

# 1 Základy informatiky a teorie informace

## 1.1 Digitální reprezentace a přenos informací

Žák dovede:

- definovat jednotky informace bit a byte a jejich násobné jednotky;
- rozlišit analogová a digitální zařízení;
- vysvětlit princip bezeztrátové a ztrátové komprese dat;
- uvést příklady typických přenosových rychlostí připojení k Internetu.
- vysvětlit princip digitalizace informace včetně určení počtu bitů pro zakódování zadaného počtu možných stavů a aplikace základního postulátu Shannonova teorému;
- převádět vzájemně čísla mezi desítkovou, dvojkovou a šestnáctkovou soustavou;
- popsat proces komunikace;
- vypočítat dobu přenosu dat po zadané datové lince.

### 1.1.1 Jednotka informace

#### Bit

Bit (z anglického binary digit - dvojková číslice; angl. bit = drobek, kousek) je základní a současně nejmenší jednotkou informace, používanou především v číslicové a výpočetní technice. Značí se malým písmenem b, např. 16 b, ale současně se může také objevit i označení bit, např. 16 bit.

1 bit reprezentuje informaci, získanou odpovědí na jednu otázku typu ano/ne, u které je apriorní pravděpodobnost obou odpovědí stejná (jinými slovy, u které nemáme žádnou předchozí informaci, která by jednu z odpovědí favorizovala). Tyto odpovědi můžeme označit binárními číslicemi 0 a 1.

#### Bit v praxi

Bit se v praxi nejčastěji objevuje jako základní jednotka kapacity paměti, tzn. jednotka množství informace, která může být v jednom okamžiku v paměti uložena. Pokud kapacitu podělíme časem, získáme přenosovou rychlost, jejíž jednotkou je tedy bit za sekundu (bit/s, podle angličtiny někdy označován bps – bit per second). Např. modem s přenosovou rychlostí 56 kbit/s je schopen každou sekundu přenést 56 kilobitů dat.

Skupina 8 bitů se nazývá Byte (čteme bajt). Při práci s bity se někdy předpony soustavy SI (kilo, mega, ...) chápou mírně odlišně (např. kilo=210 namísto 103). Nově se používají jiné předpony pro mocniny 10 (kilo (kB), mega (MB), ...) a pro mocniny 2 (kibi (KiB), mebi (MiB), ...), tzv. binární předpony.

#### Byte

Bajt, původním, anglickým zápisem byte, zřídka také slabika, je jednotka množství dat v informatice, zpravidla označuje osm bitů, tzn. osmiciferné binární číslo. Takové množství informace může reprezentovat například celé číslo od 0 do 255 nebo jeden znak. Jeden bajt je obvykle nejmenší objem dat, se kterým dokáže počítač (resp. procesor) přímo pracovat.

#### Násobky a užívané předpony

S bajty i bity se používají běžné předpony soustavy SI jako kilo-, mega-, giga- atd., např. 10 GB, 11 Mb/s. Tyto předpony však mají někdy odlišný význam, který je nutno rozlišit z kontextu.

Z technologických důvodů jsou velikosti některých počítačových pamětí obvykle násobkem nějaké mocniny dvou. Například počítač IBM PC Model 5150 měl kapacitu operační paměti 65 536 B =

64-210.[1] Pro zjednodušení se však taková paměť neoznačovala jako 65,5 kilobyte, ale pojem kilobyte se „mírně“ upravil na  $2^{10} = 1024$  a kapacita paměti se označila prostě jako 64 KB. Pro odlišení se tato jednotka zpravidla značila s velkým písmenem – KB – a neformálně se označovala jako „velké kilo“, na rozdíl od tradičně pojatého kilobajtu, tzv. „malého kila“, které se značilo malým písmenem – kb. Obdobně se později začaly užívat další upravené předpony mega-, giga- atd. (u těch už kontext podle velikosti písmena rozlišit nelze) označující 220, 230 atd.

Takto upravené předpony se používají zejména při vyjadřování velikosti polovodičových pamětí, zatímco například kapacity pevných disků používají dekadické předpony. Bez znalosti kontextu tedy může být složité určit, jaký význam byl zamýšlen, přičemž rozdíl může dosahovat až několika procent (nejasnosti kolem skutečné kapacity prodávaných pamětí dokonce vedly k právním sporům[2]).

V prosinci 1998 proto IEC vytvořila dodatek k normě IEC 60027-2 (v Česku převzatá jako ČSN IEC 60027-2), ve kterém zavedla pro počítačové jednotky nový systém označování násobků. V tomto systému bylo pro původní „velké kilo“ = 1024 B navrženo označení kibibyte a značka KiB, zatímco jednotka kilobyte (se značkou kB) označuje 1000 B, tak jak je obvyklé v soustavě SI. (Byte, 2010)<sup>1</sup>

Přehled násobných jednotek										
Jednotka	Značka	B	kB	KiB	MB	MiB	GB	GiB	TB	TiB
Kilobyte	kB	1000	1	~0,9766						
Kibibyte	KiB	1024	1,024	1						
Megabyte	MB	1 000 000	1000	~976,6	1	~0,9537				
Mebibyte	MiB	1 048 576	~1048,6	1024	1,049	1				
Gigabyte	GB	$10^9$	1 000 000	976 562,5	1000	953,7	1	~0,9313		
Gibibyte	GiB	$\sim 1,074 \cdot 10^9$	~1 073 742	1 048 576	~1073,7	1024	1,074	1		
Terabyte	TB	$10^{12}$	$10^9$	~0,9766 $\cdot 10^9$	1 000 000	~953 674,3	1000	931,3	1	~0,9095
Tebibyte	TiB	$\sim 1,1 \cdot 10^{12}$	~1,1 $\cdot 10^9$	~1,074 $\cdot 10^9$	~1 099 512	1 048 576	~1099,5	1024	~1,1	1

Binární násobky			
Jednotka	Značka	Velikost v B (byte)	Mocnina
Kibibyte	KiB	1 024	$2^{10}$
Mebibyte	MiB	1 048 576	$2^{20}$
Gibibyte	GiB	1 073 741 824	$2^{30}$
Tebibyte	TiB	1 099 511 627 776	$2^{40}$
Pebibyte	PiB	1 125 899 906 842 624	$2^{50}$
Exbibyte	EiB	1 152 921 504 606 846 976	$2^{60}$
Zebibyte	ZiB	1 180 591 620 717 411 303 424	$2^{70}$
Yobibyte	YiB	1 208 925 819 614 629 174 706 176	$2^{80}$

### 1.1.2 Analogová a digitální zařízení

<sup>1</sup> <http://cs.wikipedia.org/wiki/Byte>

Svět kolem nás se stává čím dál tím více digitálním - klasické gramodesky a magnetofonové kazety jsou nahrazovány digitálními kompaktními disky, staré analogové telefonní ústředny jsou nahrazovány digitálními ústřednami, kvalitnější rozhlasové přijímače již nemají žádný ladící knoflík, ale mají digitální ladění atd. No a o počítačích už ani nemluvě, jejich „digitálnost“ je dnes takovou samozřejmostí, že se už ani explicitně nezdůrazňuje. Ale v čem je vlastně podstata „digitálního“, a v čem se doopravdy liší analogový svět od světa digitálního?

Možná bychom byli překvapeni tím, jak obtížně se ještě i dnes, v dnešním tak digitálním světě, hledá něco co je skutečně digitální již svou bytostnou podstatou. Představme si například digitální přenosovou cestu, kterou podle našich představ „tečou“ jen samé logické nuly a logické jedničky. Ve skutečnosti tam ale žádné logické nuly ani jedničku téci nemohou - ve skutečnosti i digitální cestou „teče“ vždy něco takového, jako třeba elektrický proud, nebo tudy prochází světlo (světelný paprsek) apod. Tedy takové veličiny, které mají vždy nějaký měřitelný aktuální stav resp. hodnotu (proudu či napětí, světelnosti resp. intenzity atd.), a těchto jejich hodnot může být nekonečně mnoho. A právě to je hlavním rysem analogové veličiny - její schopnost nabývat nekonečně mnoha různých hodnot, zatímco digitální veličina může nabývat vždy jen konečně mnoha různých hodnot.

Abychom si to ale správně vyložili: **schopnost nabývat nekonečně mnoha různých hodnot**, charakteristická pro analogovou veličinu, nemusí znamenat schopnost nabývat libovolně velké, či naopak libovolně malé hodnoty. Analogová veličina může nabývat nekonečně mnoha různých hodnot i v kterémkoli omezeném a uzavřeném intervalu. Přestavíme-li si například elektrické napětí mezi 0 a 5 volty, pak i takovéto napětí může nabývat nekonečně mnoha různých hodnot větších než nula a menších než 5.

Nyní si představme „digitální“ veličinu - zůstaňme u obvyklého dvouhodnotového (tj. binárního) systému, tj. **systému předpokládajícího jen dva možné stavy**, a představujme si například logickou nulu jako situaci resp. stav, kdy napětí je někde v rozmezí od nuly voltů výše, a jako logickou jedničku napětí jiné, tedy napětí záporné. Jak jistě snadno nahlédneme, nebude nám v tomto případě příliš vadit, když na přenosové cestě bude v důsledku zkreslení a útlumu docházet k jisté deformaci přenášeného signálu - když třeba odesílatel „pustí“ do přenosové cesty 4 volty a my na druhé straně přijmeme např. jen volty 3, je to stále v pořádku, a víme, že jsme přijali logickou nulu (nebo naopak logickou jedničku, podle zvolené konvence). Kdyby ale přenášenou informací nebylo to, zda hodnota příslušné analogové veličiny spadá do určitého intervalu nebo nikoli, ale kdyby přenášenou informací byla přímo okamžitá hodnota zmíněné analogové veličiny, situace by vypadala úplně jinak - příjemce by díky nedokonalosti (zkreslení) přenosového kanálu přijal něco úplně jiného, než mu odesílatel původně posílal. A zde jsme právě u jádra věci, neboli u hlavního rozdílu mezi „analogovým“ a „digitálním“: digitální veličinu můžeme zrekonstruovat, zduplikovat či jinak zpracovat naprosto přesně, tedy s ideální přesností, zatímco u analogové veličiny nikoli. Jakékoli zpracování analogové veličiny, díky „nekonečné jemnosti“ jejích možných hodnot, je vždy zatíženo nějakou chybou. Jak velkou, to už záleží na konkrétním zařízení, kanálu či technice zpracování, ale ideální přesnosti jako u digitální veličiny nedosáhneme nikdy. Ideálnímu stavu se můžeme pouze přiblížit, a toto přiblížení je navíc hodně drahé - čím větší budou nároky na přesnost analogového zpracování, tím kvalitnější a tudíž i dražší budou muset být všechny části řetězce, kterým analogová veličina při svém zpracování prochází. Naproti tomu při digitálním zpracování bude všechno jednodušší a lacinější. No a to je také ten hlavní důvod, proč má dnes digitální svět tak „navrch“ - díky mnohem větší toleranci vůči různým nedokonalostem vychází digitální zařízení lacinější a jednodušší než analogové, digitální přenosy mohou být efektivně rychlejší než přenosy analogové atd. (Peterka, 2010)<sup>2</sup>

---

<sup>2</sup> Peterka Jiří, <http://www.earchiv.cz/a95/a533k130.php3>

### 1.1.3 Princip bezeztrátové a ztrátové komprese dat

#### Bezeztrátová komprese

Bezeztrátová komprese (=bezeztrátová komprimace; někdy bezztrátová) je jeden ze dvou základních přístupů ke kompresi dat. Jedná se o algoritmy, které dovolují přesnou zpětnou rekonstrukci komprimovaných dat, na rozdíl od ztrátové komprese, kde to možné není.

Bezeztrátová komprese se používá všude tam, kde je důležité, aby originální data a data po dekompresi komprimovaného souboru byla totožná - např. komprese textů nebo komprese čehokoli, kde je nepřipustná i sebemenší ztráta kvality.

Existuje mnoho různých formátů, které využívají bezeztrátové komprese. Například velmi populární ZIP, se kterým je možné pracovat téměř na jakékoli platformě. Dalším velmi oblíbeným formátem je RAR, tento formát je však komerční a neexistují svobodné nástroje pro vytváření archivů RAR. Komerční nástroje existují pro většinu platform, například DOS, MS Windows, \*BSD, GNU/Linux, Mac OS X.

Mezi open source komunitami jsou oblíbeny formáty gzip a bzip2, jejichž algoritmy mají tu výhodu, že nejsou patentované. Dalším méně známým, ale velice schopným kompresním algoritmem je LZMA, který je užíván programem 7-Zip.

#### Ztrátová komprese

Ztrátová komprese je způsob ukládání některých digitálních dat v počítačích. Pomocí speciálního algoritmu se zmenšuje objem dat na zlomek původní velikosti. Přitom se některé méně důležité informace ztrácejí a z vytvořených dat již nejdou zrekonstruovat.

Ztrátová komprese se nejčastěji používá pro ukládání obrazových a zvukových záznamů.

Přesto, že se část informace při ztrátové kompresi nevratně ztrácí, je tento způsob ukládání dat často velmi výhodný. Ztráta některých informací je totiž zcela vyvážena velmi výrazným zmenšením komprimovaných dat. Obvykle je tak určitá (malá) ztráta kvality vyvážena výraznou úsporou místa.

Díky ztrátové kompresi lze na CD-ROM umístit i 10 původních nekomprimovaných zvukových CD, nebo tisíce obrázků; na DVD se vejde celovečerní film, včetně mnoha zvukových doprovodů.

Takto komprimovaná data se lépe ukládají a lépe se s nimi manipuluje.

Příklady užití:

- digitální televizní vysílání
- filmová DVD
- přenášení dat přes Internet a další datové sítě
- ukládání do multimediálních přehrávačů

Ztrátová komprese je nepoužitelná v případě, kdy je potřeba uchovat přesnou kopii původních dat, například text knihy, program nebo výsledky měření.

#### Principy ztrátové komprese

Obecný přístup ztrátové komprese je jednoduchý. Po úvodním předzpracování se přeskupí nebo transformují data tak, aby bylo možno lehce oddělit důležité informace od nedůležitých. Nedůležité informace se pak potlačí mnohem více než důležité a nakonec se výsledek zkomprimuje některým z bezeztrátových kompresních algoritmů.

Algoritmus ztrátové komprese má tedy dvě podstatné části — transformace původních dat a potlačení různě důležitých dat.

Některé formáty ztrátové komprese dat: JPEG, JPEG 2000, MPEG,MP3, Vorbis, WMA, AAC, VQF

#### **1.1.4 Přenosové rychlosti připojení k Internetu(Herlický, 2010)**

Autor: Martin Herlický<sup>3</sup>

Jaké základní možnosti má dnes uživatel k tomu, aby se mohl připojit? Především je to nejznámější vytáčené připojení, linka euro ISDN, ADSL, připojení přes přípojku kabelové televize, připojení přes mobilní telefon, bezdrátové připojení využívající sítě Wi-Fi.

##### **Vytáčené připojení**

Základním a nejběžnějším způsobem připojení je spojení vytáčené. Je velmi dostupné, protože k jeho realizaci potřebuje uživatel pouze vytáčenou linku. V nabídce výrobců notebooků, jako zařízení, o které je největší zájem (už druhý rok zájem o notebooky převyšuje zájem kupujících o stolní počítače) jsou v notebookech vestavěné modemy standardu V.90/V.92, jako posledního technologického kroku, který snad už (konečně) nebude překonán, protože cesty pro připojení se už ubírají zcela jinými směry a počet vytáčených linek v domácnostech ubývá. Navíc, cena za takový druh velmi pomalého, a ne zcela dostatečně spolehlivého, připojení není právě nízká a tak o této variantě se zmiňujeme pouze z historických důvodů. Rychlost: max. kolem 40 kb/s. Cena: 1,04 Kč/min mimo špičku až cca 2,66 Kč/min ve špičce.

##### **Linka euro ISDN**

Historická vykopávka v době jejího zavádění u nás. Zatímco v Německu byla technologie euro ISDN už několik let v provozu, náš monopolní poskytovatel, Český Telecom, přemýšlel, jak tuto problematiku vůbec vyřešit. A tak v době nástupu kabelové televize a oboustranného zprůchodnění kabelových cest kabelovky (předpoklad pro fungování internetu je oboustranná průchodnost dat, kabelovky byly původně navrhovány jako jednostranně průchodné – jen pro TV signál) a prvních i dalších krůčků ADxL technologií ve světě, se realizace euro ISDN ukázala skoro jako, s prominutím, hovadina, za kterou jsme ale nakonec všichni docela dobře zaplatili.

Přípojka euroISDN se skládá ze dvou linek s kapacitou 64 kb/s, které lze ale sloučit v jednu a tak získat linku s kapacitou 128 kb/s. To je v dnešní době ale skoro úplně mimo. Samozřejmě hovoříme o metodě připojení BRI (Basic Rate Interface) neboli 2B+D, 2x 64 kb/s (dva kanály pro přenos dat, videa, zvuku) + 1x 16 kb/s (datový kanál pro přenos signalizace). Existuje i možnost připojení PRI (Primary Rate Interface) neboli 30B+D (30x 64 kb/s + 64 kb/s). V jednom okamžiku lze jeden z kanálů využít například na přenos dat a druhý na telefonování nebo první na posílání faxu a druhý na telefonování, atd. Pro případy internetu lze kanály sdružovat pro dosažení vyšší rychlosti. V případě BRI se tak dostaneme na teoreticky možnou rychlost 2x 64 kb/s, tj. na 128 kb/s, v případě PRI pak na 30x 64 kb/s, tj. cca 1,92 Mb/s. Ceny a rychlost jsou dnes opět dnes už omezujícím prvkem. Rychlost: 64 nebo 128 kb/s (BRI). Cena: 1,16 Kč/min mimo špičku, 2,18 Kč/min ve špičce (rozumí se za 64 kb/s, u připojení 128 kb/s je pak cena dvojnásobná, platí pro BRI)

##### **Kabelový modem**

Velmi populární připojení k internetu, protože patří povětšinou k těm nejstabilnějším. Hned na úvod ale řekněme, že jeho velkou nevýhodou je velmi omezená dostupnost. Takový druh připojení si můžete vybrat především na sídlištích a velkých bytových celcích, kde žijí především mladí lidé, tedy tam, kde se dá pro poskytovatele předpokládat největší možnost výtěžku. Tedy existuje zde velká penetrace zájemců. Zatažením do jednoho domu získává operátor v podstatě někdy i více, než 50

---

<sup>3</sup> <http://www.ictsecurity.cz/08/03-bezpecne-pripojeni-k-internetu/druhy-prenosove-rychlosti-a-ceny-pripojeni-k-internetu.html>

potencionálních zájemců – a to se vyplatí. V oblasti staré zástavby, rodinných domků, kde bydlí staří občané, operátoři nabídku vůbec neřeší.

Princip činnosti je jednoduchý a spočívá v podstatě v připojení speciálního kabelového modemu ke kabelové přípojce, tedy do místa, kudy se do bytu dostává televizní signál. Velkou výhodou je už jmenovaná spolehlivost a výběr z nabízených rychlostí.

Rychlost (download/upload – jedná se o asymetrické připojení, kdy se předpokládá, že více dat bude putovat směrem k vám – proto download má vyšší hodnoty, než upload): od 256/64 kb/s do 20/1,5 Mbit/s. Cena: měsíční poplatek od cca 400 Kč.

### **ADSL**

Asymetrické připojení ADSL je jedním z nejvíce oblíbených připojení na českém trhu vůbec. Je to především dáno mnohem lepší dostupností služby, než je tomu u kabelové televize. Existuje i dostatečně dobrá nabídka na trhu. Jediným problémem je, že do vašeho domu musí být zavedeno metalické vedení s vytáčenou pevnou linkou. Tak je použita jako „nosič dat“, která se transportují nad frekvenčním pásmem, vyhrazeným pro přenos hlasu. Principem je pak oddělení „hlasových dat“ do telefonu a dat jako takových v odbočovači u přípojky ADSL, vedoucího do místnosti. Spektrum je nabídky je široké, jediným limitujícím faktorem je vzdálenost od elektronické ústředny (to jdou dnes už skoro všechny), která nesmí překročit určitou mezní hodnotu. Ale pokud váš dům není v této toleranci, je možné provést měření a posoudit možnost nasazení ADSL i v místech, která jsou dále, než vyhovuje doporučení. Obvykle se ADSL připojení dodává společně v balíčku například s poskytováním televizního připojení (IPTV).

Rychlost: dnes nabídka začíná zhruba na rychlosti 64/64 kb/s a končí na 20 480/1024 kb/s

Cena: od cca 600 Kč/měsíčně až do zhruba 1150 Kč

### **Připojení přes mobilní telefon**

Je trojího druhu: GPRS, EDGE a CDMA. Jak šel život v oblasti mobilních telefonů, zůstávaly na trhu technologie, které odrážejí stav technologických znalostí.

#### **Mobily – GPRS**

GPRS se podobá klasickému vytáčenému připojení a podobně má i množství nevýhod. Rychlost se může i na úrovni 60 kb/s, ale velkou nevýhodou tohoto připojení je, že vám takovou rychlost nikdo nezaručí. Buďto to prostě půjde nebo to nepůjde a rychlost se výrazně sníží. Jedinou náplastí na bolístku u neustálého čekání na příchod bitů a bajtů do vašeho počítače nebo digitálního zařízení ale je to, že platíte za přenesený objem dat. Není to ale žádná láce – a to ani v případě, že si vyberete paušál. Nicméně máte aspoň jakous takous jistotu, že datové připojení máte k dispozici „skoro“ všude. Další výhodou je, že takové připojení podporují téměř všechny mobilní telefony na trhu. Rychlost: do 86 kb/s. Cena: od cca 200 Kč/měs.

#### **Mobily – EDGE**

EDGE je dostatečně podobné technologii GPRS, je však mladší a skoro třikrát hbitější. Dá se hovořit o i o vyšší stabilitě ve srovnání s GPRS. Ceny jsou rovněž srovnatelné s GPRS. Výběr mobilů je už dostatečně omezený, protože například Sony Ericsson zatvrzele trval na tom, že EDGE do svých přístrojů montovat nebude. Ledy v tomto roztály až teprve nedávno. Dostupnost technologie je dána především většími městy a postupně se rozlézá do všech koutů naší vlasti. Rychlost: do 480 kb/s. Cena: od 200 Kč/měs.

## Mobily – CDMA

Technologie CDMA je u nás nabízena zatím pouze u O2. Rychlosti CDMA mohou teoreticky přesahovat až 1 Mb/s, ale v praxi se průměrné rychlosti směrem k uživateli pohybují mezi 200 – 300 kb/s v závislosti na denní době. K tomu, abyste tuto legraci mohli provozovat, si musíte nejprve pořídit CDMA modem s cenou kolem 5000 Kč, pak budete s ním do jisté míry i mobilní, i když operátor poskytuje možnost připojit se jako alternativu k připojení pevnou linkou nebo přes Wi-Fi. Rychlost: 1 Gb/s, reálně do 800 kb/s. Cena: od 500 do 1200 Kč dle tarifu

## Wi-Fi

Nejoblíbenější systém hned po ADSL v našich zemích. Dle rychlosti se využívá tři standardů 802.11 b – nejnámější a nejpoužívanější, 802.11 a, velmi málo využívaný s protokolem typu OFDM a konečně nejmladším 802.11 g, který používá systém OFDM nebo DSSS (stejně jako 802.11 b).

Na webu [www.lupa.cz](http://www.lupa.cz) se dočteme o Wi-Fi (konkrétně na: <http://www.lupa.cz/clanky/802-11g-rychlejsi-wifi/>), citujeme:

Režim IEEE 802.11b používá na fyzické vrstvě metodu rozprostřeného spektra DSSS (Direct Sequence Spread Spectrum s klíčováním CCK (Complementary Code Keying). CCK mapuje čtyři bity na symbol (na 8 Mbit/s) a současně mírně zvyšuje symbolovou rychlost na 1,375 Msymbol/s, čímž se dosáhne na fyzické vrstvě maximální rychlost 11 Mbit/s.

WiFi nabízí čtyři rychlosti na fyzické vrstvě: 11 Mbit/s, 5,5 Mbit/s, 2 Mbit/s a 1 Mbit/s. Pro všechny WLAN platí, že se přenosová rychlost na fyzické vrstvě se podle situace mění: snižuje se s růstem chybovosti nebo zvyšuje při zlepšení podmínek prostředí, takže maximální rychlost jednotlivých WLAN lze předpokládat pouze na krátkou vzdálenost v prostředí bez rušivých vlivů na přenos.

11 Mbit/s je tedy maximální rychlost WiFi na fyzické vrstvě, rychlost užitečná (pro uživatelská data) je ale nižší, protože 30-40 procent teoretické kapacity spolknou režie protokolu MAC, která je např. ve srovnání s Ethernet/802.3 (také sdílené médium) u WLAN vyšší. WLAN pracují v režimu polovičního duplexu – buď data vysílají, nebo přijímají (na rozdíl od 802.3, kde je možný režim plného duplexu, tj. současně stanice může data vysílat i přijímat). Uživatelská datová rychlost u WiFi proto dosahuje maximálně 6 Mbit/s.

Pro dosažení vyšší rychlosti se u 802.11g používá ortogonální multiplex s kmitočtovým dělením OFDM (Orthogonal Frequency Division Multiplex), jedna z přenosových metod MCM (MultiCarrier Modulation), kdy se data k vysílání nejprve rozdělí do několika paralelních toků bitů o mnohem nižší bitové rychlosti. Každý z toků se používá pro modulaci jiné nosné.

Zatímco tradiční kmitočtový multiplex dělí kmitočtové pásmo do N nepřekrývajících se kmitočtových subkanálů vzájemně oddělených ochranným kmitočtovým pásmem (guard), OFDM používá překrývající se subkanály, takže kmitočtové pásmo se využívá účinněji. Přísně vzato není OFDM modulační metoda, ale metoda pro generování a modulaci více nosných současně, každé s malou částí datového toku. Jako konkrétní modulace lze pak použít jakýkoli typ digitální modulace včetně QPSK (Quadrature Phase Shift Keying), 16-QAM (Quadrature Amplitude Modulation) či 64-QAM.

Způsobem paralelního vysílání se OFDM účinně brání zkreslení při přenosu signálu různými cestami (multipath distortion), protože každý přenášený symbol trvá na dílčí nosné déle, takže se prakticky vyloučí nepříznivý dopad zpoždění signálu delší cestou. Navíc se používá více úzkopásmových nosných a jejich vzájemné rušení ovlivní jen velmi malou část signálu.

Podporovaných rychlostí u 802.11g je víc než u WiFi. Rychlosti podporované pomocí OFDM jsou následující (v závislosti na modulaci): 54, 48, 36 a 24 Mbit/s (16-QAM); 18 a 12 Mbit/s (QPSK), 9 a 6

Mbit/s (BPSK, BiPhase Shift Keying). Další rychlosti jsou v souladu s 802.11b a vyžadují použití DSSS a 11 Mbit/s; 5,5 Mbit/s; 2 Mbit/s a 1 Mbit/s.

Tolik k nejnovější technologii 802.11 g. Pásmo Wi-Fi je účinné pokryto ve větších městech, kde však mnohdy signál má nepředvídané výchyly díky rozmístění domů či paneláků. Existuje spousta operátorů, nabízejících své služby i mnohé poloamatérské sítě, které svůj účel při rozjezdu této technologie splnily svůj účel a dnes přenechaly žezlo operátorům, kteří mohou nabídnout jakous takous jistotu, že v případě výpadku bude spojení rychle a profesionálně „nahozeno“ zpět.

Existuje i spousta hot-spotů (vykrývacích vysílačů), které jsou dostupné tam, kde se shromažďuje více lidí – typickým příkladem jsou internetové kavárny nebo benzinové stanice. Tak se můžete buď zdarma, nebo za symbolický poplatek připojit a stáhnou zdarma svá data. Podobně, ovšem na ryze komerční bázi, funguje připojení k Wi-Fi například v hotelech. Setkáte se třeba i s tím, že za jednu noc a den (24 hodin) si v některých hotelích naučtují i nehorázných 20 euro, což je někdy více, než za využití pay-TV na pokoji. Bysnys ovšem chce své, hoteliéři to vědí, a tak někomu tohle řešení vyhovuje.

Wi-Fi je velmi žádané připojení. K tomu, abyste si vytvořili bezdrátovou síť doma, potřebujete Wi-Fi router, který pak pokryje váš domov signálem a tak můžete snadno připojovat například notebooky, které jsou v poslední době vybavovány Wi-Fi modulem standardu 802.11 a, b, g. Výhodou Wi-Fi je rovněž možnost nechat nad Wi-Fi fungovat nadstavbu UPnP, neboli bezdrátové připojení hudebních přístrojů. Díky tomu můžete na síť vměstnat hudební server a s ním pak spolupracovat se sdílením dat k jednotlivým hudebním stanicím. Kupodivu tento systém funguje a nejdále je s ním firma Philips u přístrojů řady WACS – Streamium. Rychlosti: 802.11 b do 6 Mb/s, 802.11 a do 25 Mbit/s, 802.11 g do 22 Mbit/s. Cena: od cca 200 Kč, dle místních podmínek

### **1.1.5 Princip digitalizace informace**

Digitalizace obecně je převod analogového (spojitého) signálu (např. hlasový projev) do digitálního tvaru (do vhodné binární soustavy) resp. do nespojitého signálu.

#### **Digitalizace televize a rozhlasu**

Od 12. května 2000 probíhalo v České republice zkušební vysílání televizního a rozhlasového signálu v digitální podobě (standard DVB-T), na který se postupně přechází místo signálu analogového, zkušební provoz přešel 21. října 2005 na řádné vysílání. Výhodou je především stálá a úplná technická kvalita, tedy výsledný signál bez známého kolísání úrovně šumu podle počasí (v oblasti se slabším signálem vysílače) i bez tzv. „duchů“ (tj. projevu interference např. v místech s výskytem silných a slabých signálů na shodném či blízkých kmitočtech). Další výhodou je zvýšení počtu televizních programů při zachování stejné šířky přenášeného pásma, namísto jednoho analogového kanálu lze přenášet ve stejné kvalitě až 5 kanálů digitálních. Nové televizory se již vyrábějí se zabudovaným zařízením na příjem digitálního signálu, u starších je nutno použít tzv. Set-top box (běžně zkracovaný na STB). Ten se zapojí mezi TV přijímač a zásuvku televizního rozvodu (resp. koaxiální kabel od antény). Přijatý digitální signál převede na signál analogový a po běžném koaxiálním kabelu (zapojeném do anténního konektoru TV) nebo pomocí kompozitního kabelu (zapojeného např. do konektoru SCART) předá do televizoru.

#### **Digitalizace kulturního dědictví**

Dne 30. září 2005 vyhlásila Evropská komise plán Digitalizace evropského kulturního dědictví, podle kterého by v dohledné době mělo být na internetu k volnému použití aspoň šest milionů knih v různých jazycích. Plán má navázat na americký Projekt Gutenberg či francouzský projekt Gallica. U



nás se problémem zabývá webová stránka ČTE! (České Texty Elektronicky!). Národní knihovna České republiky v roce 2005 získala od UNESCO ocenění za digitalizaci vzácných rukopisů.

Evropská komise v rámci programu Kultura 2007 - 2013 podpořila také projekt EOD: eBooks on Demand - A European Library Network (Elektronické knihy na objednávku), jehož náplní je digitalizace knih vydaných v letech 1500-1900 z fondů zúčastněných knihoven na vyžádání zájemce.

### **Digitalizace zvukových a obrazových nahrávek**

Digitalizace starých zvukových a obrazových nahrávek nastala zejména s prudkým rozvojem digitálního způsobu nahrávání zvuku i obrazu na konci 20. století a také díky velkému rozvoji digitálních nosičů dat, zejména pak CD disků, později DVD disků, komprimovaných formátů obecně. Digitalizace těchto analogových dat probíhá již delší dobu. (Wikipedia, Digitalizace, 2010)<sup>4</sup>

### **Určení počtu bitů pro zakódování zadaného počtu možných stavů**

Určíme mocninu dvou, která je nejbližší vyšší příslušnému číslu. Např. pro 200 možných stavů potřebujeme celkem 8 bitů, neboť  $2^8=256 > 200 > 2^7=128$ .

Také je možné počet možných stavů v desítkové soustavě převést na dvojkové číslo a pak spočítat počet bitů (číslic) tohoto čísla. Např.:  $200 = 11001000$  B má 8 bitů.

Určete, kolika bity je možné zakódovat 65 možných stavů.

Tabulka mocnin 2:

n	$2^n$
1	2
2	4
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1024
11	2048
12	4096
13	8192
14	16384
15	32768
16	65536
17	131072
18	262144
19	524288
20	1048576

---

<sup>4</sup> <http://cs.wikipedia.org/wiki/Digitalizace>

## Aplikace základního postulátu Shannonova teorému

Shannonův teorém: „Přesná rekonstrukce spojitého, frekvenčně omezeného, signálu z jeho vzorků je možná tehdy, pokud byl vzorkován frekvencí alespoň dvakrát vyšší, než je maximální frekvence rekonstruovaného signálu.“

### Shannonův teorém a vzorkovací frekvence v praxi

V praxi se tedy vzorkovací frekvence volí dvakrát větší plus ještě malá rezerva než je maximální požadovaná přenášená frekvence. V telekomunikacích je to např. 8 kHz neboť je třeba přenášet pouze signály ve standardním telefonním pásmu (od 0,3 do 3,4 kHz zaokrouhleno směrem nahoru 4 kHz). Například u záznamu na CD je to 44,1 kHz neboť průměrné zdravé lidské ucho slyší maximálně cca do 20 kHz a tudíž vzorkovací frekvence 44,1 kHz byla zvolena s určitou rezervou.

V případě použití nižší vzorkovací frekvence může dojít k tzv. aliasingu, kdy rekonstruovaný signál je výrazně odlišný od původního vzorkovaného signálu. (WikiPedia, Shannonův teorém, 2010)<sup>5</sup>

### 1.1.6 Převody číselných soustav

#### Desítková (decimální) číselná soustava

Desítková soustava zahrnuje číslice 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 a já vám budu ukazovat, jak přepočítávat čísla z jiných číselných soustav právě do této desítkové a naopak.

#### Dvojková (binární) číselná soustava

S dvojkovou soustavou se můžete nejčastěji setkat ve výpočetní technice. Dvojková soustava je založená na mocninách čísla 2 a zapisujeme ji číslicemi 0 a 1. Vezmeme si pro příklad decimální číslo 173. Jeho převod do dvojkové soustavy spočívá v neustálém dělení tohoto čísla dvojkou → číslo vydělíme 2 a pokud zůstane zbytek (1), bude hodnota 1. Pokud nebude zbytek, bude hodnota 0. Příklad:

$$173 : 2 = 86 (1)$$

$$86 : 2 = 43 (0)$$

$$43 : 2 = 21 (1)$$

$$21 : 2 = 10 (1)$$

$$10 : 2 = 5 (0)$$

$$5 : 2 = 2 (1)$$

$$2 : 2 = 1 (0)$$

$$1 : 2 = 0 (1)$$

Takže zápis decimálního čísla 173 je ve dvojkové soustavě takto: 10101101 (zapisujeme odspodu).

Teď si ještě zkusíme převod binárního čísla 1101101 do soustavy decimální. Tady budeme postupovat jiným způsobem. Jak jsem již říkal, dvojkovou soustavu počítáme pomocí mocnin 2, a to v pořadí 20, 21, 22, atd. Takže převod čísla 1101101 bude vypadat následovně:

$$1 \cdot 2^6 + 1 \cdot 2^5 + 0 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 64 + 32 + 0 + 8 + 4 + 0 + 1 = 109$$

#### Osmičková (oktalová) číselná soustava

Osmičková soustava může obsahovat cifry 0, 1, 2, 3, 4, 5, 6 a 7. Stejně jako dvojková soustava funguje na principu mocnin, ale tentokrát čísla 8. Jednoduší to bude předvést na převodu nějakého čísla, např. 594 do oktalové soustavy. Nejprve si zjistíme největší mocninu osmi, která se vejde do čísla 594, a tou je číslo 512 ( $8^3$ ). 512 se do 594 vejde pouze jednou, takže první číslice bude 1. Dále odečteme 512 od 594 a dostaneme číslo 82 a následuje odečtení mocniny  $8^2$  (64) – ta se vejde do 82

<sup>5</sup> [http://cs.wikipedia.org/wiki/Shannonův\\_theorém](http://cs.wikipedia.org/wiki/Shannonův_theorém)

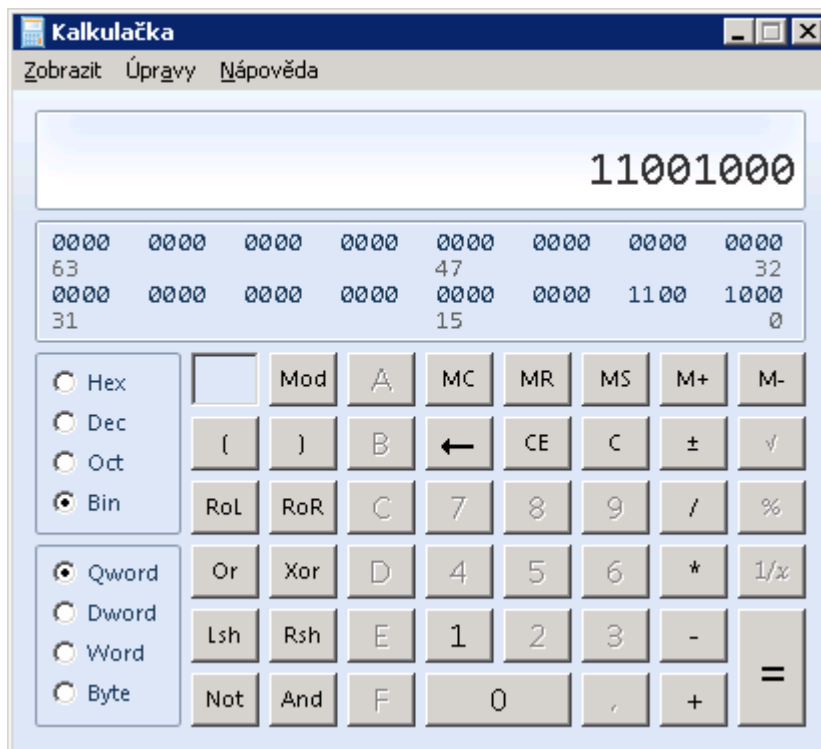
taky jen jednou (druhá číslice je také 1). Po odečtení získáme číslo 18 a následovat bude odečtení mocniny  $8^1$  (8), ovšem 8 se vejde do 18 dvakrát, tudíž 3. číslice bude 2 a po odečtení nám zbude 2. Dvojku budeme dělit mocninou  $8^0$  (1), takže 4. číslice bude rovněž 2. Výsledkem je tedy, že decimální číslo 594 je číslem 1 122 v oktálové soustavě.

### Šestnáctková (hexadecimální) číselná soustava

Šestnáctková soustava zahrnuje číslice 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 a znaky A, B, C, D, E, F a primárním číslem je 16 (respektive opět jeho mocniny). Takže vám ukáži převod čísla 6 540 do hexadecimální soustavy. Opět si najdeme největší mocninu, která se do čísla 6540 vejde a to je  $16^3$  (4 096) – to se tam vejde pouze jednou, takže 1. číslice bude 1. Po odečtení dostanete 2 444 a dělit budete mocninou  $16^2$  (256). Dvě stě padesát šest  $256$  se do 2 444 vejde dokonce devětkrát, takže 2. číslice bude 9. Teď si zjistíte zbytek, a to  $2\,444 - 9 \cdot 256 = 140$ . A 140 vydělíte mocninou  $16^1$  (16) a dostanete se na 3. číslici, na 8. Následuje výpočet  $140 - 8 \cdot 16 = 12$ . Dvanáctka už není dělitelná šestnáctí a v hexadecimální soustavě ji reprezentuje písmeno C (což je 4. „číslíce“). Decimální číslo 6 540 tedy zapíšeme v hexadecimální soustavě jako 198C. Vysvětlovat převod hexadecimálního čísla na číslo decimální opět není, myslím, nutné.

Dek.	Hex.	Bin.
0	0	0000
1	1	0001
2	2	0010
3	3	0011
4	4	0100
5	5	0101
6	6	0110
7	7	0111
8	8	1000
9	9	1001
10	A	1010
11	B	1011
12	C	1100
13	D	1101
14	E	1110
15	F	1111

Rychlé převody mezi soustavami umožňuje kalkulačka.



### 1.1.7 Proces komunikace

#### Manažerská komunikace

Komunikační proces je složen z pěti základních složek, jestliže nějaká chybí, komunikace nemůže začít ani probíhat:

- osoba, od níž určité sdělení vychází (mluvčí – komunikátor) a která může své sdělení i zakódovat,
- osoba, které je sdělení určeno a která se ho snaží dešifrovat a rozumět mu (příjemce – komunikant),
- zpráva, která je předávána (komuniké),
- zpětná vazba (zpráva o tom, že informace byla přijata),
- kontext (prostředí a situace, v které komunikace probíhá), který může zásadně měnit význam sděleného.

Kontext komunikace se může zdát nedůležitý, ale jak zásadně může změnit význam řečeného, můžeme ukázat na příkladě:

- 1) Dva podnikatelé se potkají na mezinárodním veletrhu. Jeden říká druhému: „Jsem skutečně rád, že tě tady potkávám“. Situace, za které se potkali, dává řečenému význam: těší mě, že jsi tak úspěšný, velice ti to přeji a raduji se z toho.
- 2) Dva podnikatelé se potkají a jeden říká druhému: „Jsem skutečně rád, že tě tady potkávám.“ Situace, za níž se potkali, je však zcela jiná: „Těší mě, že na tebe také došlo, zákony jsou spravedlivé, zlomyslně ti to přeji a nezakrytě se z toho raduji“. Potkali se totiž ve věznici, která se pro oba stala přechodným domovem.

Aby komunikace mohla započít, probíhat a plnit svůj účel – umožnit vzájemnou výměnu informací – musí být obě komunikující strany schopny tří základních duševních operací: schopnosti informací vnímat, zhodnotit a uchovat nebo dále předat.

K tomu, aby interakce mohla probíhat produktivně, to znamená tak, aby si komunikující vzájemně poskytovali pro ně důležité, musí být splněny tři základní podmínky:

- oba musí chtít,
- oba musí umět, a
- mít možnost, aby mohli komunikovat.

Chtít, umět a moci. To se zdá být naprosto samozřejmý požadavek, který však ne vždy je splněn. (SCOMP, Manažerské dovednosti, 2010)<sup>6</sup>

**Meziprocesová komunikace** (anglicky Inter-Process Communication, IPC) je **v informatice** sada technik pro výměnu dat mezi dvěma nebo více procesy nebo thready. Komunikace může probíhat i mezi různými počítači propojenými počítačovou sítí. IPC může být též označována jako mezithreadová komunikace (anglicky inter-thread communication) a meziaplikační komunikace (anglicky inter-application communication).

IPC techniky můžeme rozdělit na metody pro zasílání zpráv, synchronizace, sdílená paměť a vzdálené volání procedur (RPC). Metoda použitá pro IPC může záviset na datové propustnosti, latenci komunikace mezi thready a typem dat, která je potřeba přenášet a podobně.

### 1.1.8 Doba přenosu dat po zadané datové lince

Máme určit za jak dlouho se přenesou data o velikosti 100 MB po datové lince s rychlostí přenosu 512 Kb/s.

Použijme vzorec  $t = \frac{\text{velkost dat}}{\text{rychlost přenosu}}$ , čísel i jmenovatel musí být v jednotkách bity nebo Byte.

V našem případě je 512 Kb/s = 0,5 MB/s po osazení dostáváme  $t = \frac{100 \text{ MB}}{0,5 \text{ MB/s}} = 200 \text{ s}$ .

## 1.2 Informační zdroje a jejich kvalita

**Žák dovede:**

- charakterizovat informační zdroje a posuzovat vhodnost jejich použití pro daný účel;
- popsat a využívat služby poskytované knihovnami;
- vyhledat informace pomocí katalogu a pomocí fulltextového vyhledávače, rozlišovat mezi různými způsoby hledání informací;
- vysvětlit způsob fungování vyhledávače a orientovat se ve webovém vyhledávači, využívat rozšířené vyhledávání, formulovat zadání dotazu pro získání relevantních výsledků a orientovat se ve výstupu vyhledávání;
- kriticky přistupovat k informacím a ověřovat informace z různých zdrojů, posoudit relevanci a kvalitu informačního zdroje;
- využívat a vytvářet metadata (metainformace);
- používat myšlenkové mapy pro organizaci pojmů a vztahů mezi nimi;
- vysvětlit princip a přínosy digitalizace reálných objektů, virtualizaci reálných objektů a míst.

### 1.2.1 Informační zdroje a jejich charakteristika

---

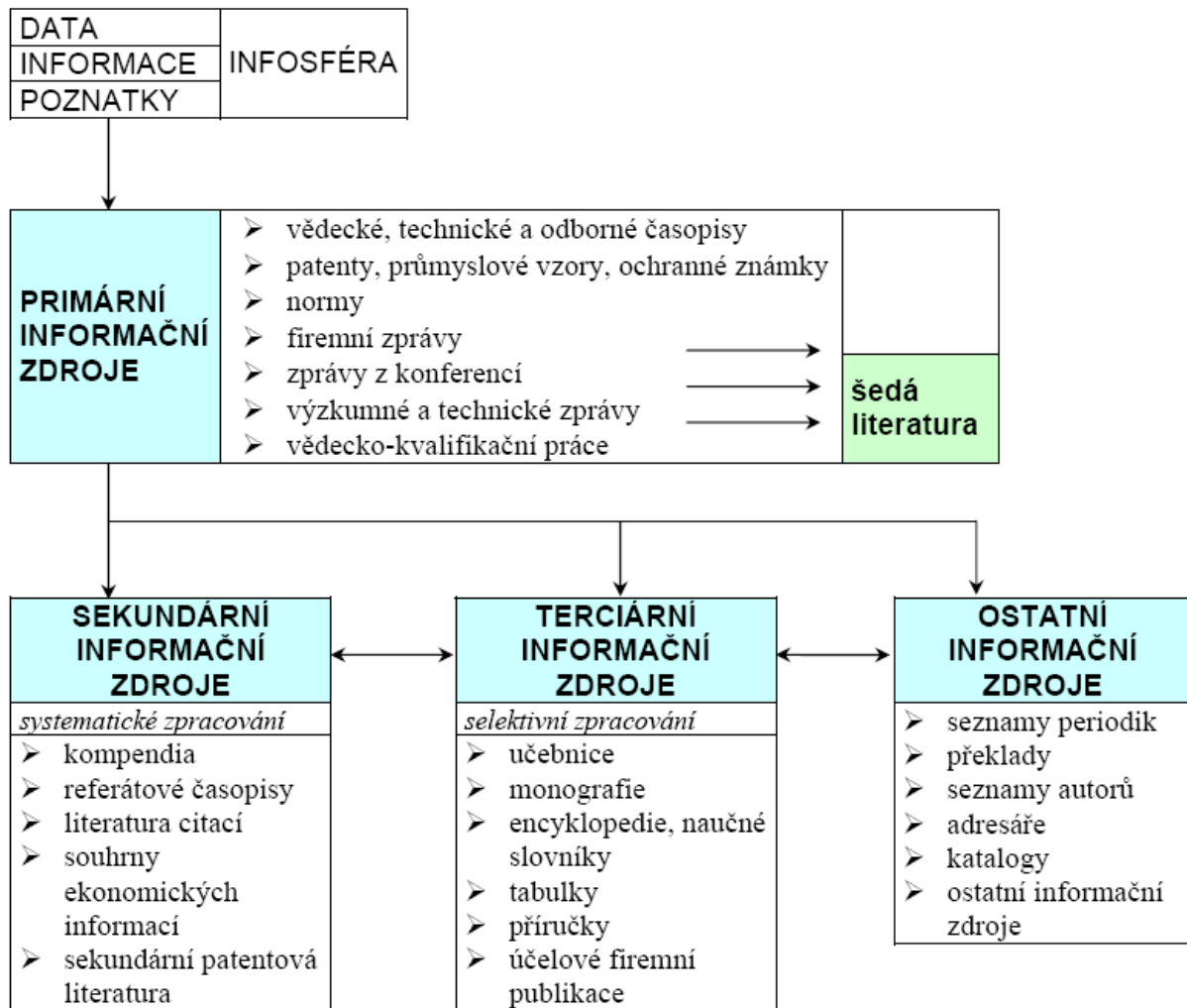
<sup>6</sup> SCOMP,

[http://www.scomp.cz/demo/scomp/eurad/Manazerske\\_dovednosti\\_soci%C3%A1lni\\_komunikace\\_demo/P1002/Page.htm](http://www.scomp.cz/demo/scomp/eurad/Manazerske_dovednosti_soci%C3%A1lni_komunikace_demo/P1002/Page.htm)

## Dokument

Základním informačním zdrojem je dokument, informační pramen sestávající z nosiče informací a množiny dat nebo informací, které jsou na nosiči fixované a formálně i obsahově uspořádané.

V současnosti se tak tímto pojmem nazývají všechny informační zdroje bez ohledu na jejich fyzikální podstatu, formu, obsah a způsob prezentace. Jsou to hmotné jednotky vznikající libovolným zaznamenáním dat či informací za účelem uschování informací (archivace) v čase a jejich přenosu (distribuce) v prostoru.



**Primární informační zdroje** tvoří původní prameny, obsahující bezprostřední nová sdělení o výsledcích odborné a vědecké práce. Nejdůležitějšími dokumenty jsou články v různých odborných časopisech a patenty.

- vědecké, technické a odborné časopisy;
- patenty, průmyslové vzory a ochranné známky;
- normy;
- interní firemní informace;
- zprávy z vědeckých a odborných setkání;
- výzkumné a technické zprávy;
- vědecko-kvalifikační práce.

**Sekundární informační zdroje** představují systematicky zpracované poznatky uvedené v primárních informačních zdrojích. Cílem sekundárních zdrojů je převést tyto původní informace a poznatky do zkrácené podoby a zpřístupnit je formou neperiodických či periodických publikací. Podle rozsahu zpracování dělíme tyto zdroje na:

- kompendia;
- referátové časopisy;
- literaturu citací;
- sekundární patentovou literaturu;
- souhrny ekonomických informací;
- externí firemní informace.

Třetí skupinou (terciární literatura, **terciární informační zdroje**) zpracování poznatků a informací je soubor publikací, přinášející již vyhodnocené informace o primárních nebo sekundárních zdrojích. Obecně bývají tyto poznatky zpracovány selektivně za konkrétním účelem a jsou zpravidla zhuštěné. Zahrnujeme zde:

- učebnice;
- monografie;
- encyklopedie a naučné slovníky;
- tabulky;
- příručky;
- účelové firemní publikace.

**Ostatní informační zdroje.** Zde řadíme především elektronické informační zdroje, které za poslední desetiletí nabyly velkého významu. Patří mezi ně:

- profesionální online systémy, zejména databázová střediska např. DIALOG, DataStar, STN International, LEXIS/NEXIS, ORBIT•QUESTEL, OVID, atd.;
- offline systémy např. firemní intranety;
- optická média (CD-ROM atd.);
- internet;
- kombinace výše jmenovaných. (Informační zdroje, 2010)<sup>7</sup>

### 1.2.2 Služby poskytované knihovnami

Knihovny zabezpečují základní veřejné knihovnické a informační služby vyplývající z knihovního zákona a další speciální veřejné a knihovnické služby.

#### Prezenční výpůjčky

Knihovna poskytuje veškeré materiály ze svého fondu pouze k prezenčnímu studiu do studovny. Mimo prostory studovny si mohou vypůjčit knihy, časopisy i jiné materiály na svá pracoviště pouze pracovníci muzea.

#### Meziknihovní výpůjční služba

Knihovna zprostředkovává výpůjčky dokumentů z jiných knihoven v ČR prostřednictvím meziknihovní výpůjční služby. Zároveň knihovna zpřístupňuje knihovní dokumenty ze svého fondu prostřednictvím meziknihovních služeb i pro jiné knihovny na základě žádanky MVS.

---

<sup>7</sup>[http://fame.utb.cz/mtp/infozdroje/prednasky/ppt/prednaska\\_2.ppt](http://fame.utb.cz/mtp/infozdroje/prednasky/ppt/prednaska_2.ppt)

## **Reprografické služby**

Knihovna poskytuje reprografické služby. Pracovníci knihovny jsou oprávněni určit, které materiály mohou být z kopírování trvale vyloučeny, zejména z důvodu jejich dalšího poškození, je-li zhotovení kopie v rozporu s právními předpisy, jde-li o vzácné tisky apod.

## **Informační služby a bibliografické služby**

Knihovna poskytuje veřejnosti ústní, telefonické i písemné informace bibliografického a faktografického charakteru, informace o katalozích, bázích, fondech a o jejich dostupnosti.

## **Rešeršní služby**

Knihovna vyhotovuje rešerše na zadané téma s využitím všech katalogů, kartoték a pramenů dostupných v knihovně včetně využití databází přístupných též na internetu.

## **Nejvýznamnější české automatizované knihovnické systémy**

### **LANius**

Tento knihovnický systém vyvíjí firma ArrowSys v úzkém kontaktu s Okresní knihovnou v Táboře, kde je také testován a využíván. Vzniká speciálně pro používání ve veřejných knihovnách či knihovnách podobného typu jako knihovna v Táboře, jejichž fond obsahuje maximálně 150 000 svazků. Má výhodu v tom, že neobsahuje nevyužitelné části, které často komplikují přehlednost a použitelnost programu v běžné praxi. Základní záznam je konstruován s ohledem na ty údaje, které jsou potřebné pro vyhledávání ve veřejných knihovnách.

LANius řeší řadu základních i speciálních činností, které na sebe navazují nebo jedna druhou ovlivňují. Společná pro ně je práce s určitou množinou dat (informací). Protože určité skupiny prací jsou úzce časově či obsahově blízké, jsou sdružovány a zpracovávány jedním programem, který se nazývá modul.

Každý z modulů se chová jako samostatný program se všemi potřebnými volbami. Je důležité, že všechny moduly pracují nad společnou datovou základnou a jsou tedy úzce svázány. To znamená, že změníme-li data v jednom modulu, tak se změna projeví na příslušných místech i v ostatních modulech. Přitom je možné používat jeden či více modulů samostatně s možností postupné kompletace systému.

LANius je v současné době také poskytovat nové služby komunikace mezi knihovnami celé naší republiky. Spojení je řešeno pomocí propojení v síti Internetu.

V roce 1995 vznikl základ souborného katalogu SKAT, který umožňuje knihovnám uspokojovat náročné požadavky uživatelů. Dnes obsahuje více než 70 000 titulů naučné literatury z více než padesáti velkých veřejných knihoven ČR. Každý uživatel knihovnického systému LANius se tak může podílet na jeho tvorbě a využívat uložené informace.

Systém LANius je dnes nainstalován ve většině veřejných knihoven. Spolupracující knihovny vytvořily základ katalogů svých fondů, který je zpřístupněn na Internetu a čtenáři z různých míst ČR si mohou vyhledávat informace o jednotlivých knihách a jejich uložení ve fondech veřejných knihoven.

### **CLAVIUS**

Knihovnický systém Clavius začal vznikat na konci roku 1997, kdy firma ArrowSys započala s vývojem zcela nového knihovnického systému ve spolupráce se sdružením KAVKa se sídlem v Uherském Hradišti. Počátkem roku 1998 byl dokončen návrh struktury a ukládání dat a začal vlastní vývoj jednotlivých modulů.



System Clavius není přímým následníkem systému LANius. Tým programátorů začal úplně znova budovat nový moderní systém s odlišnou strukturou záznamu, která je nyní úplně v souladu se standardy a také v budoucna je připravena akceptovat jejich změny.

Součástí nového systému bude plná podpora prezentace dat uložených v na WWW serveru a komunikace s okolím prostřednictvím elektronické pošty. Velkou předností systému je přizpůsobivost konkrétním požadavkům uživatele ve všech směrech. Díky vizuálnímu přístupu v režimu nastavení jsou úpravy formulářů velmi snadné a jednoduché.<sup>8</sup>

### 1.2.3 Vyhledání informací pomocí katalogu a pomocí fulltextového vyhledávače

#### Předmětové katalogy

- „ruční“ budování
- hierarchicky uspořádaný seznam kategorií
- vyhledávání dle témat

#### Výhody

- obecnost vyhledávání
- kvalita a relevance
- hierarchická struktura

#### Nevýhody

- malé rozsahy katalogů
- náročnost na čas
- neaktuálnost

#### Vyhledávače

- automatizovaný sběr dat
- plné texty (fulltext)
- vyhledávání pomocí klíčových slov (jednoduché X pokročilé vyhledávací rozhraní)

#### Výhody

- konkrétnost vyhledávání
- nenáročné na čas
- aktuálnost zdrojů
- velikost databází

#### Nevýhody

- různé GUI
- úplnost vyhledávání
- relevance a kvalita nalezeného

#### Vyhledávání na www, autor: Petr Macek ([spravce@the.cz](mailto:spravce@the.cz)) 14.11. 2001

Internet, to jsou tisíce serverů, na kterých je obrovské množství informací. Podle odhadů je na www serverech uloženo více než 2 miliardy www stránek a denně přibývají tisíce dalších. Jak v takovém množství informací najít to co požadujeme? K tomu vznikly a stále vznikají specializované servery - vyhledávače.

---

<sup>8</sup> [http://www.phil.muni.cz/~klimsova/skola/projekt\\_1.html](http://www.phil.muni.cz/~klimsova/skola/projekt_1.html)

Vyhledávače patří k nejnavštěvovanějším www serverům, například yahoo.com obslouží denně více než 400 000 uživatelů. Protože vyhledávače jsou tak moc využívány, fungují na velmi rychlých a výkonných počítačích. Jedná se většinou o víceprocesorové počítače s výkonnými, převážně RISC procesory. Počítače jsou sdruženy do tzv. clusterů (čti clástrů; cluster = sestava počítačů, které spolupracují: rozdělují si mezi sebou úkoly, při výpadku jednoho stroje přebírají jeho úkoly ostatní, ...). Data se ukládají na velmi rychlé SCSI harddisky (sdružené do diskových polí (RAID)).

### **1.2.3.1 Druhy vyhledávačů**

Rozeznáváme dva druhy vyhledávání: katalogové a fulltextové. Každé z nich má své výhody a nevýhody. Mnoho vyhledávačů kombinuje obě metody. Typ vyhledávání značně ovlivňuje výsledek hledání.

#### **1.2.3.1.1 Katalogové vyhledávání**

Při katalogovém vyhledávání se využívá katalog, v kterém jsou www stránky hierarchicky zaříděny podobně jako v knihovně. Při vyhledávání se prochází jednotlivé úrovně až do té doby, než se najde oblast, která nás zajímá. Pak jsou zobrazeny odkazy na příslušné stránky. Katalogové vyhledávání se používají v situacích, kdy hledáme nějaké informace týkající se určité oblasti. Kromě ručního procházení kategorií můžete také zadat klíčové slovo. Podle něj se pak v katalogu vyhledává klíčové slovo v názvech stránek a v jejich popiscích.

Do katalogového vyhledávače se většinou stránky zadávají ručně. Zadává se jméno stránky a popis, který musí co nejvěrněji popsat obsah stránky. Před vložením do databáze vyhledávače ještě údaje a funkčnost stránky prověří pracovník vyhledávače. Pokud je vše v pořádku, odkaz je umístěn do příslušné kategorie (včetně popisku).

#### **1.2.3.1.2 Fulltextové vyhledávání**

Pro vkládání do fulltextového vyhledávače stačí zadat jen adresu stránky. Stránku pak zindexuje tzv. robot = uloží do databáze všechna slova obsažená ve vkládané stránce. Při fulltextovém vyhledávání se porovnávají uživatelem zvolená klíčová slova s těmi v databázi. Následně se zobrazí všechny stránky, které obsahují hledané klíčové slovo (nebo slova).

Jak je popsáno výše: údaje o stránkách vkládají do databází lidé. V mnoha případech se používají roboti. Ti procházejí internet a sami informace o prohledaných stránkách ukládají. Kvalita výsledné databáze závisí na schopnostech prohledávacích robotů.

### **1.2.3.2 Katalog nebo fulltext?**

Použití katalogu nebo fulltextu závisí na tom, co hledáme. V některých případech je vhodné obě skupiny vyhledávačů použít současně. Katalogové služby většinou lépe vystihují obsah stránek (pokud jsou v katalogu správně řazeny - dostáváme opravdu odkazy na stránky, které požadujeme). Fulltext je vhodný, pokud hledáme informace nezapadající do žádné kategorie (třeba úzce specifická oblast) nebo nějaká konkrétní přesně definovaná informace.

Zajímavou službou jsou vyhledávací centrály. Ty umožňují pohodlně z jedné stránky vyhledávání ihned v několika vyhledávacích službách (vyhledávač sám vytvoří dotaz pro ostatní servery a zeptá se jich). Zadaný dotaz se zpracuje na více vyhledávacích a výsledek všech zobrazí.

### **1.2.3.3 Tvorba dotazů pro vyhledávací služby**

Vyhledává se podle tzv. klíčových slov. Při jejich vkládání do vyhledávače je nutné si dávat pozor na několik věcí:

- rozlišování velkých a malých písmen. Některé servery velikost rozlišují jiné ne (case sensitive)

- diakritika může být také problém. Existuje několik druhů kódování češtiny (Latin 2,1250, Mac , ...). Vyhledavače pak většinou z textu diakritiku odstraní (např: znaky ě,é nahradí za e, to samé provedou se všemi ostatními znaky). Takhle se ošetří i prohledávaná data v databázi a pak se slova porovnávají.

Mnoho uživatelů vyhledává dost neefektivně, při vyhledávání informací zadávají jen jedno klíčové slovo. Vyhledavač pak vrací odkazy na tisíce stránek s daným slovem (např. při zadání slova sex na [www.altavista.com](http://www.altavista.com) mi bylo vráceno 15 584 335 odkazů). Uživatel pak neprochází všechny vrácené stránky, ale max. pár desítek. A protože stránek denně přibývá, stává se vyhledávání s jedním klíčovým slovem stále méně účinné. Příklad. Na dotaz dvou slov (sex a foto) už altavista našla "jen" 81 540 stránek.

Prohledavače se stále zdokonalují (např. možnost vyhledat podobné odkazy, filtrování výsledku vyhledání podle kategorie do které spadá klíčové slovo a pod.), ale chyba není na straně vyhledavače, ale u uživatele. Není praktické zadávat jen jedno klíčové slovo. Pro efektivní vyhledávání je nutno používat složitějších dotazů s více klíčovými slovy a operátory.

### 1.2.3.4 Operátory

Pomocí operátorů se vytvářejí dotazy pro vyhledávání. Dalším pomocníkem při vyhledávání jsou i zástupné znaky (tzv. wildcards)

""

- vyhledávají se stránky obsahující frázi uvedenou v uvozovkách
- u některých serverů je možné uvozovky nahradit závorkami ( ) nebo [ ]

Příklad. na dotaz "zlatý řez" se hledají pouze dokumenty, kde se vyskytuje tato fráze ( a ne jen slova zlatý a řez)

#### AND (a)

- spojuje slova v dotazu
- dá se většinou nahradit znakem '+' nebo '&', na českých serverech i písmenkem 'a' (ne na všech)

Příklad. na dotaz Neruda & dílo se hledají dokumenty, kde se vyskytují obě slova tj. mělo by najít stránky dílech Jana Nerudy

#### OR (nebo)

- vyhledává dokumenty ve kterých je obsaženo alespoň jedno ze zadaných slov
- dá se nahradit znakem '|'

Příklad. na dotaz Česko OR Slovensko se hledají dokumenty obsahující oba dva termíny nebo jen jeden z nich

#### NOT (ne)

- při hledání vyloučí dokumenty, v nichž je obsaženo dané slovo
- dá se nahradit znakem '-' nebo '!'

Příklad. na dotaz televize NOT Nova se hledají všechny televize kromě Novy.

#### NEAR

- hledá slova která jsou poblíž sebe

Příklad: slovo1 NEAR slovo2 vyhledá dokumenty, kde se vyskytuje slovo1 a do určitého počtu slov od něho se vyskytuje i slovo2 Počet slov mezi hledanými slovy je možné upravit pomocí znaku "/" a čísla (slovo1 near/10 slovo2)

#### **FAR**

- opak NEAR, hledá slova která jsou od sebe vzdálená Vzdálenost mezi slovy je možné opět upravit pomocí "/"

#### **Zástupné znaky (wildcards)**

Většina vyhledávacích služeb podporuje tzv. zástupné znaky (wildcards). Ty se používají i například při práci se soubory a při vyhledávání obecně. Jedná se o znaky, které zastupují jeden nebo více znaků v hledaném slově.

Standardně se používá :

\* (hvězdička) jako zástupný znak pro více znaků

? (otazník) jako zástupný znak pro jeden znak

Příklady :

sou?ek - může najít např. výrazy souček, soudek, ...

Mar\* - může najít nepřeborné množství slov, začínajících na Mar (Mars, Martin, Martina,...)

Je možné použít i oba zástupné znaky v jednom slově nebo i vícekrát v jednom slově.

#### **1.2.3.5 Vyhledavače**

<b>České</b>	<b>Cizí</b>
<a href="http://www.seznam.cz">www.seznam.cz</a> Obrovský katalog + fulltext. Vyhledávání lidí a emailů. Další služby (SMS brána, email, zpravodajství, ...)	<a href="http://www.google.com">www.google.com</a> Žádné doplňkové služby - jen fulltext. vyhledávání. Ale za to skvělé vyhledávání
<a href="http://www.atlas.cz">www.atlas.cz</a> Katalog i fulltext + spousta dalších služeb	<a href="http://www.altavista.com">www.altavista.com</a> Podporuje složité vyhledávací příkazy, bohužel databáze je aktualizována pomalu
<a href="http://www.centrum.cz">www.centrum.cz</a> Jeden z nejoblíbenějších vyhledavačů, má katalog i fulltext	<a href="http://www.yahoo.com">www.yahoo.com</a> Katalog i fulltext
<b>Další vyhledávače</b>	
<a href="http://www.redbox.cz">www.redbox.cz</a> <a href="http://www.najdi.to">www.najdi.to</a> <a href="http://www.quick.cz">www.quick.cz</a>	<a href="http://www.aol.com">www.aol.com</a> <a href="http://www.goto.com">www.goto.com</a> <a href="http://www.lycos.com">www.lycos.com</a>

#### **1.2.4 Způsob fungování vyhledávače**

## Co jsou vyhledávače

Třeba Google, Jyxo, Morfeo, fulltext Seznamu nebo třeba Yahoo.com. Servery, které mají velkou databázi stránek s informacemi o tom, která stránka internetu obsahuje jaké slovo. Protože takové vyhledávače umějí hledat v celém textu stránek, říká se jim fulltextové. Takových serverů existují řádově stovky. Nepleťte si prosím fulltextové vyhledávače s katalogy (Yahoo, katalog Seznamu), ačkoli poslední dobou (2005) se to k sobě dost blíží.

Programům, které procházejí web a schraňují pro vyhledávače informace, se říká roboti, boti nebo crawleři. Na české stránky přichází nejvíce uživatelů přes fulltext Seznamu. Světově nejdůležitější vyhledávač je Google. Údaje uvedené níže platí ale pro všechny fulltextové vyhledávače.

## Jak pracují vyhledávače

Ve třech krocích:

- 1) Roboti (crawleři) sbírají data
- 2) Pak se to zpracuje do databáze (indexování)
- 3) Nakonec se to zpřístupní návštěvníkům, aby mohli klást dotaz (webmodul).

## Robot prochází internetem

Většinou někde začne, třeba v nějakém katalogu. Najde odkaz a sleduje ho. Nalezenou stránku si stáhne pro zaindexování, zároveň hledá další odkazy, které případně opět sleduje, aby si to zase postahoval.

Z toho vyplývá, že robot může odhalit pouze stránku, na kterou vede nějaký odkaz. Musí přitom jít o klasický. Pokud na vaše stránky nebudou roboti chodit, žádný čtenář ty stránky nemůže najít fulltextovým vyhledávačem.

V praxi ale roboti najdou jenom část stránek, které existují, protože to prostě všechno nestíhají.

- Aby se nezdržovali blouděním, mají ve zvyku sledovat odkazy v témž webu jen do určité úrovně (1 až 2) nebo do určitého počtu.
- Roboti si pamatují, kde už byli, aby tam nemuseli znovu. Čas od času se robot na stránky vrací, aby se podíval, jak se změnil.
- Roboti skoro vůbec nestahují pomocné soubory jako obrázky nebo stylopisy.
- Někteří roboti se prý nedokáží nebo nechtějí prokousat rámovou strukturou.<sup>9</sup>

### 1.2.5 Relevance a kvalita informačního zdroje

*Jonák Zdeněk, Převzato z <http://clanky.rvp.cz/clanek/o/z/120/JAK-ROZPOZNAT-VEROHODNOST-RELEVANCI-SPOLEHLIVOST-INFORMACI-A-INFORMACNICH-PRAMENU.html/>*

Jedním z charakteristických rysů informační společnosti je dramatický, obtížně zvládnutelný nárůst informací. Publikační exploze vypovídá o tom, že věda dospěla do stadia, kdy svými aplikacemi v oblasti průmyslových a informačních technologií zasahuje do stále většího počtu oblastí společenského života.

Ve světě vychází ročně 70 000 odborných a vědeckých časopisů, což představuje 100 000 stránek za den ve více než 65 jazycích. Denně je publikováno 6000 - 7000 vědeckých článků, přičemž počet publikovaných článků se každých 5,5 roku zdvojnásobuje. V současnosti žije 80 - 90 % všech vědců v dějinách lidstva. (Vymětal, 2000) Internetové prohlížeče obsahují řádově miliardy webových stránek a zabezpečují služby desítkám miliónů uživatelů.

---

<sup>9</sup> <http://www.jakpsatweb.cz/vyhledavace.html>

Publikační a informační exploze má ale i své negativní stránky. Stále větší měrou narůstá počet informací a informačních zdrojů, které nejsou zárukou zmnožení dosavadního poznání. Nejenže mu mnohdy neprospívají, někdy dokonce škodí.

Vzdělávací instituce i rodiče stojí před problémem ochránit mladou generaci před zavádějícími poznatky, ale i před zhoubným vlivem eticky závadných stránek. Často ani soudný člověk (natožpak dítě školou povinné) se nedokáže v množství informací orientovat a při nedostatku času spolehlivě odlišit informace relevantní od nerelevantních, ověřené od neověřených, pravdivé od nepravdivých. Zvláště dospívající jedinec není, bez předchozího varování, často připraven předpokládat, že by veřejně kontrolovaná média mohla poskytovat falešné informace.

### **1.2.5.1 Spolehlivé nebo nespolehlivé?**

Pokusíme se alespoň ve stručnosti poukázat na nástroje, metody či alespoň intuitivní mechanismy, které napomáhají kvalifikovat kvalitu informačních zdrojů a s určitou pravděpodobností usnadňují dosažení vyšší míry spolehlivosti při jejich selekci.

Seriózní vědecké články a publikace by měly splňovat následující kritéria.

#### **1. Odkaz<sup>1</sup> na zdroj poznatků**

Pokud vědecký článek odkazuje na prameny postrádá, může čtenář považovat text za úvahu, fejeton nebo jinak popularizující článek, v horším případě za pokus přiživit se na výsledcích práce skutečného badatele.

Čtenář, který chce proniknout k jádru věci, by se nikdy neměl spoléhat na jediný informační zdroj, ale ověřit, zda odkaz (pokud existuje) je citován přesně nebo ho autor článku upravil ve svůj prospěch. Cílevědomý čtenář by měl usilovat o nalezení sítě odkazů směřujících k určitému prvotnímu zdroji (jakým je např. v oblasti teorie relativity Einstein), od něhož se ostatní teorie odvíjejí, a od tohoto zdroje dospět k jeho současným důsledkům. Internet díky síti hypertextových odkazů tuto možnost plně podporuje.

#### **2. Garance spolehlivosti zdroje**

Důležitou indicií dokumentu je již samo jméno autora a instituce, kterou reprezentuje. Prostým nahlédnutím do encyklopedie, slovníku typu Who is Who či dotazem odborné autoritě lze věrohodnost autora ověřit. Existuje vědní odvětví - scientometrie -, které používá exaktní metody k identifikaci faktorů, jež by mohly automaticky (u intelektuálně jinak nezvládnutelného množství informací) odlišit kvalitní informace od nekvalitních. Scientometrie (též infometrie, naukometrie apod.) představuje metodu stanovení kvality vědecké práce, která se opírá o jednoduchý intuitivní předpoklad, že čím více si vědecká obec určité publikace všímá a cituje ji, tím větší hodnotu publikace má. "Hodnota" dokumentu je sumou hodnot dokumentů, které na daný dokument odkazují, dělenou počtem odkazů vedoucích z dokumentu. Pro získání vysokého hodnocení musí na daný dokument vést buď velké množství odkazů z málo významných dokumentů nebo stačí i menší množství z významných dokumentů. Spolehlivost informačního zdroje lze odvodit z míry citovanosti<sup>2</sup>, která se zjišťuje tzv. impaktovým faktorem (IF) časopisu, ve kterém je informace zveřejněna. Hodnota IF udává, kolikrát byl článek v určitém časopisu citován během prvních dvou let. Snahou vědce by mělo být publikování v nejkvalitnějších časopisech s maximálním IF. (Ukázka časopisů s uvedením zmíněných ukazatelů viz tabulka <http://www.geocities.com/bioreference2003/impact2003-1.htm>)

Dalším faktorem, který umožňuje odlišit relevantní informace od méně relevantních, je stáří informace. Tento ukazatel vychází ze zkušenosti, že užitná hodnota informace klesá v závislosti na čase. Příčinou je dynamika vývoje vědy a překonávání starších poznatků novými. V souvislosti s exponenciálním růstem informací (každé 2 až 4 roky se v průměru objem informací zdvojnásobí) se

mluví o zrychlení tempa stárnutí informací. Stárnutí se projevuje v souvislosti s tzv. informačním cyklem, tj. rychlostí toku informace od jejího vzniku v hlavě autora ke vzniku rukopisu, článku či knihy v nakladatelství až po její distribuci na knižním trhu a zpracování v knihovnicko-bibliografickém procesu. Cílem je tento cyklus zkracovat. V elektronickém publikování se čas mezi vznikem rukopisu a jeho publikováním daří zkracovat na minimum.

Průměrná hodnota životnosti informace (poločas stárnutí) se zjišťuje statistickým rozbořem souborů informačních zdrojů. V roce 1960 bylo pravděpodobně poprvé použito analogie mezi průběhem vzrůstu vědecké literatury a průběhem rozpadu radioaktivních látek. V obou případech je možné při měření průběhu události použít pojmu "poločas". Informační poločas je čas, který uplyne od zveřejnění informace do doby, kdy se v literatuře objeví polovina informací o této tematické. Protože tento čas lze jen těžko měřit, zjišťuje se zpravidla tzv. zpětný poločas, tj. čas, za který se od určitého data vyskytne ještě polovina citací. Tento čas ve skutečnosti představuje dobu, během níž zestárla polovina publikované literatury, nebo také čas, ve kterém kulminuje využívání určitého informačního pramene.

Informační věda zjistila některé zákonitosti chování informací, které umožňují exaktně odlišit relevantní zdroje informací od nerelevantních. Jedním z nich je Bradfordův zákon, který popisuje rozptyl dokumentů v určité vědní oblasti. Rozptylem rozumíme vlastnost informací projevující se tím, že relevantní informace se vyskytují v různých dokumentech, kontextech a v různé formě. Bradfordův zákon stanoví závislost mezi celkovým počtem dokumentů (nejčastěji časopisů) a rozložením článků týkajících se jisté tematiky. Ze zákona vyplývá, že maximální počet relevantních článků je soustředěn v minimálním počtu časopisů tvořících tzv. jádro oboru.

Znalost tohoto "jádra" představuje další z možností, jak odlišit informace neověřené od ověřených. Existují oborové "žebříčky" časopisů, uspořádaných podle stupně blízkosti k jádru.

### **3. Architektura struktury uspořádání myšlenek a argumentů v článku**

Ta může signalizovat výskyt zjevných nesrovnalostí a protimluv ve výrociích a tvrzeniích. Prvotním vodítkem je intuice čtenáře, která ho díky jeho zkušenostem a nabytým vědomostem dovede varovat před očividně nespolehlivými informacemi. Škola zásobuje žáka spolehlivými a zaručenými poznatky, které jsou fixovány v encyklopediích a učebnicích. Ve svém soukromém životě, v kontaktu s vrstevníky, dospělými a médií nabývá však žák zkušenosti, jež jsou často v rozporu se získanými poznatky. I neucelené školní poznatky se však mohou stát základnou, o níž se žák může opřít. Rozpory s fakty textu a znalostmi získanými v přírodovědě, přírodopise, fyzice apod. jsou navzdory své neúplnosti prvními vodítky při rozlišování věrohodné a nevěrohodné informace.

### **4. Pravidlo o ekonomii myšlení (Occamova břitva)**

Bez tohoto kritéria by neexistoval spolehlivý a rozhodující způsob, jak si vybrat mezi jednotlivými hypotézami. Occamova břitva uplatňuje názor, že pokud existují dvě nebo více vysvětlení nějakého fenoménu, mělo by být zvoleno to nejjednodušší.

Příznakem teoreticky málo přínosných prací bývá často neúčelné rozpitvávání a zamlžování. Uvedené pravidlo doporučuje zabývat se výlučně tím, co je podstatné, nutí autory opírat se doložitelná fakta a používat podložené argumenty.

Důležitost ekonomického myšlení se projevuje významnou měrou v oblasti ICT zejména při algoritmizaci a tvorbě programů. Neúspěšně vytvořené programy se při obrovských kapacitách toků dat a množství operací uplatňovaných nad těmito daty projevují měřitelnými finančními a časovými ztrátami.

### **1.2.5.2 Ověřování spolehlivosti informací na internetu**

Žák se často již v období předškolního a elementárního vzdělávání setkává s obrovským množstvím textových, obrazových a zvukových informací uložených na nejrůznějších multimédiích. Internet přinesl do publikační činnosti naprostou svobodu uveřejňovat na celosvětové síti cokoliv. Autor může navíc na tomto médiu cokoli ze dne na den změnit. Zdánlivě neexistuje možnost, podle níž lze s určitou pravděpodobností odlišit informace kvalitní od nekvalitních. Přesto lze nastavit určitá pravidla proti tlaku komerce a pokleslé vědy. Internetovým prohlížečům pomáhají k řešení tohoto problému poznatky řady věd - lingvistiky, kognitivní vědy, umělé inteligence apod. Na následující adrese lze zjistit statistické údaje o nejvýznamnějších webových prohlížečích.

Jednodušší vyhledávače vyhledávané dokumenty uspořádávají do struktury vytvořené na základě analýzy frekvenční struktury textu. Stránka, jejíž frekvenční struktura se největší měrou blíží struktuře dotazu, tzn. slova formulovaná v dotazu se maximálně shodují se slovy v textu, je vybrána do souboru vyhledaných dokumentů.

Nejinteligentnější vyhledávače však používají metody opírající se jak o hlubší lingvistickou analýzu textu, tak o rutinizaci a kvantifikaci intuitivního předpokladu, že nejvíce odkazů směřuje na známé a oceňované stránky. Např. internetový prohlížeč Google řadí vyhledané stránky do pořadí nejen podle shody zadaného dotazu a frekvenční struktury textu. Významným ukazatelem, který rozhoduje o pořadí významnosti vyhledaného textu mezi ostatními, je počet odkazů, které vedou na danou stránku z jiných stránek. Tuto úlohu plní funkce PageRank.<sup>3</sup>

### **1.2.5.3 Neetické chování tvůrců webových stránek**

Někteří tvůrci stránek se pokoušejí obelstít hodnotu indexovacích systémů prohledávačů tím, že opakují stejné obsahy na několika stránkách. To zvyšuje údaj o frekvenci výskytů, a tím i pravděpodobnost, že obsah jejich stránek bude upřednostněn. Určitý počet webových stránek uloží naprosto stejný, duplicitní obsah na různých stránkách nebo doménách. Kdyby vyhledávače neuměly rozeznat, že jde o stejný obsah, uživatelé by mohli být klamáni, protože by ve výsledcích dostávali nadbytečný počet naprosto stejných stránek.

Vyhledávače mají proto algoritmy, kterými zjišťují, zda jde o stejný obsah. Jsou prý schopné poznat, že se stránka liší třeba jen navigací. Prohledávač Google např. duplicitní stránky indexuje tak, že si vybere jednu stránku z mnoha stejných a ostatní ignoruje. Webmasteři používají ke zmatení vyhledávačů i skrytý text obsahující klíčová slova, neviditelné odkazy a jiné metody. Google z těchto důvodů přistoupil k penalizaci takových praktik. Podobné kvalitní služby jako zmiňovaný Google nabízí i některé další - alltheweb.com a náš jyx.cz.

### **Závěr**

Odhalování neseřízných přístupů k pravdivému poznání skutečnosti nemá vždy tak průkazné prostředky, jaké byly uvedeny v našem článku "Mimikry", které používají někteří autoři, dokáží skryt pravdivé poznání tak úspěšně, že oklamán může být i seriózní vědec.

Chtěli bychom od jednoznačných a snadno průkazných dokladů, které jsme předvedli, přistoupit k případům, v nichž jsou vědecké pravdy přehlušeny komerčními zájmy jedinců (popř. celých korporací), kteří jsou ve svém zájmu ochotni dokonce obracet výsledky vědy proti ní samotné. Obrana proti těmto přístupům vyžaduje ze strany zastánců vědy trpělivější a radikálnější přístupy.

### **Literatura**

Česká terminologická databáze knihovnictví a informační vědy (TDKIV)

[sigma.nkp.cz/F/?func=file&file\\_name=find-a&local\\_base=ktD](http://sigma.nkp.cz/F/?func=file&file_name=find-a&local_base=ktD)

Kritické myšlení [www.fzs-chlupa.cz/o-skole/charakter-vyuky/metody-vyuky/kriticke-mysleni/](http://www.fzs-chlupa.cz/o-skole/charakter-vyuky/metody-vyuky/kriticke-mysleni/)



VYMĚTAL, J. (2000) Současné informační prostředí. Vesmír, s. 335-336.

#### 1.2.5.4 Vysvětlivky v textu

**1 Odkaz na zdroj poznatků** (bibliografická citace) představuje formalizovaný údaj o dokumentu, který autor bezprostředně použil při přípravě své práce.

2 Faktor vyjadřující informační hodnotu, oblíbenost dokumentu. Stanoví se metodou citační analýzy zjišťující vztahy citujícího a citovaného dokumentu. Slouží jako podklad při citačním mapování vědy pomocí konstrukce citační sítě a praktické využití nalézá rovněž při rozhodování o profilování a optimalizaci knihovnických fondů.

3 Page Rank je skalární číslo přiřazené každé stránce. Vyjadřuje něco jako věrohodnost nebo důležitost stránky, dosahuje hodnot od 0 do 1. Google si Page Rank počítá (zjednodušeně řečeno) podle toho, kolik a jak důležitých stránek na tu počítanou stránku odkazuje. Existuje na to vzoreček. Nainstalováním Google Toolbar lze zjistit jakou hodnotu PageRank (na škále 0-1) má vámi vyhledaná stránka.<sup>10</sup>

#### 1.2.6 Využití a vytváření metadat (metainformace)

*Autor: Srna Robert, <http://computerworld.cz/software/metadata-ve-firemnich-informacnich-systemech-4640>*

Metadata jako data o datech představují významný prvek pro získání informací či dokonce znalostí. V tomto článku si přiblížíme jejich pozici a specifika ve firemních IS.

Binární reprezentace informací bez jejich interpretace nemá pro člověka žádnou hodnotu. Metadata jsou proto nutnou podmínkou (nikoliv však dostačující) pro získání informací nebo dokonce znalostí a slouží k vysvětlení významu čísel nebo třeba popisují stav či procesy. Využívání metadat ve společnostech přispívá obecně k vyšší efektivitě jejich procesů, ať už se jedná o proces, kdy předáváme práci kolegovi, nebo o postup přípravy testovacího prostředí či o optimalizaci zákaznického centra.

Ve všech případech existuje potřeba pro správné rozhodnutí interpretovat existující nebo právě publikovaná data. Podíváme-li se do praxe, pak zaměstnanci call centra obtížně zjišťují typ telefonu zákazníka či druh komponent tvořících produkční prostředí - metadata tak vlastně popisují pracovní proces při předávání agendy.

Velmi známým příkladem metadat může být atribut tabulky (jedná se o tabulku, která spravuje databázový server) – třeba atribut [datum fakturace] vysvětluje v tabulce [faktury] význam konkrétní hodnoty [17.3.2009]; viz příklad tabulky.

?	?	?	?	?	datum fakturace	?	?
902483028	2009	PR	748374	17.3.2009	17.3.2009	31.3.2009	21227
673284872	2009	PS	132641	17.3.2009	17.3.2009	17.3.2009	877

Je nutné rozlišovat několik typů metadat. V první řadě je členíme na strukturovaná a nestrukturovaná, což má význam už jenom proto, že podle odhadů je velká většina metadat nestrukturovaných. Mezi strukturované patří např. datové modely ERD nebo formuláře účetní aplikace či zdrojový kód. Jako příklad těch nestrukturovaných mohou posloužit například zákony, smlouvy.

<sup>10</sup>Jonák Zdeněk, <http://clanky.rvp.cz/clanek/o/z/120/JAK-ROZPOZNAT-VEROHODNOST-RELEVANCII-SPOLEHLIVOST-INFORMACI-A-INFORMACNICH-PRAMENU.html/>

Dále jde o rozlišení na technická a obchodní metadata – to zase vypovídá o obecnstvu, které je udržuje a užívá. Nikoho nepřekvapí, že strukturovaná metadata se velmi často protínají s těmi technickými.

Metadata mají několik zvláštností – vyskytují se všude, a to jak na formulářích aplikací, tak v e-mailech, a pochopitelně i v samotných hlavách jedinců. Velmi často má jeden termín více významů – např. termín [účet] má jiný význam pro prodejní oddělení a pro oddělení finanční. Často není možné určit zodpovědného vlastníka metadat, protože ta patří celé firmě – namísto vlastníka se proto určuje tzv. steward.

### **Struktura metadat**

Jako všechny informace i metadata je možné organizovat. Organizace přispívá k pořádku a zejména k jejich vyššímu využití. Pokud se uživatel dotazuje například na databázi umělců a zajímají ho osoby žijící na vybraném území, bude pro něj užitečné, když bude geografická organizace metadat odpovídat jeho dotazu. Jestliže tak bude hledat v jednom okresu a geografická hierarchie nebude pokrývat úroveň okresů, nemůže být jeho dotaz efektivně zodpovězen.

Struktuře metadat se věnuje celá řada standardů, počínaje ERD a UML až po RDF-OWL. Cenným pomocníkem jsou tzv. case nástroje. K hlavním funkcionalitám těchto prostředků patří design metadat a jejich export.

Pod strukturováním metadat je třeba vidět především vytváření hierarchických vazeb mezi objekty (objekt se v UML nazývá [Element], v RDF se objekt nazývá [Resource]). Např. geografické objekty lze uspořádat jako Země -> Okres -> Město. Dále jsou důležité specifikace formátu dat (v jakém formátu je číslo, datum...) či jazyka, v němž je popis uveden, a také definice omezení (constraints) – například to, že datum narození nemůže být novější než to aktuální.

### **Přínosy správy metadat**

Správa v tomto případě v sobě zahrnuje sběr metadat, jejich uspořádání a distribuci. Pro příklad jsou k dispozici pouze metadata databáze a uživatelské aplikace, která zpřístupňuje data z databáze pro koncové uživatele. Ta prvně jmenovaná popisují databázové objekty a transformace. Aplikační zase definují objekty a transformace na aplikační vrstvě, tedy například to, co se v aplikaci děje, než se čísla dostanou na formulář. Že to někomu připadá jako fikce? Ještě to sice není ideální stav – bylo by žádoucí metadata z obou vrstev propojit – přesto ani takový stav není dnes běžnými prostředky dosažitelný.

Sběr metadat totiž v podstatě znamená určitou duplikaci skutečnosti – její míra záleží na účelu, jakému mají sloužit. Jiný detail potřebuje rozhodnutí, které aplikace čtou informace z jaké databáze, a jiný stupeň abstrakce vyžaduje uživatel reportu, jemuž „nesedí“ data. Míra detailu ovlivňuje také náklady vlastnictví metadat.

### **Hlavní přínosy správy metadat spočívají v:**

- 1) porozumění datům. Bez nich nemají data hodnotu. Správná interpretace informací je nutnou podmínkou pro identifikaci těch nečistých a pro jejich opravu. Příkladem může být vysvětlení termínu „tržba“ na prodejním reportu. Představuje hodnota částku vč. DPH, zahrnuje slevy či prodeje v ekonomické skupině?
- 2) podílu na efektivitě procesů. Metadata poskytují různou míru abstrakce – umožňují měnit úroveň detailu, tak aby odpovídala potřebám - na rozdíl od zdrojového kódu, který obsahuje pouze tu nejnižší míru detailu. Mezi zajímavé procesy založené na metadatech patří impact analýza (závislosti mezi objekty metadat jsou klíčové pro hledání dopadů změn; v IT se typicky řeší, co se rozbije, pokud změníme určitou proceduru, v obchodních odděleních se

často zkoumá, odkud se berou data) a efektivita využití aktiv (analýza závislostí hardwaru a softwaru pomůže s větší jistotou připravit testovací prostředí, implementovat virtualizaci atp.).

- 3) podpoře integrace. Výměna dat z různých systémů bývá založena na tzv. rozhraní, což jsou metadata, která popisují, jakým způsobem systém komunikuje s okolím.
- 4) sdílení znalostí. Některé nástroje umožňují uživatelům bezprostřední reakci na informaci. Tu je možné změnit např. formou diskuse či revize, nebo si vyžádat od moderátora či stewarda určité vyjádření. Příkladem může být správa zmiňovaného termínu „tržba“ ve firemním wiki slovníku.
- 5) realizaci centrálního vyhledávání. Je naprosto nutné podporovat různé strategie pro přístup k metadatům. Správné kanály jejich distribuce vedou jistě ke snížení nákladů na doručování a tisk. Určité skupiny uživatelů ocení notifikace změn například pomocí RSS. Fulltextové vyhledávání by mělo podporovat práci se synonymy a se strukturami metadat – v již zmíněném geografickém příkladu bych měl mít možnost vyhledat všechny pobočky ve státě, přestože na detailní úrovni sleduji pobočky na úrovni měst. Vyhledávání stavu informace k určitému datu je dalším často opakovaným požadavkem spolu s porovnáváním změn.
- 6) řízené centrální péči o metadata. Centrální zálohování ochrání informace, které mohou být snadno ztraceny na lokálních discích uživatelů. Jiná forma pečování o metadata je řízení přístupů k nim. S tím, jak jejich objem roste, nabývá na významu i omezování přístupu k nim.

Jako každý projekt musejí i metadatové projekty přinášet společnosti užitek. Finanční vyjádření obchodních dopadů se často opírá o zvýšení účinnosti takových procesů, které jsou ve podniku dobře zmapované (práce zákaznického centra, help desk...).

### **Technické prostředky správy**

Základní prostředky pro sběr strukturovaných metadat tvoří různé case nástroje, které podporují doménovou oblast a často ji i normalizují. Základní funkcí těchto nástrojů je tvorba a správa modelů (modely na zelené louce, reverzování, import, verzování a porovnávání rozdílů), vyhledávání závislostí (impact analýza), generování kódu nebo jeho kostry, export modelu, plánování práce a její záznam či validace modelů, analýzy kódu a práce s pravidly.

Oblast nestrukturovaných metadat pokrývají různé tzv. collaboration nástroje, které kladou důraz hlavně na bezpečnou a snadnou komunikaci, jako je zabezpečený přístup, podpora různých přístupových kanálů (desktop a mobile verze), multimediální obsah a integrace nástrojů do operativních aplikací, fulltextové vyhledávání, plné verze dokumentů nebo jejich vlákna, centrální zálohování nebo strukturování (často formou obsahů a rejstříků nebo tagů; u těch je zajímavé zmínit tzv. folksonomy, kdy strukturu určují ostatní uživatelé, nikoli sám autor dat/textu...)

Na těchto dvou základních skupinách je patrný velmi podstatný rozdíl v normalizaci obsahu. Dosažení normalizované struktury je pro technická metadata podstatné pro jejich využití v dalších fázích technických projektů, jako jsou tvorba zdrojového kódu, odstranění chyb či návrh dle bloků a návrhových vzorů (patterns).

S nestrukturovanými metadaty mohou dnes pracovat hlavně lidé, což je nevýhodou v technických projektech, kde je žádoucí vysoké zapojení softwaru. Procesu strukturalizace původně nestrukturovaných metadat se věnuje čím dál tím více pozornosti. Hlavní motivací této pozornosti je získání dalších informací a znalostí vedoucích velmi často k získání konkurenční výhody.

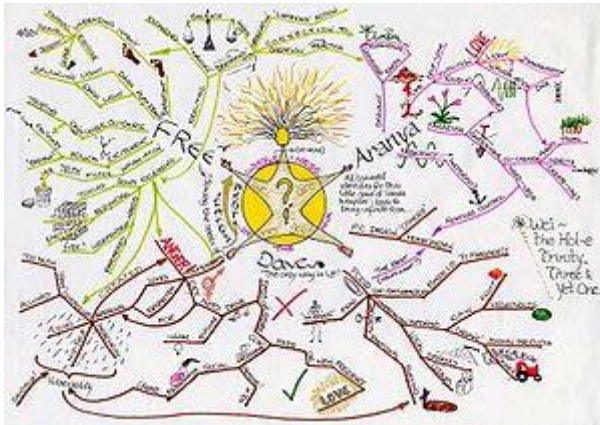
Zvláštní skupina prostředků se zabývá sdílením a integrací metadat. Významné instituce se snaží definovat formáty výměny anebo principy integrace. Například OMG (Object Management Group) má formát XMI, který slouží zejména pro export a import UML modelů z case nástrojů.

W3C oblast metadat zase přispívá celou sadou konceptů pro sémantický web – jedná se o standardy pro strukturalizaci a pro integrace. Mnoho nástrojů business intelligence se věnuje sdílení a publikaci metadat – bohužel jde velmi často o řešení vytvořené pro dané BI prostředí, které nemá zanedbatelné celkové náklady na vlastnictví (TCO).

Autor pracuje jako konzultant ve firmě Adastra.  
autor Robert Srna

### 1.2.7 Myšlenkové mapy

Myšlenková mapa (někdy také mentální mapa) je graficky uspořádaný text doplněný obrázky s vyznačením souvislostí. Po staletí byla využívána k učení, pamatování, grafickému zobrazení nebo řešení problémů. Jedním z prvních uživatelů byl Porfyrios z Tyru.



#### Jak tvořit mapu

Tradiční postup tvorby mapy, který je však třeba pozměnit, aby nejlépe vyhovoval individuálním záměrům.

- 1) Začněte ve středu papíru hlavním námětem
- 2) Využijte obrázků, symbolů, kódů
- 3) Vyberte hlavní témata a zdůrazněte pomocí velkých a malých písmen
- 4) Využijte barev
- 5) Vytvořte si svůj osobní styl tvorby myšlenkových map.<sup>11</sup>

**FreeMind** je svobodný software pro tvorbu myšlenkových map, licencovaný pod GNU GPL. Je napsaný v Javě. Mezi nejvýznamnější funkce aplikace FreeMind patří:

- Rozbalování a zabalování větví dokumentu
- Export do HTML, XHTML, PNG, JPEG, SVG, PDF, Flash
- Prohlížení nativního formátu .mm ve webovém prohlížeči pomocí java applet pluginu
- Ikony, vázané k uzlům
- Obláčky kolem větví
- Grafické odkazy spojující uzly
- Hledání, včetně omezení na jednotlivé větve

<sup>11</sup> [http://cs.wikipedia.org/wiki/My%C5%A1lenkov%C3%A1\\_mapa](http://cs.wikipedia.org/wiki/My%C5%A1lenkov%C3%A1_mapa)

- Uzly mohou odkazovat na web i lokální soubory

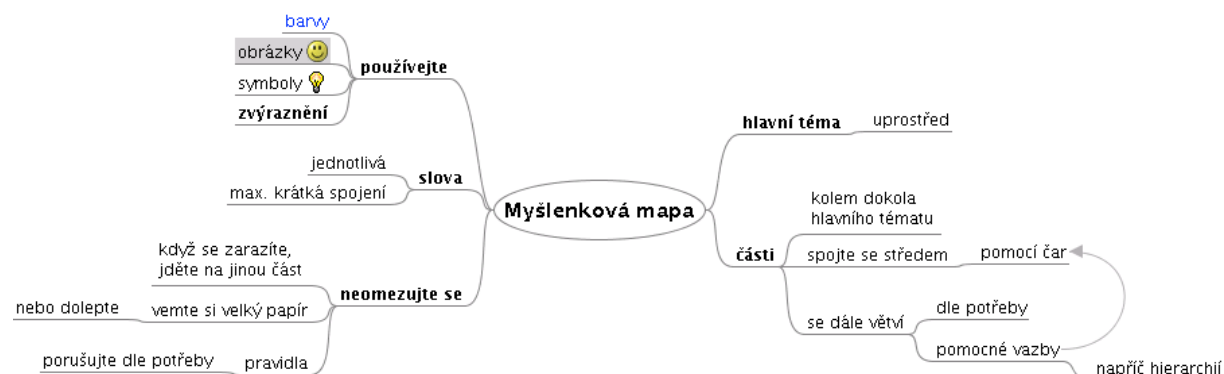
Myšlenkové mapy se hodí pro strukturované psaní poznámek, jako nástroj při analýze a řešení problémů nebo pomůcka při studiu (při tvorbě mapy pochopíte strukturu dané problematiky a podle mapy si lze studované téma také zopakovat). Pokud se stejně jako já pohybujete v IT, možná vás zaujme, že myšlenkové mapy mohou zafungovat jako docela dobrá forma odlehčené specifikace ("light-weight specification"). Myšlenkové mapy jsou informačně hutnější než běžný text, přesto se s nimi v mnoha případech pracuje lépe než s textem.

### Jak na to

Pojďme se podívat na to, jak se myšlenkové mapy tvoří. Pravidla jsou jednoduchá:

- hlavní téma (název) umístíme do středu;
- jednotlivé části pak rozmísťujeme kolem středu a spojíme je se středem čarami;
- části se mohou dále větvit podle potřeby;
- pomocnými čarami nebo šípkami můžeme vyjádřit vztahy, které jdou napříč hierarchií;
- v mapě bychom měli používat barvy, obrázky, symboly a různé další způsoby zvýraznění (míru určitě odhadnete sami, cílem samozřejmě není mapu znepřehlednit);
- je účelné používat buď jednotlivá slova nebo krátká slovní spojení - prostě se soustřeďte na strukturu problému a souvislosti jednotlivých částí, neplývejte silami na formulování celých vět;
- během práce na mapě se neomezujte - vezměte si velký papír nebo si přilepte další list (cílem je omezit počet faktorů, které brání soustředění a "rozmachu" kreativity).

Tato pravidla můžeme zobrazit přímo v myšlenkové mapě takto:



12

## 1.2.8 Princip a přínosy digitalizace reálných objektů, virtualizace reálných objektů a míst

### Virtualizace reálných objektů

Současná technologie je připravena podpořit reálné aplikace **virtuální reality** v mnoha oborech lidské činnosti. Postupný rozvoj této oblasti odkrývá nové problémy, které vyžadují základní výzkum. Specificky se jedná o problematiku jednotlivce a virtuální reality (užitelnost virtuální reality, vhodné postupy, rychlost a kvalita), problematiku skupiny uživatelů ve virtuálním prostředí (formy virtuální spolupráce, vzdálená spolupráce), práce s rozsáhlými datovými soubory (virtualizace reálného světa,

<sup>12</sup> [http://toncar.cz/Clanky/myslenkove\\_mapy1.html](http://toncar.cz/Clanky/myslenkove_mapy1.html)

úpravy modelů) a kombinace virtuálního a reálného světa obecně (projekce, video, specifické vstupní a výstupní technologie).

Při pořizování modelů objektů, které chceme přenést z reálného světa do VR, se pracuje s velkými objemy původně nestructurovaných dat, které je nutné uspořádat do formy vhodné pro další zpracování. V této oblasti je nutné řešit výkonné algoritmy tvorby povrchových a bodových modelů a jejich úpravy do formy, která umožní řešit základní úlohy modelování, zobrazování, kolize atd. v reálném čase. Další oblastí, která má bezprostřední vazbu na virtuální realitu, je vizualizace. Vizualizace se dotýká klíčových otázek sémantiky dat, interpretace jejich obsahu, případně odhalování nových poznatků ve vizualizovaných datech. Spojení vizualizace s VR nabízí možnosti interaktivního zkoumání prezentované informace a získávání nových poznatků v konkrétní aplikační oblasti.

Klíčové metody a algoritmy používané v počítačové grafice a ve virtuální realitě jsou postupně přenášeny do specializovaného HW a ovlivňují tak návrh nových architektur. Souběžně se rozvíjí nová kategorie algoritmů s vysokým stupněm paralelizace.<sup>13</sup>

---

<sup>13</sup> <http://www.fi.muni.cz/research/graphics/vr.xhtml>