavřené

Jak pracuje elektronická konference:

Celý provoz konference řídí **specializovaný program**, který běží na nepřetržitě spuštěném počítači – serveru. Nejpoužívanější jsou programy označené jako **majordomo**, **listserv**, nebo **listproc**. Tento program zpracovává zprávy, které došly na tzv. **administrativní** nebo **distribuční** adresu konference.

Administrativní adresa umožňuje účastníkům:

- přihlášení
- odhlášení
- dočasné odhlášení
- získání seznamu účastníků dané konference
- získání seznamu konferencí, které daný server provozuje

Administrativní adresu najdete nejčastěji v některém z následujících tvarů:

listserv@počítač.firma.cz

→ např. listserv@cesnet.cz

Na administrativní adresu se posílají pouze příkazy pro server.

Distribuční adresa

- je adresou, na kterou účastníci konference posílají své příspěvky

Jak se do konference přihlásit:

1. Na administrativní adresu např.

listserv@cesnet.cz

pošlete jednoduchý dopis, který bude obsahovat jedinou řádku:

subscribe název konference jmeno prijmeni

Jak se z konference odhlásit

Na administrativní adresu např. listserv@cesnet.cz pošlete dopis s jedinou řádkou

unsubscribe nazev konference

Příkazy:

Přihlášení – subscribe,

Odhlášení – unsubscribe/ signoff

Popis příkazů - help

Seznam účastníků - who název konference

Dočasné pozastavení zaslání příspěvků – set nomail nazev konference

Obnovení zaslání – set mail nazev konference

Pro experimentování je možné využít například tyto zkušební konference

administrativní adresy	jejich názvy	Distribuční adresy
listserv@cesnet.cz	demo-1	demo-1@cesnet.cz
listserv@cesnet.cz	cs-test	cs-test@cesnet.cz
majordomo@list.zcu.cz	Test	test@list.zcu.cz

B. Diskusní skupiny

- rozdíl: u elektronických konferencí „každý dostává všechno“, u diskusních skupin „nikdo nedostává nic“

Diskusí skupiny fungují na stejném principu jako web. chat, ale všechny příspěvky zůstávají na serveru uloženy a lze je kdykoliv zobrazit

C. Další komunikační služby

MSN Messenger: <http://messenger.msn.cz/>

SKYPE: <http://www.skype.com/>

Yahoo! Messenger: <http://messenger.yahoo.com/>

ICQ: <http://www.icq.com>

Google Talk: <http://www.google.com/talk/>

Literatura:

1. BERNERS-LEE, T. *Information Management: A Proposal* [online]. March 1989, May 1990. [cit. 2010-04-17]. Dostupné z: <http://www.w3.org/History/1989/proposal.html>.
2. BUSH, V. *As We May Thing* [online]. July 1945, Atlantic Magazine [cit. 2010-04-02]. Dostupné z: <http://www.theatlantic.com/magazine/archive/1969/12/as-we-may-think/3881/>
3. *Čeština* [online]. [cit. 2010-04-17]. Dostupné z: <http://www.cestina.cz>
4. *Doug Engelbart Institute* [online]. c2008. [cit. 2010-04-02]. Dostupné z: <http://dougengelbart.org/>
5. NEJEDLÝ, J. – VRÁNOVÁ, D. *Výklad zákona o některých službách informační společnosti*. 1. vyd. Vyškov: TSM, 2005. 82 s. ISBN 80-239-5020-7.
6. PETERKA, Jiří. *Galerie archivu aneb Muzeum českého Internetu od roku 1994 do roku 1999*. [online] 2008. [cit. 2010-04-17]. Dostupné z: <http://galerie.earchiv.cz/gopher.php3>
7. *TechTrain* [online]. cTechnology Training Team. [cit. 2010-04-17]. Dostupné z: <http://www.techtrainteam.com/services/olt/internet/ftp-0.html#gopherheading>

4. ÚVOD DO STRATEGIE VYHLEDÁVÁNÍ V INTERNETU

Internet lze přirovnat k obrovské knihovně, do které někdo uložit mnoho klenotů světového vědění, praktických příruček, návodů i zábavného čtení a stejně tak mnoho vysloveně brakové literatury.. Hlavně ale v takové knihovně zlikvidoval všechny kartotéky, rejstříky či jiné druhy členění, všechny časopisy a knihy naházet na jednu hromadu. Abychom se v této velké hromadě vyznali, potřebujeme vyhledávací stroje, vyhledávací služby, které nám pomohou problém vyřešit. S určitou nadsázkou lze konstatovat, že v Internetu je sice všechno, ale je těžké to najít. Naproti tomu v knihovnách je jenom něco, ale to něco je systematicky uspořádáno, a tudíž to lze nalézt jednodušeji. Je taktéž potřeba si uvědomit, že Internet sám žádné informace neprodukuje, je pouze přenosovou infrastrukturou, která umožňuje pracovat s informacemi.

J. Peterka [1998] uvádí několik důvodů, proč je práce s informacemi v Internetu tak komplikovaná

- **neexistence povolovacího a schvalovacího řízení**

Snad největší odlišností dnešního Internetu od klasických médií je jeho velká dostupnost pro publikující subjekty. Chce-li někdo vysílat vlastní televizní programy, musí založit vlastní televizní společnost, získat potřebnou licenci, vybavit televizní studia, získat televizní vysílače pro šíření signálu atd. Chce-li někdo vydávat noviny, musí založit redakci, najít tiskárnu, získat inzerenty, opatřit si potřebné vydavatelské povolení atd. Naproti tomu publikovat na Internetu může každý. Nákladová i jiná náročnost publikování vlastních informací na Internetu je nesrovnatelně nižší než u všech ostatních médií. V Internetu navíc platí, že ten, kdo chce zpřístupnit vlastní informace jiným lidem, se nemusí nikoho ptát, zda to smí udělat, ani nemusí nechat schválit způsob, jakým to chce udělat. Na jedné straně je to velmi demokratické, na druhé straně se na Internetu objevují věci, které hraničí se zákonem.

- **neexistence katalogizační povinnosti**

Neexistence jakéhokoliv schvalování má vedle svých etických a morálních důsledků i velmi významné důsledky technického rázu – nikdo totiž zveřejňovatele neovlivňuje ani neřídí v tom, jakým konkrétním způsobem mají své informace zveřejňovat a jaké přitom mají dodržovat konvence. Chybí mechanismus, který by do zveřejňování informací vnesl určitý ráz.

- **problém věrohodnosti informací**

Protože neexistuje jakékoliv schvalovací a povolovací řízení, odvíjí se od toho problém posuzování věrohodnosti informací. Jak se tedy stavět k informacím na Internetu a jak je hodnotit? Možným vodítkem je identita zdroje informací neboli toho, kdo informace na Internet pouští – zda jde o stránku renomované firmy či jde o stránku pod kterou se nikdo nepodepsal.

- **problém kvality informací**

Vzhledem k obrovské kvantitě informací na Internetu vyvstává problém s kvalitou informací

- **neexistence oznamovací povinnosti**

Internet nemá žádného vlastníka, tudíž neexistuje nikdo, kdo by udržoval centrální evidenci přijatých informací. Existují jen dílčí evidence.

- **neexistence absolutní vyhledávací služby**

Neexistence jakékoliv centrální evidence a neexistence oznamovací povinnosti mají velmi významné důsledky pro všechny uživatele, kteří chtějí na Internetu něco najít. Neexistuje-li žádná závazná centrální evidence, pak zde neexistuje žádná vyhledávací služba, která by přesně dokázala odpovědět na otázku, zda něco konkrétního na Internetu existuje či ne. Závěr: pokud uživatel hledá konkrétní informaci a nenalezne ji pomocí jednoho vyhledávače, neznamená to, že informace na Internetu není. Je třeba použít jiný vyhledávač a porovnat výsledky.

4.1 Jak vyhledávat - tipy pro snazší hledání

1. Buďte přesní, tzn. používejte více klíčových slov. Vyhledávací enginy pracují lépe, když zadáte kombinaci více slov, která vás neodvedou na irelevantní stránky. Požadavek "armstrong nasa apollo moon landing" vás přivede k výsledku snadněji, než když budete zkoušet slova jedno po druhém.
2. Přečtěte si instrukce. Každý vyhledávač má vlastní syntaxi. Vložíte-li dotaz ve správné podobě, najdete to co hledáte, jinak se budete probírat hromadou nikam nevedoucích odkazů.
3. Používejte Boolean. Některé vyhledávače vám umožní vkládat dotazy s pomocí booleanovské logiky.
4. Buďte pokročilí. Mnoho serverů nabídne možnost pokročilého vyhledávání s více hesly a kategoriemi. To se vám může hodit, pokud zadáváte dotaz příliš obecný.
5. Pokud neuspějete napoprvé, zkuste jiný vyhledávač.
6. Uložte si stránky. Jakmile najdete stránky, které hledáte, zase je neztrat'te. Přinejmenším si je uložte do záložky, ale pokud se jedná o užitečné informace, nebojte se je zazálohovat na pevný disk

4.1.1 Hlavní kroky vyhledávací strategie

Pro úspěšné hledání na Internetu je důležité osvojení si základní rešeršní strategie, která se skládá z následujících kroků:

1. Slovní formulace tématu (souvislou větou či několika větami)
2. Výběr vhodných klíčových slov, jenž vyjadřují hledané téma, včetně synonym a příbuzných výrazů
3. Předběžná formulace dotazu, včetně rozhodnutí o tom, zda použít jednoduché hledání nebo pokročilé metody hledání,
4. Výběr vyhledávacího nástroje
5. Prostudování nápovědy zvolené vyhledávací služby
6. Formulace dotazu v souladu s funkcemi dané vyhledávací služby
7. Průzkum a zhodnocení výsledků vyhledávání
8. Nejste-li spokojeni s výsledkem hledání, měla by následovat úprava (oprava) dotazu a opětné zhodnocení výsledků vyhledávání
9. Podle potřeby zvolit jinou vyhledávací službu

4.2 Vyhledávací dotazy

V rámci vyhledávání dělají uživatelé často chybu, že se zaměřují pouze na vyhledávání v tzv. režimu „simple“. Je velmi důležité využít všechny možnosti rešeršních strategií.

4.2.1 Jednoduché dotazy

Jednoduché dotazování nabízí všechny vyhledávací stroje a je označován jako „simple query“. Předností jednoduchých dotazů je to, že uživatel si nemusí nic pamatovat, nic se učit. Nevýhodou je širší výsledků, vyhledávací stroj vrací tisíce odkazů, přičemž většina z nich nesouvisí s informačním požadavkem uživatele.

A. Klíčová slova

Velmi důležité pro úspěch hledání je správná volba klíčových slov. Jednou z největších chyb, které uživatelé při sestavování dotazu dělají je použití nedostatečného počtu klíčových slov. V průměru používají uživatelé dvě klíčová slova na dotaz. Klíčová slova lze kategorizovat:

- stopslova (předložky, spojky atd., tedy slova, která nemají pro vyhledávání žádný informační význam, ve většině případů je vyhledávací stroje ignorují.
- objekty (podstatná jména)
- akce (slovesa)
- modifikátory (přídavná jména, příslovce, přísudky)

Jak postupovat při správné volbě klíčových slov?

- Je třeba být co nejvíce konkrétní.
 - co nejvíce zúžit význam slova
 - používat fráze, popř. celé věty
 - používat více klíčových slov
- První slovo je určující
 - discography Beatles nebo Beatles discography? První dotaz spíše přinese seznam různých stránek o Beatles, kde se také může vyskytovat jejich diskografie. Druhý dotaz přinese stránky různých diskografií různých kapel, kde se rovněž může vyskytovat diskografie Beatles. Při zadávání dotazu musíme mít jasno, co je pro nás důležitější.
- Je třeba používat neobvyklá, signifikantní slova a jejich kombinace
 - Když zadáme slovo nebo frázi, které je velmi běžné, příliš často se vyskytující, vrátí se nám milióny výsledků. Například když chceme informace o skupině the Frames a zadáme toto slovo do počítače – slovo je v angličtině často používané, je tedy lepší zadat do vyhledávání jména členů kapel, jména skladeb apod.

Další tipy

- Nepoužívat synonyma v jednom vyhledávání (jde o jednoduché vyhledávání). Je to nebezpečné hlavně tehdy, pokud nevíme, jaký operátor má daný vyhledávač jako předdefinovaný (default operator).
- Listovat dál, jsou-li výsledky slibné.
- Zkoušet to stále znovu, jedno z pravidel Internetu zní – na Internetu je toho podstatně více, než najdete.
- Hledat na správném místě - není dobré vždy oslovovat vyhledávače. Pokud například hledáme nějaký výrobek a cenu k němu, je lepší se podívat do obchodního domu nebo aukčního centra.
- V dokumentu používejte Ctrl+F (příklad: hledáte slovo trombocytopenie, objeví se velmi dlouhý dokument, ve kterém je „někde“ toto slovo. Při použití této zkratky nemusíme pročítat celý dokument, ale kurzor se nám objeví u daného slova).
- Jak najít synonyma - používání synonym je velmi důležité. Hledáme např. slovo „jazyk“, ale na stránkách může být slovo „řeč“, „mluva“. Pokud nás žádné synonymní slovo nenapadá, použijeme pomůcku Wordu. Najedeme kurzorem na slovo, u kterého chceme znát synonyma a stiskneme klávesu Shift+F7, objeví se Thesaurus.

B. Možnost zkráceného zápisu – hvězdičková konvence

Jednou z prvních chyb při formulování dotazu je používání pouze přesného znění slov. Přitom možnost zkráceného zápisu vyhledávaného slova patří mezi základní funkce vyhledávacích strojů. Zástupný znak (hvězdička) následuje po několika počátečních písmenech (kořenu) termínu (musí být alespoň tři písmena). Zástupný znak nahrazuje libovolný počet jakýchkoliv znaků. Zástupný znak je velmi výhodné používat i v případech, kdy nevíme, jak se správně hledané slovo píše, popřípadě u slov s delším kořenem. Zkracování má i své nevýhody. Slabinou je, že zástupný znak nahrazuje libovolný počet jakýchkoliv znaků. Např. na zadání kořene card budou nalezeny dokumenty, které obsahují nejen card a cards, ale i cardboard (lepenka), Cardiff nebo cardinal (podstatný, kardinální), cardigan (pletená vesta)

C. Fráze

Jednou z cest přesného vyhledávání je používání frází. Fráze jsou tvořeny kombinací slov, které musejí být přítomny v dokumentu přesně v tom pořadí, v jakém je zadal uživatel. V dotazu se fráze zapisují pomocí uvozovek, např. „finanční účetnictví“.

Fráze je vhodné používat vždy, když se jedná o ustálené spojení či sousloví, např. Pražské jaro, Berlínská zeď, solar systém apod.

Možnosti zápisu frází u Alta Visty:

"John Lennon"

John-Lennon

John;Lennon

John:Lennon

John,Lennon

4.2.2 Pokročilé dotazy

Booleovské operátory

Georgie Boole dostal v roce 1844 za svůj článek o operátorech zlatou medaili Královské společnosti. Booleova algebra je základem pro strukturované techniky vyhledávání, je taktéž jedním ze stavebních kamenů moderních počítačů. Základním východiskem Boolovy logiky je teorie množin. Operátor AND je ekvivalentní operaci průniku množin, operátor OR odpovídá sjednocení množin. Pro názorné vyjádření se používají Venovy diagramy.

AND – termíny uvedené na obou stranách operandu se musejí vyskytovat současně někde v dokumentu. Operátor AND lze s výhodou použít ke zřetězování požadovaných termínů nebo frází do jednoho dotazu, přičemž pro zařazení dokumentu do výsledku hledání musí být splněn požadavek současného výskytu všech.

OR – v dokumentu se musí vyskytovat aspoň jeden z uvedených termínů. Operátor OR není ekvivalentní prostému součtu. Dokumenty, které obsahují obě zadané fráze, jsou ve výsledku započítány jen jednou. Jednoduchou úvahou se lze dopracovat k tvrzení, že zaručená velikost výsledku nemůže být nikdy menší než počet dokumentů odpovídající nejčetnějšímu z uvedených termínů nebo frází. Operátor OR lze s výhodou využít ke zřetězení požadovaných termínů nebo frází do jednoho dotazu, přičemž pro zařazení dokumentu do výsledku musí být splněna podmínka přítomnosti libovolného z nich. S operátorem OR je nutno zacházet velmi opatrně. Je vhodný v situacích hledání pomocí pojmů obdobného nebo stejného významu (synonyma). Nadměrné použití operátoru OR vede k nezvládnutelnému množství nalezených dokumentů.

AND NOT – dokument, který obsahuje termín uvedený na pravé straně operandu, je vyloučen z výsledku. Operátor AND NOT vylučuje dokumenty, které obsahují zadaný termín nebo frázi. Týká se jen termínu nebo fráze, které jsou bezprostředně uvedeny za operátorem.. Pokud si představíme za vyhledávanými termíny nebo frázemi množiny dokumentů, ve kterých se vyskytují, pak operátor AND NOT má charakter rozdílu obou množin. Operátor AND NOT musíme používat velmi obezřetně, aby nedocházelo k vyloučení relevantních dokumentů.

NEAR – má podobný význam jako AND. ale oba termíny se musejí vyskytovat v určité vzdálenosti (obvykle měřeno ve slovech) od sebe. Podmínka „blízko sebe“ je vyhodnocována jako určitá maximální vzdálenost mezi oběma termíny. U většiny vyhledávacích strojů je max. vzdálenost nastavena na deset slov. Hlavní přednost operátoru NEAR spočívá ve vzdálenostním omezení výskytu hledaných termínů, což obvykle znamená z hlediska struktury dokumentu jednu větu nebo odstavec. Tímto způsobem jsou vyloučeny rozsáhlé dokumenty s odkazy na vše možné i nemožné, ale z hlediska přímého hledání specifických informací mají nulovou hodnotu. Je nutné ovšem zmínit i určité nevýhody operátoru NEAR.

Např. díky dotazu „knowledge management“ NEAR „information retrieval“ mohou být vyloučeny dokumenty, ve kterých se v jednom odstavci rozebírá problém vyhledávání informací a teprve až v odstavci následujícím se hovoří o managementu znalostí. Je velmi obtížné předvídat, jaký způsob psaní, zacházení s pojmy aj. zvolil autor. Operátor NEAR je zpřísněním operátoru AND. Operátor NEAR lze doporučit v situacích, kdy si je uživatel jist spolehlivosti vzdálenosti hledaných termínů v rámci věty, resp. odstavce.

BEFORE – má podobný význam jako NEAR s tím, že první termín (na levé straně) tohoto operátoru se musí nacházet v dokumentu vlevo od druhého termínu.

AFTER – má podobný význam jako NEAR s tím, že první termín (na levé straně) tohoto operátoru se musí nacházet v dokumentu vpravo od druhého termínu.

Závorky – používají se k ovlivnění pořadí, ve kterém mají být vyhodnocovány uvedené operátory. Zpravidla platí, že nejdříve vyhodnocují vnořené operátory a operátory na stejné úrovni se vyhodnocují zleva doprava Každý výraz uzavřený do závorek se označuje jako boolovský výraz. (Celý dotaz je chápán vždy jako boolovský výraz nezávisle na tom, zda jsou nebo nejsou kolem něho umístěny závorky). Tento přístup je stejný, jako se uplatňuje v matematice při vyhodnocování algebraických výrazů.

Použití filtrů

Filtry umožňují zjemňovat výsledky vyhledávání. Funkce filtrů spočívá v omezení prostoru dokumentů, ve kterém je dotaz vyhledáván. Filtry jsou podporovány různými vyhledávacími stroji různě. Velmi dobře promyšlenou soustavu filtrů má AltaVista, viz následující příklady.

Filtry vyhledávacího systému AltaVista

Filtr uzlu

Filtr uzlu umožňuje omezit prostor vyhledávání na jeden konkrétní uzel nebo na specifickou podmnožinu uzlů podle struktury jejich URL. URL vyjadřuje adresu umístění libovolného informačního zdroje v prostředí Internetu.

Příklad: host:vodafone..cz

Filtr data

Filtry data jsou užitečné při hledání informací, které byly publikovány v určitém období. Zadání datového filtru je u většiny vyhledávacích strojů obvykle jednou z částí formuláře pro pokročilé vyhledávání.

Filtr anchor – (kotva) vyhledávají se dokumenty, které obsahují zadané slovo nebo frázi v textu hypertextového odkazu. Například: anchor:“Click here to visit AltaVista“ budou nalezeny dokumenty, v nichž je uvedený text odkazem, tj.vyhledá stránky s touto frází v textu hyperodkazu.

Filter applet – jsou nalezeny dokumenty se specifikovaným appletem v programovacím jazyce Java. Například: applet:Butterfly vyhledá stránky obsahující jméno Java appletu Butterfly, který je uveden v tagu Applet. Applet = program napsaný v Javě. JAVA – programovací Jazyk vyvinutý firmou SUN. Hl. odlišností Javy je to, že program napsaný

v JAVĚ (tzv. applet) neběží přímo na skutečném počítači, ale na počítači virtuálním, který je samozřejmě aplikován jako specializovaný program na počítači skutečném.

Filtr domain – omezení vyhledávaných dokumentů pouze na řečenou doménu. například domain:cz – budou nalezeny jen dokumenty z České republiky.

Filtr host – slouží k nalezení dokumentů na určitém serveru: například zadáním zadáním: host:www.oavm.cz.cz jsou vyhledány pouze dokumenty uložené na serveru Obchodní akademie a VOŠ Valašské Meziříčí

Filtr image – jsou vyhledány dokumenty, které obsahují obrázek uvedeného jména souboru. Příklad: image:spider.gif

Filtr link – vyhovují dokumenty, z nichž směřuje odkaz na specifikované URL, například pomocí zadání link:www.oavm.cz budou nalezeny, které někde ve svém těle mají odkaz na server OA VM - tímto způsobem se dá zjišťovat citovanost.

Filtr text – zadaný text se musí vyskytovat v textu dokumentu (s vyloučením obrázků, odkazů a URL)

Filtr title – slovo nebo fráze musejí být uvedeny v názvu dokumentu.

Filtr url – hledají se dokumenty jejichž URL specifikace obsahuje zadaný text. Na zadání: url:obchod budou nalezeny dokumenty na různých serverech, které obsahují slovo obchod ve jménu serveru, adresářové cestě, jméně souboru, čili kdekoli v URL adrese.

Je velmi důležité si uvědomit, že každý vyhledávací systém má svou vlastní vyhledávací strategii, která se od dob AltaVisty pro uživatele velmi zjednodušila.

Literatura:

1. PETERKA, J. *Budoucnost práce s informacemi v Internetu. PragueInternet World, 1998* [online]. [cit. 2010-02-22]. Dostupné z: <http://www.earchiv.cz/l311/index.php3>
2. *Search Strategies for Search Engines* [online]. Poslední aktualizace 29. 3. 2008 [cit. 2010-02-22]. Dostupné z: <http://webliminal.com/search/search-web05.html>

5. KLASIFIKACE VYHLEDÁVACÍCH SLUŽEB

V Internetu existuje obrovské množství vyhledávacích služeb. Při zpracování rešeršního požadavku je potřeba se zamyslet nad tím, která služba je pro řešení ta nejvhodnější. Níže je uvedena základní klasifikace.

5.1 Seznamy zdrojů (Resource List)

Seznamy zdrojů jsou monotematicky zaměřené evidence. V současnosti neexistuje žádný seznam seznamu zdrojů, je velmi těžké se v nich orientovat.

Například:

Denní přehled článků o digitálním vysílání
<http://www.digimedia.cz/zdroje/>

Denní přehled článků ze světa mobilní techniky
<http://www.mobilar.cz/zdroje/>

Přehled zdrojů Johna Decembra
<http://www.valeriemates.com/december/resources.html>
<http://www.december.com/john/>

5.2 Předmětové katalogy (Webové portály)

Předmětové katalogy jsou obrovské, polytematicky zaměřené databáze. Dříve byly předmětové katalogy udržovány výhradně ručně. Vzhledem k uspokojování informačních potřeb si kladou 2 cíle [SKLENÁK, 2001, s. 254]:

- snaží se o strukturovaný pohled na informační zdroje v Internetu
- snaží se odstranit nevýhodu robotů, kteří pracují mechanicky – dokumenty chápou jen jako soubory jednotlivých slov, jejich obsahu nerozumí, takže netuší co indexují.

Výhody:

- obecné vyhledávání
- záruka kvality

Nevýhody

- rozsah katalogů
- obecnost popisů
- struktura kategorií
- časová náročnost
- platnost odkazů

Nejvýznamnější české předmětové katalogy (horizontální, všeobecné)

[ATLAS.CZ](#)

[Caramba!](#)

[Centrum.cz](#)

[VOLNÝ – NAJDI.TO](#)

[Seznam](#)

[QUICK.CZ](#)

Nejvýznamnější zahraniční webové portály (horizontální)

Yahoo www.Yahoo.com

Open Directories Project <http://www.dmoz.org>

Galaxy <http://www.galaxy.com>

LookSmart <http://www.looksmart.com/>

About.com <http://home.about.com/index.htm>

GoGuides <http://www.goguides.org>

JoeAnt <http://www.joeant.com>

V současné době existuje a stále narůstá počet tzv. vertikálních portálů (vertical portals, vortals), které jsou zaměřeny na určitou skupinu uživatelů (např. tzv. *community portals* slouží uživatelům z určitého regionu) nebo jsou orientovány tematicky, opět s ohledem na zájmy určité skupiny uživatelů.

Příklady:

[ChemWeb](#) - poskytuje přístup k informacím z oblasti chemie a příbuzných oborů. Jedná se o snadný přístup k abstraktům, zprávám z konferencí, newsletterům, knihám. Je možné se zdarma registrovat.

[Medscape](#) – informace z oblasti lékařství, přehledové články, recenze knih, sloupky expertů, vzdělávání pacientů atd. Registrace zdarma.

[CNET](#) – the source for computing and technology

[Linux Online](#)

[Minerals Engineering International Online](#) – The Internet Magazine for Mineral Processors and Extractive Metallurgists

[PHPDeveloper.org](#)

[PSIgate](#), Physical Sciences Information Gateway

5.3 Vyhledávače (Search Engines)

Vyhledávací nástroje jsou velkým pomocníkem při lokalizaci informací v Internetu. Jejich používání je velmi jednoduché, většinou nabízí vyhledávání v režimu „simple“ i „advanced“.

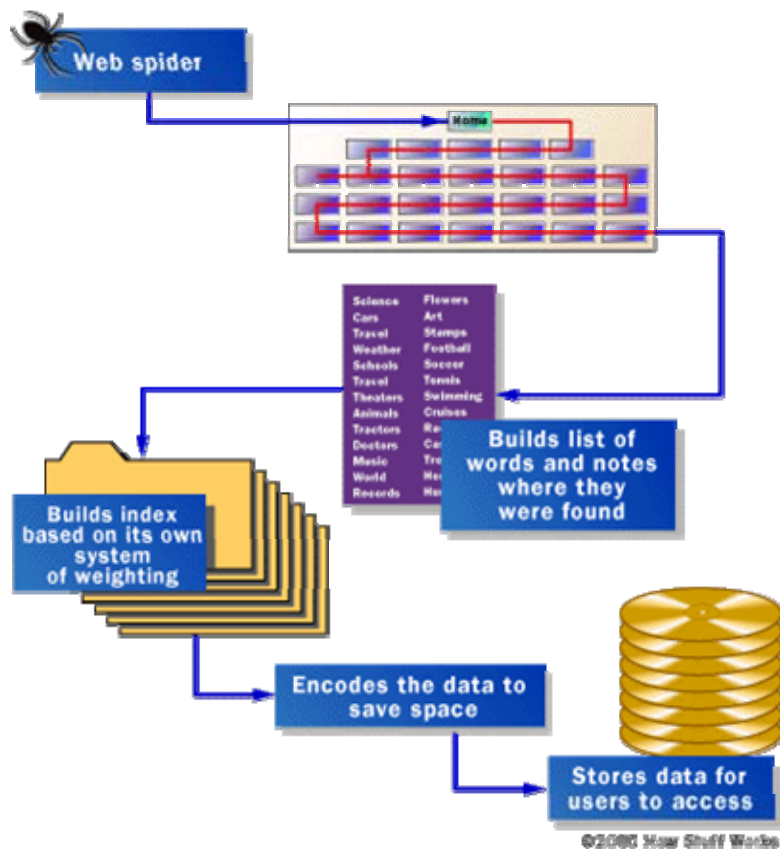
Vyhledávací systémy tvoří tři komponenty:

- robot
- index
- vlastní vyhledávací stroj

Robot - v průběhu milisekund prohledá webové stránky a v závislosti na zadaném klíčovém slově poté zobrazí na obrazovku seznam stránek, které nejlépe odpovídají podmínkám zadaných uživatelem.

Index - výsledky z vyhledávání pavoukem jsou zpracovány do indexové stránky. V podstatě to je stránka, která vznikne z jednotlivých webových stránek, jenž obsahují zadavatelem vložené klíčové slovo, nebo vyhledávací termín. Vyhledávací stroje používají algoritmus, což je ve své podstatě matematický soubor určitých norem. Tím je každá stránka uložená na světovém webu, testována ve vztahu ke klíčovým slovům, nebo vyhledávacím termínům, které byla použity pro vyhledávání. Do nedávna platilo pravidlo, čím vyšší hustota klíčových slov v obsahu stránky, tím vyšší zařazení ve výpisech vyhledávačů. Jenomže to dnes již neplatí (i když řada SEO expertů na tom stále trvá). Google (a s většina dalších vyhledávačů) nyní využívá tzv. sémantické vyhledávání, t.j. zhodnocení stránky je přímo závislé na tom, jak obsah stránky odpovídá zadanému klíčovému slovu (nebo termínu). Ke každému slovu, které je součástí indexu připojen seznam odkazů na WWW dokumenty. Struktura uchovávaných údajů v indexu závisí na konkrétním vyhledávacím stroji, resp. na jeho dotazovacím jazyku.

Vlastní vyhledávací stroj – vlastní uživatelské rozhraní, ve kterém uživatel zadává dotazy.



Zdroj: FRANKLIN, C. [2]

Příklady SE:

Google <http://www.google.com>
 AltaVista <http://www.altavista.com>
 HotBot <http://www.hotbot.com>
 Lycos <http://www.lycos.com>
 Gigablast <http://gigablast.com>
 Nothern Lihgt <http://www.nothernlight.com>
 MSN <http://www.msn.com>
 BAIDU <http://www.baidu.com>

5.4 Metavyhledávače (Metasearch Engines)

Metavyhledávač nemají svou vlastní databázi dokumentů, ale pokládají dotaz jiným vyhledávačům, tzn. že automaticky a simultánně provedou vyhledávání v několika systémech. Většina metavyhledávačů zkombinuje výsledky, odstraní duplicitní nálezy a předloží více odkazů, než bychom dostali z jakéhokoliv jednotlivého serveru.

Metavyhledávače lze řadit do dvou skupin podle toho, kde jsou instalovány. První skupinu tvoří metavyhledávače instalované na serverech a zdarma permanentně přístupné klientům (MetaCrawler, Mamma apod.). Druhou skupinou jsou metavyhledávače určené k instalování v počítačích klientů (desktop). Tyto nástroje jsou poskytovány za úplatu. Zpravidla jsou nabízeny ve dvou verzích: levnější *PLUS* a dražší a vybavenější *PRO*. Umožňují přípravu dotazů v režimu off-line a bývají vybaveny prostředky pro správu již hotových rešerší (výběry, třídění podle dalších kritérií, automatické aktualizace apod.). Mají předem definovaný seznam portálů, katalogů a vyhledávačů (řádově stovky), z nichž si uživatel může vybrat, které chce do hledání zapojit. Do skupiny metavyhledávačů typu desktop patří například Copernic.

Výhody:

- U se nemusí učit rešeršní strategii všech vyhledávacích systémů
- Nemusí sledovat vznik nových systémů
- Vyšší úplnost vyhledávání

Nevýhody:

- uživatel je ochuzen o speciální možnosti vyhledávání, které nemohou být zobecněny
- uživatelé přicházejí o informace reklamního charakteru (???)
- POZOR: Metavyhledávač převezme z každého paralelního systému jen 10-50 hitů.

Příklady:

Vivisimo <http://vivisimo.com>

Kartoo <http://www.kartoo.com>

Metacrawler <http://www.metacrawler.com>

Mamma <http://www.mamma.com>

Ixquick <http://www.ixquick.com>

Fazzle <http://www.fazzle.com>

5.5 Další typy vyhledávacích nástrojů

Další vyhledávací služby jsou specializovány na prohledávání určitého informačního prostoru:

- podle typu aplikace nebo služby,
- podle typu informací,
- podle geografického hlediska – lokální služby například shromažďují pouze informace o zdrojích dostupných v určité zeměpisné či jazykové oblasti.

Hledání osob nebo osobních domovských stránek

<http://anywho.com>

<http://www.bigfoot.com>

<http://www.worldemail.com>

Firemních informací

<http://centrumfirem.centrum.cz//search/index.php>

<http://yp.yahoo.com/>

Software, IT

<http://www.download.com/?legacy=cnet>

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

<http://www.pegasoft.cz/index.php>

Hledání obrázků

<http://gallery.yahoo.com/>

<http://www.picsearch.com>

Časopisy a časopisecké články

<http://journalseek.net/index.htm>

Knihovní katalogy

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

5.6 Virtuální knihovny

Databáze těchto služeb jsou vytvářeny odborníky, kteří informace nejen sbírají, ale také vybírají, popisují a hodnotí na základě určitých kritérií.

Použití virtuálních knihoven je vhodné zvláště tehdy, hledáte-li **kvalitní zdroje** na určité téma.

Příklady virtuálních knihoven

[About.com](#)

[Internet Public Library](#)

[World Wide Web Virtual Library](#)

5.7 Neviditelný web

Jak uvádí Mgr. Tkačiková [6], kromě veřejně přístupných a indexovatelných webových dokumentů síť Internet umožňuje přístup k obrovskému množství dalších informací, pro něž se v literatuře i na webu vžil termín „**neviditelný web**“ (angl. invisible web), někdy také „hluboký (skrytý) web“ (angl. deep web).

- takové informace, jež jsou uloženy v databázích a **generují se dynamicky** až na základě interakce uživatele se systémem (např. online katalogy knihoven nebo bibliografické báze dat

- takové informace k nimž je přístup **chráněn heslem** a jsou dostupné jen autorizovaným uživatelům, často pouze na komerční bázi (viz např. [elektronické verze časopisů](http://www.sciencedirect.com/), <http://www.sciencedirect.com/>)

Rozdíl mezi „viditelným“ a „neviditelným“ webem ilustrují dva příklady informačních zdrojů (první orientovaný na problematiku stavebního inženýrství, druhý zaměřený na oblast výchovy a vzdělávání):

„viditelný Web“: [iCivilEngineer.com](http://www.icivilengineer.com)

„neviditelný Web“: [Civil Engineering Database](#)

„viditelný Web“: [Educator's Reference Desk Resource Guides](#)

„neviditelný Web“: [ERIC Database](#)

Počet dokumentů, které nemohou být vyhledány prostřednictvím tradičních vyhledávacích nástrojů, je až pětinasobně vyšší než ta část informačního prostoru Webu, ke které poskytují přístup vyhledávací stroje.

Zdroje takového typu (především nejrozšířenější báze dat) registrují některé specializované vyhledávací služby, např. Complete Planet nebo Direct Search

Služba Elsevier Science Scirus - vyhledává informace na „viditelném“ i „neviditelném“ webu. Umožňuje kromě jiného vyhledávání informací o publikovaných časopiseckých článcích a dalších odborných publikacích z různých zdrojů (zahrnuje i prohledávání plných textů článků v ScienceDirect).

5.8 Sociální sítě

Systémy sociálních sítí jsou v podstatě kombinací specializované webhostingové služby a specializovaného vyhledávače. Uživatel si vyplní svůj strukturovaný profil a hned po té může hledat a být nalézán. mít k dispozici strukturovaná data v přesně stanoveném formátu a položky kategorizované do přesných číselníků.

Mezi sociální weby v ČR patří např. Lidé.cz, Spolužáci.cz, Líbím se ti.cz, apod. Ze zahraničních je určitě nejznámější a nejrozšířenější Facebook a Twitter.

Sociální sítě se stávají stále více populárnějšími. O jejich nepřehledném množství je možné se přesvědčit na adrese:

http://en.wikipedia.org/wiki/List_of_social_networking_websites

Literatura:

1. *European Search Engines, Directories and Lists* [online]. c1990 – 2010 Searchenginesoftheworld.com. [cit. 2010-03-12]. Dostupné z: http://www.netmasters.co.uk/european_search_engines/
2. FRANKLIN, C. *How Internet Search Engines Work* [online]. c1998 – 2010 HowStuffWorks, Inc. [cit. 2010-03-21]. Dostupné z: <http://computer.howstuffworks.com/internet/basics/search-engine1.htm>
3. GREŠKOVÁ, M. *Vyhľadavanie v Internete* [online]. Blog o vyhľadavání. [cit. 2010-03-12]. Dostupné z: <http://vyhladavanie.blogspot.com/>
4. *Search Engine Guide: the small business guide to search marketing* [online]. c 1998 – 2010 K. Clough, Inc. Dostupné z: <http://www.searchengineguide.com/searchengines.html>
5. *SearchEngineWatch.com* [online]. c2010 Incisive Interactive Marketing. Dostupné z: <http://searchenginewatch.com/>
6. SKLENÁK, V. *Data, informace, znalosti a Internet*. 1. vyd. Praha: C.H. Beck, 2001. 507 s. ISBN 80-7179-109-0.
7. TKAČÍKOVÁ, D. *Co je neviditelný web* [online]. Ústřední knihovna Vysoké školy báňské. [cit. 2010-03-21]. Dostupné z: <http://knihovna.vsb.cz/kurzy/hledani/04.html>

6. GOOGLE.COM

Vyhledávač Google je v současné době bezesporu nevíce využívaným vyhledávacím nástrojem v Internetu. (viz statistiky <http://searchenginewatch.com/>). Při vymýšlení názvů pro svůj vyhledávací nástroj se Larry Page a Sergey Brin inspirovali anglickým termínem „googol“, který vymyslel Milton Sirott. Termín byl popularizován v knize „Mathematics and the Imagination“ autorů Edwarda Kasnera a Jamese Newtona. „*Now her eis the name of a wery large numer „Googol“ . Most people say: „ A Googol is so large, that you cannot namet or talk about it, it so large that i tis infinite.* [KASNER, 2001, s. 20]. Jde o matematický výraz, který označuje číslo 10 na 100, tj. jednička se 100 nulami. Tvůrci Googlu jej tak pojmenovali, aby vystihli, že jejich vyhledávač bude pracovat v podstatě s nepředstavitelným množstvím informací dostupných na Internetu.

6.1 Chronologie klíčových událostí

Březen 1995 – první setkání Sergeye Brina a Larryho Page

Leden 1996 - prosinec 1997 - Page a Brin pracují na systému BackRub,

7. září 1998 - Se základním kapitálem milion dolarů získaným od rodičů, přátel a "zázračných" investorů (Andy Bechtolsheim Sun Microsystems Inc.) Brin a Page zastavují svá studia a se čtyřmi svými zaměstnanci zakládají společnost Google. Sídlo - v Mountain View v Silicon Valley jižně od San Franciska v Kalifornii; areál je přezdíván Googleplex

Únor - červen 1999 - Google získal investici 25 milionů USD od venture kapitálových fondů Sequoia Capital a Kleiner, Perkins, Caufield&Byers. Zástupci obou fondů míří i do vedení společnosti.

Květen 2000 - Vyhledávač provádí 18 milionů operací denně a stává se nejpoužívanější službou na síti.

6. srpna 2001 - Novým výkonným ředitelem Google byl jmenován Eric Schmidt, bývalý šéf Novellu a technologický ředitel Sun Microsystems. Brin a Page se stali prezidenty firmy pro technologie, respektive pro produkty.

Leden 2002 - Firma začíná prodávat vyhledávací aplikace, které umožňují hledat informace bez vlastní intranetové sítě.

Březen 2002 - Google spouští betaverzi programu Google News, jenž z informací ze serverů vytváří vlastní zpravodajství.

Květen - srpen 2002 - Společnost otevírá kancelář v Paříži, rozšiřuje tak své dosavadní zastoupení, které má v Londýně, Tokiu, Hamburku a Torontu. Google zároveň podepisuje smlouvu o technologické a inzertní spolupráci.

Září 2002 - Google rozšířil prodej inzerce podle klíčových slov na Velkou Británii, Francii, Německo a Japonsko.

Leden 2003 - Google se dohodl na akvizici společnosti Pyra Labs, tvůrce úspěšného webového systému Blogger.

Březen 2004 – Google spouští Gmail

6. ledna 2004 - Yahoo nebude déle používat Google jako hlavní vyhledávač a plánuje vyvinout svůj vlastní systém.

29. dubna 2004 - Google požádal americkou Komisi pro cenné papíry o primární emisi akcií v hodnotě zhruba 3,3 miliardy USD.

13. srpna 2004 - Google ukončil přijímání objednávek na své akcie a zahajuje jejich aukci.

2006 je Google oficiálně zastoupen pobočkou v České republice.

2006 Google kupuje youTube.com za 1.65 miliard US dolarů

2008 youTube.cz (ČR v elitní skupině 22 zemí)

Září 2008 – nový prohlížeč Google Chrome

6.2 Vyhledávací strategie Google.com

- Předdefinovaný operátor: AND
- Google nerozlišuje malá a velká písmena
- Google rozpozná fráze uvozené v uvozovkách
- OR – jediný booleovský operátor, který Google detekuje
- Znaménko +.
(Google má předdefinován operátor AND. V případě, že některá slova nelze označit za klíčová, tzn. systém je automaticky vylučuje, ignoruje, je třeba před toto slovo zadat znaménko +. Například chceme vyhledat drama Romeo a Julie, Google vyloučí písmeno „a“, tudíž výsledky nemusí být dobré. Objeví se například Romeo Og Julie, což je dánsky. Proto použijeme operátor AND formou znaménka +, tedy Romeo +Julie.)
- Znaménko -.
Google nepoužívá operátor NOT, ale znaménko mínus. Znaménko musí být napsáno těsně před slovem, nesmí tam být mezera.

Výsledky

V seznamu jsou obvykle zobrazeny třířádkové výseky ze stránek.

- jsou-li před výpisem tři tečky (...), pak se ukázaná část textu nalézá někde uvnitř stránky, nikoliv na jejím začátku. V tomto případě je potřeba využít klávesové zkratky Ctl+F a zadat znovu jedno z vyhledávaných slov, kurzor se přenesne na hledané slovo
- někdy spatříme v odkazech položku odsazenou doprava. Obě položky, odsazená i neodsazená nad ní jsou podobné, mají podobnou hlavní doménu. Google vám tímto chce naznačit souvislost a říct „Tohle není jen jediná stránka o hledaném subjektu, ale asi nějaký větší web, kde o tom, co hledáte pojednává více stránek. Pod odsazeným odkazem navíc vidíte odkaz „**Další stránky z xxx.xxx.cz**“, klepnutím na tento odkaz si můžeme stránky projet. Po otevření tohoto odkazu se nám ve vyhledávacím poli objeví **site:normale.buty.cz**, je to filtr, který říká, že následující slova se mají hledat jen na dané webové stránce

Filtry v textovém zadání

Filtry: **cache**, **link**, **related**, **info**, **spell** se **nedají** používat v kombinaci se standardním vyhledávacím příkazem, jsou to specifické funkce Google, použitelné jen samy o sobě.

Filtry: **site**, **allintitle**, **intitle**, **allinurl**, **inurl** jsou tzv. **modifikátory** dotazu, upravují vlastní vyhledávací dotaz

cache:URL

Příkaz použijeme, pokud chceme zjistit uloženou verzi některé webové stránky. Hodí se v případě, kdy si například zobrazíme stránku a ta je nedostupná.

Příklad: cache:www.oavm.cz

link:URL

Zjišťuje, na kterých internetových stránkách se nalézá **odkaz** na hledanou adresu.

Příklad: link:www.oavm.cz

Velmi dobrý filtr pro zjišťování citovanosti

related:URL

Příkaz zobrazí stránky, které jsou této stránce podobné, výsledky nebývají příliš efektivní, někdy spíše kuriózní. Málo používaná funkce

info:URL

Příkaz zobrazí informace o dané adrese, ovšem velmi chabě, příkaz je téměř nepoužitelný.

site:URL

Zadáme-li tento výraz do vyhledávacího příkazu, jsou výsledky vyhledávání omezeny pouze na danou adresu a adresy v ní vnořené. Tento výraz používáme také při omezování prohledávání třeba jen na jednu doménu, např. cz

Příklad: vyučující site:www.oavm.cz

Glen Hansard site.cz

Velmi dobré jako rešeršní nástroj v elektronických novinách

„strukturální fondy“ site:idnes.cz

allintitle:vyhledávací výraz

Hledá **všetchna** slova hledaného výrazu v titulku webové stránky

Příklad: allintitle:obchodní akademie

intitle:vyhledávací výraz

Hledá alespoň jedno hledané slovo v titulku webové stránky

Příklad: intitle:obchodní škola

allinurl:vyhledávací výraz

Hledá **všetchna** slova hledaného výrazu v adrese webové stránky

Příklad: allinurl:maturita prijimacky

inurl:vyhledávací výraz

Hledá alespoň jedno hledané slovo v adrese webové stránky

Příklad: inurl:maturita prijimacky

Filetype:formát dokumentu klíčové slovo

Hledá zadané klíčové slovo v určitém formátu dokumentu

Příklad: filetype:ppt internet

Případová studie:

1. Chceme najít stránky ve formátu pdf o žábě ropuše obecné kdekoli na Internetu. Nevíme, jak se řekne ropucha v různých sv. jazycích, použijeme latinu.

„bufo bufo“ filetype:pdf

Objeví se asi stránky ve všech světových jazycích. Dokumenty chceme omezit na české a slovenské sity:

„bufo bufo“ filetype:pdf site:cz OR site:sk

2. Přejeme si získat ropuší stránky v dalších formátech, např. .doc. .xls

„bufo bufo“ filetype:pdf OR filetype:doc OR filetype:xls

3. Chceme nalézt stránky o ropuchách, ale *nechceme*, aby se ve výsledcích objevil formát DOC a PDF

„bufo bufo“ -filetype:pdf – filetype.doc

Řazení výsledků

Google řadí výsledky vyhledávání podle tzv. PageRanku. PageRank je algoritmus pro ohodnocení důležitosti webových stránek. Algoritmus využívá strukturu hypertextových odkazů odkazů jako vzájemné “doporučování” stránek, ne nepodobné hodnocení vědeckých prací podle počtu citací. Na rozdíl od sledování počtu citací ale dovádí tento princip ještě dál: hodnocení stránky se nepočítá z prostého počtu odkazů, které na ni vedou, ale bere se v úvahu i hodnocení odkazujících stránek. Velmi zjednodušeně se dá říci, že Google řadí své výsledky podle **četnosti a citovanosti**. Google vrátí výsledky seřazené tak, jako kdyby náš dotaz zněl asi takto: “*ukáž mi stránky, které se danými slovy zabývají nejvíce, nejčastěji, přednostně před jinými, a které jsou natolik dobré, že na ně odkazuje hodně dalších stránek.*”

Vzoreček výpočtu PageRangu

$$PR(A) = (1-d)/m + d * (PR(T1)/C(T1) + ... + PR(Tn)/C(Tn))$$

PR(A) - PageRank stránky A

T1 až Tn – PageRanky stránek, které odkazují na stránku A

d - dampening faktor (nastavený pravděpodobně na 0,85),

m - celkový počet zaindexovaných stránek

C(T) - počet odkazů vedoucích ze stránky T

Mezi ostatní součásti hodnocení patří **jazykové modely** (schopnost pracovat s frázemi, synonymy, diakritikou, pravopisnými chybami atd.), **modely vyhledávacích dotazů** (nejde jen o jazyk, ale o to, jak jej dnes lidé používají), **časové modely** (pro některé dotazy je nejlepším výsledkem 30 minut stará stránka, zatímco pro jiné je lepším výsledkem stránka, která je prověřená časem) a personalizované modely (ne všichni chtějí to samé).

(Poznámka: Gmail má taktéž velmi precizně propracovanou vyhledávací strategii – viz <http://mail.google.com/support/bin/answer.py?hl=en&answer=7190>)

Služby Google v češtině:

[AdWords Editor](#), [Blogger](#), [Dokumenty Google](#), [Gmail](#), [Google AdSense](#), [Google AdWords](#), [Google Analytics](#), [Google Apps](#), [Google Chrome](#), [Google Desktop](#), [Google Earth](#), [Google Pack](#), [Google Reader](#), [Google Scholar](#), [Google Sky](#), [Google Toolbar](#), [Hlášky Google](#), [iGoogle](#), [Kalendář](#), [Google](#), [Mapy Google](#), [Nástroje pro webmastery](#), [Picasa - obrázkový editor](#), [Překladač Google](#), [Skupiny Google](#), [Slovník Google](#), [Vlastní vyhledávání Google](#), [Webová alba Picasa](#), [Weby Google](#), [YouTube](#), [Zprávy Google](#)

Zdroj: *Official Google Blog Česká publika* [online], [4]

Vyhledávací strategie Gmail

Operátor	Definice	Příklad
from:	Používá se pro určení odesílatele	Příklad - from:lenka
to:	Používá se pro určení příjemce	Příklad - to:david Význam - Všechny zprávy, které byly zaslány Davidovi (vámi nebo někým jiným)
subject:	Hledání slov v předmětu zprávy	Příklad - from:lenka OR from:david Význam - Zprávy od Lenky nebo od Davida
OR	Hledání zpráv odpovídajících termínu A nebo termínu B	
-	Používá se pro vyloučení zpráv z hledání	Příklad - večere -kino Význam - Zprávy obsahující slovo „večere“, ale neobsahující slovo „kino“
label:	Hledání zpráv podle štítku* *Neexistuje vyhledávací operátor pro zprávy bez štítků	Příklad - from:lenka label:přátelé Význam - Zprávy od Lenky označené štítkem „přátelé“
has:attachment	Hledání zpráv s přílohou	Příklad - from:david has:attachment Význam - Zprávy od Davida obsahující přílohu
filename:	Hledání přílohy podle názvu nebo typu	Příklad - filename:úkoldofyziky.txt Význam - Zprávy s přílohou nazvanou „úkoldofyziky.txt“
		Příklad - label:pracovní filename:pdf Význam - Zprávy se štítkem „pracovní“ a obsahující soubor PDF v příloze
in:spam in:trash in:inbox	Najde zprávu v konkrétní "složce"	Příklad - in:spam in:trash in:inbox Význam - vyhledá zprávy označené jako spam, zprávy v koši, zprávy v doručené poště
In:anywhere	Hledá zprávy všude, tj. včetně spamu a koše	Příklad - in:anywhere has:attachment Význam - hledá zprávu s přílohou včetně spamu a v koši
" "	Používá se pro hledání přesné fráze* *Velká písmena nehrají roli	Příklad - "zkusím štěstí" Význam - Zprávy obsahující frázi „zkusím štěstí“ nebo „Zkusím štěstí“
language: lang:	Hledá zprávu napsanou v určitém jazyce	Příklad - lang:cs
is:starred is:unread is:read is:chat	Další upřesnění hledání podle typu, označení a stavu zprávy	is:starred - hledá zprávy s hvězdičkou is:unread - hledá nepřečtené zprávy is:read - hledá přečtené zprávy is:chat - hledá uložené zprávy z chatu

cc:	Hledá podle příjemce kopie	cc:jan- zprávy, jejichž kopii dostal Jan
bcc:	Hledá podle příjemce skryté kopie	bcc:jan- zprávy, jejichž skrytou kopii dostal Jan
after:	Odesláno po datu ve tvaru RRRR/MM/DD	after:2007/01/01 - zprávy odeslané po 1. lednu 2007
before:	Odesláno před datem ve tvaru RRRR/MM/DD	before:2007/01/01 - zprávy odeslané před 1. lednem 2007.

Literatura

1. *Google Guide: Making Searching Even Easier* [online]. [cit. 2010-04-21]. Dostupné z: <http://www.googleguide.com/>
2. ISKRA, J. *Google:vyhledávání, Gmail, Google Talk a další*. 1. vyd. Praha: Computer Press, 2006. 143 s. ISBN 80-251-1043-5.
3. KASNER, E., Newton, J. *Mathematics and the Imagination*“. Mineola, N.Y. : Dover Publications, 2001. 380 s. ISBN 0486417034.
4. *Official Google Blog Česká publika* [online]. c2010. [cit. 2010-04-21]. Dostupné z: <http://google-cz.blogspot.com/>.