

Jela Steinerová  
Mirka Grešková  
Jana Ilavská

**Vyhľadávanie  
informácií  
a organizácia  
poznania  
v elektronickom  
prostredí**

Filozofická fakulta Univerzity  
Komenského v Bratislave  
Katedra knižničnej a informačnej vedy

Jela Steinerová  
Mirka Grešková  
Jana Ilavská

# Vyhľadávanie informácií a organizácia poznania v elektronickom prostredí

Učebný text bol vypracovaný  
v rámci riešenia úloh KEGA  
3/7275/09 Informačné štúdiá  
v podmienkach web 2.0 a nových  
technológií (INWENT) a VEGA  
1/0429/10 Akademické informačné  
prostredie – modelovanie  
z hľadiska informačnej ekológie

Technická a odborná redakcia:  
Helena Ondriašová,  
Irena Lányiová

ISBN 978-80-89236-80-0  
EAN 9788089236800

**Elektronický učebný text**  
**Verzia 1.0**  
**Bratislava, Stimul 2010**

## Abstrakt

Text sumarizuje teoretické a praktické východiská vyhľadávania a organizácie informácií v elektronickom prostredí. Cieľom je predstaviť konceptuálne východiská, praktické prístupy a príklady konkrétnych nástrojov vyhľadávania a organizácie digitálnych informácií.

Objasňuje sa vzťah medzi základnými kategóriami informačnej vedy – informačným správaním, vyhľadávaním informácií a informačným prieskumom. Sú vysvetlené rôzne dimenzie vyhľadávania informácií. Prístupy k vyhľadávaniu sú charakterizované v časti venujúcej sa informačnému horizontu, informačným a prieskumovým stratégiám, ale aj konkrétnym odporúčaniam na postupy efektívneho vyhľadávania. Táto časť premostuje konceptuálne východiská s praktickými prejavmi vyhľadávania pomocou dostupných vyhľadávacích systémov v elektronickom prostredí. Po predstavení princípov a prístupov vyhľadávania informácií sú konkretizované typy nástrojov na vyhľadávanie informácií.

Elektronické informačné zdroje sprostredkované databázovými centrami sa vyznačujú vysokou kvalitou. V samostatnej časti sú rozobrané nielen konkrétne databázové zdroje, ale aj vybrané knižničné katalógy. Okrem overených a selektovaných zdrojov majú dnes používatelia k dispozícii rôznorodé vyhľadávacie nástroje. Sú charakterizované najmä aktuálne trendy v oblasti vyhľadávateľov v kontexte webu 2.0, sociálneho vyhľadávania, vizualizácie a umelej inteligencie.

Fundamentálny pojem informačnej vedy, relevancia, je rozobraná z používateľského a systémového pohľadu. Relevancia je ďalej charakterizovaná z vývojového a konceptuálneho hľadiska. Nový model relevancie založený na interaktívnosti, nelineárnosti, spolupráci a vizualizácii poukazuje na rozdiely pri posudzovaní relevancie v tradičnom a elektronickom prostredí.

Informačné správanie tvorí jadro súčasného výskumu informačnej vedy. Sú opísané viaceré dimenzie informačného správania ľudí v elektronickom prostredí. Taktiež sú priblížené kognitívne a afektívne aspekty informačného správania a jeho prejavy v kontexte sociálnych sietí.

Časť venovaná kategorizácii a klasifikácii otvára problematiku organizácie poznania. Je priblížená aj téma reprezentácie znalostí s dôrazom kladeným na kognitívne a kultúrne východiská reprezentácie znalostí.

Organizácia poznania sprostredkúva rôzne prístupy a nástroje na usporiadanie informácií a dát. Okrem tradičných systémov organizácie poznania sú predstavené najmä tie, ktoré sú aktuálne v elektronickom prostredí. Patria k nim sémantické siete, ontológie, tematické mapy, konceptuálne a myšlienkové mapy, folksonómie.

Text uzatvára časť prepájajúca teoretické aspekty digitálnych knižníc z pohľadu ich štruktúry a ekológie s príkladmi konkrétnych aplikácií. Tiež sú priblížené aktuálne trendy digitálnych knižníc a informačného prieskumu.

Publikácia syntetizuje kľúčové prístupy informačného prieskumu, vyhľadávania a organizácie informácií a informačného správania. Prináša opis konceptuálnych východísk vybraných oblastí informačnej vedy prepája ich s príkladmi nástrojov a postupov vyhľadávania a organizácie informácií.

# Obsah

<b>Úvod</b>	<b>6</b>
<b>1 Vyhľadávanie informácií a informačný prieskum</b>	<b>8</b>
1.1 Vzťahy pojmov informačné správanie, vyhľadávanie informácií, informačný prieskum	8
1.2 Základné zložky vo vyhľadávaní informácií	10
1.3 Princípy a procesy vyhľadávania	12
1.4 Kognitívizmus a doménová analýza	13
1.5 Zmeny vo vyhľadávaní pod vplyvom vývoja informačných technológií	15
1.6 Interaktívne modely informačného prieskumu a kolaboratívny prieskum	17
1.7 Od vyhľadávania k integrovaným systémom	21
1.8 Zhrnutie	21
Literatúra	22
<b>2 Informačný horizont a prieskumové stratégie</b>	<b>24</b>
2.1 Informačný horizont	24
2.2 Informačné stratégie	27
2.3 Prieskumové stratégie	28
2.4 Druhy prieskumových stratégií	29
2.4.1 Systémový prístup k informačným stratégiám	29
2.4.1.1 Stratégie Booleovho modelu/Booleovské indexovanie	29
2.4.1.2 Stratégie vektorového priestorového modelu	31
2.4.1.3 Stratégie pravdepodobnostného modelu/pravdepodobnostný prieskum	32
2.4.1.4 Stratégie zhukového modelu	34
2.4.1.5 Tezaurus	35
2.4.2 Používateľský prístup k prieskumovým stratégiám	37
2.4.2.1 Postup pri vyhľadávaní	37
2.4.2.2 Prax vyhľadávania	38
2.4.2.3 Zameranie vyhľadávania a predchádzajúce poznatky	39
2.5 Ako efektívne vyhľadávať informácie – princípy vyhľadávania	41
2.6 Zhrnutie	42
Literatúra	42
Zoznam webových sídel	43
<b>3 Vyhľadávanie v elektronických informačných zdrojoch</b>	<b>44</b>
3.1 Informačné zdroje Web of Knowledge a SCOPUS	45
3.1.1 Spôsoby vyhľadávania	47
3.1.2 Vyhľadávanie ohlasov a citácií	53
3.1.3 Práca s výsledkami vyhľadávania	58
3.2 Knižničné katalógy	61
3.2.1 WorldCat	61
3.2.2 Projekt KIS3G	62
3.2.3 Katalógy Univerzity Komenského v Bratislave	62
3.3 Zhrnutie	64
Literatúra	65
Zoznam webových sídel	65
<b>4 Vyhľadávanie v internete</b>	<b>67</b>
4.1 Modely webových systémov	67
4.2 Web 2.0	69
4.3 Knižnica 2.0	72
4.4 Webové vyhľadávacie nástroje	74
4.4.1 Východiská a typológia tradičných vyhľadávačov	74
4.4.2 Sociálne vyhľadávanie	75
4.4.2.1 Sociálne siete	75
4.4.2.2 Nástroje s podporou sociálneho vyhľadávania	76
4.5 Prístupy umelej inteligencie	80
4.5.1 Personalizácia	80
4.5.2 Odporúčanie informácií	81
4.5.3 Inteligentní informační agenti	82
4.5.3.1 Spracovanie prirodzeného jazyka	82
4.6 Vizúálne vyhľadávanie	82
4.6.1 Klasifikácia	82
4.6.2 Organizácia výsledkov	83
4.6.3 Doplnkové funkcie	84
4.7 Zhrnutie	85
Literatúra	86
Zoznam webových sídel	87
<b>5 Relevancia</b>	<b>90</b>
5.1 Všeobecné vlastnosti relevancie	90
5.2 Systémový pohľad na relevanciu	91
5.3 Používateľský pohľad na relevanciu	93



5.4 Základné koncepcie a typy relevancie v informačnej vede.....	94
5.5 Vývoj názorov na relevanciu v informačnej vede.....	96
5.6 Relevancia v sieťovom elektronickom prostredí.....	96
5.7 Využitie výsledkov výskumov relevancie v praxi.....	99
5.8 Ako posudzovať relevanciu odborných/vedeckých informácií.....	100
5.9 Zhrnutie.....	101
Literatúra.....	102
<b>6 Informačné správanie v elektronickom prostredí</b> .....	<b>105</b>
6.1 Problémy informačného správanie mladých ľudí v elektronickom prostredí.....	106
6.2 Typy vyhľadávania informácií v elektronickom prostredí.....	107
6.3 Štádiá vyhľadávania informácií.....	107
6.4 Informačné štýly.....	108
6.5 Informačné správanie a sociálne siete.....	108
6.6 Emócie pri informačnom správaní.....	109
6.7 Zhrnutie.....	111
Literatúra.....	112
Zoznam webových sídel.....	113
<b>7 Kategorizácia a klasifikácia ako základ vyhľadávania informácií</b> .....	<b>114</b>
7.1 Kategorizácia a klasifikácia ako myšlienkové procesy.....	114
7.2 Systémy klasifikácie a kategorizácie.....	115
7.3 Vplyv informačných štruktúr na vyhľadávanie informácií.....	117
7.4 Sémantické vzťahy.....	119
7.5 Reprezentácie poznania.....	120
7.6 Kultúrne odlišnosti v reprezentáciách.....	122
7.7 Zhrnutie.....	123
Literatúra.....	123
<b>8 Organizácia poznania v elektronickom prostredí</b> .....	<b>125</b>
8.1 Pohľad do histórie.....	125
8.2 Od tradícií k metadáta.....	126
8.3 Zoznamy termínov a klasifikačné systémy.....	127
8.4 Delenie nástrojov podľa zložitosti štruktúr.....	128
8.5 Organizácia poznania v praxi.....	131
8.6 Organizujeme sami.....	132
8.7 Hľadanie významu/sémantický web.....	133
8.8 Sémantické siete.....	134
8.8.1 Vznik sémantických sietí.....	134
8.8.2 Gigantická sémantická sieť.....	135
8.8.3 Strokový preklad.....	136
8.8.4 Sémantické vyhľadávanie informácií.....	137
8.9 Ontológie.....	137
8.10 Tematické mapy.....	141
8.10.1 Príklady aplikácie tematických máp.....	144
8.10.2 Výhody využitia tematických máp.....	145
8.11 Konceptuálne a myšlienkové mapy.....	146
8.11.1 Príklady online nástrojov na tvorbu máp.....	147
8.12 Folksonómie.....	150
8.12.1 Podpora folksonómii v aplikáciách založených na webe.....	150
8.13 Zhrnutie.....	154
Literatúra.....	155
Zoznam webových sídel.....	156
<b>9 Digitálne knižnice</b> .....	<b>159</b>
9.1 Stručná história.....	159
9.2 Definovanie pojmu.....	161
9.3 Vlastnosti a problémy digitálnych knižníc.....	162
9.4 Štruktúra digitálnej knižnice podľa Manifestu digitálnych knižníc.....	164
9.4.1 Univerzum digitálnych knižníc – základné pojmy.....	164
9.4.2 „Bohatá“ funkčnosť – rôzne prezentácie, základní aktéri.....	165
9.4.3 Referenčný rámec digitálnych knižníc.....	166
9.5 Ekologické aspekty digitálnych knižníc.....	167
9.6 Zložky informačnej ekológie digitálnych knižníc.....	169
9.7 Trendy vývoja digitálnych knižníc a informačného prieskumu.....	169
9.8 Zhrnutie.....	173
Literatúra.....	174
Zoznam webových sídel.....	176
<b>Zoznam ilustrácií</b> .....	<b>178</b>

# Úvod

Vyhľadávanie informácií a informačný prieskum sú jadrom odboru knižničná a informačná veda. Na praktickej úrovni sú spojené nielen s profesionálnym vyhľadávaním informácií (rešeršovaním), ale aj s každodenným informačným správaním človeka. Na teoretickej úrovni sa opierajú o dlhú tradíciu výskumov, modelov a experimentov. V praxi ide o rýchlo sa meniacu a dynamickú oblasť, ktorú ukotvujú najmä intelektuálne základy vyhľadávania a využívania informácií.

Tento učebný text vznikol ako reakcia na zmeny v princípoch vyhľadávania spôsobené najnovším vývojom elektronického prostredia. Je tiež výsledkom vzdelávania informačných profesionálov a prednášok v rámci predmetov Vyhľadávanie informácií a Informačný prieskum na KKIV FiFUK. Jeho cieľom je poskytnúť úvod do štúdia vyhľadávania informácií, informačného správania, organizácie informácií aj praktického rešeršovania. Tematicky pokrýva všetky významné témy, ktoré sú súčasťou európskych študijných programov v oblasti vyhľadávania informácií, informačného prieskumu, informačného správania a organizácie informácií. Štruktúra každej kapitoly obsahuje okrem východísk a príkladov aj záverečné zhrnutie základných téz, literatúru a zoznam webových sídiel. Tri autorky prepojili v texte tri uhly pohľadu na problematiku: princípy a modely, informačnú prax a nové nástroje spracovania informácií.

V učebnom texte sa tak prepájajú teoretické aj praktické aspekty vyhľadávania a organizovania informácií v elektronickom prostredí. Nanovo interpretujeme pojmy vyhľadávanie informácií, informačný prieskum a informačné správanie. Vysvetľujeme problematiku informačného horizontu a informačných a prieskumových stratégií. Vo svetle histórie a analýzy základných modelov vyhľadávania informácií formulujeme princípy efektívneho vyhľadávania a posudzovania relevancie informácií.

Osobitne sa zameriavame na vyhľadávanie informácií v elektronických zdrojoch a v internete. Pritom okrem základných princípov prinášame mnoho príkladov významnejších systémov, katalógov a služieb, ktoré reprezentujú súčasné trendy vo vyhľadávaní a organizovaní informácií v elektronickom prostredí, najmä v internete. Stručne vysvetľujeme aj princípy sociálneho a vizuálneho vyhľadávania a uplatnenie umelej inteligencie vo vyhľadávaní informácií.

Na základe vlastných výskumov sme sformulovali kapitoly o informačnom správaní v elektronickom prostredí a relevancii z používateľského hľadiska. Vyhľadávanie informácií je vždy spojené s ich organizáciou. Preto sú v texte podrobnejšie spracované aj kapitoly o kategorizácii a klasifikácii ako predpokladoch práce s informáciami a o organizácii informácií v elektronickom prostredí.

Organizácia poznania v elektronickom prostredí sa v poslednom období rozvíja mimoriadne intenzívne na profesionálnej aj každodennej úrovni. Vzniká množstvo nových nástrojov na kategorizovanie informácií. Okrem prenosu tradičných nástrojov na organizáciu informácií (ako napríklad

klasifikačné systémy a tezaury) do elektronického prostredia upozorňujeme na účinnosť nových nástrojov na príkladoch sémantických sietí, tematických máp, ontológií či pojmových máp. Na záver sumarizujeme poznatky o digitálnych knižniciach ako o zložitých sociotechnických systémoch a významnom trende vyhľadávania a práce s informáciami v rôznych kontextoch.

V interpretácii pojmov a koncepcií zdôrazňujeme hľadisko informačnej vedy ako vedy o komunikovaní, spracovaní a využívaní informácií v spoločenských kontextoch. V modernom prístupe tu prepájame tradície informačného prieskumu orientovaného na informačné systémy s tradíciami poznania správania človeka pri využívaní informácií. Tento integračný prístup považujeme za základný trend rozvoja informačnej teórie aj praxe, najmä z hľadiska vplyvov a možností elektronického informačného prostredia.

Učebný text môžu využiť nielen študenti odboru informačné štúdiá, ale aj profesionáli z praxe a záujemcovia o ďalšie vzdelávanie v oblasti vyhľadávania a organizácie informácií. Veríme, že texty zaujmú aj študentov, učiteľov a profesionálov z príbuzných informačných odborov (informatika, informačný manažment, vzdelávanie, komunikačné disciplíny a pod.). Odporúčania na praktické vyhľadávanie a organizáciu informácií môžu pomôcť aj študentom iných oborov pri rozvíjaní informačnej gramotnosti a chápaní intelektuálnych princípov práce s informáciami. Poznanie týchto princípov môže napokon podporiť personálny informačný manažment profesionálov v rôznych odboroch ľudskej informačnej činnosti a v každodennej praxi.

Jela Steinerová

# 1

## Vyhľadávanie informácií a informačný prieskum

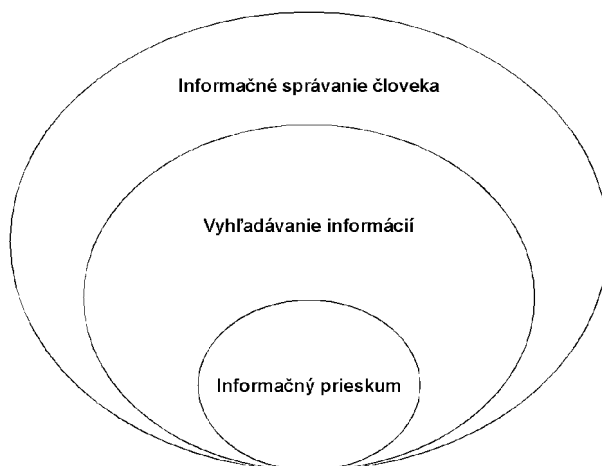
### 1.1

#### **Vzťahy pojmov informačné správanie, vyhľadávanie informácií, informačný prieskum**

Vzťahy pojmov informačné správanie, vyhľadávanie informácií a informačný prieskum v súčasnosti charakterizujeme ako navzájom prepojené a v sebe obsiahnuté. Vytvárajú hierarchiu od najvšeobecnejšieho k špecifickému. V terminológii informačnej vedy nie sú tieto pojmy jednoznačne rozlíšené a interpretujú sa po novom pod vplyvom výsledkov nových výskumov informačného správania človeka a digitálnych knižníc. Niektorí autori dokonca hovoria o „obrate“ v informačnej vede (Ingwersen, Järvelin 2005). Tu sa prepájajú sociálne východiská s východiskami informatiky a integruje vyhľadávanie informácií s interakciou s prieskumovým systémom. Vyhľadávanie informácií má základ v sociálnych vedách a výskumoch používateľov, informačný prieskum je založený skôr na prístupe informatiky, počítačových vied. Aj v oblasti vyhľadávania informácií vidíme tradície vývoja knižničnej a informačnej vedy. Na jednej strane je to orientácia na používateľa a knižnice, klasifikáciu a kategorizáciu (založená Jesse H. Sherom), na druhej strane tradícia informačných systémov a informatiky (založená Gerardom Saltonom).

V prvej línii sa prehľbuje poznanie kognitívnych a sociálnych základov vyhľadávania informácií, v druhej línii sa zdokonaľujú inteligentné technológie na reprezentáciu poznania a interakciu so systémami. Prepája sa kognitívna a fyzikálna paradigma informačného prieskumu (Steinerová, 1996). Rozvíja sa kognitívny informačný prieskum a skúmajú sa rôzne kontexty informačného správania človeka (Steinerová 2005).

Zjednodušené znázornenie súvisiacich pojmov informačné správanie človeka, vyhľadávanie informácií a informačný prieskum si môžeme naznačiť na obrázku 1.1.



Obr. 1.1 Vzťahy pojmov informačné správanie, vyhľadávanie informácií a informačný prieskum (upravené podľa ložiskového modelu T. Wilsona) (Steinerová 2005)

*Informačné správanie* človeka je viacúrovňová integrovaná ľudská aktivita vyplývajúca z adaptácie človeka na informačné prostredie. Obsahuje aktívne zložky (cieľové hľadanie, formulovanie dotazov) aj pasívne zložky (napríklad sledovanie médií). Jeho základným cieľom je hľadanie zmyslu v rôznych komunikačných aktoch a interakciách. Skladá sa z neurofyziologickej, senzomotorickej, kognitívnej, emocionálnej a sociálnej zložky. Teda obsahuje pocity, myšlienky a činy ľudí, ktorí spracúvajú informácie. Informačné správanie možno deliť podľa rôzneho účelu spracovania informácií. Rozlišuje sa informačné správanie pri vyhľadávaní informácií, informačné správanie pri organizovaní informácií, informačné správanie pri využívaní informácií, informačné správanie pri tvorbe informácií, informačné správanie pri posudzovaní relevancie informácií (relevantné správanie) a pod. Podrobnejšie spracovanie problematiky informačného správania je v publikácii Steinerová 2005.

*Vyhľadávanie informácií* je súčasťou aktívneho informačného správania človeka. V informačnej vede sa mu venuje najvyššia pozornosť, je to základ praktickej informačnej činnosti aj modelov v teórii. Vyhľadávanie informácií je proces, ktorý je založený na hľadaní informácií prostredníctvom informačných zdrojov a na interakcii človeka s informačným prostredím. Najvýraznejšie sa v súčasnosti vyhľadávanie prejavuje pri využívaní internetu a v laickom vyhľadávaní prostredníctvom webových vyhľadávačov sa rozšírilo medzi široké vrstvy obyvateľstva. Podrobnejšie je vyhľadávanie rozpracované napr. v publikácii Makulová 2002.

*Informačný prieskum* má v informačnej vede dlhú tradíciu. Objavuje sa v informačnej vede už od 40. rokov 20. stor. Informačný prieskum možno charakterizovať v širšom a užšom zmysle slova. V širšom zmysle je to komplex procesov spojených s hľadaním informácií z informačných zdrojov zodpovedajúcich informačnej požiadavke a informačnej potrebe. Obsahuje metódy a techniky vyhľadávania informácií zodpovedajúcich formulovanej požiadavke používateľa prostredníctvom výpočtovej techniky a s využitím informačných jazykov reprezentujúcich obsah dokumentov. Informačný prieskum sa vyvinul ako osobitná disciplína informačnej vedy. V najužšom zmysle je to interakcia človeka s informačným prieskumovým systémom. V informačnom prieskume je dôležité reprezentovanie, ukladanie a výber informácií z pamäte. Podrobnejšie sú história, modely, koncepcie informačného prieskumu, prieskumová stratégia a taktika a hodnotenie efektívnosti vyhľadaných informácií spracované v publikácii Steinerová 1996.

Informačný prieskum ako termín prvýkrát použil Calvin Mooers v roku 1951. Vannevar Bush zas už v roku 1945 predvídal osobný informačný systém Memex, ktorý by fungoval na princípe automatizácie a umožnil efektívne vyhľadávanie a využívanie informácií. Najväčší rozmach tejto disciplíny nastal práve v 50. rokoch 20. storočia. V tom čase sa začalo publikovať množstvo vedeckej literatúry, ktoré bolo potrebné organizovať, spracovať a vyhľadávať. V začiatkoch vývoja mohli informačný

prieskum vykonávať informační profesionáli a školení referenční knihovníci, prípadne odborníci na informačné technológie. Neskôr sa vyhľadávanie informácií rozšírilo medzi široké vrstvy laických používateľov a postupne sa z neho stala dôležitá súčasť informačnej gramotnosti.

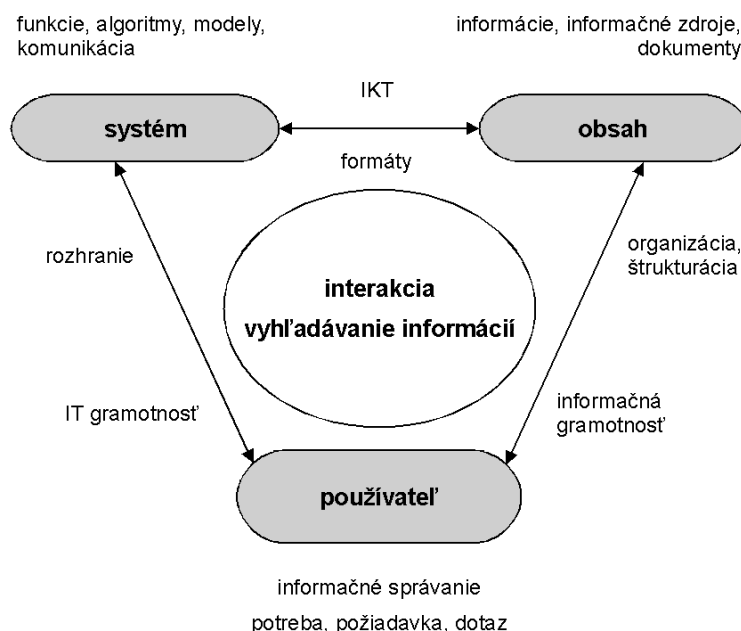
V akademickom zmysle je informačný prieskum *získavanie informácií z neštruktúrovaných dokumentov (najmä textov) z veľkých súborov uchovávaných v počítačových systémoch s cieľom odpovedať na informačnú požiadavku a uspokojiť informačnú potrebu*. Neštruktúrované dokumenty (dáta) sú také, ktoré nemajú pre počítačový systém vopred danú jasnú štruktúru na spracovanie významu. Protikladom sú štruktúrované údaje v relačných databázach, v ktorých je v tabuľkovej podobe presne vyjadrený zoznam objektov s jednoznačne priradenými vlastnosťami či hodnotami. Prirodzene, každý text je do určitej miery štruktúrovaný. Preto v súčasnosti vznikajú aj značkovacie jazyky na reprezentáciu informácií v textoch, napr. HTML a XML. Od tejto štruktúry závisí aj informačný prieskum a vývoj sémantického webu.

## 1.2

### Základné zložky vo vyhľadávaní informácií

Informačný prieskum sa začal uplatňovať najmä v oblasti knižničných dokumentov so zameraním na vedecké informácie. Postupne sa stal súčasťou všetkých informačných profesií, napríklad žurnalistiky, medicíny, práva. V súčasnosti sa práca s informáciami na profesionálnej aj na laickej úrovni prenáša najmä do sieťového elektronického informačného prostredia (internet a World Wide Web). Dôraz sa kladie najmä na efektívne nástroje organizácie poznania v elektronickom prostredí a budovanie digitálnych knižníc ako komplexných systémov poskytovania služieb a produktov v elektronickom prostredí.

Vyhľadávanie informácií, informačné správanie a informačný prieskum prepájajú systémy, obsah a ľudí. Tieto tri zložky sú potrebné na to, aby bolo možné realizovať spracovanie a využívanie informácií. Ich vzťahy sú znázornené na obrázku 1.2.



Obr. 1.2 Situácia základných zložiek pri vyhľadávaní informácií/informačnom prieskume

Obrázok znázorňuje situáciu, v ktorej vstupujú do vzájomného pôsobenia (interakcie) človek, informačné zdroje a informačné systémy. Človek je aktívny aktér, ktorý sa vyznačuje informačným správaním. V tejto situácii má určitú informačnú potrebu, ktorú vyjadruje v informačnej požiadavke. Pri preklade do jazyka systému vzniká dotaz. Človek – používateľ na prístup k informačným zdrojom/ obsahu využíva informačný prieskumový systém. Tento systém obsahuje funkcie a vnútorné vyhľadávacie modely a algoritmy, ktoré ovplyvňujú možnosti vyhľadávať informácie. Systém je založený na informačných a komunikačných technológiách (počítače a komunikačné siete). Systém komunikuje s človekom prostredníctvom špeciálnej zložky – tzv. rozhrania, ktoré umožňuje dorozumenie medzi prirodzeným jazykom človeka a jazykom systému. Na takéto porozumenie je potrebný určitý stupeň technologickej gramotnosti (ovládanie technológií). Na strane obsahu sú informačné zdroje a dokumenty, ktoré sú určitým spôsobom organizované buď v tradičných knižniciach a informačných jazykoch alebo špeciálne v elektronickom prostredí. Na efektívne využívanie informácií z informačných zdrojov prostredníctvom procesov vyhľadávania a prieskumu musí človek disponovať informačnou gramotnosťou.

Východiskové situácie vyhľadávania informácií a informačného prieskumu opísali vo svojich modeloch napríklad N. Belkin (model ASK), F. Lancaster, C. J. van Rijsbergen, P. Ingwersen, T. Saracevic, M. Batesová, C. Kuhlthauová, D. Ellis, B. Hjørland, K. Järvelin, P. Vakkari, T. Wilson a i. (Steinerová 1996, Steinerová 2005).

Zo základnej situácie vyhľadávania informácií a informačného prieskumu vyplýva, že tu existujú ďalšie dôležité pojmy, ktoré je potrebné vysvetliť. 1. Na strane človeka sú to informačné potreby, informačné požiadavky, dotazy a relevancia. 2. Na strane systému je to informačný prieskumový systém, rozhranie, informačný priestor, interakcia. 3. Na strane obsahu je to informácia, poznatok, organizácia poznania.

1. Informačná potreba je identifikovaná medzera v poznatkoch, ktorá býva súčasťou širších potrieb človeka (biologické, kognitívne, sociálne). Môže viesť k informačnému správaniu a vyhľadávaniu informácií až po formulovanie požiadavky.

Informačná požiadavka je formulovaná informačná potreba. Vzniká ako stav pocítovanej zamerania v určitom čase, ktorý používateľ predkladá systému alebo zdroju informácií.

Dotaz je transformácia formulovanej požiadavky do jazyka systému. Môže ho formulovať buď používateľ alebo sprostredkovateľ (knihovník, rešeršér). Cieľom je prehľadať informačný priestor prostredníctvom algoritmov systému. Relevancia je vzťah medzi informačnou požiadavkou a vyhľadanými informácií. Ide o hodnotenie informácií z hľadiska témy, užitočnosti a vhodnosti, ktoré vykonáva používateľ v kontexte svojej informačnej potreby a situácie v určitom čase. Môže sa meniť v čase (je dynamická) a v rôznych mierach subjektívna (stav poznania) aj objektivizovaná (systémom reprezentácie poznania).

2. Informačný prieskumový systém je informačný systém vytvorený s cieľom vyhľadávať informácie, ktoré sú pre používateľa hodnotné a užitočné (relevantné). Prepája informačné zdroje v technologickom prostredí a funkcie na komunikáciu s používateľom (rozhranie). Informačný priestor obsahuje obsah – informačné zdroje a informačné objekty štruktúrované podľa technologických možností organizácie a reprezentácie v informačnom systéme. Rozhranie je mechanizmus sprostredkovania komunikácie medzi človekom a systémom. Často sa používa termín používateľské rozhranie. Interakcia naznačuje viacstrannú výmenu informácií medzi zložkami vyhľadávania informácií (človek, systém, obsah). Interaktívny informačný prieskum je vzájomné pôsobenie medzi hlavnými zložkami pri vyhľadávaní informácií ako používateľ, rešeršér, socio-organizačný kontext, prostredie informačného systému, obsah, rozhranie. Interaktívnosť sa všeobecne chápe ako schopnosť systému komunikovať s používateľom prostredníctvom dialógu pri vyhľadávaní informácií.

3. Informáciu môžeme zjednodušene vysvetliť ako komunikovaný a zaznamenaný poznatok, ktorý mení stav poznania príjemcu. Poznatok je výsledkom poznávacej činnosti človeka, individuálne chápanie objektov založené na skúsenosti, kognícii (myšlienkovvej činnosti) aj emóciách a intuícii. Organizácia poznania je usporiadanie a reprezentácia informácií a poznatkov s cieľom ich efektívneho vyhľadania. Je to aj disciplína zaoberajúca sa metódami štruktúrácie poznania vrátane jeho kategorizácie a klasifikácie, transformácie, reprezentácie a prezentácie (vizualizácie). Reprezentácia poznania je potom spôsob vyjadrenia poznatkov a vzťahov medzi nimi v pamäti prieskumového systému.

Všetky pojmy spája nová paradigma (nové zameranie výskumov) informačnej vedy. Podstatnou črtou novej paradigmy informačnej vedy je postupné zaostrovanie ohniska pozornosti na zložité výsledky poznávacej činnosti človeka – na *poznatky*. Mnohí odborníci už dávnejšie vysvetľujú podstatu komunikovania poznatkov prostredníctvom vzájomne prepojených a premenlivých foriem poznania – od dát cez informácie k poznatkom až po akcie, rozhodnutia a múdrosť. Zjednodušene povedané – využitím informácií v kontextoch vznikajú poznatky. Úloha kontextu pri informačnom prieskume a vyhľadávaní informácií sa zdôrazňuje v posledných rokoch ako kľúčová (Ingwersen, Järvelin 2005). Organizujú sa aj konferencie, ktoré majú kontext vo svojom názve (napr. IRiX, Information Retrieval in Context, alebo ISIC, Information Seeking in Context). Kontext možno vysvetliť ako úlohu, riešený problém či projekt, každodennú situáciu. Podľa Ingwersena a Järvelina (2005) kontext vytvárajú jednotlivé zložky situácie vyhľadávania a slúžia pre seba ako kontexty vo vzájomnej interakcii. Rozlišuje sa sociálny, organizačný, kultúrny a systémový kontext. Kontext sa pritom dynamicky vyvíja.

Do hry vstupujú aj príbuzné disciplíny, ktoré skúmajú premeny poznania. Informačná veda sa tak zblíži s manažmentom poznatkov ako disciplínou vynárajúcou sa z praxe využívania informácií v organizáciách. Súvisiacimi disciplínami sú aj kognitívne vedy, znalostné inžinierstvo či inteligentné prieskumové systémy.

### 1.3 Princípy a procesy vyhľadávania

Princípy vyhľadávania informácií závisia od toho, ako sa pristupuje k zdrojom a informáciám. Tento prístup možno rozdeliť na fyzický a intelektuálny. Fyzický prístup je zameraný na určenie miesta zdrojov a informácií: Intelektuálny prístup je spojený s interpretáciou informácií a poznatkov zo zdrojov. Intelektuálny prístup reprezentuje kognitívna paradigma informačnej vedy. Sústreďuje sa na myšlienkové procesy pri spracovaní a využívaní informácií, používateľa a obsah zdrojov a informácií.

Knižničné a informačné služby by mali podporovať práve intelektuálny prístup k informáciám. Pri organizovaní knižníc sa to nezdôrazňuje dostatočne. Aj vyhľadávanie informácií v knižnici alebo internetovom vyhľadávači je vlastne proces hľadania významu a zložitej intelektuálnej konštrukcie. V teórii aj vlastnom empirickom výskume to dokázala napríklad aj C. C. Kuhlthauová, významná autorka z oblasti vyhľadávania informácií a informačných služieb (Kuhlthau 1993). Uplatňuje psychologickú teóriu osobnej konštrukcie Georgea Kellyho. Predpokladá, že vyhľadávanie informácií je proces konštrukcie, ktorá sa začína neurčitou a úzkou. Dokazuje to aj prostredníctvom empirického prieskumu procesu vyhľadávania a využívania informácií študentmi. Konštrukcie sú myšlienkové schémy, predstavy a modely.

Informačný prieskumový proces sa prejavuje v troch oblastiach, emocionálnej (pocity), kognitívnej (myšlienky) a fyzickej (aktivity). Informácia nie je vec, hotový produkt, správna odpoveď, ale skôr proces učenia a zmeny konštrukcií. Vyhľadávanie informácií je potom hľadanie zmyslu, v ktorom si človek formuje osobné hľadisko (na problém). V tomto zmysle je aj relevancia informácií konštrukciou a dynamickým procesom.



Kuhlthauová vytvorila dôležitý model informačného prieskumového procesu (Kuhlthau 1993), ktorý sa skladá zo šiestich štádií. Tento model je znázornený na obrázku 1.3.

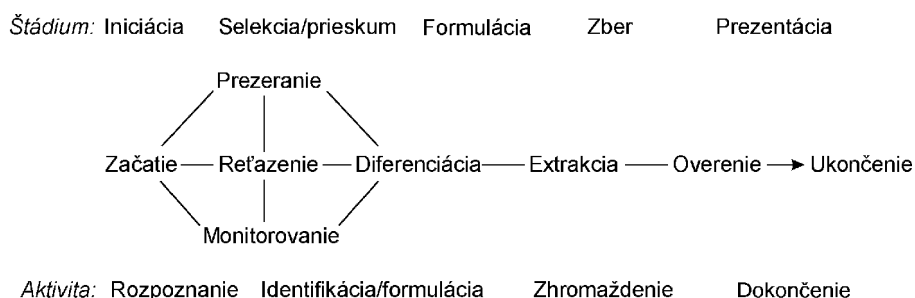
Štádiá	Iniciácia	Výber	Objavovanie	Formulácia	Zhromažďovanie	Prezentácia
Pocity	Neistota	Optimizmus	Zmätok, frustrácia, pochybnosti	Jasnosť	Pocit kontroly, sebaistota	Úľava, spokojnosť alebo sklamanie
Myšlienky	Nejasné			Jasnejšie	Zvýšený záujem	Sústredené
Činnosti	Vyhľadávanie východiskových informácií		Vyhľadávanie relevantných informácií		Vyhľadávanie pertinentných informácií	
Vhodné úlohy	Rozpoznanie	Identifikácia, výskum	Identifikácia, výskum	Formulácia	Zber	Dokončenie

Obr. 1.3 Model informačného prieskumového procesu podľa Kuhlthau (Ingwersen, Järvelin 2005)

Model naznačuje, že ľudia hľadajú informácie rôzne v rôznych štádiách procesu vyhľadávania. V každom štádiu sú určené pocity, myšlienky, aktivity a vhodné úlohy. Tieto štádiá sú:

1. začatie (iniciácia), znamená rozpoznanie informačnej potreby pri riešení problému
2. selekcia (identifikácia a výber témy)
3. explorácia (hľadanie a štúdium informácií)
4. formulácia ohniska (vyjadrenie a štruktúrovanie problému)
5. zhromaždenie pertinentných (subjektívne vhodných) informácií pre ohraničenú tému
6. prezentácia (ukončenie hľadania, využitie informácií na riešenie úlohy, referovanie o téme)

Rozdelenie procesov a krokov pri vyhľadávaní spracovali aj ďalší autori, najmä F. Lancaster (Steinerová 1996) a David Ellis (Steinerová 2005). Porovnanie štádií Ellisa a Kuhlthauovej je znázornené na obrázku 1.4.



Obr. 1.4 Štádiá vyhľadávania informácií (Ellis – Kuhlthau) (podľa Wilson, 1999)

## 1.4 Kognitivismus a doménová analýza

Približne od začiatku 80. rokov 20. stor. sa v teórii informačného prieskumu presadila kognitívna paradigma (Steinerová 1996). Reprezentujú ju najmä práce P. Ingwersena, N. Belkina, C.C. Kuhlthauovej. Spája ich zdôrazňovanie kognitívnych (psychických) procesov ako primárnych pri vyhľadávaní

informácií. Autori predpokladajú, že vyhľadávanie informácií je individuálny intelektuálny proces spracovania a využívania informácií. Niektorým autorom sa však tento prístup zdal príliš individualistický, zameraný na jednotlivca mimo sociálnych vzťahov alebo domén (riešených problémov a tém). Preto v rámci kognitívnej paradigmy vznikol aj protikladný názor na vysvetľovanie podstaty vyhľadávania informácií. Ide o tzv. doménovú analýzu, ktorú predstavujú autori ako B. Hjørland a A. Albrechtsenová (Hjørland 1997).

Rozdiely medzi kognitivismom a doménovou analýzou sú v tom, čo autori považujú za prioritu pri interpretácii vyhľadávania informácií. Ak kognitivismus kladie dôraz na pochopenie potrieb používateľa ako jednotlivca, doménová analýza vychádza z pochopenia potrieb používateľa v sociálnom kontexte. Kognitivismus sa zaoberá najmä sprostredkovaním medzi používateľom a tvorcom prostredníctvom psychického porozumenia potrebe. Doménová analýza zdôrazňuje, že informačný prieskumový systém treba začať budovať v určitej disciplíne alebo doméne (tému sprostredkovanej terminológiou vedného odboru). Doménová analýza zdôrazňuje kolektívny rozmer poznania v spoločných pojmoch, ktoré sa používajú na sprostredkovanie poznania vo vede, výskume, vzdelávaní alebo organizácii.

Kognitivismus sa zameriava na jednotlivca (používateľa), disciplína sa tu berie do úvahy ako kognitívna štruktúra osobnosti. Doménová analýza zdôrazňuje výskum rôznych disciplín, prípadne ich vzájomné porovnanie (napr. terminológia technických a spoločenských vied).

Kognitivismus sa inšpiruje kognitívnymi vedami a umelou inteligenciou, najmä pri interakcii človeka a systému. Doménová analýza zdôrazňuje teóriu vedy a sociológiu poznania.

Kognitivismus v praxi kladie dôraz na vonkajšiu stránku informačného prieskumového systému, najmä používateľské rozhrania. Doménová analýza sa sústreďuje na predmetovú reprezentáciu poznania, ktorá je vnútornou stránkou prieskumového systému (reprezentovaný obsah).

Obidva prístupy sa navzájom dopĺňujú. Kognitivismus dôkladne skúma individuálne spracovanie informácií, pamäťové procesy a mentálne modely a štruktúry. Doménová analýza sa zaujíma o vedeckú komunikáciu, dokumenty a databázy v určitých disciplínach.

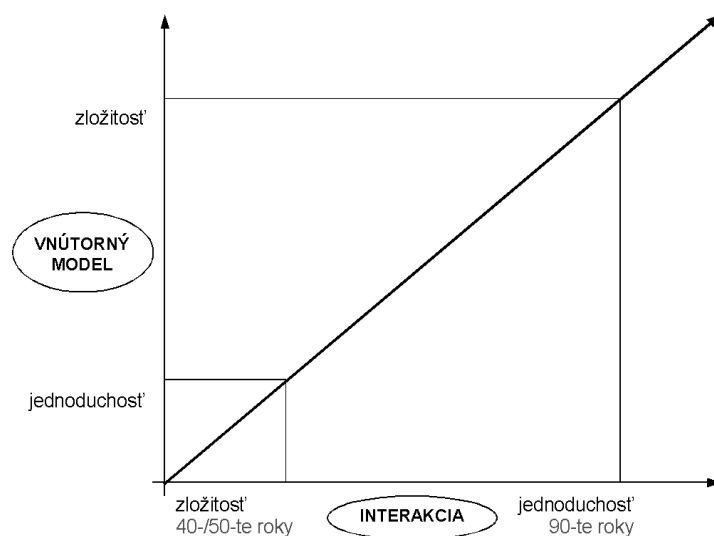
Doménová analýza sa stáva východiskom na definovanie typu tzv. **predmetového vyhľadávania informácií**/predmetového prieskumu (Hjørland 1997). Podľa Hjørlanda je predmetový prieskum vyhľadávanie neznámych dokumentov, ktorých obsah môže prispieť k riešeniu konkrétneho problému alebo uspokojiť informačnú potrebu. Za predmet dokumentu považuje epistemologický potenciál dokumentu, t.j. to, ako vyjadruje poznanie v pojmoch. Aj keď interpretácia týchto pojmov môže byť subjektívna, určitý stupeň objektívnosti sa v predmetovom prieskume dosahuje stavom vedeckého poznania, najmä ustálenou terminológiou. Explicitné informácie pre predmetový prieskum sú napríklad klasifikačné kódy, identifikátory, deskriptory, predmetové heslá. Boli vytvorené priamo na uľahčenie prieskumu. Medzi implicitné možno zaradiť slová z názvov, abstraktov, citácie, slová z textu, názov a renomé vydavateľstva. Neboli vytvorené priamo pre prieskum, ale sú užitočnými náznakmi významu informácií.

V **predmetovom prieskume** sa kladie dôraz na doménu (vedeckú disciplínu, riešený problém). Vo vzťahu jazyka a myslenia sa pritom vytvárajú pojmy na základe ľudskej komunikácie a spolupráce. Dokazuje sa, že tradičný informačný prieskum bol problematický preto, lebo hľadal presnú zhodu, nepracovalo sa s významom. V kognitívnej paradigme a doménovej analýze je dôležitý práve jazyk sprostredkujúci významy a poznanie v pojmoch. Na individuálnej úrovni sa formuluje požiadavka na základe predchádzajúceho stavu poznania a vzdelania. Na kolektívnej úrovni je schopnosť formulovať požiadavku spojená so stavom poznania v disciplíne a doméne, čo súvisí aj s organizáciou vedeckej kooperácie.

## 1.5

### Zmeny vo vyhľadávaní pod vplyvom vývoja informačných technológií

Rozvoj informačných technológií významne ovplyvnil možnosti vyhľadávania informácií. V histórii vývoja si môžeme všimnúť určitú súvislosť medzi zložitosťou a prepracovanosťou vnútorných mechanizmov reprezentácie a vyhľadávania na jednej strane a medzi interakciou so systémom. To naznačujú *vývojové modely informačného prieskumu* (Steinerová 1998). V začiatkoch vývoja boli mechanizmy vyhľadávania pomerne jednoduché a interakcia so systémom pomerne zložitá. S prieskumovými systémami komunikovali najmä školení programátori a rešeršéri, ktorí ovládali jazyk systému. Prieskumové systémy pracovali na veľkých sálových počítačoch, ku ktorým mali prístup „zasvätení“ odborníci. S rozšírením osobných počítačov do kancelárií a všetkých oblastí života spoločnosti sa situácia zmenila. Interakcia sa zjednodušila, začala sa vytvárať a podporovať informačná gramotnosť ako ovládanie počítačových kancelárskych systémov. Najväčší rozmach nastal od začiatku 90. rokov 20. stor. pri rozšírení internetu do domácností a na pracoviská. Ovládanie prieskumových informačných systémov je jednoduché, vyhľadávacie systémy sa čoraz viac približujú sociálnym a myšlienkovým procesom a včleňujú nielen do odborných činností, ale aj do každodenného života človeka. Dnes môžeme konštatovať, že interakcia so systémami je pomerne jednoduchá a nové prieskumové systémy z hľadiska výstavby svojich modelov čoraz prepracovanejšie (sofistikovanejšie) a inteligentnejšie. Tento vývoj je znázornený na obrázku 1.5.



Obr. 1.5 Vývoj vzťahu medzi mechanizmami vyhľadávania a interakciou systému a človeka

Podobný vývoj zaznamenávame aj pri zdokonaľovaní nástrojov *reprezentácie poznania*. Od prírodných hierarchických štruktúr klasifikačných systémov sa postupne cez indexačné systémy, slovníky a plné texty prechádza k inteligentným nástrojom reprezentácie poznania ako sú napríklad asociatívne siete, produkčné pravidlá, rôzne pojmové štruktúry, rámce a pod.

Vývoj *mechanizmov vyhľadávania* možno tiež interpretovať v duchu ich zdokonaľovania v priebehu 50-ročného historického vývoja. V začiatkoch boli algoritmy jednoduché, skôr založené na mechanických prístupoch. Neskôr sa začali spracúvať texty, ktoré si vyžadovali zložitejšie prístupy k novej štruktúre a automatizovanému spracovaniu. Začali sa uplatňovať najmä logicko-množinové a lingvistické poznatky (Booleov model). Zdokonalila sa spätná väzba. Ďalej sa rozpracovali

možnosti práce s neurčitou (pravdepodobnostný, vektorový priestorový model) a s prepájaním príbuzných termínov (zhlukovanie). Veľkou zmenou bolo uplatnenie asociatívnych princípov reprezentácie v hypertextoch. To sa stalo významným princípom fungovania webových stránok a ovplyvnilo aj možnosti stratégie prezerania pri vyhľadávaní. Najnovšie mechanizmy vyhľadávania kombinujú rôzne princípy založené na výskumoch umelej inteligencie, nových nástrojov reprezentácie poznania a sémantickom webe (Steinerová 1998).

Vývoj *náborov na relevanciu* v informačnej vede aj v praktických prieskumových systémoch kopíroval celkový vývoj vyhľadávania informácií. Od funkcie formálnej zhody sprostredkovanej mierami odozvy a presnosti, cez funkciu spätnej väzby (relevančná spätná väzba v systéme SMART) po relevanciu ako logicko-lingvistický vzťah (hierarchia, následnosť, vzor, asociácia, synonymia, inklúzia, prienik, zjednotenie, negácia). V súčasnosti sa aj predstava o relevancii zmenila. Relevancia je kognitívno-sociálna kategória, ktorá je spojená so zmenou stavu poznania človeka a s užitočnosťou pre riešenie problému v kontexte.

Zmeny vo vyhľadávaní v súčasnosti môžeme charakterizovať prostredníctvom niekoľkých zásadných funkcií informačných prieskumových systémov.

1. *Kolaborácia*. Možnosti spolupráce pri formulovaní dotazu zlepšujú výsledky vyhľadávania. Spoločné hodnotenie, komentovanie, vyjadrovanie prostredníctvom pojmov, kategorizovanie a klasifikovanie v elektronickom prostredí pridáva vyhľadávaniu nový sociálny rozmer. Ak sa pôvodne predpokladalo, že vyhľadávanie je najmä individuálny intelektuálny proces, nové systémy dokazujú, že sociálne kontakty môžu byť užitočné aj pri samotnom informačnom prieskume.
2. *Interaktívnosť*. Na revolučnú zmenu pri kontakte s prieskumovým systémom upozornil Ingwersen už v r. 1992 (interaktívny informačný prieskum). Možnosť okamžitej reakcie, upravovanie dotazu, nelineárnosť sa stali prirodzenou súčasťou vyhľadávania informácií. Kľukatú cestu znázornila vo svojom modeli M. Batesová. Interaktivita je v súčasnosti podmienkou efektívneho informačného prieskumu, ako na úrovni komunikácie človek-systém, tak aj na úrovni komunikácie človek-človek v elektronickom prostredí (Steinerová 1996).
3. *Personalizácia a podpora používateľa*. Čím viac podpory pre používateľa systém integruje, tým môže byť prieskum úspešnejší. Filozofiou takýchto systémov je „splnomocniť“ používateľa, t.j. dať mu možnosť samostatne sa pohybovať v informačnom priestore a samostatne riadiť svoje vyhľadávanie a využívanie informácií. Prieskumový systém môže podporiť používateľa pri plánovaní, organizovaní stratégie prieskumu, pri orientácii v pojmoch, pri formulovaní dotazu (oprave chýb, preklade do jazyka systému), ale aj pri organizácii výsledkov a pri posudzovaní relevancie (radenie výsledkov podľa vypočítanej pravdepodobnosti relevancie s dotazom). Na úrovni rozhrania sa systém môže prispôbovať preferenciám foriem, formátov, prípadne vizualizáciou sa približovať náladám používateľa.
4. *Pridávanie kontextu*. Novšie prieskumové systémy umožňujú vytvárať pridanú hodnotu z hľadiska zaznamenania vzniku informačnej potreby. Je možné prepojiť vyhľadávanie s kontextom riešenej úlohy. Systém môže tiež podporiť rôzne informačné štýly, prípadne aj kognitívne a učebné štýly. Na obohatenie informácií kontextom možno využiť aj kolaboratívne funkcie ako editovanie, anotovanie, neformálne poradenstvo, vytváranie komunitnej/kolektívnej pamäte. Napríklad elektronické referenčné služby typu ASK... (spýtajte sa...) podporujú spoločné využívanie predchádzajúcich výsledkov, zaznamenanie a využitie výsledkov už raz formulovaného dotazu, vytváranie báz znalostí z poskytnutých referenčných informácií.

## 1.6

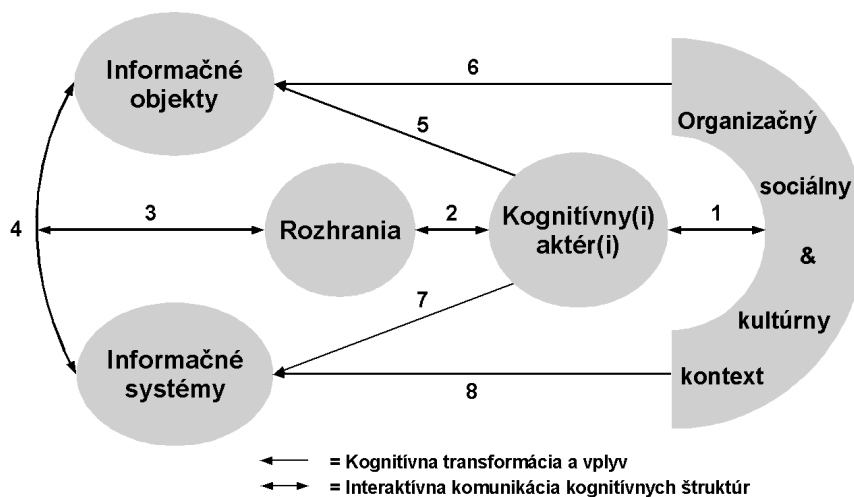
### Interaktívne modely informačného prieskumu a kolaboratívny prieskum

Interakcia je vzájomné pôsobenie medzi rôznymi zložkami prieskumovej situácie. Otázky interakcie rozpracoval v teórii informačného prieskumu najmä Peter Ingwersen vo svojej monografii o interaktívnom informačnom prieskume (Ingwersen 1992). Z hľadiska jednotlivých zložiek zúčastnených na interakcii možno v prieskume rozlíšiť interakciu človek-človek a interakciu človek-systém. Interakciu človek – systém skúma disciplína Human-Computer Interaction (HCI) (Steinerová 1996). Príbuznou disciplínou je aj interakcia človeka s agentom (Human Agent Interaction).

Interaktívny informačný prieskum priamo nadväzuje na kognitívny rámec využívania informácií. Vznikol pod vplyvom vývoja informačných technológií a ich rozšírením z tradičných laboratórnych podmienok medzi široké vrstvy obyvateľstva. Niektorí autori v tejto súvislosti hovoria o kognitívnej, interaktívnej a relevantnej revolúcii vo vyhľadávaní informácií. Princípom interaktívneho vyhľadávania je nový pohľad na to, kto riadi samotný proces vyhľadávania. Používateľ sa zmenil z pasívneho príjemcu informácií na aktívneho kognitívneho aktéra. Tým sa vyhľadávanie približuje myšlienkovým procesom človeka. Podľa zmeny kognitívneho stavu sa menia požiadavky a hodnotí sa relevancia. Túto situáciu zachytávajú aj viaceré interaktívne modely informačného prieskumu, prípadne vyhľadávania informácií. Konceptuálne modely sú dôležité, pretože sprehľadňujú prieskumovú situáciu, jednotlivé zložky a súvislosti medzi nimi.

Medzi interaktívne modely informačného prieskumu možno zaradiť modely P. Ingwersena, T. Saracevica, M. Batesovej. Interaktívny prieskum však modelovali aj N. Belkin (epizodický model), Amanda Spinková a iní autori.

Na obrázku 1.6 je model interaktívneho vyhľadávania informácií.



Obr. 1.6 Interaktívne vyhľadávanie informácií (podľa Ingwersen, Järvelin 2005)

Interaktívne vyhľadávanie vychádza z kognitívneho prieskumu, z individuálnych stavov poznania, ktoré sú v interakcii s informačným prostredím. Ingwersen v predchádzajúcich modeloch naznačil, ako sa stav neurčitosti prepája s problémovým priestorom, aktuálnym stavom poznania, pracovným priestorom. Pritom všetky tieto zložky sú integrované v kognitívnom modeli a v interakcii s modelom okolitého sveta. V problémovom priestore sa nachádzajú situácie umožňujúce hľadať zaujímavý článok, prípadne prehľad literatúry. Vo vedeckej práci môže ísť aj o rôzne spôsoby nadväzo-

vania na výsledky kolegov, buď citovaním a vytváraním zoznamov citácií, alebo pripomienkami či oponovaním voči výsledkom iných autorov.

Ingwersen tiež skúmal podrobnejšie interakciu medzi objektmi (texty, dokumenty) a prostredím informačného prieskumového systému (indexačný a prieskumový jazyk, databázová štruktúra, spôsoby spätnej väzby, stratégia selekcie dokumentov). V poslednom modeli (na obrázku 1.6) spolu s Järvelinom pridali dôležité kontexty informačného prieskumu: sociálne, organizačné a kultúrne prostredie.

Ingwersen, Belkin aj ďalší autori zdôraznili, že hlavným princípom vyhľadávania informácií a informačného prieskumu je zameranosť (intencionalita). Vyhľadávanie a prieskum vždy vychádzajú z kontextov ako 1. pracovná úloha alebo každodenné úlohy a záujmy človeka, 2. vnímané prieskumové úlohy (myšlienkové, emocionálne a fyzické aktivity pri riešení pracovných a každodenných úloh).

V modeli na obrázku 6 je naznačená kognitívna transformácia (zmena kognitívnych stavov) postupujúca od kognitívneho aktéra, prípadne viacerých aktérov (tímu) smerom k informačným objektom a informačnému systému rozdelenému na samotnú technologickú zložku, algoritmy a logickú štruktúru. Zo *sociálnych, organizačných a kultúrnych* kontextov sa ako dôsledok interakcií smerom ku kognitívnemu aktérovi vynárajú pracovné alebo každodenné úlohy, ktoré vedú k potrebám a požiadavkám na vyhľadávanie informácií. Významnú úlohu v modeli predstavuje práve interface ako rozhranie na komunikáciu človeka so systémom a informačným objektmi. To je základ interaktívnej komunikácie, kde prebieha interakcia kognitívnych štruktúr (väzby 1, 2, 3, 4). Väzba 1 naznačuje procesy sociálnej interakcie medzi aktérmi a aktuálnymi sociálno-kultúrnymi a organizačnými kontextami. Ďalšie interakcie sa odohrávajú medzi reprezentáciami poznania v systéme a existujúcimi informačnými objektmi. Vertikálna interakcia medzi objektmi a systémami vytvára základ prieskumového systému, pričom sa tu interakcie odohrávajú aj na najnižšej lingvistickej, znakovkej úrovni. Kognitívne a emotívne transformácie (5, 7) sa vyskytujú pri vytváraní nových informácií individuálnym aktérom a vyplývajú aj z kontextov ovplyvňujúcich systémy a informačné objekty v čase (6, 8). Model znázorňuje mnohonásobné interakcie v prieskumovom procese. Pritom nepoukazuje na postupnosť procesov. Okrem možného dialógu a interpretácie model v interakciách zdôrazňuje význam kontextov a kultúrnej podmienenosti prieskumových interakcií. Model naznačuje aj možnú prácu v kolektíve spoločným prispôsobovaním sa prostrediu, pričom vzniká kolaboratívny informačný prieskum (CIR).

**Kolaboratívne vyhľadávanie a informačný prieskum** skúmajú autori už dlhšie ako spoluprácu ľudí pri aktívnom hľadaní informácií s cieľom riešiť špecifickú úlohu. Môže ísť o spoluprácu pri formulovaní požiadavky, uchovávaní dotazov a ich viacnásobnom využití, pri vytváraní stratégií prieskumu aj pri posudzovaní relevancie. Zvyčajne je súčasťou širšej spolupráce ľudí pri využívaní informácií (kolaboratívneho informačného správania). **Kolaboratívne informačné** správanie možno definovať ako také činnosti, v ktorých komunikujú dvaja alebo viacerí aktéri s cieľom identifikovať informácie na riešenie problému alebo splnenie úlohy. Zahŕňa identifikáciu problému, analýzu informačnej potreby, formuláciu dotazu, prieskumové interakcie, hodnotenie, prezentáciu výsledkov a aplikáciu informácií na riešenie problému (Talja, Hansen 2005). Napríklad Fidelová a kol. (Fidel et al. 2000) modelujú aktéra s jeho kognitívnym stavom, ktorý vykonáva činnosti (aktivity) v určitej doméne, v organizácii. Rôzne aktivity sú spojené s riešením konkrétnej úlohy, typom úlohy, stratégiami riešenia a sú ovplyvnené sociálnymi interakciami. Spoluprácu možno rozdeliť na spoločnú tvorbu dotazov, kolaboratívne filtrovanie, kolaboratívnu sociálnu navigáciu (synchronickú a asynchronickú) a delenie sa o informácie.

**Delenie sa o informácie** v akademickom prostredí bližšie charakterizovali aj štúdie informačnej vedy (Foster 2006, Talja, Hansen 2005). Obsahuje široké spektrum informačných aktivít, od náhodného získavania informácií až po cieľavedomé spoločné formulovanie dotazov a informačných problémov. Podľa iných kritérií existuje strategické „zdieľanie“ pri formovaní stratégie, paradigmatické „zdieľanie“ pri spoločnom postupe, viacstranné komunikovanie medzi mladšími a staršími výskumníkmi a sociálne delenie sa o informáciami, ktoré prispieva k budovaniu komunity.

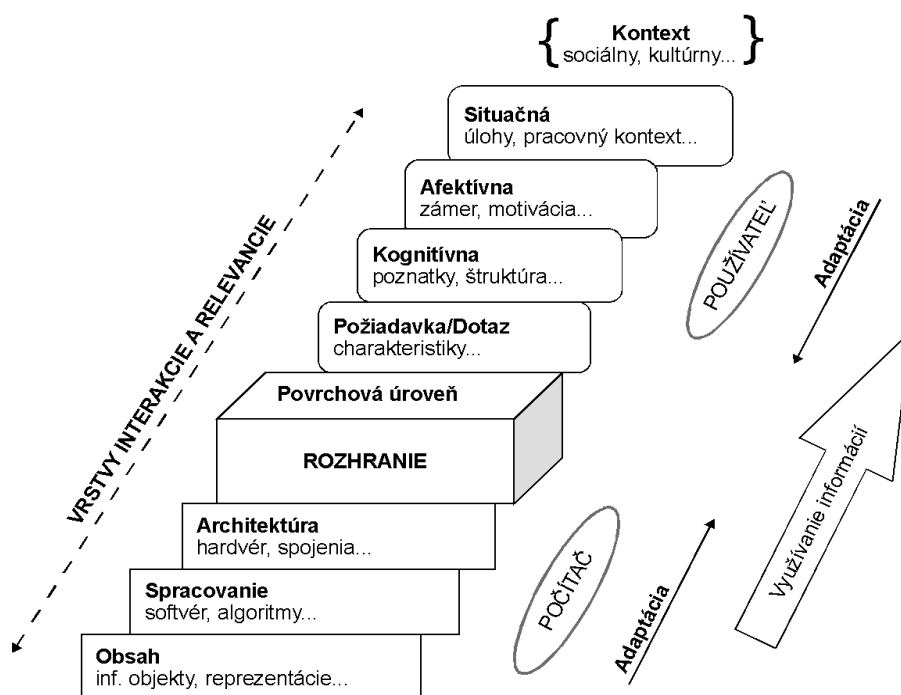
Kolaboratívne informačné správanie sa delí na asynchrónne a synchrónne, vzdialené a v spoločnom mieste, voľne alebo pevne prepojené činnosti, plánovaná a neplánovaná spolupráca, spolupráca vo vnútri skupiny alebo medzi skupinami, priama a nepriama spolupráca, koordinované a diferencované aktivity (Talja, Hansen 2005).

Kolaboratívny prieskum a vyhľadávanie sú dôležité aj pri vedeckej kolaborácii a budovaní kolaboratórií v elektronickom priestore (Sonnenwald 2006, Foster 2006). Kolaboratívne informačné správanie obsahuje nielen informačný účel, ale aj sociálny účel. V skupinách sa prelínajú rôzne roly informačného správania človeka ako akvizičný pracovník, referenčný pracovník, overovateľ, filtrovač (editor), analytik, organizátor, indexátor, administrátor, manažér. Kolaborácia sa odvíja v takých štádiách ako založenie projektu, formulácia vízie a cieľov, udržiavanie vzťahov, učenie a komunikácia až po ukončenie spolupráce (definícia úspechu, rozširovanie výsledkov). Osobitné spôsoby kolaborácie a zdieľania informácií sa rozvíjajú v rámci sociálnych sietí v internete.

Spolupráca pri vyhľadávaní a informačnom prieskume môže obsahovať: vzájomné využívanie (zdieľanie) rovnakej informačnej potreby, prieskumových stratégií, výsledkov prieskumu, vyhľadávaných informačných objektov, ďalšie spracovanie vyhľadávaných informačných objektov (interpretácia, filtrovanie, syntéza, archivácia potenciálne užitočných informácií v skupinových repozitároch).

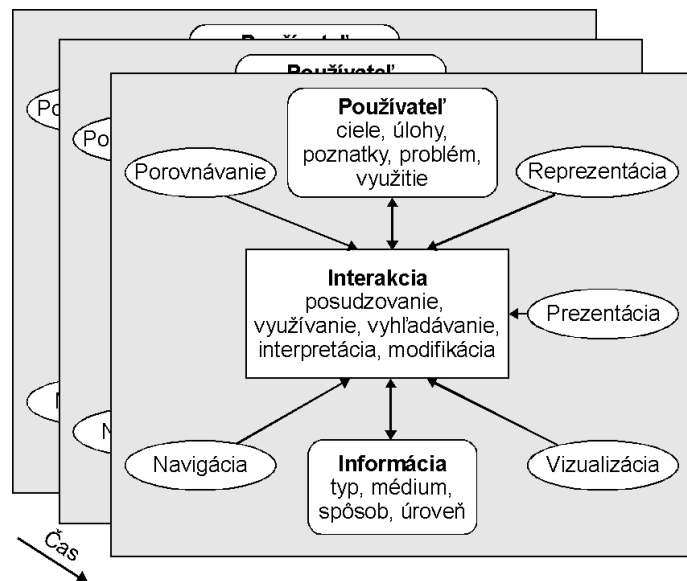
Ďalší interaktívny model informačného prieskumu predstavil Tefko Saracevic (Saracevic 1997) (Steinerová 2005). Ide o model rozvrstvenej interakcie (obrázok 1.7).

Model znázorňuje sériu rôznych interakcií medzi systémom a človekom v rôznych vrstvách. Na úrovni systému sú to technológie, programy (algoritmy) a rozhranie. Na úrovni človeka je to prostredie, situácia, stav poznania a požiadavky používateľa. Jednotlivé vrstvy týchto zložiek sa navzájom ovplyvňujú. Prítom existuje dôležitý proces adaptácie, prispôbovania človeka a systému svojmu prostrediu v procese využívania informácií.



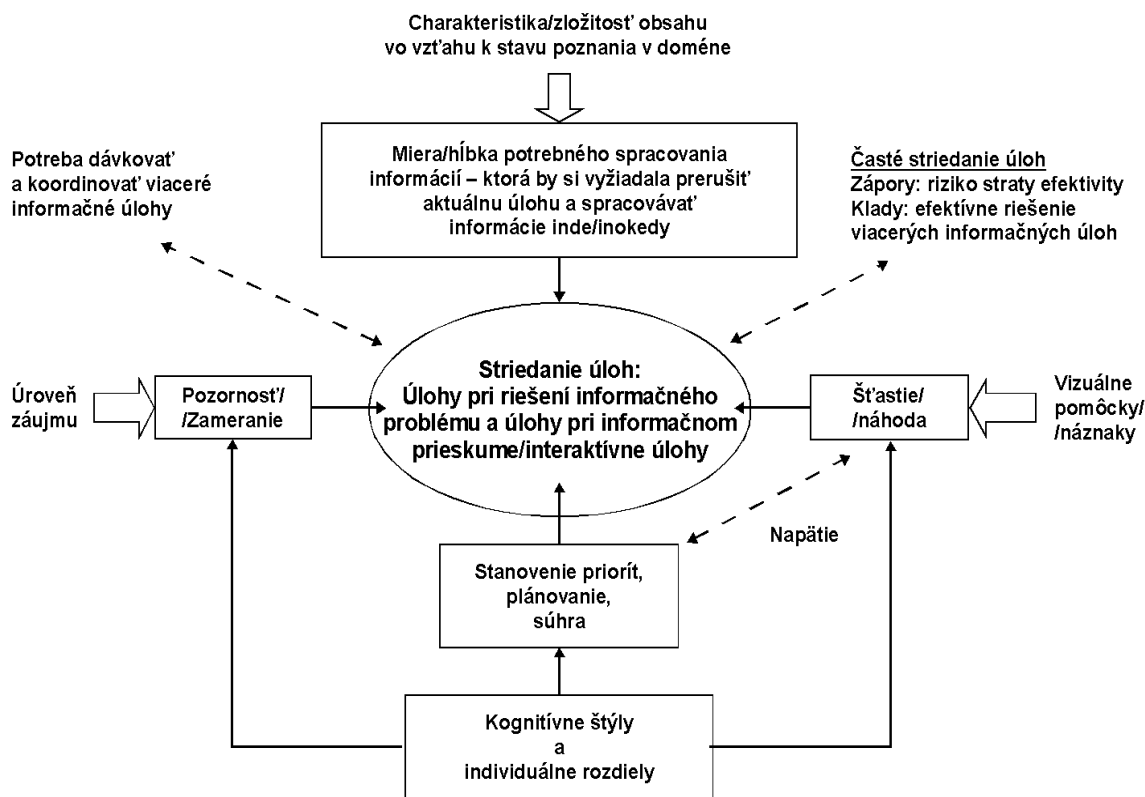
Obr. 1.7 Saracevicov model rôznych vrstiev interakcií pri vyhľadávaní a využívaní informácií

Belkinov epizodický model naznačuje interakciu ako sériu epizód v priebehu vyhľadávania (obrázok 1.8).



Obr. 1.8 Epizodický model N. Belkina (podľa Belkin 1996)

Model Amandy Spinkovej (Spink, Cole 2005b) zdôrazňuje dôležitosť multi-taskingu (paralelného spracovania viacerých úloh v informačnom prieskume) ako typickú vlastnosť interaktívneho prieskumu. Model je znázornený na obrázku 1.9.



Obr. 1.9 Model interaktívneho informačného prieskumu ako multitaskingu



## 1.7

### Od vyhľadávania k integrovaným systémom

Aj keď sme sa doteraz zaoberali najmä prieskumovými systémami, v reálnom elektronickom informačnom prostredí sú zvyčajne súčasťou komplexných integrovaných knižnično-informačných systémov prípadne iných integrovaných systémov (napríklad Content Management Systems, CMS, Web Content Managers a pod.). Prieskumové systémy či vyhľadávače sú prepojené s viacerými zdrojmi (funkcia metavyhľadávania, integrácia databáz), s filtrovaním, archivovaním. Prieskumový systém sa tiež čoraz viac prispôsobuje používateľovi (funkcia personalizácie), dôležité sú príjemné používateľské rozhrania. Pritom sa čoraz väčší dôraz kladie na funkcie prezentácie a vizualizácie, ktoré aj pri manipulácii s pojmami môžu podporiť pochopenie témy a súvislostí. Z hľadiska hodnotenia relevancie je dôležitá aj funkcia validácie. To znamená overovanie pôvodu zdroja, dôveryhodnosti a spoľahlivosti informácií.

Ak profesionálne prieskumové systémy (špecializované databázy a online katalógy knižníc) možno považovať za spoľahlivé a overené zdroje, zodpovedne získavané a organizované informačnými profesionálmi, vo webovom prostredí takáto kontrola chýba. Preto niektoré vyhľadávače poskytujú možnosť validácie výsledkov vyhľadávania.

V tomto vývoji sa vyhľadávacie systémy integrujú s inteligentnými agentmi, sofistikovanými systémami, ktoré poskytujú pridanú hodnotu k vyhľadávaniu informácií (napríklad agent Copernic zameraný na vyhľadávanie vedeckej literatúry).

K ďalším dôležitým funkciám pridaným k vyhľadávaniu patrí sémantický rozmer a kolaboratívny rozmer. Sémantický rozmer súvisí so snahami lepšie reprezentovať význam informácií. Tu vznikajú nové nástroje organizácie poznania v elektronickom prostredí (ontológie, taxonómie, tematické mapy, pojmové mapy, inteligentné tezaury a i.). Vytvára sa sémantický web, ktorý uplatňuje inteligentné nástroje na reprezentáciu významu informácií. Novšie nástroje sa približujú prirodzeným procesom kategorizácie a klasifikácie. Zlepšujú sa nástroje na opis elektronických zdrojov (metadáta). Okrem toho sa vyhľadávanie orientuje na kontext (súvislosti), pracuje s anotáciami, komentármi, spoločným hodnotením. Tak vzniká kolaboratívny rozmer vyhľadávania. Môže ísť o „zdieľanie“, diskusiu, pomoc pri vyhľadávaní, formulovaní, hodnotení informácií. Podpora komunity môže byť dôležitá ako pri organizácii výskumu, tak aj pri laickom vyhľadávaní a informačnom správaní (napríklad podpora pacientov s rovnakou diagnózou, podpora komunity so spoločným problémom).

Významnou modernou funkciou prieskumových systémov je aj podpora tvorivosti človeka a objavovanie nových súvislostí a trendov. Napríklad systémy dolovania a analýzy dát a textov (data mining, text mining) môžu nájsť spoločné a odlišné charakteristiky zdrojov a odhadnúť tendencie vývoja. Systémy môžu podporiť používateľa pri skladaní „mozaiky“ zo zdrojov („snippets“ – útržky nájdených textov, napr. systém NetSnippets). Podpora používateľa smeruje až k tvorbe informačných produktov, napríklad pri správe zdrojov, tvorbe bibliografických zoznamov, analýzach, publikovaní, ale aj tvorbe textov).

## 1.8

### Zhrnutie

Informačný prieskum, vyhľadávanie informácií a informačné správanie sú vzájomne súvisiace pojmy. Vyhľadávanie informácií sa spája so sociálnymi vedami a poznaním používateľa. Informačný prieskum vychádza najmä z možností spracovania informácií informačným prieskumovým systémom. Základnú situáciu vyhľadávania informácií a informačného prieskumu tvoria interakcie medzi tromi zložkami: systémom, človekom, obsahom (dokumenty, zdroje).

Medzi dôležité pojmy na pochopenie vyhľadávania informácií patria informačná potreba, informačná požiadavka, informačný prieskumový systém, organizácia poznania, informácia, poznatok,

relevancia, kontext. Okrem informačnej vedy sa vyhľadávaním zaoberajú aj také disciplíny ako informatika, manažment poznatkov, umelá inteligencia a pod.

Princípy a procesy vyhľadávania ho delia na fyzický a intelektuálny prístup. Vyhľadávanie možno rozdeliť na niekoľko krokov, od začatia (formulovania ohniska záujmu) po interpretáciu. Postupnosť krokov pri vyhľadávaní naznačujú modely D. Ellisa, C. C. Kuhlthauovej a i. Rôzne interpretácie procesu vyhľadávania závisia od toho, či zdôrazňujeme myšlienkové procesy jednotlivca (kognitivismus) alebo terminológiu disciplín a domén (problémových oblastí) (doménová analýza). Medzi dôležité modely informačného prieskumu patria interaktívne modely (Ingwersen, Järvelin, Saracevic, Batesová, Belkin, Spinková). Najnovšie sa skúma aj kolaboratívny informačný prieskum a kolaboratívne informačné správanie.

V histórii vývoja informačného prieskumu možno určiť mnoho zmien pod vplyvom rozvoja informačných technológií. Vývojové modely ukazujú, že sa na jednej strane vnútorné algoritmy zdokonaľovali a na druhej strane sa zjednodušovala interakcia so systémami. Najnovšie trendy vo vyhľadávaní zdôrazňujú kolaboráciu, interaktívnosť, personalizáciu a kontext. Vyhľadávanie je integrované v širších systémoch zameraných na analýzu informácií a podporu tvorivej informačnej činnosti používateľa.

## Literatúra

BELKIN, Nicholas J. 1996. Intelligent Information Retrieval: Whose Intelligence? In *ISI '96: Proceedings of the Fifth International Symposium for Information Science* [online]. Konstanz : Universitätsverlag Konstanz, 1996. S. 25 – 31. Dostupné na internete: <<http://comminfo.rutgers.edu/~belkin/articles/isi96-paper.pdf>>.

ELLIS, David. 1990. *New Horizons in Information Retrieval*. London : LA, 1990. 138 s.

FIDEL, R. et al. 2000. Collaborative Information Retrieval (CIR). In *The New Review for Information Behaviour Research*. 2000. Studies of information seeking in context. Cambridge : Taylor Graham, 2000, vol. 1, s. 235 – 247.

FOSTER, Jonathan. 2006. Collaborative Information Seeking and Retrieval. Chapter 8. In Ed. CRONIN, Blaise. *ANNUAL Review of Information Science and Technology*. Medford : Information Today, 2006. ISSN 0066-4200, vol. 40, s. 329 – 356.

HJØRLAND, Birger. 1997. *Information Seeking and Subject Representation : an Activity-Theoretical Approach to Information Science*. Westport : Greenwood Press, 1997. 213 s. ISBN 0-313-29893-9.

HJØRLAND, Birger. 2000. Information Seeking Behaviour : What Should a General Theory Look Like? In *The New Review of Information Behaviour Research*. London : Taylor Graham, 2000, vol. 1, s. 19 – 33.

INGWERSEN, Peter. 1992. *Information Retrieval Interaction* [online]. London : Taylor Graham, 1992 [cit. 2007-07-10]. 246 s. Dostupné na internete: <<http://www.db.dk/pi/iri/>>.

INGWERSEN, Peter – JÄRVELIN, Kalervo. 2005. *The Turn : Integration of Information Seeking and Retrieval in Context*. Dordrecht : Springer, 2005. 448 s. ISBN 1-4020-3850-X.

KUHLTHAU, Carol C. 1993. *Seeking Meaning : A Process Approach To Library And Information Services*. Norwood : Ablex, 1993. 199 s.

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

MANNING, Christopher D. – RAGHAVAN, Prabhakan – SCHÜTZE, Heinrich. 2009. *Introduction to Information Retrieval* [online]. Cambridge University Press, 2009 [cit. 2010-01-11]. Dostupné na internete: <<http://informationretrieval.org/>>.

SARACEVIC, Tefko. 1997. The stratified model of information retrieval interaction : Extension and application. In *Proceedings of the American Association for Information Science* [online]. 1997, vol. 34 [cit. 2003-09-30]. s. 313 – 327. Dostupné na internete: <<http://www.scils.rutgers.edu/~tefko/ProcASIS1997.doc/>>.

SONNENWALD, D. H. 2006. Scientific Collaboration : A Synthesis of Challenges and Strategies. In *ARIST*. 2007. Draft, January 2006.

SPINK, Amanda – COLE, Charles. 2005a. Chapter 1. Introduction : New Directions in Human Information Behavior. In Eds. SPINK, A., COLE, Ch. *New Directions in Human Information Behavior*. Dordrecht : Springer, 2005. s. 3 – 10. ISBN 1-4020-3667-1.

SPINK, Amanda – COLE, Charles. 2005b. Chapter 8. Multitasking and Co-ordinating Framework for Human Information Behavior. In Eds. SPINK, A., COLE, Ch. *New Directions in Human Information Behavior*. Dordrecht : Springer 2005, s. 137 – 154. ISBN 1-4020-3667-1.

STEINEROVÁ, Jela. 1996. *Teória informačného prieskumu*. Bratislava : CVTI SR, 1996. 262 s. ISBN 80-65165-58-9.

STEINEROVÁ, Jela. 1998a. *Tvorba informačných produktov : nové prístupy informačnej vedy*. Bratislava : CVTI SR, 1998. 130 s. ISBN 80-85165-73-2.

STEINEROVÁ, Jela. 2005a. *Informačné správanie : Pohľady informačnej vedy*. Bratislava : CVTI SR, 2005. 189 s. ISBN 80-85165-90-2.

TALJA, Sanna – HANSEN, Preben. 2005. Information Sharing. Chapter 7. In Eds. SPINK, A., COLE, Ch. *New Directions in Human Information Behavior*. Dordrecht : Springer, 2005. s. 113 – 134. ISBN 1-4020-3667-1.

WILSON, Tom D. 1999. Models in information behaviour research. In *Journal of Documentation* [online]. 1999, vol. 55, no. 3 [cit. 2003-09-30]. Dostupné na internete: <<http://informationr.net/tdw/publ/papers/1999JDoc.html>>. ISSN 0022 0418.

# 2

## Informačný horizont a prieskumové stratégie

### 2.1 Informačný horizont

Informačný horizont môžeme definovať ako súbor informačných zdrojov, ktoré má človek k dispozícii pri vyhľadávaní informácií. Vynára sa ako vejár možností, ktorou cestou postupovať pri uspokojovaní informačnej potreby.

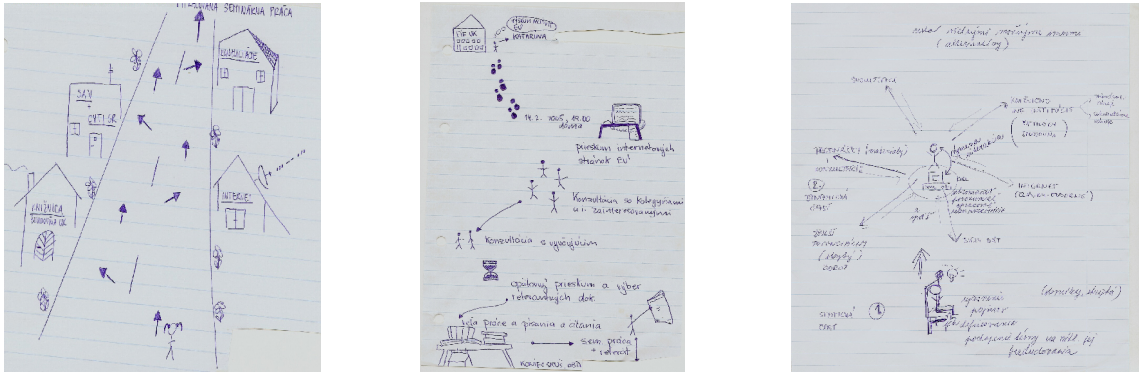
Informačný horizont vychádza z rámca výskumu informačného správania (Sonnenwald, Willemuth, Harmon 2001, Savolainen 2008). Informačný horizont je súčasťou orientačného štádia informačného správania človeka. Funkciou štádia orientácie je spresnenie témy alebo problému a postupné formulovanie informačnej požiadavky na základe informačnej potreby. Medzi ďalšie funkcie orientačného štádia patrí formulovanie a selekcia pojmov. Pochopenie základných pojmov z určitej oblasti a témy je dôležité kvôli možnému úspechu vyhľadávania. Pojmová štruktúra problematiky zabezpečuje intelektuálnu orientáciu. Je to ako návšteva doteraz neznámej krajiny alebo mesta, po ktorých nás naviguje mapa. Túto mapu si vytvárame zo základných pojmov.

Medzi základné informačné zdroje pri vytváraní informačného horizontu s cieľom orientovať sa pri vyhľadávaní informácií patria najmä slovníky a encyklopédie. Často ich hľadáme v elektronickej forme v prostredí webu.

Informačné zdroje v informačnom horizonte chápeme v čo najširšom zmysle slova. Každý človek má svoj vlastný informačný horizont. Môže obsahovať osoby a sociálne siete, vyhľadávacie nástroje, webové sídla, knižnice, databázy, systémy na experimentovanie a pozorovanie, dokumenty v rôznych formách a formátoch (knihy, časopisy, videozáznamy, zborníky). Pri vytváraní mapy informačného horizontu hrajú dôležitú úlohu odporúčania expertov (učitelia, školitelia) a kolegov. Interakcie s inými ľuďmi poukazujú na hodnotu zdrojov v našom informačnom horizonte a menia ho. Napríklad ak učiteľ odporúča určitý zdroj študentovi, pozitívne ovplyvní jeho informačný horizont.

Okrem „referenčných“ informácií a pojmov možno do informačného horizontu človeka začleniť aj rozhodnutia a činnosti, ktoré vykonávame pri vyhľadávaní informácií. Súčasťou sú aj údaje o tom, kedy a prečo využívame informačné zdroje a vzájomné vzťahy medzi zdrojmi. Z hľadiska informačných štýlov sú v informačnom horizonte dôležité preferencie pri spracovaní informácií a hodnoty, ktoré človek pripisuje informačným zdrojom. V informačnom horizonte sú začlenené aj vplyvy kontextov a situácií na spracovanie informácií.

Pre študentov môže byť zaujímavé vytvoriť si a aj graficky znázorniť informačný horizont ešte pred riešením informačného problému. Príklady informačných horizontov sú znázornené na obrázku 2.1.

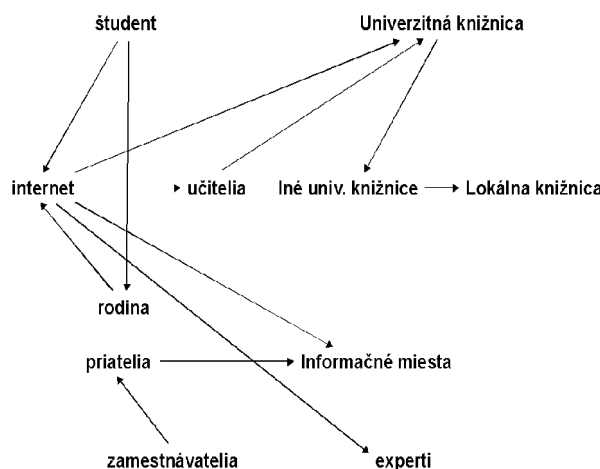


Obr. 2.1 Príklady informačných horizontov (študenti KKIV)

Informačný horizont obsahuje aj „informačné miesta“ – prostredia, kde sa vymieňajú informácie. V sieti zdrojov majú svoje funkcie noviny, časopisy, médiá, katalógy, databázy, knižnice, ale aj experti, rodina a priatelia. Ak si chceme vytvoriť informačný horizont ako začiatok vyhľadávania, mali by sme opísať špecifické situácie, ktoré nás podnecujú k hľadaniu informácií. Najdôležitejší je **kontext** obsahujúci študijnú úlohu alebo vedecký problém či projekt. Podrobnejšie si treba určiť najmä typy vhodných informácií, typy informačných zdrojov, odpovedať na otázku prečo práve konkrétne typy zdrojov a informácií môžu riešiť náš informačný problém. Tiež je vhodné určiť si postupnosť využitia informačných zdrojov a poradie dôležitosti typov informácií. Často sa využíva delenie informácií na *dôležité*, *menej dôležité*, *okrajové*. Pri určovaní a znázorňovaní informačného horizontu nám pomáha najmä nedávna skúsenosť s využívaním informácií a zdrojov. V pamäti máme zaznamenané úspešné stratégie, ktoré sme použili pri riešení predchádzajúcich informačných problémov, najmä spôsob využitia a prípadné chyby, ktorým sa chceme vyhnúť. Preto pri metóde opisu informačného horizontu výskumníci žiadajú účastníkov, aby opísali aj situácie, keď bolo ťažké /alebo ľahké nájsť informácie a kedy boli spokojní, prípadne nespokojní pri hľadaní informácií.

Pri grafickom znázornení horizontu je dôležité naznačiť aj postavenie samotného človeka medzi možnými zdrojmi pri riešení informačného problému. V opise je užitočné si určiť kedy a prečo používateľ využije konkrétny zdroj, poradie zdrojov, prípadne aj to, či niektorá služba aktívne ponúka ďalšie možné zdroje.

Ako príklad možno uviesť grafické znázornenie siete informačných zdrojov študentov podľa výskumu (Sonnenwald, Wildemuth, Harmon 2001) na obrázku 2.2.



Obr. 2.2 Informačný horizont (podľa Sonnenwald, D., Wildemuth, B., Harmon, G. L. 2001)

Autori výskumu dokázali, že na začiatku sa používajú najmä médiá, katalógy, internet, rodina, priatelia, zamestnávateľ a rôzne (vzdialené) knižnice. Pre hlbšie zameranie pozornosti prichádzajú na rad učitelia, univerzitná knižnica, experti, časopisy a informačné miesta. Pri ukončení riešenia informačného problému sa napokon najčastejšie používa lokálna knižnica (napríklad katedrová) s konkrétnymi dokumentmi na štúdium.

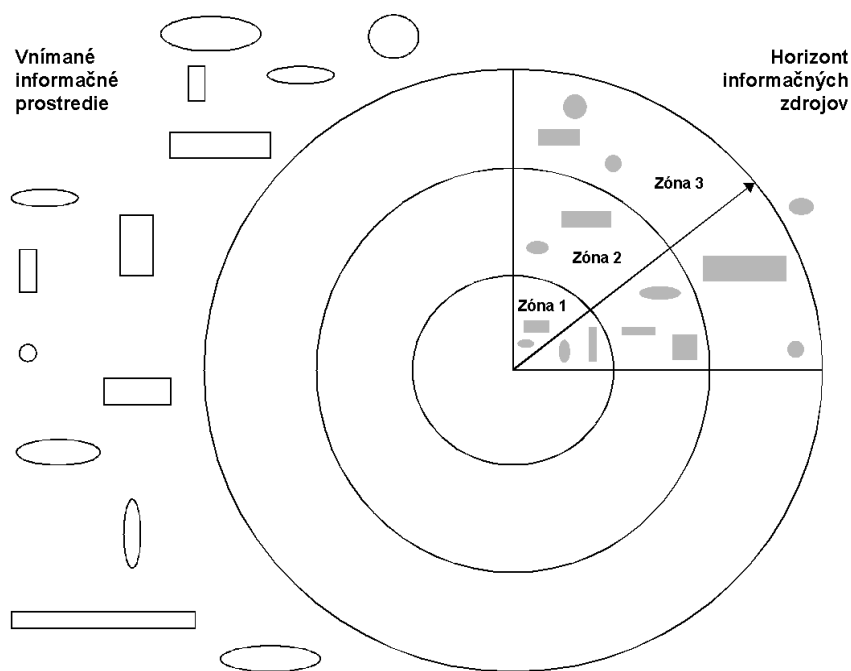
Informačné zdroje v informačnom horizonte sa kategorizujú aj na *počítačové zdroje*, *odporúčajúce zdroje*, *zdroje na zúženie problému a finálne zdroje*. Kritériom je nielen časový postup pri využívaní jednotlivých druhov zdrojov, ale aj postup od orientačného (referenčného) štádia vyhľadávania k hĺbkovému, zameranému a analytickému.

Pri každodennom informačnom správaní definuje Reijo Savolainen (2008) *horizonty informačných zdrojov*. Tieto horizonty vznikajú v širšom kontexte tzv. vnímaného informačného prostredia. Človek pri vytváraní informačného horizontu pristupuje k informačným zdrojom selektívne a pri tom posudzuje možnú relevanciu informačného zdroja. Savolainen rozlišuje v horizonte informačných zdrojov tri zóny:

1. zóna – najviac preferované informačné zdroje,
2. zóna – informačné zdroje sekundárneho významu,
3. zóna – periférne informačné zdroje.

Informačný horizont je tu vlastne uhlom pohľadu pri plánovaní vyhľadávania a využívania informácií. Môžeme ho nazvať aj „*myšlienkovým okom*“.

Zjednodušené znázornenie informačného horizontu informačných zdrojov a zón preferencií zdrojov podľa Savolainena (2008) je na obrázku 2.3. Konštrukcia informačného horizontu závisí od kontextu. Napríklad ak je naším cieľom monitorovanie aktuálnych udalostí, potom v prvej zóne sa objavujú noviny, televízia, prípadne spravodajské informácie na internete.



Obr. 2.3 Horizont informačných zdrojov (podľa Savolainen 2008)

Na základe výskumov informačného správania možno informačné horizonty rozdeliť na relatívne stabilné horizonty, ktoré odzrkadľujú stabilné preferencie človeka pri hodnotení zdrojov v rôznych situáciách a na dynamické horizonty, ktoré sú závislé od problému a situácie.

## 2.2 Informačné stratégie

Informačné stratégie sú spojené s prístupom k riešeniu problémov a postupnosťou vyhľadávania, využívania a tvorby informácií. Informačná veda a informačný prieskum rozlíšili niekoľko druhov informačných stratégií podľa možností systémov a podľa vzorcov informačného správania.

Informačné stratégie sa od prieskumovej stratégie odlišuje širším významom. Informačná stratégia je plánovanie riešenia informačného problému, ktoré je často intuitívne. Prieskumová stratégia je plánovanie samotného procesu vyhľadávania informácií, najmä výber nástroja na vyhľadávanie a možných relevantných zdrojov, formulovanie požiadavky a dotazu, výber pojmov a viacnásobné hodnotenie výstupov. V užšom zmysle je prieskumová stratégia súbor algoritmov, ktoré sa používajú v systéme na vyhľadávanie relevantných informácií.

Informačné stratégie na úrovni používateľa spájame so stratégiami riešenia problémov. Takto určuje štádiá informačného správania Wilson (1999) podľa štádií riešenia problému. 1. Určenie problému (*Aký mám problém*), 2. Plánovanie postupu riešenia (*Ako budem postupovať pri riešení tohto problému?*), 3. Realizácia riešenia (*Toto je riešenie problému*), 4. Kontrola a hodnotenie účinnosti riešenia, spätné väzby (*Spresnenie formulácie a riešenia problému*).

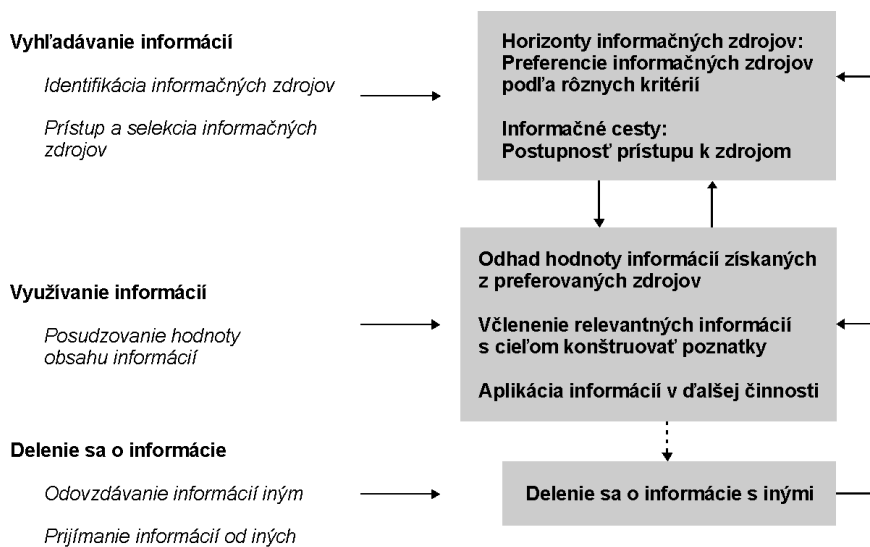
Informačné stratégie možno deliť aj podľa fáz riešenia problému. Medzi základné fázy patria: orientačno-analytická fáza, strategicko-operačná fáza, synteticko-overovacia fáza (Ruisel 2004).

V *orientačno-analytickej fáze* sa snažíme určiť problém. Najdôležitejšie je vnímať a formulovať problém. Analyzujú sa počiatočné údaje a štruktúra problému. Vo vyhľadávaní sa to prejavuje ako hľadanie kľúčových termínov na vyjadrenie problému a ich vzťahov. Súčasťou tejto fázy je určenie cieľa vyhľadávania, najmä na čo sa majú informácie použiť. Pri výbere správnych termínov sa oddeľujú dôležité údaje od menej dôležitých, dopĺňujú sa niektoré údaje v systéme témy.

*Strategicko-operačná fáza* sa vyznačuje uplatňovaním poznatkov o postupoch, ktoré máme uložené v pamäti a ktoré možno použiť pri riešení problému. Napríklad heuristické pravidlá uplatňujeme pri rozhodovaní, v akom poradí budeme vyhľadávať jednotlivé časti témy a ako prepojíme termíny. V tejto fáze sa najprv vytvára globálny smer hľadania, potom sa vytvárajú čiastkové návrhy (hypotézy) a nakoniec sa formuluje konečný návrh. Často sa pracuje s rôznymi alternatívami riešenia. Táto fáza je vo vyhľadávaní najužšie spojená s myslením človeka, ktoré je založené nielen na logických zákonoch a algoritmoch. Dôležitú úlohu aj vo vyhľadávaní informácií zohráva aktivita človeka vedená motiváciou, záujmom, hodnotovou orientáciou alebo typom osobnosti.

*Synteticko-overovacia fáza* sa spája s hodnotením alternatív riešenia a informácií. Dôležitou súčasťou je overovanie informácií ako hodnotenie alternatív podľa dostupných informácií. Základným procesom pri vyhľadávaní je hodnotenie relevancie informácií. Pritom sa uplatňuje kritické hodnotenie, rozhodovanie, výber alternatív. Problémom sú nepresné informácie, napríklad na niektorých internetových stránkach, skreslené informácie a interpretácie a pod. Emócie a motivácia podporujú, ale môžu aj brániť pri posudzovaní relevancie a overovaní informácií a riešení. Často sa stáva, že ľudia uprednostňujú prvé informácie (napríklad prvé tri stránky pri vyhľadávaní vo webe).

V informačnej vede sa skúma aj každodenné vyhľadávanie a informačné správanie človeka. Savolainen (2008) vypracoval konceptuálny rámec každodenných informačných postupov/praktík. Do tohto rámca zaraďuje vyhľadávanie informácií, využívanie informácií a delenie sa o informácie. Rámec každodenných informačných postupov je naznačený na obrázku 2.4.



Obr. 2.4 Rámec informačných postupov (podľa Savolainen 2008)

V súvislosti s vyhľadávaním informácií súčasná informačná veda rozlišuje ešte niekoľko zaujímavých pojmov. Termín *informačná cesta* naznačuje postupnosť, v ktorej ľudia plánujú pristupovať (alebo aktuálne pristupujú) k informačným zdrojom v informačnom horizonte. *Informačné pole* zas predstavuje typické usporiadanie informačných stimulov, ktorým je jednotlivец denne vystavený. Ľudia sa snažia stabilizovať svoje informačné pole pri selekcii typu zdroja média a pri usporiadaní prvkov informačného prostredia tak, aby nad nimi mali kontrolu. *Informačná akcia* je proces, v ktorom jednotlivец uskutočňuje zmysluplné činy vo vzťahu k informáciám a poznatkom s cieľom niečo dosiahnuť (Savolainen 2008). V širšom zmysle je to *informačné správanie*, najmä aktivity vyvolané mentálnymi stavmi alebo udalosťami pri spracovaní informácií. Informačné správanie môžeme definovať ako *integrovanú ľudskú činnosť pri spracovaní a využívaní informácií, ktorú možno rozdeliť na rôzne úrovne (neurofyziologickú, senzomotorickú, kognitívnu, afektívnu, sociálnu, kultúrnu, integračnú)* (Steinerová 2005).

## 2.3 Prieskumové stratégie

*Prieskumová stratégia* je v širšom zmysle definovaná ako *voľba krokov* a metód, ktoré si používateľ určuje pri riešení informačného problému. Prvým krokom je potom špecifikovanie informačnej požiadavky. Určujeme najmä termíny, ktoré použijeme pri formulovaní požiadavky. Súčasťou prieskumovej stratégie je aj špecifikovanie typu a formátu dokumentov, jazykové, časové a geografické hľadisko. Prieskumová taktika je konkrétny postup nájdenia cesty k dokumentu a informáciám.

V užšom zmysle slova je prieskumová stratégia spojená s *vnútorným modelom* prieskumového systému. Prieskumová stratégia je v systémovom pohľade vlastne algoritmus, ktorý vypočíta stupeň podobnosti medzi dotazom a množinou dokumentov. Miera tejto podobnosti môže byť rôzna. V systémovom prístupe sa počíta s tým, že čím častejšie sa termíny nachádzajú v dokumente, tým je dokument relevantnejší. Problémom je dvojznačnosť termínov, ktoré sú vlastné prirodzenému jazyku, najmä synonymia a homonymia. Jeden pojem môže byť opísaný rôznymi termínmi, rovnaký termín môže mať viac sémantických definícií. Pri zjemňovaní dotazu sa väčšina algoritmov snaží pridávať alebo odstraňovať termíny z pôvodného dotazu.



## 2.4 Druhy prieskumových stratégií

Prieskumové stratégie možno deliť na základe vnútorných modelov prieskumových systémov (systémové hľadisko) a podľa postupnosti členenia problému (používateľské hľadisko).

### 2.4.1 Systémový prístup k informačným stratégiám

Z hľadiska systémového sa rozlišujú tieto informačné stratégie:

1. Stratégie Booleovho modelu
2. Stratégie vektorového priestorového modelu
3. Stratégie pravdepodobnostného prieskumu
4. Stratégie zhlukového modelu

#### 2.4.1.1 Stratégia Booleovho modelu/Booleovské indexovanie

Booleov model je najrozšírenejším modelom pri spracovaní textových štruktúrovaných údajov. Pracuje na základe invertovaných indexov (súborov slov, termínov). Pri formulovaní dotazu používateľ používa logické operátory na kombinovanie termínov, ktoré reprezentujú množiny dokumentov. Booleov model sa preto nazýva aj logickým a množinovým modelom. V tradičnom Booleovom modeli však systém vyhľadáva všetko alebo nič, neodhaduje sa stupeň relevancie. V rozšírenom Booleovom modeli je možné triedenie výsledkov podľa stupňa/pravdepodobnosti relevancie. Základom je priradovanie váh termínom reprezentujúcim dotaz a dokument. Do triedenia dokumentov sa začleňujú váhy termínov. Na základe váh termínov sa vypočíta koeficient podobnosti.

Z pohľadu používateľa je stratégia Booleovho modelu takou stratégiou, pri ktorej kombinujeme kľúčové slová ako reprezentácie množín dokumentov. Kľúčové slová v informačnej požiadavke sa môžu kombinovať prostredníctvom operátorov AND, OR, NOT.

*Logické AND (súčin, prienik)*

Ak v požiadavke kombinujeme dve kľúčové slová prostredníctvom operátora AND, ide o logický súčin. Možno ho znázorniť ako prienik množín. V tom prípade systém vyhľadá len tie dokumenty, v ktorých sa vyskytujú obidve kľúčové slová. Výsledok prieskumu sa zužuje.

Napríklad Digitálne knižnice AND archivovanie

*Logické OR (súčet, zjednotenie)*

Ak sa slová spoja operátorom OR (logický súčet), systém vyhľadá tie dokumenty, ktoré obsahujú buď jedno alebo druhé kľúčové slovo. Výsledok prieskumu sa rozširuje. Znázorňuje sa to zjednotením množín.

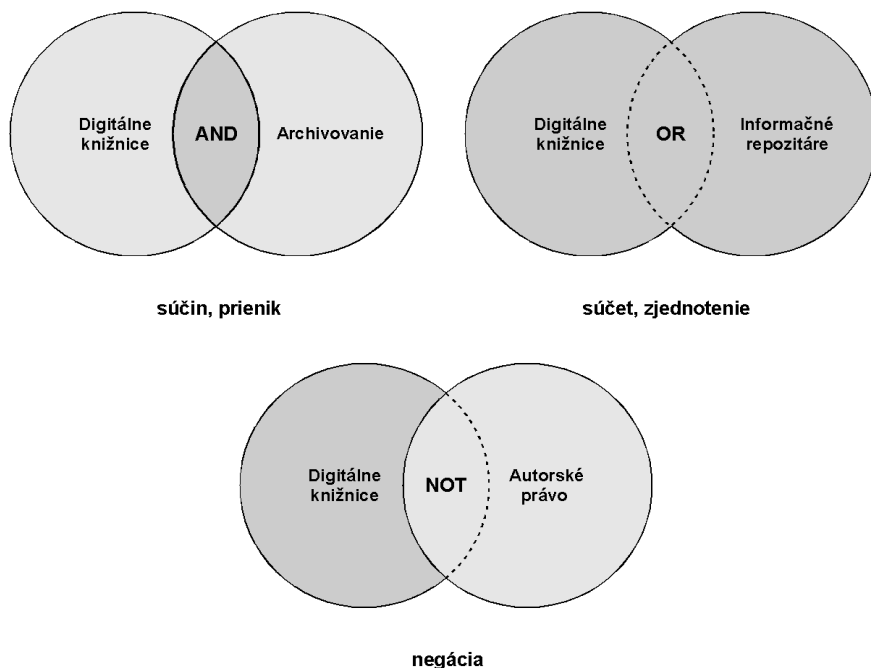
Napríklad Digitálne knižnice OR informačné repozitáre

*Logické NOT (negácia)*

Pri negácii sa používa operátor NOT. Systém v tomto prípade vylúči tie záznamy o dokumentoch, ktoré obsahujú označené kľúčové slovo. Záleží na poradí kľúčových slov. Výsledok prieskumu sa zužuje.

Napríklad Digitálne knižnice NOT autorské právo

Grafické znázornenie týchto prípadov je na obrázku 2.5.



Obr. 2.5 Grafické znázornenie prípadov formulácie dotazov v Booleovom modeli

#### *Proximitné operátory (operátory blízkosti)*

Tieto operátory počítajú s vyhľadávaním vzájomne blízke pojmy. Najčastejšie sú to operátory NEAR, ADJACENT (ADJ), FOLLOWED BY

Vyhľadávajú sa informácie a dokumenty, ktoré obsahujú pojem a za ním nasledujúci termín.

Napríklad základná NEAR škola

Vyhľadávajú sa blízke, príahlé, susediace pojmy

ŠKOLA (základná, stredná, vysoká, atď.)

Napríklad Johann NEAR Bach, Gates NEAR Microsoft

#### *Krátenie termínov/truncation*

Pri formulovaní dotazu sa často používa krátenie termínov, keď sa termín na vyhľadávanie skrúti na koreň slova a systém vyhľadáva všetky možné tvary podľa koreňa slova. Prípony alebo koncovky sú nahradené zástupným znakom.

Napríklad pri vyhľadávaní dokumentov spojených s témou filozofia sa prostredníctvom formulácie FILOZ\$ vyhľadajú dokumenty/stránky obsahujúce termíny *filozofia, filozofický, filozofovanie atď.*

Výsledok vyhľadávania sa rozširuje.

#### *Zástupné znaky/maskovanie (wild cards)*

Zástupné znaky umožňujú nahradzovať určité časti slov/termínov pri vyhľadávaní. Najčastejšie sa používajú \*, ?, \$, %, ale aj iné znaky. Znak môže zastupovať predpony, prípony, prípadne aj vnútorné vzorce znakov (napríklad pri chemických zlúčeninách).

Výsledok vyhľadávania sa rozširuje.

Napríklad wom?n, \*a\*e\*i\*, col\*r, behavi\*r

#### *Vyhľadávanie prostredníctvom fráz*

Používa sa pri bližšom špecifikovaní dotazu. Fráza je slovné spojenie, ktoré sa hľadá tak, že sa obe slová musia vyskytovať v presnom poradí a v uvedenom tvare. Najčastejšie sa vyhľadáva uvedením v úvodzovkách.

Napríklad „informačná ekológia“.

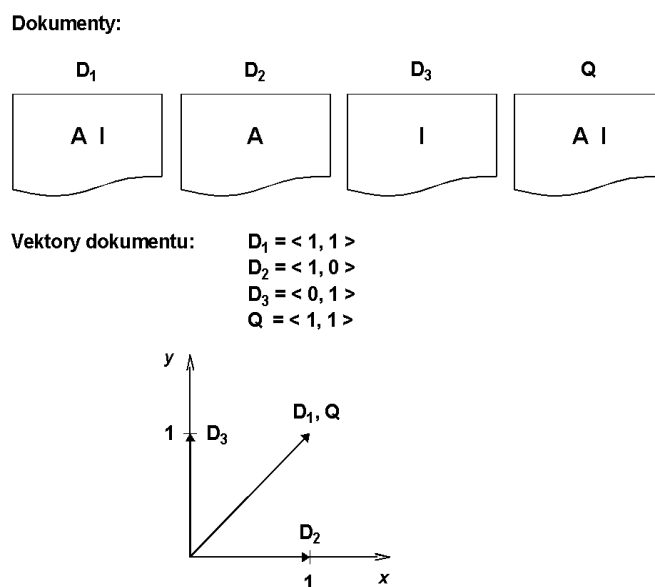
Výsledok vyhľadávania sa zužuje.

### 2.4.1.2

## Stratégie vektorového priestorového modelu

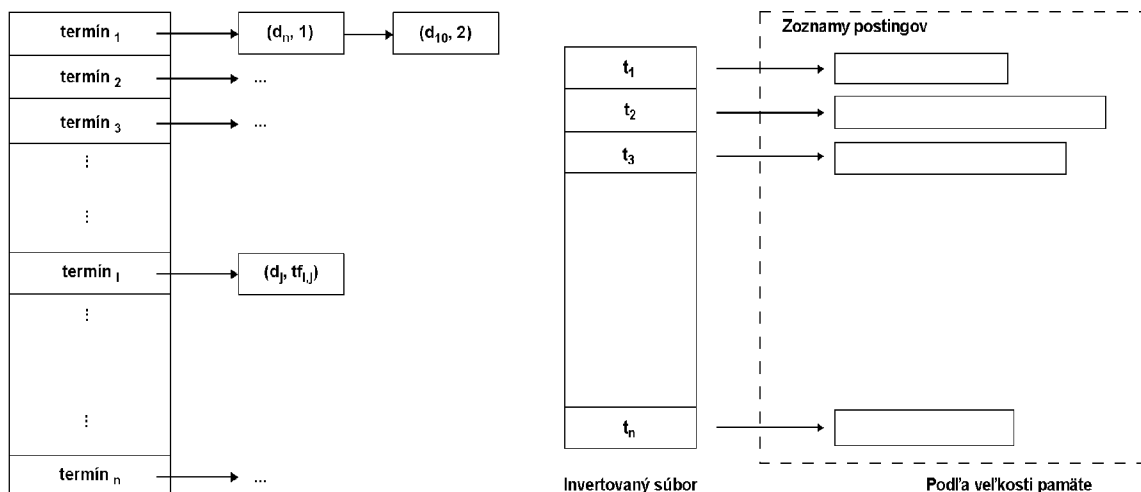
Prvýkrát tento model definoval Gerard Salton v systéme SMART v 60. a 70. rokoch 20. stor. Predpokladá sa, že význam dokumentu je sprostredkovaný použitými slovami, pričom tieto slová v dokumente (termíny) možno reprezentovať prostredníctvom vektora. Dotaz a každý dokument sú reprezentované ako vektory v priestore termínov. Vypočíta sa miera podobnosti (koeficient) medzi týmito dvoma vektormi. Meranie koeficientu podobnosti sa používa dodnes pri experimentovaní s textami v rámci výskumu TREC (Text Retrieval Conference). V tomto modeli môže používateľ priradiť termínom váhu na označenie toho, že jeden termín je dôležitejší ako druhý. Priradovanie váh môže byť aj automatizované. Tradičnou metódou na určenie blízkosti dvoch vektorov je využitie veľkosti uhla medzi nimi. Mierou zhody medzi reprezentáciami dokumentu a dotazu je tu „koeficient podobnosti“ (veľkosť uhla vektoru).

Na obrázku 2.6 je znázornený vektorový priestorový model, ktorý počíta s umelým jazykom s dvoma termínmi (A a I). Všetky dotazy (Q) a dokumenty (D) môžu byť reprezentované v dvojrozmernom priestore týmito dvoma termínmi. Koeficient podobnosti možno vypočítať ako vzdialenosť medzi dotazom a vektormi (Grossman, Frieder 2004).



Obr. 2.6 Vektorový priestorový model (podľa Grossman Frieder 2004)

Pri implementácii vektorového priestorového modelu a iných prieskumových stratégiách sa využíva tzv. invertovaný index. Index je štruktúra, ktorá obsahuje záznam pre každý termín 1-n. Každý termín reprezentuje záznam obsahujúci odkazy (posting list), identifikátor dokumentu a frekvenciu termínu. V praxi bývajú štruktúry zložitejšie. Jednoduché príklady invertovaného indexu sú na obrázkoch 2.7a a 2.7b.



Obr. 2.7a, 2.7b Príklady invertovaného indexu (podľa Grossman, Frieder 2004)

### 2.4.1.3

#### Stratégie pravdepodobnostného modelu/pravdepodobnostný prieskum

V pravdepodobnostnom modeli ide o aplikáciu teórie pravdepodobnosti na počítanie relevancie dokumentu voči dotazu. Pre každý termín v súbore sa vypočíta pravdepodobnosť, že sa termín objaví v relevantnom dokumente. Pre termíny, ktoré vyhovujú funkcii porovnania (zhody) medzi dotazom a dokumentom, sa vypočíta miera podobnosti ako kombinácia pravdepodobnosti každého z vhodných, zhodujúcich sa termínov.

Pri riešení pravdepodobnostného modelu sa historicky vyvinuli dva prístupy.

Pôvodný prístup sformulovali Maron a Kuhns. Tento prístup aplikuje vzorce využívania na predvídanie relevancie. Druhý prístup využíva každý termín v dotaze ako náznak, či je dokument relevantný. Najvýznamnejšími autormi druhého prístupu sú Steven Robertson a Karen Sparck Jonesová. Viac informácií o histórii informačného prieskumu je v publikáciách ako Steinerová 1996, Grossman Frieder 2004, Ingwersen, Järvelin 2005 a i.

Väčšina riešení pri pravdepodobnostnom prieskume vychádza z toho, že sa počíta váha termínov na základe frekvencie ich výskytu v relevantných a nerelevantných dokumentoch.

Pravdepodobnostných stratégií existuje niekoľko. Váhy jednoduchých termínov sú založené na využívaní pravdepodobnostného princípu triedenia – rankingu (PRP – probability ranking principle). Dokumenty sa zoradujú na základe odhadu pravdepodobnosti ich relevancie k dotazu. Dôležitá je prítomnosť alebo neprítomnosť termínu v dokumente. Ďalšie riešenie pravdepodobnostného modelu je odhad pravdepodobnosti, že sa termín v dokumente vyskytne v relevantnom dokumente (nebinárny nezávislý model). Pri počítaní relevancie sa berie do úvahy veľkosť a dĺžka dokumentu.

Predstavme si napríklad súbor desiatich dokumentov, v ktorých jeden dokument obsahuje termín „červená“ raz a dokument dva obsahuje tento termín v desiatich výskytoch. Obidva dokumenty 1 a 2 sú relevantné a osem ďalších je nerelevantných. V jednoduchom modeli váh termínov je vypočítaná relevancia 0.2, pretože „červená“ sa objavuje v 2 z 10 relevantných dokumentov.

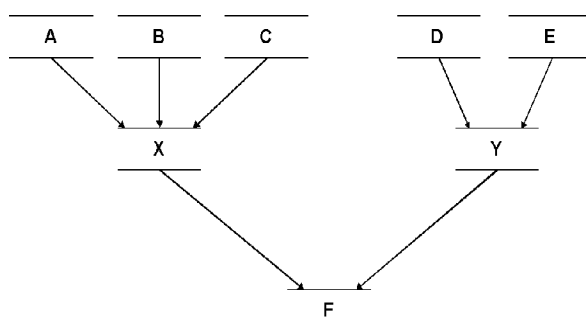
Pri nebinárnom nezávislom modeli sa počíta osobitná pravdepodobnosť pre frekvenciu každého termínu. Termín sa môže vyskytovať v jednom dokumente raz. V tom prípade je pravdepodobnosť 0.1, pretože termín sa vyskytol v dokumente 1 raz. Termín sa môže v jednom z desiatich dokumentov vyskytnúť aj desaťkrát. V tom prípade sa osobitne vypočíta aj pravdepodobnosť výskytu termínu desaťkrát, pretože sa termín vyskytol desaťkrát v jednom z desiatich dokumentov.

Pravdepodobnostné modely zvyčajne fungujú na základe *odhadu parametrov a predpokladoch nezávislosti*. Pracuje sa so súbormi údajov, ktoré umožňujú prostredníctvom testov odhadnúť vhodné parametre (Grossman, Frieder 2004).

V stratégiách a algoritmoch sa používajú aj inferenčné siete, neurónové siete, jazykové modely, sémantické siete, genetické algoritmy a tezaury.

*Inferenčné siete* sa používajú pri pravdepodobnostnom modeli. Ide o mechanizmus používaný na vyvodenie relevancie dokumentu vo vzťahu k dotazu. Nazýva sa aj *Bayesiánska sieť*. V dokumente sa hľadá „dôkaz“, na základe ktorého sa vyvodzuje relevancia. Vyvodenie relevancie sa tu používa ako koeficient podobnosti. Základným princípom je určenie známych vzťahov a ich využitie na vyvodenie ďalších vzťahov.

Príklad jednoduchej inferenčnej siete je na obrázku 2.8. Predpokladáme, že udalosti A, B, C spôsobujú udalosť X a udalosti D a E spôsobujú udalosť Y. Na základe toho môžeme z udalostí X a Y vyvodit' udalosť F. V prieskume sa to aplikuje na inferenčnú sieť dokument – termín – dotaz. V troch vrstvách sa modeluje sieť dokumentov, sieť termínov/pojmov a sieť dotazov. Prítom pri porovnávačej funkcii je dôležité, že zhoda sa nehľadá priamo s termínmi v dotaze, ale dokumenty sa musia zhodovať s pojmami, ktoré reprezentujú tieto termíny. Na základe toho vzniká relevantné triedenie.



Obr. 2.8 Inferenčná sieť (podľa Grossman, Frieder 2004)

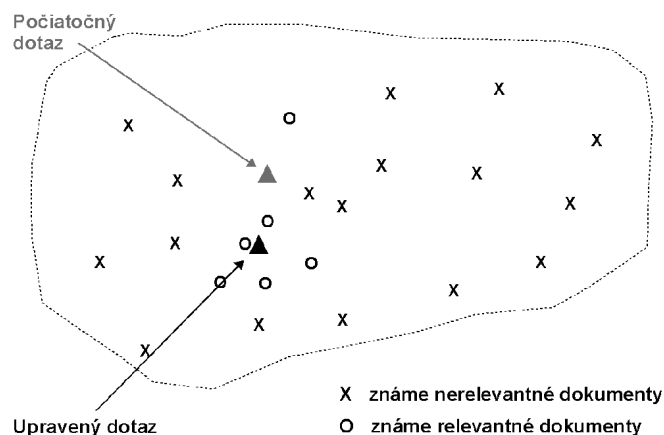
#### *Jazykové modelovanie*

Na analýzu textu a jeho možných slovných sekvencií sa používajú štatistické jazykové modely. Je to typ pravdepodobnostných modelov. Úspešne sa použili najmä pri rozpoznávaní reči.

#### *Neurónové siete*

Neurónové siete sa používajú na implementáciu vektorového modelu aj pravdepodobnostného modelu. Neurónové siete sa skladajú z uzlov a väzieb. V uzloch sa nachádzajú vstupné a výstupné hodnoty. K väzbám sa priradujú váhy, počítajú sa hodnoty vstupov a výstupov. V sieťach sú vrstvy, ktoré aktivujú dáta v rôznych fázach. Podľa odzvy skúšobných dát sa modifikujú hodnoty väzieb, čím sa neurónové siete učia. Váhy dokumentových uzlov určujú mieru relevancie dokumentu k danému dotazu. Tak sa vytvára aj algoritmus *relevantnej spätnej väzby* ako základ fungovania prieskumových systémov.

**Relevantná spätná väzba** sa využíva aj pri vektorom modeli aj pri pravdepodobnostnom modeli. Jej podstata spočíva v tom, že používateľ upravuje dotaz vo viacnásobných prieskumoch. Používateľ označuje, ktoré dokumenty sú relevantné a na základe tohto výberu sa pridávajú ďalšie termíny a dotaz sa rozširuje (expanded query). Alternatívne riešenie predpokladá, že najvyššie radené dokumenty sú relevantné a pôvodný dotaz sa modifikuje automaticky. Prvýkrát bola relevantná spätná väzba definovaná v systéme SMART G. Saltona. Algoritmus relevantnej spätnej väzby spracoval Rocchio. Príklad je na obrázku 2.9. (Manning et al. 2009).



Obr. 2.9 Posun vektoru dotazu ako reakcia na spätnú väzbu, aplikácia Rocchioho algoritmu

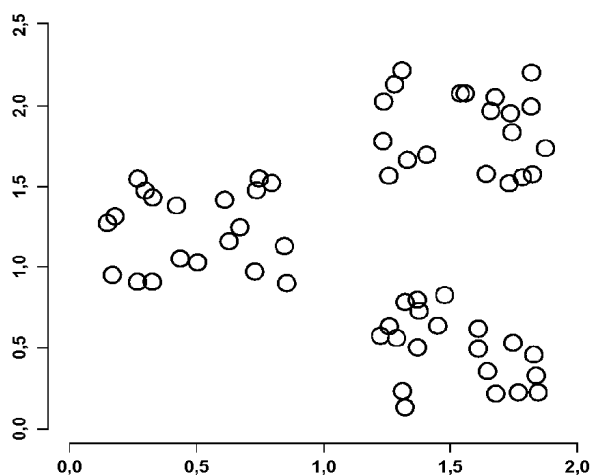
#### 2.4.1.4

### Stratégie zhlukového modelu

Zhlukovanie dokumentov zužuje priestor vyhľadávania tým, že sa podobné dokumenty prepájajú do skupín (zhlukov). Napríklad dokumenty možno prepojiť do zhlukov podľa témy – lekárske, právnické, finančné. Pri vyhľadávaní sa dotazy usmerňujú (automaticky alebo manuálne) práve do tematicky relevantného zhluku.

Existujú rôzne algoritmy zhlukovania – napríklad hierarchické, matricové a pod. Známý Rocchiov zhlukovací algoritmus z roku 1971 uplatňoval v systéme SMART aj zhlukovanie. Počítal s hustotou dokumentov v zhlukoch a s centrálnym dokumentom v zhluku.

Príklad zhlukovania je na obrázku 2.10.



Obr. 2.10 Zhlukovanie (podľa Manning et al. 2009)

Ak sa navzájom podobné zhluky organizujú hierarchicky, potom sa v priebehu prieskumu vstupuje do uzlov, v ktorých sa vyhodnocuje porovnávacia funkcia. Prehľadávanie možno riadiť zhora dolu alebo zdola hore v stromovej štruktúre.

Nástroje, ktoré sa často používajú pri zabezpečovaní funkcií jednotlivých modelov, sú sémantické siete, genetické algoritmy a tezaury. Existujú aj ďalšie metódy automatického spracovania textov.

*Sémantické siete* umožňujú reprezentovať poznatky pojmami, ktoré sú navzájom prepojené rôznymi vzťahmi. Graficky je to množina uzlov a hrán. Hrany sú nazvané podľa typov vzťahov. Informácie o danom uzle (napríklad vlastnosti ako *farba*, *veľkosť* a pod.) sa uchovávajú v štruktúre nazývanej **rámec**. Vzniká tak rámcová reprezentácia poznania, ktorú definoval Marvin Minsky. Jednoduchý príklad rámca: *ruža – má farbu / červená, má výšku / 50 cm, je / kvet*. Sémantické siete rozpracoval najmä Roger Schank, ale aj iní autori. Snažia sa riešiť problém, keď termíny v dotaze nezodpovedajú termínom v dokumentoch, aj keď sú relevantné. Meria sa aj *sémantická vzdialenosť* medzi termínmi. Pri vyhľadávaní je to výhodné pri hľadaní slovných spojení a operátorov blízkosti. Mnoho pojmov je však ťažké reprezentovať, najmä ak sa týkajú časových a priestorových aspektov pojmov.

**Genetické algoritmy** vznikli na základe princípu evolúcie a dedičstva. Princípy genetickej selekcie sú aplikované do oblasti riešenia problémov (vytvorenie populácie, pokračovanie v ďalších generáciách). Neurónové siete a genetické algoritmy sa používali najmä na strojové učenie, neskôr na radenie dokumentov (ranking).

Latentné sémantické indexovanie umožňuje riešiť problém, keď sú viaceré termíny významovo rovnaké, ale nemajú rovnakú syntax. Odhad vlastností „latentnej sémantiky“ je určený v maticiach termínov. Modely neostých množín umožňujú rozšíriť dotazy na relevantné rady.

Novšie systémy využívajú zložitejšie metódy automatického lingvistického spracovania textov. Príklady systémov sú WordNet a Cyc, v ktorých sa modeluje reprezentácia poznania v sofistikovaných sémantických sieťach, rámcoch a tezauroch.

WordNet je systém obsahujúci lexikálnu databázu anglického jazyka. Princípom sú kognitívne synonymá vytvorené z podstatných mien, prídavných mien, prísloviak a slovies (tzv. synsets). Tieto pojmy sú pospájané sémantickými vzťahmi. Výslednú sieť si možno prezerat' prostredníctvom prehliadača a umožňuje navigáciu pri vyhľadávaní.

Cyc je projekt, ktorý využíva nástroje umelej inteligencie na modelovanie každodenného informačného správania s cieľom vytvoriť zložitú ontológiu. Ontológia by mala byť súčasťou bázy znalostí s cieľom priblížiť sa myšlienkovým procesom človeka. Projekt sa rozvíja od roku 1984 pod vedením Douglasa Lenata. Niektoré časti projektu sú sprístupnené ako OpenCyc

#### 2.4.1.5

#### **Tezaurus**

Tezaurus je riadený slovník termínov a v informačnom prieskume má za úlohu riešiť problémy sémantiky. Riadený slovník je súbor pojmov usporiadaných do určitej štruktúry s určenými vzťahmi medzi nimi.

Definuje najmä vzťahy hierarchie, ekvivalencie (synonymie, homonymie) a asociácie. Základom tezauru je deskriptorový odstavec, ktorý najčastejšie určuje pre pojem širší pojem (Broader term, BT), užší pojem (Narrower term, NT) alebo príbuzný pojem (related term, RT). Asociatívne vzťahy určujú aj také vzťahy ako časť-celok, antonymia (protiklady), definícia.

Tezaurus sa v informačnom prieskume používa od 60. rokov na rozširovanie dotazu. Pôvodné termíny z dotazu sa vyhľadávajú v tezaure a dotaz sa rozšíri o pojmy obsiahnuté v deskriptore tezauru (nadradené, podradené, synonymá, asociatívne pojmy). Napríklad ak je v dotaze termín „*pes*“, prostredníctvom tezauru sa doraz rozšíri na všetky druhy psov (*psovité šelmy*).

Tezaurusy sa vytvárajú buď intelektuálne alebo automatizovane. Vytvárajú sa matice podobnosti termínov (termín-termín) a matica termín-dokument. Tezaurusy dobre fungujú v tematicky ohraničených doménach (disciplínach) s ustálenou terminológiou, ako napríklad fyzika (INSPEC Thesaurus), medicína (systém Medline, tezaurus MeSH), poľnohospodárstvo (EUROVOC), letectvo (NASA Thesaurus for Aeronautics).

Príklad klasického tezauru je na obrázku 2.11. Príklad automatickej tvorby tezauru je na obrázku 2.12.

CANTEENS BT Catering areas	COOKING APPLIANCES BT Catering equipment BT Domestic appliances NT Cookers NT Microwave ovens	Kitchen Equipment USE CATERING EQUIPMENT
CATERING NT Catering operations RT Catering areas RT Catering equipment RT Catering personnel RT Food RT Meals	Cooks USE CHEFS	KITCHENS BT Catering areas
CATERING AREAS NT Bars (licensed) NT Canteens NT Kitchens NT Restaurants RT Catering	DAIRY PRODUCTS BT Food	LICENSING LAWS RT Bars (licensed)
CATERING EQUIPMENT SN Includes food mixers, cooking ware, tableware, etc. UF Kitchen equipment NT Cooking appliances RT Catering RT Dishwashers	DIET RT Menus	MANAGEMENT NT Catering management
	DISHWASHERS RT Catering equipment	MANAGERS NT Catering managers
	DOMESTIC APPLIANCES NT Cooking appliances	MEALS RT Catering RT Food dishes RT Menus
	Drinks USE BEVERAGES	MEAT BT Food
	Fast food USE CONVENIENCE FOOD	MENUS RT Diet RT Meals

Obr. 2.11 Príklad tezauru (podľa Aitchison, Gilchrist, Bawden 1987 – 1997)

Slovo „tezaurus” pôvodne z gréčtiny znamená „poklad”. Už v 19. storočí vznikol Thesaurus of English Words and Phrases ako slovník organizovaný podľa významu. V informačnej vede vytvoril základ pre tezaury Mortimer Taube ako systém termínov (unitermov) v roku 1951. Unitermy obsahovali záznamy pre každý termín. Každý záznam unitermu obsahuje jedno slovo a čísla spojení s ďalšími slovami. Vytvoril sa tak základ na americký abecedný prístup k budovaniu tezaurov a invertovaných indexov prieskumových systémov.

Pri tvorbe tezaurov možno rozlíšiť v histórii dve línie. Prvú predstavuje americký pragmatickejší prístup (predmetové alfabeticke indexovanie vytvorené Charlesom A. Cutterom a zdokonaľovanie unitermov M. Taubeho). Druhá predstavuje európska línia založená na Ranganathanovej fazetovej klasifikačnej teórii z 30. rokov 20. stor. Tu sa kategórie využili na organizáciu pojmov v doméne.

V 60. rokoch 20. storočia sa začali rozvíjať súbory autorít ako abecedné zoznamy slov, pričom každé slovo obsahovalo zoznam súvisiacich slov.

Význam tezauru možno vidieť v dvoch pohľadoch. 1. Štruktúrny pohľad zdôrazňuje dynamiku sémanticky prepojených termínov a vzťahy reprezentujúce špecifickú doménu. 2. Funkčný pohľad zdôrazňuje funkciu kontroly terminológie a použitia tezauru na preklad z prirodzeného jazyka dokumentov, indexátorov a používateľov do informačného jazyka.

Tezaurus tak možno definovať ako špecializovaný normalizovaný postkoordinovaný jazyk používaný na dokumentačné účely, v ktorom sú lingvistické jednotky syntakticky a sémanticky prepojené.

Norma ANSI/NISO Z 39.19-1993 definuje tezaurus ako riadený slovník usporiadaný v poradí a štruktúre tak, aby boli jasne zobrazené a identifikované vzťahy ekvivalencie, homonymie, hierarchie a asociácie. Najnovší vývoj naznačuje zdokonaľovanie tezaurov v tzv. inteligentných a pojmových tezauroch, ktoré zlepšujú reprezentáciu obsahu dokumentov prostredníctvom inteligentných technológií. Umožňujú spracovanie a využívanie vzťahov medzi pojmi v hypertextových prepojeniach alebo iných reprezentáciách poznania. Známe medzinárodné tezaury sú napr. tezaurus ERIC pre oblasť vzdelávania alebo tezaurus v umení a architektúre Art and Architecture Thesaurus (AAT). Nevýhodou tezaurov je ekonomická náročnosť na ich vypracovanie a údržbu.



word	ten nearest neighbors
absolutely	absurd whatsoever totally exactly nothing
bottomed	dip copper drops topped slide trimmed slig
captivating	shimmer stunningly superbly plucky witty
doghouse	dog porch crawling beside downstairs gaze
Makeup	repellent lotion glossy sunscreen Skin gel p
mediating	reconciliation negotiate cease conciliation p
keeping	hoping bring wiping could some would othe
lithographs	drawings Picasso Dali sculptures Gauguin
pathogens	toxins bacteria organisms bacterial parasite
senses	grasp psyche truly clumsy naive innate awl

Obr. 2.12 Príklad automatického vytvárania tezauru (Manning et al. 2009)

**Najvýznamnejšie prieskumové stratégie** používané v algoritmických modeloch sú Booleov model, vektorový priestorový prístup, zhukový prístup a pravdepodobnostná stratégia. V nich sa uplatňujú sémantické siete, tezaury, neurónové siete, generické algoritmy aj algoritmy lingvistických analýz. Podrobnejšie informácie k týmto stratégiám si možno nájsť napr. v Grossman Frieder 2004, Manning, Raghavan, Schütze 2008 a v prácach ďalších autorov z oblasti informačného prieskumu (Ingwersen a Järvelin, Robertson, Sparck Jonesová, Meadow et al., Baeza-Yates a Ribeiro-Neto, Willet, Smeaton, van Rijsbergen, Sanderson, Tang, Spinková), umelej inteligencie (napr. Minsky, Lenat) a digitálnych knižníc (Borgmanová, Fox, Marchionini a i.).

## 2.4.2

### Používateľský prístup k prieskumovým stratégiám

#### 2.4.2.1

##### Postup pri vyhľadávaní

Z hľadiska postupu pri vyhľadávaní sú najčastejšie prieskumové stratégie tieto:

- Stavanie blokov
- Postupné lámanie
- Citačné/perličkové reťazenie

*Stavanie blokov (stavebné kamene)* znamená, že problém sa rozčlení na čiastkové problémy. Z toho formulujeme viac dotazov. Jednotlivé výsledky pre jednotlivé bloky sa nakoniec pospájajú do konečnej odpovede.

Napríklad *Aplikácia autorského práva do digitálnych knižníc a elektronického publikovania*

Kľúčové pojmy: 1. *Autorské právo*; 2. *Digitálne knižnice*; 3. *Elektronické publikovanie*

Dotazy: digitálne knižnice AND autorské právo

autorské právo AND elektronické publikovanie

Výsledná odpoveď: spojenie výsledkov dvoch dotazov

*Postupné lámanie (osekávanie)* je stratégia prieskumu, v rámci ktorej sa z rozsiahleho súboru dokumentov postupne odstraňujú jednotlivé prvky. Dotaz sa zužuje. Pritom sa využívajú napríklad užšie termíny, logický súčin, negácia.

Napríklad *Vplyv informačného správania na komunikáciu vedeckých informácií v rôznych kultúrach*

Dotaz: informačné správanie AND vedecká komunikácia AND kultúrne rozdiely

Ak získame ako odozvu príliš veľa dokumentov, odstránime napríklad operátorom NOT jeden typ dokumentu – NOT knihy

*Citačné reťazenie (rastúca perla)* je stratégia, ktorá sa používa aj pri štúdiu dokumentov. Jej podstata spočíva v tom, že sa nájde jeden vysoko relevantný dokument, ktorý obsahuje odkazy (citácie) na ďalšie dokumenty. V informačnom prieskume to znamená, že ak sa nájde jeden relevantný dokument, na ďalšie vyhľadávanie sa použijú termíny použité na indexovanie pôvodného relevantného dokumentu. Proces možno viackrát opakovať.

Napríklad: Na dotaz využívanie informácií bol nájdený dokument Kari, J. (2007). Conceptualizing the personal outcomes of information. *Information Research*, 12(2) paper 292 [Available at <http://InformationR.net/ir/12-2/paper292.html>]

Na ďalšie vyhľadávanie sa použije termín „information outcomes“ vybraný ako kľúčové slovo z dokumentu, alebo sa sledujú spojenia citovaných dokumentov, napr. odkaz na dokument Taylor, Robert S. (1991). Information use environments. *Progress in Communication Sciences*, X, 217-255.

#### *Priorita najšpecifickejšej fazety*

Vyhľadávanie sa začína určením najužšieho kľúčového pojmu. Na základe odozvy sa prieskum ďalej rozširuje na príbuzné kľúčové pojmy.

Dotaz autorské právo AND elektronické publikovanie

Ak je odpoveďou na tento doraz príliš málo dokumentov, rozšíri sa téma všeobecnejšími pojmami, napr. Autorstvo AND publikovanie

#### *Priorita najnižšieho počtu odkazov*

Prieskum sa začína pojmom, ktorý má najmenší počet odkazov a končí prvým záznamom, ktorý je prijateľný.

Dotaz: „Informačná ekológia“

Odpoveď: Davenport, Thomas H., Prusak, Lawrence. 1997. *Information Ecology*. Mastering the Information and Knowledge Environment. New York: Oxford University Press, 1997. 255 s. ISBN 0-19-511168-0.

#### 2.4.2.2

### **Prax vyhľadávania**

Z hľadiska praxe vyhľadávania najznámejšie informačné stratégie vo vyhľadávaní zvyčajne kombinujú tieto postupy:

#### *Prezeranie (browsing)*

Vychádza z myšlienkovej činnosti pri zrakovom vnímaní vizuálnych podnetov. V hypertextových prepojeniach sa sledujú súvislosti. Prezeranie patrí k exploratívne a intuitívne vyhľadávaniu, pretože umožňuje sledovať hypertextové odkazy. Tým je možné aj náhodné získavanie informácií a objavovanie nových súvislostí. Prezeranie je založené na schopnosti vnímať podnety a pohybovať sa v informačnom priestore.

Prezeranie bližšie definovala Marcia Batesová (Bates 2008). Prezeranie je podľa nej aktivita zapojenia sa do zaznamenania objektov možného záujmu. Podľa toho sa človek rozhoduje, či bližšie preskúma daný objekt (fyzické alebo iné reprezentácie), príjme ho alebo odmietne. Prezeranie je kognitívny a behaviorálny výraz prieskumného správania založený na zvedavosti. Obsahuje štyri kroky: 1) zaznamenanie scény, 2) vizuálne alebo fyzické priblíženie objektu, 3) bližšie preskúmanie objektov, 4) fyzické alebo pojmové nadobudnutie skúmaného objektu.

#### *Monitorovanie*

Monitorovanie je stratégiou sledovania určitej tematickej oblasti, prípadne vybraných zdrojov. V odbornej oblasti experti často sledujú niekoľko vybraných odborných časopisov, ktoré sa zaoberajú ich oblasťou záujmu. Monitorovanie informačného prostredia môže byť cieľovo riadené, ale aj náhodné a špecificky nezamerané. Autori zistili, že až 80 % poznatkov človek absorbuje svojou vedomou prítomnosťou v informačnom prostredí tým, že vníma podnety a monitoruje sociálny kontext a fyzické prostredie.

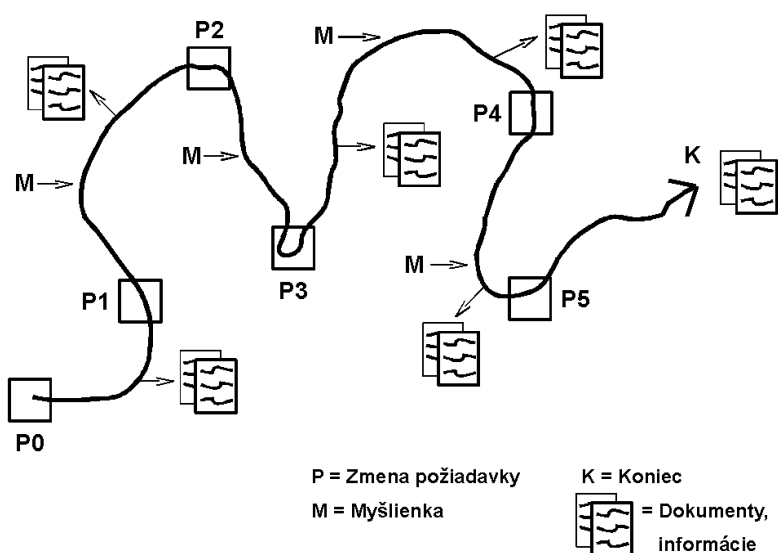
### Krátky prieskum

Krátky prieskum znamená, že sa prepoja dve kľúčové slová logickými operátormi (Booleov model).

### Berrypicking/zber lesných plodov

Postupné objavovanie a skúmanie informačného priestoru podľa modelu Marcie Batesovej. Model platí najmä pre vyhľadávanie v elektronickom prostredí. Kombinuje sa vyhľadávanie, prezera- nie, monitorovanie a rôzne iné stratégie. Z rôznych zdrojov a vyhľadávačov sa postupne skladá mozaika informácií. Postupuje sa nelineárne, kľukatou cestou. Dotaz sa mení pod vplyvom nových in- formácií. Cesta nie je priamočiara. V internete a pri plnotextovom vyhľadávaní získavame často množstvo (státisíce) výsledkov, ktoré sú nadbytočné (redundantné). Pri konštrukcii významu sa tu preto na ceste vyhľadávania môže objaviť aj veľa slepých uličiek. Model berrypicking podľa M. Bate- sovej je znázornený na obrázku 2.13.

Človek tu využíva viac informačných stratégií a vždy pod vplyvom relevantných informácií pre- formuluje dotaz, pretože vidí tému v novom svetle. Metafora zbierania čučoriedok v lese výstižne charakterizuje pohyb človeka krížom-krážom v informačnom prostredí, keď sa zastaví pri každom zdroji (kríku), ktoré sú roztrúsené v informačnom prostredí. Z jedného zdroja si vyberie kúsok infor- mácií, pri inom sa zastaví a odloží informáciu a postupne mení stav poznania.



Obr. 2.13 Model vyhľadávania informácií M. Batesovej – berrypicking/zbieranie lesných plodov

#### 2.4.2.3

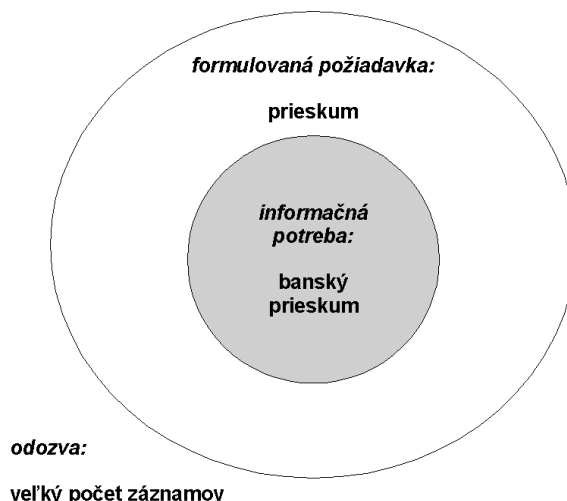
### Zameranie vyhľadávania a predchádzajúce poznatky

Z hľadiska zamerania vyhľadávania a predchádzajúcich poznatkov sa rozlišujú tieto prieskumové stratégie:

#### Všeobecné vyhľadávanie

Používateľ vie o danej téme málo informácií a formuluje dotaz jednoduchým termínom. Vyberá si všeobecné zdroje na orientáciu, najmä všeobecné vyhľadávače na internete (napr. Google). Problémom môže byť príliš veľká odozva pri vyhľadávaní, napríklad veľa odkazov na rôzne stránky. Často chýba kontext hľadaného termínu.

Graficky si to možno znázorniť na obrázku 2.14.

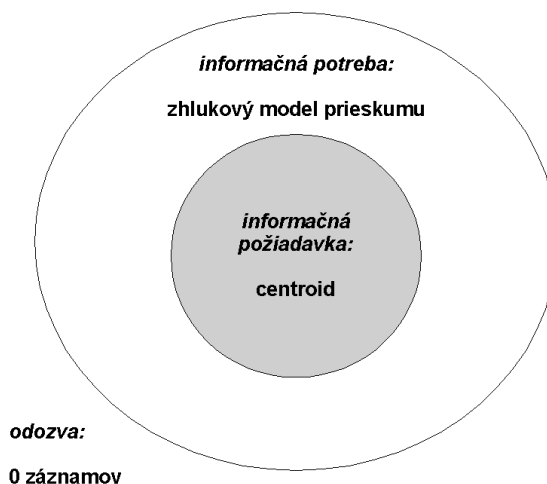


Obr. 2.14 Všeobecné vyhľadávanie, široká formulácia dotazu

### *Špecifické vyhľadávanie*

Používateľ pozná tému veľmi dobre. Vyberá si konkrétne zdroje, o ktorých vie, že v nich môže nájsť požadované informácie (napríklad špecifické predmetové portály). Možno hovoriť o vyhľadávaní v doméne – určitej problémovej alebo tematickej oblasti. Presnosť vyhľadávania môže byť vyššia. Môže sa však stať, že k špecifickému dotazu sa nenájde žiadna informácia. V tomto prípade všeobecné vyhľadávače nie sú vhodné. Google neumožňuje spresňovať dotaz v kontexte určitej domény a nepodporuje používateľa pri konštrukcii dotazu na základe nejasej informačnej potreby.

Graficky to znázorňuje obrázok 2.15.



Obr. 2.15 Špecifické vyhľadávanie, úzka formulácia dotazu

### *Postupné vyhľadávanie (Inkrementálne vyhľadávanie)*

Používateľ vyhľadáva informácie v jednotlivých krokoch. Postupne si skladá odpoveď na dotaz kombináciou viacerých zdrojov. Dotazy sa viackrát opakujú, často sú súčasťou celého cyklu vyhľadávania.

Napríklad *Zákony o slobodnom prístupe k informáciám v západnej Európe za posledných 10 rokov*  
 Kľúčové slová: *slobodný prístup k informáciám, západná Európa, 1999 – 2009, zákony.*

### *Kategoriálne vyhľadávanie*

Ide vlastne o prezeranie v jednotlivých kategóriách, najmä v predmetových adresároch (ako Yahoo! a pod.). Informácie sú roztriedené logicky a prehľadne, čo umožňuje znižovať kognitívnu záťaž. Kategórie tiež podporujú používateľa pri tematickej orientácii.

### *Reťazové vyhľadávanie*

Znamená používanie viacerých kľúčových slov naraz. Vyhľadávanie sa zjednodušuje. Často sa však odozva rozširuje (napríklad použitie viacerých synonym slov v rôznych jazykoch).

Napríklad `maliny+raspberries+Himbeeren`

### *Dvojkrokové vyhľadávanie*

#### *Vyhľadaj a skoč*

Ak záleží na rýchlosti a presnosti vyhľadaných informácií, použije sa táto stratégia na určenie pravdepodobne najrelevantnejších záznamov vo vyhľadanom súbore.

#### *Vyhľadaj a zorad'*

Pri vyhľadávaní s odozva automaticky zoradí podľa najrelevantnejších záznamov. Tu záleží na čase odozvy. V reálnej praxi sa tieto stratégie najčastejšie navzájom kombinujú.

Stratégie vyhľadávania sú často spojené s konkrétnym vyhľadávacím nástrojom.

Napríklad Google umožňuje jednoduché vyhľadávanie, implicitné vyhľadávanie (už pri formulácii dotazu je implicitne včlenený operátor AND), vyhľadávanie podľa polí (určí sa text, ktorý sa má nachádzať v názve, v adrese, v texte stránky, napr.: `allintitle:TEXT`, `allinurl:TEXT`, `allintext:TEXT`).

**Google** umožňuje vyhľadávať aj frázy (formulované v úvodzovkách) a rôzne manipulácie s termínmi, napríklad:

- fráza „president of the United States“
- vylúčenie termínov: znak „-“: „univerzita -fakulta“
- zabránenie vylúčeniu všeobecných slov: znak „+“: „know+how“
- nahradenie rôznych synonym: znak „~“: „~ food“.

Osobitné nároky na vyhľadávanie sa spájajú s vyhľadávaním obrázkov, máp a videozáznamov.

Pri vyhľadávaní vedeckých informácií sa v praxi často používa okrem špecializovaných databáz alebo katalógov vedeckých a akademických knižníc aj špecializovaný vyhľadávač **GoogleScholar**. Umožňuje vyhľadávať podľa autora, názvu, odborov, druhov časopisov a článkov, pritom články sa radia podľa relevancie. Dotaz sa formuluje prostredníctvom naznačenia poľa (aspektu, užšieho kritéria), v ktorom sa vyhľadáva. Napríklad formulácia dotazu na vyhľadávanie autora: `autor: „a einstein“`.

Trendy vyhľadávania informácií sa prejavujú aj v nových vyhľadávačoch a nových firemných spoluprákach. Microsoft a Yahoo! sa spájajú a snažia sa vytvoriť konkurenčný vyhľadávač voči Google – Bing ([www.bing.com](http://www.bing.com)). Vždy sa objavujú nové nástroje, ktoré sa stávajú „módou“. Preto je potrebné sledovať zmeny a nové funkcie a princípy vyhľadávania.

## 2.5

### **Ako efektívne vyhľadávať informácie – princípy vyhľadávania**

Každá situácia vyhľadávania informácií je jedinečná a vnorená v kontexte (téma, problém, úloha, projekt). Napriek tomu možno sformulovať niekoľko odporúčaní na vyhľadávanie informácií.

1. Pred vyhľadávaním sa snažíme pochopiť tému a určiť základné pojmy, ktorými možno tému alebo problém vyjadriť.
2. Určí sa možný informačný horizont, najmä vzdialené a blízke in-

- formačné zdroje. 3. Selektujú sa informačné zdroje a určí sa postup pri ich využití na vyhľadávanie (informačná cesta). 4. Vytvorí sa predbežná orientácia v pojmoch témy a predpokladaná predstava o možnej odozve. 5. Vyberú sa možné prieskumové stratégie podľa informačnej požiadavky v kontexte. 6. Sformuluje sa informačná požiadavka a vyjadrí sa v kľúčových slovách, prípadne iných termínoch. 7. Podľa odozvy sa viacnásobne prezerá informačný priestor. 8. Neustále sa vyhodnocuje odozva a monitoruje informačný priestor. 9. Kombinujú sa prieskumové stratégie. 10. Odporúča sa postupovať od všeobecnej prieskumovej stratégie k špecifickej, od orientácie v problematike k hlbšej analýze informácií.
2. Vyhľadávanie informácií súvisí úzko aj s našou informačnou gramotnosťou. Opiera sa aj o štádiá riešenia informačného problému. Model šiestich štádií pri riešení informačného problému (Eisenberg, Berkowitz 2009) určuje tento postup. 1. štádium: *definovanie úlohy*, obsahuje 1.1 definovanie informačného problému, 1.2 identifikovanie informačnej potreby, 2. štádium: *určenie informačných stratégií* vyhľadávania, obsahuje 2.1 určenie všetkých možných zdrojov, 2.2 výber najvhodnejších zdrojov, 3. štádium: *určenie lokácie a prístupu*, obsahuje 3.1 lokalizáciu zdrojov (fyzicky, intelektuálne), 3.2 nájdenie informácií v zdrojoch, 4. štádium: *využitie informácií*, obsahuje 4.1 aktivity (čítanie, písanie, počúvanie, pozeranie, dotyk a i.), 4.2 extrakcia relevantných informácií, 5. štádium: *syntéza*, obsahuje 5.1 organizácia rôznych zdrojov, 5.2 prezentácia informácií, 6. štádium: *hodnotenie*, obsahuje 6.1 posúdenie produktu (efektívnosť), 6.2 hodnotenie procesu (účinnosť).

## 2.6 Zhrnutie

Informačný horizont definujeme ako súbor informačných zdrojov, ktoré má človek k dispozícii pri riešení informačného problému. Najčastejšie obsahuje prioritné preferované zdroje, zdroje sekundárneho významu a periférne zdroje.

Informačná stratégia predstavuje prístup k riešeniu informačného problému. Súvisí s fázami riešenia problémov (orientačno-analytická, strategicko-operačná, synteticko-overovacia). Prieskumová stratégia v širšom zmysle predstavuje plánovanie procesu vyhľadávania informácií, je voľbou krokov pri riešení informačného problému. Prieskumová stratégia v užšom zmysle je súborom algoritmov používaných systémom pri vyhľadávaní informácií a opiera sa o vnútorný model informačného prieskumu.

Prieskumové stratégie z hľadiska systémového sa delia na stratégiu Booleovho prieskumu, stratégiu pravdepodobnostného prieskumu, stratégiu vektorového priestorového modelu, stratégiu zhlukového prieskumu.

Z používateľského hľadiska sa rozlišujú prieskumové stratégie podľa a) postupnosti vyhľadávania (stavanie blokov, postupné lámanie, perličkové reťazenie, priorita najšpecifickejšej fazety, priorita najnižšieho počtu odkazov), podľa b) praxe pri vyhľadávaní (prezeranie, krátky prieskum, monitorovanie, zber „lesných plodov“) a podľa c) zamerania a predchádzajúcich poznatkov (všeobecné, špecifické, postupné, kategoriálne, reťazové, dvojkrokové vyhľadávanie). Princípy vyhľadávania obsahujú postupnosť krokov od definovania problému, pochopenia témy a orientácie v pojmoch až po kombinácie stratégií, analýzu a hodnotenie informácií.

## Literatúra

ANSI/NISO Z 39.19-1993. Usmernenia pre stavbu, formát a manažment viacjazyčných tezaurov.

Aitchison, J., Gilchrist, A., Bawden, D. 1987 – 1997. *Thesaurus construction : a practical manual*. London : Aslib. S. 160.

BATES, Marcia J. 2007. What is browsing – really? : A model drawing from behavioural science research. In *Information Research* [online]. October 2007, vol. 12, no. 4, paper 330 [cit. 2010-01-11]. Dostupné na internete: <<http://InformationR.net/ir/12-4/paper330.html>>.

EISENBERG, Michael – BERKOWITZ, Bob. 2009. *The Big6™ Skills* [online]. Information and Technology Skills for Student Achievement [cit. 2010-01-11]. Dostupné na internete: <<http://www.big6.com/>>.

GROSSMAN, David A., FRIEDER, Ophir. 2004. *Information Retrieval : Algorithms and Heuristics*. Sec. ed. Dordrecht : Springer, 2004. 332 s. ISBN 1-4020-3004-5.

LOWE, Carrie A., EISENBERG, Michael B. 2005. Big6™ Skills for Information Literacy. In Eds. FISHER, Karen E., ERDELEZ, Sanda, McKECHNIE, Lynne. *Theories of Information Behavior*. Medford : Information Today, 2005, s. 63 – 68. MANNING, Christopher D., RAGHAVAN, Prabhakan, SCHÜTZE, Hinrich. 2009. *Introduction to Information Retrieval* [online]. Cambridge University Press, 2008 [cit. 2010-01-11]. Dostupné na internete: <<http://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf>>.

RUISEL, I. 2004. *Inteligencia a myslenie*. Bratislava : Ikar, 2004. 432 s. ISBN 80-551-0766-1.

SAVOLAINEN, Reijo. 2008. *Everyday Information Practices : A Social Phenomenological Perspective*. Lanham (Maryland) : The Scarecrow Press, 2008. 233 s. ISBN 978-0-8108-6111-4.

SONNENWALD, Diane H., WILDEMUTH, Barbara M., HARMON, Gary L. 2001. A research method to investigate information seeking using the concept of information horizons : an example from a study of lower socio-economic students' information seeking behaviour. In *The New Review of Information Behaviour Research*. 2001, vol. 2, s. 65 – 86.

STEINEROVÁ, Jela. 1996. *Teória informačného prieskumu*. Bratislava : CVTI SR, 1996. 262 s. ISBN 80-65165-58-9.

STEINEROVÁ, Jela. 2005a. *Informačné správanie : Pohľady informačnej vedy*. Bratislava : CVTI SR, 2005. 189 s. ISBN 80-85165-90-2.

WILSON, T. D. 1999. Models in information behaviour research. In *Journal of Documentation* [online]. 1999, vol. 55, no. 3 [cit. 2003-09-30]. Dostupné na internete: <<http://informationr.net/tdw/publ/papers/1999JDoc.html>>. ISSN 0022-0418.

## Zoznam webových sídel

Bing  
[www.bing.com](http://www.bing.com)

Google  
[www.google.sk](http://www.google.sk)  
[www.google.com](http://www.google.com)

Google Scholar  
<http://scholar.google.sk/>  
<http://scholar.google.com/>

OpenCyc  
<http://www.opencyc.org/doc>

WordNet  
<http://wordnet.princeton.edu>

# 3

## Vyhľadávanie v elektronických informačných zdrojoch

Elektronické informačné zdroje (ďalej aj EIZ) sú informačné zdroje uchovávané a sprístupňované v elektronickej forme, obvykle sú dostupné prostredníctvom rôznych technológií distribúcie elektronických dát ako napr. CD-ROM, DVD a iné, v prostredí lokálnych počítačových sietí a internetu. V súčasnosti sa používateľ obvykle rozhoduje, či si informácie nájde na internete prostredníctvom rôznych druhov vyhľadávačov alebo uprednostní vyhľadávanie v databázach či databázových centrách, kde sú informácie zhromaždené na jednom mieste a sú určitým spôsobom organizované a spravované. Jednotlivé databázy majú zvyčajne k dispozícii sofistikovaný systém vyhľadávania, tezaury, fazetové triedenie či rôzne možnosti na ďalšiu prácu s vyľadanými informáciami.

Myšlienka využitia prostriedkov výpočtovej techniky na účely vyhľadávania relevantných informácií sa dostala do popredia po publikovaní článku Vannevara Busha „As we may think“ v roku 1945. Výpočtová a telekomunikačná technika priniesla knižniciam a iným informačným inštitúciám možnosť automatizovaného prístupu k veľkému objemu informácií nielen v rámci vlastných informačných zdrojov (knižničné katalógy), ale aj v externých bázach dát v databázových centrách. Prvé databázové centrá začali vznikať v 50. a 60. rokoch 20. storočia, napr. TYMNET, DATEX-P, NEXTEL, OCLC. Začiatkom 70. rokov minulého storočia bol do prevádzky uvedený databázový systém DIALOG spoločnosti Lockheed v Palo Alto (USA). V priebehu 80. rokov vznikali ďalšie databázové centrá, napr. CAS ONLINE na báze referátového časopisu Chemical Abstracts, Institute of Scientific Information (ISI) začal sprístupňovať databázy citačných indexov Science Citation Index a Social Science Citation Index.

Elektronické informačné zdroje možno rozdeliť podľa

- charakteru na:
  - *bibliografické* – obsahujúce bibliografické informácie o publikovaných časopiseckých článkoch, konferenčných príspevkoch, knihách a pod., často sú doplnené aj o abstrakty,
  - *faktografické* – obsahujúce konkrétne faktografické informácie, patria sem napr. štatistiky, adresáre firiem, matematické, fyzikálne či chemické tabuľky,
  - *plnotextové* – obsahujúce pôvodné dokumenty v úplnom texte,
- typu obsiahnutých informácií na:
  - *primárne* – pôvodné informácie a dáta (napr. faktografické alebo plnotextové databázy, firmená literatúra a pod.),
  - *sekundárne* – informačné zdroje popisujúce primárne databázy (napr. bibliografické databázy, knižničné katalógy, a pod.),
  - *terciárne* – zdroje, ktoré obsahujú predovšetkým informácie o existencii sekundárnych informačných zdrojov (napr. databáza katalógov, webové portály a pod.),
- dostupnosti na:
  - *voľne dostupné*,



- *licencované* – produkty renomovaných vydavateľstiev a iných producentov informácií dostupné na základe predplatného, či inej dohody určujúcej časové obmedzenie prístupnosti, resp. iné podmienky ako napr. oprávnených používateľov, počet konkurenčných prístupov a pod. V porovnaní s voľne dostupnými zdrojmi u licencovaných databáz sa predpokladá vyššia dôveryhodnosť, kvalita, presnosť a aktuálnosť informácií.

Elektronické informačné zdroje sú pre používateľov výhodné najmä z hľadiska rýchlosti a pohodlnosti prístupu. Na rozdiel od tlačенých materiálov umiestnených napr. v knižnici, nie sú viazané na určité fyzické miesto ani otváracie hodiny. Navyše informácie v elektronickej forme môžu byť používateľom poskytované na základe ich individuálnych potrieb a preferencií, keďže väčšina databáz má k dispozícii aj rôzne personalizačné nástroje. Efektívne využívanie elektronických informačných zdrojov kladie na používateľov určité nároky, najmä na ich informačnú gramotnosť, vyhľadávacie schopnosti a zručnosti a aspoň základnú znalosť angličtiny či iných svetových jazykov. Informačne gramotný používateľ by mal byť schopný potrebné informácie nielen objaviť a získať, ale aj kriticky vyhodnotiť.

### 3.1 Informačné zdroje Web of Knowledge a SCOPUS

Databázové produkty na platforme *ISI Web of Knowledge* (<http://isiknowledge.com>) spoločnosti Thomson Reuters (nástupca ISI – Institute for Scientific Information) a databáza *SCOPUS* (<http://www.scopus.com>) spoločnosti Elsevier, B. V. patria k najvýznamnejším informačným zdrojom v oblasti vedeckých informácií. Oba informačné zdroje sú licencované a sprístupňujú predovšetkým bibliografické, citačné a abstraktové informácie vysokej kvalitatívnej úrovne a nástroje na analýzu a manažment vyhládaných informácií. Vo veľkej miere sa využívajú na hodnotenie vedy a výskumu na individuálnej, inštitucionálnej i regionálnej úrovni. Databázy možno využiť aj na analýzu trendov a predikciu vedeckého vývoja, skúmanie vzorcov vedeckej komunikácie, mapovanie štruktúry vedy, jej zmien a pod.

Licencované databázové produkty v rámci Web of Knowledge (ďalej aj WoK) sú:

- **Web of Science (WoS)**, obsahuje:
  - 3 časopisecké citačné indexy: SCI – *Science Citation Index*, SSCI – *Social Science Citation Index*, AHCI – *Arts and Humanities Citation Index*,
  - 2 citačné indexy z konferenčných zborníkov: CPCI-S – *Conference Proceedings Citation Index – Science*, CPCI-SSH – *Conference Proceedings Citation Index – Social Sciences & Humanities*,
  - 2 chemické databázy: IC – *Index Chemicus*, CCR-Expanded – *Current Chemical Reactions*,
- **Current Contents Connect (CCC)**:
  - Poskytuje prístup ku kompletným bibliografickým informáciám z viac ako 8 000 svetových vedeckých časopisov (tzv. karentované časopisy) a z viac ako 2 000 kníh. Databáza ponúka aj prístup k vedeckým webovým stránkam, obsahuje približne 10 615 000 záznamov (údaj k októbru 2009). Je denne aktualizovaná.
  - Obsahuje 7 edícií: ABES – *Agriculture, Biology & Environmental Science*; SBS – *Social & Behavioral Sciences*; CM – *Clinical Medicine*; LS – *Life Sciences*; PCES – *Physical, Chemical&Earth Sciences*; ECT – *Engineering, Computing&Technology*; AH – *Arts & Humanities*,
  - Ďalej obsahuje 2 kolekcie: BC – *Business Collection*, EC – *Electronics&Telecommunications Collection*.

Zistiť, či je zdrojový dokument karentovaný, možno aj priamym vyhľadávaním – je potrebné vybrať položku „Publication Name“ a zadať štandardizovaný skrátený názov daného časopisu, resp. slovo z názvu. Ak systém nenájde žiadne záznamy, je pravdepodobné, že časopis karentovaný nie je.

- **Journal Citation Reports (JCR):**
  - obsahuje 2 edície: Science, Social Sciences,
  - nástroj bibliometrických alebo scientometrických metód, slúži predovšetkým na hodnotenie vedeckých časopisov a následne na hodnotenie autorov v nich publikujúcich. V rámci tejto databázy možno okrem kompletných vydavateľských údajov časopisu zistiť aj rôzne indikátory výkonnosti časopisov – Impact Factor (faktor dopadu/vplyvu) či Immediacy Index (index bezprostrednej odozvy) a ďalšie. V rámci JCR možno sledovať aj zmeny názvov časopisov.
- **Essential Science Indicators (ESI),**
  - ESI predstavuje jedinečný a komplexný nástroj zameraný na hodnotenie vedy. Hlavnými indikátormi výstupov, resp. produktivity sú publikované časopisecké články. Na meranie významu a dopadu ESI používa jednak celkové počty citácií, ako aj priemerné počty citácií na jednu prácu. Celkový počet citácií indikuje celkový význam práce vedca, priemer citácií na článok ukazuje vážený význam/dopad, nazývaný aj impact. Je dôležité zdôrazniť, že ESI sleduje len indexované časopisy v rámci Thomson Scientific (nezahŕňa iné zdroje).
  - ESI obsahuje predovšetkým prehľady najcitovanejších autorov/vedcov, hodnotenia inštitúcií (univerzity, korporácie, vládne výskumné laboratóriá a pod.), hodnotenia krajín a hodnotenia časopisov. Databáza obsahuje stručné návody k analýzám dát, vysvetlivky, tabuľky, grafy. Unikátnou črtou je zoznam výskumných oblastí nazývaný Research Fronts, ktorý obsahuje algoritmicky odvodené oblasti súčasnej vedy, v rámci ktorých je výskum najintenzívnejší. Údaje sú aktualizované každé 4 mesiace.
- **Biological Abstracts,**
- **BIOSIS Previews,**
- **Derwent Innovations Index,**
- **CABI: CAB Abstracts and Global Health,**
- **Food Science and Technology Abstracts,**
- **Inspec,**
- **MEDLINE,**
- **Zoological Records.**

Jednotlivé databázové produkty sú licencované samostatne a autorizovaným používateľom [1] sa v rámci WoK zobrazia len tie, ku ktorým majú predplatený prístup [2].

Databáza SCOPUS bola sprístupnená v roku 2004, jej vývoj však prebiehal už od roku 2002. SCOPUS, na rozdiel od databázových produktov spoločnosti Thomson Reuters zameraných predovšetkým na anglosaskú literatúru (najmä z proveniencie USA a Veľkej Británie), si dal za cieľ pokryť aj iné geografické oblasti, najmä Európu a Áziu. Databáza je budovaná v spolupráci s 21 inštitúciami z celého sveta a je testovaná viac ako 300 vedcami. V súčasnosti SCOPUS obsahuje 21 miliónov patentov (z patentových úradov: United States Patent and Trademark Office, European Patent Office, Japan Patent Office, World Intellectual Property Organization, United Kingdom Intellectual Property Office), 286 miliónov posudzovaných webových stránok (prehľadávané prostredníctvom vyhľadávača Scirus). SCOPUS je obsahovo založený na nasledovných databázach: *Medline*, *Compendex*, *Embase*, *Geobase*, *Bio-base* a publikáciách z vydavateľstva Elsevier. Zaujímavosťou tohto zdroja je, že tu možno vyhľadať aj články v tlači.

### 3.1.1 Spôsoby vyhľadávania

Informačné zdroje na platforme WoK a databáza SCOPUS ponúkajú obdobné spôsoby vyhľadávania. K dispozícii je *jednoduché vyhľadávanie* (obrázok 3.1 a obrázok 3.6) podľa autora, jeho pracoviska, názvu článku, zdrojového dokumentu, roku vydania, predmetu a pod., pričom jednotlivé termíny možno kombinovať prostredníctvom logických operátorov – AND (a zároveň – vo vyhľadaných záznamoch sa musia nachádzať všetky termíny), OR (alebo – záznamy musia obsahovať aspoň jeden z daných termínov), NOT (okrem – vyhľadané záznamy musia obsahovať prvý zadaný termín, ale nesmú obsahovať druhý). SCOPUS v rámci jednoduchého vyhľadávania ponúka možnosť kombinácie 2 rôznych termínov, WoK dáva možnosť kombinácie až 25 termínov. *Pokročilé vyhľadávanie* (obrázok 3.5 a obrázok 3.13) umožňuje používateľom zadať vyhľadávací dotaz prostredníctvom špecifického selekčného jazyka založeného na pravidlách booleovskej algebry. Je možné vyhľadávať aj v zoznamoch zdrojových dokumentov.

Vyhľadávanie v rámci databázových produktov WoK je menej kompaktné v porovnaní s databázou SCOPUS najmä preto, že WoK obsahuje niekoľko rôznorodých databáz a k požadovanej informácii sa dá dostať rôznym spôsobom.

V rámci WoK (vyhľadávací nástroj *ScientificWebPlus*) i databázy SCOPUS (vyhľadávač *Scirus*) je možné paralelne prehľadávať aj kredibilné vedecké webové stránky (prehľadávať sa najmä webové sídla vedeckých a akademických inštitúcií). SCOPUS navyše automaticky prehľadáva aj niekoľko patentových databáz.

Pri formulácii vyhľadávacieho dotazu sú k dispozícii nasledovné znaky rozšírenia, symboly pre dopĺňanie, nahrádzanie znakov (truncation):

\* (hviezdička) – náhrada za žiaden až ľubovoľný počet znakov

(napr.: „gene\*“ → gene, genetics, generation...),

? (otáznik) – náhrada za presne jeden ľubovoľný znak na konkrétnom mieste slova

(napr.: „en?oblast“ → endoblast, entoblast),

§ (dolár) – náhrada za žiaden alebo presne jeden znak na konkrétnom mieste – k dispozícii len v rámci WoK (napr.: „colo§r“ → color, colour)

Znaky rozšírenia môžu byť použité uprostred alebo na konci slova, nemali by sa uvádzať na začiatku slova.

Databázy ďalej ponúkajú:

- vyhľadanie určitej frázy, slovného spojenia:

WoK – výraz sa musí zapísať do *úvodzoviek*, napr.: „global warming“.

SCOPUS rozlišuje medzi presným a voľným vyhľadávaním. Pri presnom vyhľadávaní je nutné vyhľadávané termíny uviesť do množinových zátvoriek, napr.: {laser technology}. V tomto prípade sa hľadá exaktný tvar a zohľadňujú sa všetky znaky, vrátane znakov rozšírenia, ktoré sú však brané ako bežné znaky. Napr. výsledok pre {laser-technology} bude iný ako vo vyššie uvedenom príklade, pretože sa bude hľadať aj zhoda pre znak „-“. Pri hľadaní {laser technology?} sa vyhľadajú opytovacie vety s uvedeným termínom. Voľné vyhľadávanie fráz vyhľadá aj odlišné tvary frázy. Fráza sa v tomto prípade uvádza v úvodzovkách, napr. "laser technology". Interpunkcia (napr. „-“) je ignorovaná. Je možné využiť aj funkciu znakov rozšírenia a hľadá sa aj množné číslo termínu. Vyhľadávané termíny možno zadávať aj s diakritikou.

- využitie operátorov blízkosti:

WoK – pomocou operátora SAME sa vyhľadajú iba také záznamy, ktoré budú dané termíny obsahovať v rámci jednej vety nezávisle od ich poradia (pod vetou sa vo všeobecnosti rozumie akýkoľvek text oddelený bodkou), napr.: „psychology SAME child“.

SCOPUS – W/ „číslo“, PRE/ „číslo“, napr.: zadaním „Edison W/10 inventor“ možno vyhľadať dokumenty, kde sú uvedené termíny od seba vzdialené maximálne 10 slov; výrazom

„Edison PRE/3 inventor“ možno vyhľadať dokumenty, kde sú uvedené termíny v danom poradí od seba vzdialené maximálne 3 slová. Logický operátor OR a operátory blízkosti nie je možné navzájom kombinovať (napr. výberový dotaz „Edison W/10 inventor OR businessman“ by nefungoval).

The screenshot shows the SCOPUS search interface. The 'Basic Search' tab is active. The search bar contains 'Search for:' with a dropdown menu open showing various search fields like 'Article Title, Abstract, Keywords', 'All Fields', 'Authors', 'First Author', 'Source Title', 'Article Title', 'Abstract', 'Keywords', 'Affiliation', 'Language', 'ISSN', 'CODEN', 'DOI', 'References', 'Conference', 'Article Title, Abstract, Keywords, Authors', 'Phy. Chemical Name', and 'SOC. CAS Number'. The 'Limit to' section includes 'Date Range' (inclusive) with options for 'Published' (All years to Present) and 'Added to Scopus in the last 7 days'. The 'Subject Areas' section includes 'Life Sciences (> 3,400 titles)' and 'Health Sciences (> 5,300 titles)'. The 'AND' operator is selected in the dropdown menu.

Obr. 3.1 SCOPUS – možnosti vyhľadávania

The screenshot shows the SCOPUS search results page. The results are displayed in a table with columns: Document (sort by relevance), Author(s), Date, Source Title, and Cited By. The first result is 'The Quadrantid meteor shower 1997-2004: Activity and mass distribution' by Pupillo, G., Porubčan, V., Cevolani, G., Bortolotti, G., Franceschi, G., Trivellone, G., Zigo, P., Hajduk, A. The second result is 'Summer streams of the Taurid meteor complex' by Porubčan, V., Zigo, P., Pecina, P. The 'Cited By' column shows 2 citations for the first result and 0 for the second. The 'Display' section shows 20 results per page.

**Annotations:**

- Document (sort by relevance):** Výber záznamu zaškrtnutím
- Abstract + Refs:** Zobrazenie zdrojového dokumentu
- View at Publisher:** Zobrazenie podrobnejších informácií, ak má inštitúcia predplatený prístup k danému časopisu - možnosť zobrazenia úplného textu
- Cited By:** Počet citácií
- Display:** Definovanie počtu zobrazovaných záznamov na stránku (dolná časť obrazovky s výsledkami)

Obr. 3.2 SCOPUS – výsledky vyhľadávania

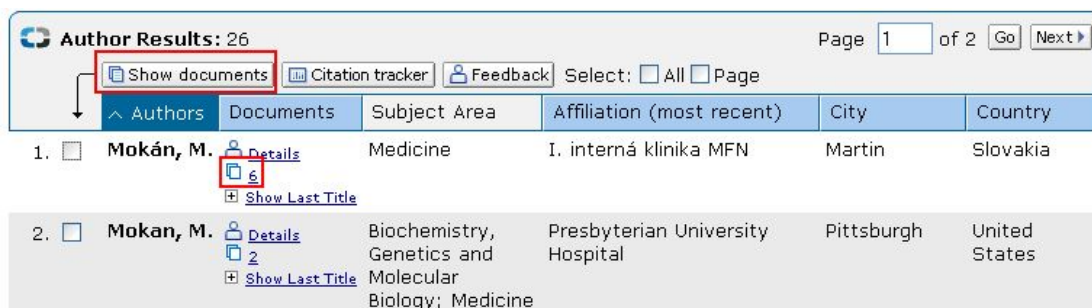
Kliknutím na tlačidlo „Sources“ je možné prehliadať abecedný zoznam oblastí a zdrojových dokumentov, ktoré SCOPUS indexuje. Zdrojové dokumenty je tu tiež možné vyhľadávať podľa názvu, čísla ISSN alebo vydavateľa.



Obr. 3.3 SCOPUS – prehľad spracovávaných zdrojových dokumentov

Vzhľadom na to, že mená autorov môžu mať rôznu formu zápisu, SCOPUS ponúka špecifické vyhľadávanie autorov („Author Search“). Jeho výsledkom je prehľad rôznych foriem hľadaného mena spolu s ďalšími informáciami ako napr.: oblasť publikovania, či inštitúcia, ktoré slúžia na presnejšiu identifikáciu autora. Následne možno zobraziť dokumenty, kde sa meno autora vyskytuje vo vybranej forme kliknutím na počet dokumentov pri mene alebo vybraním viacerých foriem a kliknutím na „Show documents“. Vyhľadávanie autorov je dostupné aj zo základného vyhľadávania zvolením položky „Author“.

Prostredníctvom odkazu „Details“ v rámci výsledkov vyhľadávania autora systém poskytne kontextové informácie o danom autorovi (pracovisko, počet publikácií, počet citácií, h-index, počet spoluautorov, atď.).



Obr. 3.4 SCOPUS – výsledky vyhľadávania podľa autora

Vyhľadávanie inštitúcií („Affiliation Search“) umožňuje nájsť rôzne formy názvu inštitúcií a dokumenty, ktorých autori pracovali/pracujú v týchto inštitúciách. S vyhľadávaním inštitúcií sa pracuje podobne ako s vyhľadávaním autorov. Pri nájdenom názve sú zobrazené aj ďalšie alternatívne formy. K danému názvu možno ešte dodatočne vyhľadať ďalšie formy kliknutím na „Find unmatched affiliations“. Vyhľadávanie inštitúcií je dostupné aj zo základného vyhľadávania zvolením položky „Affiliation“.

**Search for:** ? Search Tips | [View list of all codes](#)

AUTHLASTNAME (Zigo) AND AFFIL (COMEN\*)

**Add to search:** [Author name](#) Search Clear

**Operators**

AND  
OR  
AND NOT  
PRE/  
W/  
**Codes**

ABS  
AFFIL  
AFFILCITY  
AFFILCOUNTRY  
AFFILORG  
ALL  
ARTNUM

As you type Scopus offers code suggestions. Double click or press "enter" to add to advanced search.

**Code:** AFFIL  
**Name:** Affiliation

When searching the AFFIL field, you can specify if you want all of your search terms to be found in the same affiliation.

AFFIL is a combined field that searches the following author address fields: AFFILCITY, AFFILCOUNTRY, and AFFILORG.

[more info](#)

**Advanced search examples:**  
ALL("heart attack") AND AUTHOR-NAME(smith)  
TITLE-ABS-KEY(\*somatic complaint wom?n ) AND PUBYEAR AFT 1993

**Výber položky vyhľadávania, po kliknutí sa v pravej časti zobrazí vysvetlivka k danej položke, dvojklikom sa výber presunie do vyhľadávacieho okna, kde je potrebné následne doplniť termín**

Obr. 3.5 SCOPUS – pokročilé vyhľadávanie



ISI Web of Knowledge<sup>SM</sup> Take the next step

All Databases **Select a Database** Web of Science Additional Resources

Search Search History Marked List (0)

**ALL DATABASES** **Výber databázy** **Položky/typy vyhľadávania**

Search for:

Example: oil spill\* mediterranean in Title

Example: O'Brian C\* OR OBrian C\* in Topic

Example: Cancer\* OR Journal of Cancer Research and Clinical Oncology in Author

AND AND OR NOT

Booleovské operátory

Add Another Field >>

Možnosť pridať ďalšie vyhľadávania (max. 25)

Pomôcky => abecedné registre

Search Clear

Limit to: Latest 5 years

Časové obmedzenie vyhľadávania

Obr. 3.6 WoK – vyhľadávanie

**Publication Name Index** **Vyhľadávacia pomôcka - register zdrojových dokumentov**

Use the Browse and Find features to locate Publication Name titles to add to your query.

Click on a letter or type a few letters from the beginning of the name to browse alphabetically by title.  
Example: plant to jump to entries beginning with PLANT

biologia **Move To** 0 9 A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Enter text to find titles containing the text.  
Example: \*botan\* to find BOTANICAL REVIEW and GEOBOTANY

**Find**

Kliknutím na vybrané písmeno sa zobrazí abecedný zoznam zdrojových dokumentov

Možnosť presunúť sa/vyhľadať zadaný časopis

Records to Add to Query

3455	Add	BIOLOGIA
1	Add	BIOLOGIA 10 ANOS DE PE

Page Range: B 1 LYMPHOCYTES IN B CELL NEOPLAS LEPIDOPTERAN

Records to Add to Query

32	Add	B 1 LYMPHOCYTES IN B CELL I
33	Add	B 2007 FORMAL SPECIFICATIO
32	Add	B 2007 FORMAL SPECIFICATIO

Obr. 3.7 WoK – vyhľadávacie pomôcky (abecedný prehľad zdrojových dokumentov)

Výsledky sú implicitne zoradené podľa dátumu vydania. Je možné ich zoradiť aj podľa ďalších kritérií (počet citácií, relevancia, prvý autor, zdrojový dokument, rok vydania).

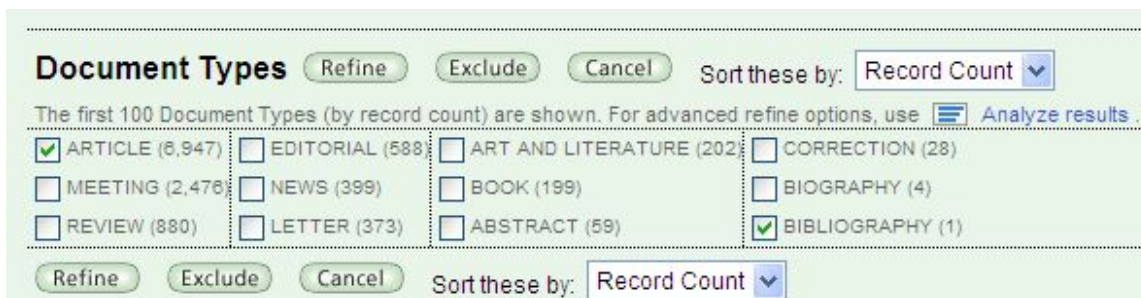
Obr. 3.8 WoK – výsledky vyhľadávania

Výsledky vyhľadávania možno ďalej zužovať, ladit' („Refine Results“) na základe rôznych kritérií (predmet, autor, rok vydania, jazyk dokumentu a pod.). Ladenie vyhľadávania umožňuje obmedziť vyhľadanie na vybrané termíny („refine“, resp. „limit to“), či naopak umožňuje vybrané termíny z množiny výsledkov vylúčiť („exclude“).

<p><b>Subject Areas</b> <span>Refine</span></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> ENVIRONMENTAL SCIENCES &amp; ECOLOGY (3,848)</li> <li><input type="checkbox"/> METEOROLOGY &amp; ATMOSPHERIC SCIENCES (3,391)</li> <li><input type="checkbox"/> ENGINEERING (1,558)</li> <li><input type="checkbox"/> GEOLOGY (1,187)</li> <li><input type="checkbox"/> AGRICULTURE (1,117)</li> </ul> <p>more options / values...</p>	<p>→ špecifikovanie predmetnej oblasti („Subject Areas“, napr.: strojárstvo, geológia, medicína, atď...), zvyčajne sa na stránke zobrazia 3 predmety, viac možno zobrazit' kliknutím na odkaz „more options / values“)</p>
<p><b>Document Types</b></p> <p><b>Authors</b> <span>Refine</span></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> CHAPIN, FS (32)</li> <li><input type="checkbox"/> PEARCE, F (32)</li> </ul>	<p>→ výber určitých autorov („Authors“)</p>

Obr. 3.9 WoK – ladenie vyhľadávania





Obr. 3.10 Ladenie vyhľadávania

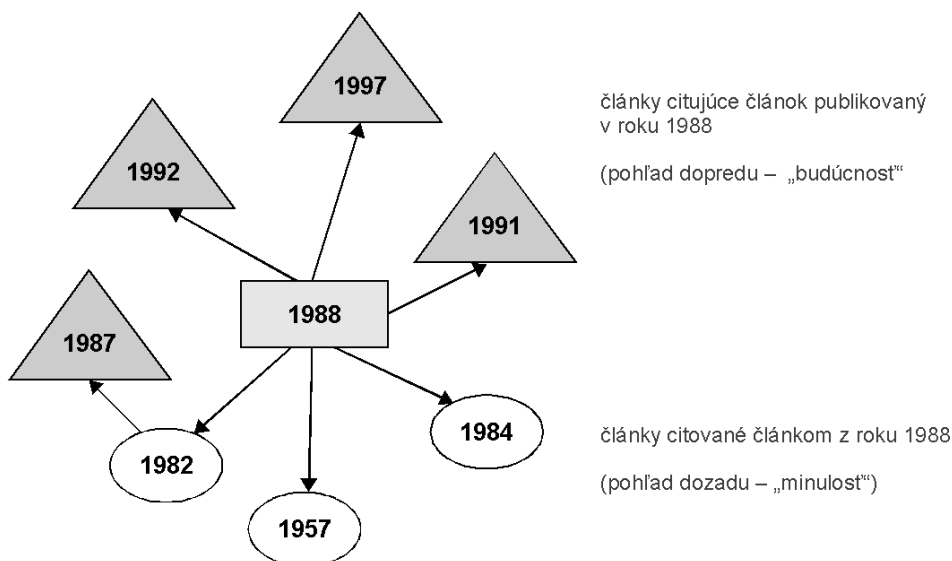


Obr. 3.11 SCOPUS – doladenie vyhľadávania

### 3.1.2

## Vyhľadávanie ohlasov a citácií

Informačné zdroje WoK a SCOPUS sú pre používateľov užitočné najmä vzhľadom na možnosť vyhľadania citácií a ohlasov na publikované práce. Sledovanie citácií poskytuje pohľad v čase dopredu i dozadu. Mnohí vedci a autori sa pri vyhľadávaní sústreďujú na referencie a citácie, ktoré autor použil pri písaní článku, t. j. čie články citoval. Často si však neuvedomujú, že by mohlo byť dôležité či aspoň zaujímavé pozrieť sa v čase aj dopredu a zistiť, kým bol daný článok citovaný. Sledovanie citácií umožňuje získať pohľad na význam práce vedcov, inštitúcií, krajín. Poskytuje informácie o vedcoch, či inštitúciách zaoberajúcich sa rovnakou problematikou a tým otvára priestor na potenciálnu spoluprácu, umožňuje sledovanie práce kolegov i konkurentov. Identifikuje publikácie vzťahujúce sa k rovnakej, resp. podobnej téme, či významné, rozhodujúce a prelomové práce v danej oblasti. Mapuje vznik určitej myšlienky a jej vývoj v čase.



Obr. 3.12 Sledovanie citácií (podľa Bartošek 2000)

V SCOPUS-e sú informácie o citáciách integrované, WoK ponúka špecializovanú databázu zameranú na sledovanie citácií – *Web of Science* (ďalej aj WoS), ktorá je týždenne aktualizovaná. Oba zdroje umožňujú nastavenie automatického zasielania notifikácií o nových citáciách vybraných článkov prostredníctvom elektronickej pošty.

All Databases | Select a Database | Web of Science | Additional Resources

Search | Cited Reference Search | **Advanced Search** | Search History | Marked List (0)

**Web of Science® – now with Conference Proceedings**

**Advanced Search.** Use 2-character tags, Boolean operators, parentheses, and set references to create your query. Results appear in the Search History at the bottom of the page.

Example: TS=(nanotub\* SAME carbon) NOT AU=Smalley RE #1 NOT #2 [more examples](#) | [view the tutorial](#)

**AU=Devinsky AND OG=COMEN\* AND PY=200\***

**Search**

Current Limits: [Change Limits and Settings](#)  
 Timespan=All Years.  
 Databases=Science Citation Index Expanded (SCI-EXPANDED); Social Sciences Citation Index (SSCI); Arts & Humanities Citation Index (A&HCI); Conference Proceedings Citation Index- Science (CPCI-S); Conference Proceedings Citation Index- Social Science & Humanities (CPCI-SSH)

Field Tags	Booleans
TS=Topic	AND
TI=Title	OR
AU=Author	NOT
GP=Group Author	SAME
ED=Editor	
SO=Publication Name	
PY=Year Published	
CF=Conference	
AD=Address	
OG=Organization	
SG=Suborganization	
SA=Street Address	
CI=City	
PS=Province/State	
CU=Country	
ZP=Zip/Postal Code	
FO=Funding Agency	
FG=Grant Number	
FT=Funding Text	

Obr. 3.13 WoS – pokročilé vyhľadávanie

## Search History

Set	Results	Combine Sets	Delete Sets
#8	74 AU=Devinsky AND OG=COMEN* Databases=SCI-EXPANDED, SSCI, A&HCI Timespan=All Years	<input type="radio"/> AND <input type="radio"/> OR <input type="button" value="Combine"/>	<input type="button" value="Select All"/> <input type="button" value="Delete"/>

Obr. 3.14 WoS – pokročilé vyhľadávanie – zobrazenie výsledkov

Vyhľadávanie citácií vo WoS sa realizuje v rámci „Cited Reference Search“. Systém ponúka možnosť vyhľadať citácie prác podľa mena citovaného autora („Cited Author“) a názvu citovaného zdrojového dokumentu („Cited Work“ – potrebné zadať slovo zo štandardizovanej skratky názvu časopisu – pozri „journal abbreviation list“). Vyhľadávanie podľa názvu citovaného článku nie je k dispozícii. Vyhľadávanie možno ďalej obmedziť rokom publikovania citovaného článku („Cited Year(s)“).

Vyhľadávanie citácií v prvom kroku vráti zoznam záznamov (citovaných prác) spĺňajúcich dané kritériá vyhľadávania. V stručnom prehľade sa zobrazia autorské údaje a skrátené názvy zdrojových dokumentov/citovaných prác (úplné názvy možno zobrazit kliknutím na „SHOW EXPANDED TITLES“), rok publikovania citovaných prác, umiestnenie a počet citujúcich prác. Ak sa pred menom autora nachádzajú tri bodky (...) znamená to, že ide o spoluautora práce (nie prvého autora). Vyhľadávanie citácií sekundárnych autorov je limitované na predplatenú retrospektívu.

Príklad vyhľadávania článkov citujúcich autora „Zigo P“.

**Cited Reference Search.** Find the articles that cite a person's work

**Step 2 of 2:** Select cited references and click "Finish Search."

Select the references for which you wish to see the citing articles, then click the "Finish Search" button.

Hint: Look for cited reference variants (sometimes different pages of the same article are cited or papers are cited incorrectly).

CITED REFERENCE INDEX Počet citovaných dokumentov

References: 1 - 7 of 7 Page 1 of 1 Go

Select Page Select All\* Clear All Finish Search

Select	Cited Author	Cited Work [SHOW EXPANDED TITLES]	Year	Volume	Page	Article ID	Citing Articles **	View Record
<input type="checkbox"/>	ZIGO P	1982 CESK ON K	1983		267		1	
<input type="checkbox"/>	ZIGO P	FAC PHILOS U COMENIA	1996				1	
<input type="checkbox"/>	ZIGO P	LOVINOBANA	1987				1	
<input type="checkbox"/>	...Zigo P	NUOVO CIMENTO C	2004	27	213	DOI 10.1393/ncc/i2004-10015-1	1	<a href="#">View Record</a>
<input type="checkbox"/>	...Zigo P	PLANET SPACE SCI	1995	43	747		2	<a href="#">View Record</a>

Pokiaľ sa citujúci dokument nachádza vo WoK, je možné ho zobrazit kliknutím na "View Record", ak sa nenachádza táto funkcia nie je aktívna.

Obr. 3.15 Vyhľadané citácie (krok 1)

V prípade, že citované práce sú obsiahnuté vo Web of Science, systém ponúkne odkaz na zobrazenie týchto záznamov („View Record“).

## THE BOLOGNA-MODRA FORWARD-SCATTER RADAR EXPERIMENT - REFLECTION PROPERTIES OF METEOR TRAINS

Full Text Print E-mail Add to Marked List Save to EndNote Web more options

**Author(s):** HAJDUK A, HAJDUKOVA M, GAJDOS S, KOSTECKY P, ZIGO P, CEVOLANI G, GRASSI G, TRIVELLONE G

**Source:** PLANETARY AND SPACE SCIENCE **Volume:** 43 **Issue:** 6 **Pages:** 747-749 **Published:** 1995

**Times Cited:** 2 **References:** 5 **Počet prác, ktoré sú citované týmto článkom**

**Abstract:** Observational results on the reflection properties of meteor trains obtained by forward-scatter bistatic radar over the baseline Bologna-Modra (about 700 km), as well as some characteristics of the Lyrid shower meteors in 1993, are presented here. **Počet prác, ktoré citujú tento článok**

The observations following the shower radiant motion show clearly that the orientation of the baseline of the forward-scatter system is the dominating factor affecting the flux of the recorded meteors. Even the long-enduring meteor trains which were proved to be independent in a backscatter system of the specular reflection conditions, are observed herewith to be strongly dependent on the baseline orientation with a maximum rate corresponding to the radiants having the baseline as azimuth.

**Document Type:** Article **Link na zobrazenie záznamov citujúcich článkov**

**Language:** English **[ view all 2 citing articles ]**

**Cited by: 2**  
This article has been cited 2 times (Web of Science).  
Meisel DD, Richardson JE. Statistical properties of meteors from a simple, passive forward scatter system. PLANETARY AND SPACE SCIENCE 1-2 107-112 1973.  
Cevolani G, Gabucci FG, Hajduk A, et al. Baseline effect on the forward-scatter radar reflection from meteor trains. NUOVO CIMENTO DELLA SOCIETA ITALIANA DI FISICA C-GEOPHYSICS AND SPACE PHYSICS 3 447-450 1993.

Obr. 3.16 Zobrazenie záznamu citovaného článku

Citujúce práce možno zobraziť aj označením relevantných citovaných prác (zaškrtnutím štvorčeka pri danej položke, výberom všetkých prác zobrazených na stránke – tlačidlo „Select Page“, či výberom všetkých nájdených citovaných prác – tlačidlo „Select All“) a následným kliknutím na tlačidlo „Finish Search“.

CITED REFERENCE INDEX  
References: 1 - 7 of 7

Select Page Select All\* Clear All  
Finish Search

Select	Cited Author	Cited Work [SHOW EXPANDED TITLES]	Year
<input checked="" type="checkbox"/>	ZIGO P	1982 CESK ON K	1983
<input checked="" type="checkbox"/>	ZIGO P	FAC PHILOS U COMENIA	1996

Obr. 3.17 Vyhľadanie citácie (krok 2) – zobrazenie vybraných záznamov

Web of Science® – now with Conference Proceedings

**Results** Author=(zigo p)  
Timespan=All Years. Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH. Scientific Web Plus BETA View Web Results >>

Results: 16 Page 1 of 2 Go Sort by: Latest Date

Print E-mail Add to Marked List Save to EndNote Web Analyze Results  
Save to EndNote, RefMan, ProCite more options Create Citation Report

**Refine Results**  
Search within results for [ ] Search  
Subject Areas Refine  
 ASTRONOMY & ASTROPHYSICS (11)

1. Title: Variation of sporadic meteor activity during the 23rd cycle of solar activity  
Author(s): Porubcan V, Zigo P, Cevolani G, et al.  
Source: NUOVO CIMENTO DELLA SOCIETA ITALIANA DI FISICA B-GENERAL PHYSICS RELATIVITY ASTRONOMY AND MATHEMATICAL PHYSICS AND METHODS Volume: 124 Issue: 1 Pages: 69-75 Published: JAN 2009  
Times Cited: 0

Obr. 3.18 Vyhľadávanie citácií – zobrazenie výsledkov

Nájdené dokumenty možno podrobiť analýze kliknutím na položku „Create Citation Report“ (nad zoznamom výsledkov).



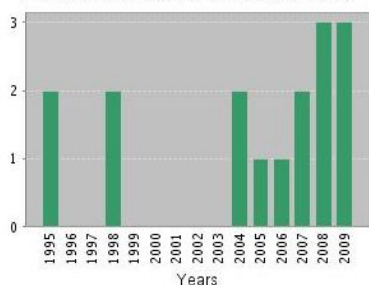
<< Back to previous results list

### Citation Report Author=(zigo p)

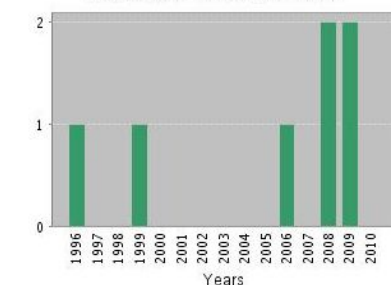
Timespan=All Years. Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH.

This report reflects citations to source items indexed within Web of Science. Perform a Cited Reference Search to include citations to items not indexed within Web of Science.

#### Published Items in Each Year



#### Citations in Each Year



Results found: 16

Sum of the Times Cited [?]: 7

[View Citing Articles](#)

[View without self-citations](#)

Average Citations per Item [?]: 0.44

h-index [?]: 2

Results: 16

Page 1 of 2 Go

Sort by: Times Cited

	2006	2007	2008	2009	2010	Total	Average Citations per Year
Use the checkboxes to remove individual items from this Citation Report or restrict to items processed between 1985 and 2010 Go	1	0	2	2	0	7	0.47
<input type="checkbox"/> 1. Title: The Geminid meteor shower of 1996-2003 from forward-scatter observations: Activity and mass distribution Author(s): Pupillo G, Porubcan V, Bortolotti G, et al. Source: RIVISTA DI SCIENZE DELLA TERRA E DELL'ATMOSFERA	0	0	1	1	0	2	0.40

Obr. 3.19 Citačná analýza

Vyhľadanie dokumentov citujúcich daného autora v SCOPUS-e je možné realizovať v rámci jednoduchého vyhľadávania („Basic Search“) – je potrebné zvoliť položku vyhľadávania „References“ a zadať meno autora.

V rámci výsledkov vyhľadávania možno kliknutím na tlačidlo „References“ zobraziť referencie (použitú literatúru) vybraných záznamov, tlačidlo „Cited by“ zobrazí záznamy dokumentov, ktoré citujú vybrané záznamy. Chronologický prehľad citácií možno získať prostredníctvom funkcie „Citation tracker“.

Output Citation tracker Add to list Download References Cited by Select:  All  Page 1 to 10

Document (sort by relevance)	Author(s)	Date	Source Title	Cited By
1. <input checked="" type="checkbox"/> On studying information seeking methodologically: The implications of connecting metatheory to method	Dervin, B.	1999	Information Processing and Management 35 (6)	55

1 Cited Documents <a href="#">save to list</a>	Citations						
	<2007	2007	2008	2009	subtotal	>2009	total
<b>Total</b>	39	13	2	1	16	0	55
1 <input type="checkbox"/> 1999 On studying information seeking ...	39	13	2	1	16		55

Display 25 Documents 1 to 1

Obr. 3.20 SCOPUS – Citation tracker

Viac možností na vyhľadanie citácií poskytuje pokročilé vyhľadanie („Advanced Search“), je možné hľadať podľa nasledujúcich kritérií:

- meno citovaného autora (REFAUTH),
- názov citovaného článku (REFTITLE),
- názov zdrojového dokumentu (REFSRCTITLE),
- rok vydania (REFPUBYEAR),
- čísla strán (REFPAGE),
- ktorékoľvek z vyššie spomenutých (REF).

Vyhľadane citačné údaje prostredníctvom vyhľadávača Scirus nie sú integrované v službe Citation tracker.

### 3.1.3 Práca s výsledkami vyhľadávania

Výsledky vyhľadávania sa dajú nielen prezerat', ale možno s nimi aj ďalej pracovať. Je možné ich triediť (zostupne i vzostupne) podľa rôznych kritérií (rok vydania, počet citácií, názov, prvý autor, zdrojový dokument), či podrobiť analýze z rôznych hľadísk (najčastejši (spolu)autori, inštitúcie, medzinárodná spolupráca, typ dokumentov, jazyk, rok vydania, zdrojové dokumenty).

Vybrané nájdené záznamy sa dajú vytlačiť, poslať elektronickou poštou, pridať do rešerše („Marked List“, „Add to list“), exportovať, uložiť na lokálny počítač v rôznych formátoch (html, jednoduchý text, a iné), spracovať cez softvér na vytváranie a správu referencií (napr. EndNote, RefWorks).

Oba informačné zdroje ponúkajú rôzne ďalšie rozšírené funkcie. K tomu je však potrebné, aby sa používateľ zaregistroval a vytvoril si vlastný personalizovaný používateľský profil. K dispozícii sú napríklad funkcie:

- uloženie uskutočnených vyhľadávanií a opätovný návrat k nim,
- nastavenie automatického zasielania *notifikácií na uskutočnené vyhľadávania* (priebežné upozornenia na nové články spĺňajúce dané kritériá vyhľadávania) a *citačných alertov* (priebežné upozornenia na nové citácie na vybrané články zasielané na zadanú e-mailovú adresu),
- vytvorenie vlastných zoznamov časopisov určených na podrobnejšie prehľadávanie a nastavenie notifikácií na zasielanie obsahov nových čísel týchto časopisov,

**Step 1. Select the fields to include in the output.** [Reset to Defaults](#) [Save as Defaults](#)

<input checked="" type="checkbox"/> Author(s)	<input checked="" type="checkbox"/> Title	<input checked="" type="checkbox"/> Source	<input type="checkbox"/> abstract*
<input type="checkbox"/> language	<input type="checkbox"/> document type	<input type="checkbox"/> keywords	<input type="checkbox"/> addresses
<input type="checkbox"/> cited references*	<input type="checkbox"/> cited reference count	<input checked="" type="checkbox"/> times cited	<input type="checkbox"/> publisher infor
<input checked="" type="checkbox"/> ISSN	<input type="checkbox"/> source abbrev.	<input type="checkbox"/> page count	<input type="checkbox"/> IDS number
<input type="checkbox"/> subject category			

\*Selecting these items will increase the processing time.

**Výber údajov, ktoré sa majú exportovať zo záznamu**

**Step 2. Select an option.** **Výber formátu zobrazenia údajov - s označením polí, bibliografický**

Field Tagged	<a href="#">Format for Print</a>	E-mail records to: ilavska@vili.uniba.sk
Field Tagged	<a href="#">Save to File</a>	Return e-mail (optional):
	<a href="#">Save to EndNote, RefMan, or other reference software</a>	Notes (optional): Resers
	<a href="#">Save to My EndNote Web</a>	Plain Text <input type="checkbox"/> E-mail <input type="checkbox"/>

Automatically delete selected records from the Marked List after output is complete.

Obr. 3.21 WoK – nastavenie formátu záznamu

## COMPACT COMBUSTION SYSTEMS USING A COMBINATION OF TRAPPED VORTEX AND HIGH-G COMBUSTOR TECHNOLOGIES

[Print](#)
[E-mail](#)
[Add to Marked List](#)
[Save to EndNote@Web](#)
[Save to EndNote@, RefMan, ProCite](#)

Author(s): Zelina J (Zelina, J.)<sup>1</sup>, Anderson W (Anderson, W.)<sup>1</sup>, Koch P (Koch, P.)<sup>1</sup>, Shouse DT

Cited by: 0

This article has been cited 0 times (from Web of Science).

[Create Citation Alert](#)

Obr. 3.22 WoK – nastavenie citačného alertu

87 records. Author=(DEVINSKY F)

Rank the records by this field:	Analyze:	Set display options:	Sort by:
Author <b>Country/Territory</b> Document Type Institution Name	upTo 500 records.	Show the top 10 Results. Minimum record count (Threshold): 2	<input checked="" type="radio"/> Record count <input type="radio"/> Selected field

Analyze

---

Use the checkboxes below to view the records.  
**Note:** The number of records displayed may be greater than the listed Record Count if the original set contained more records than the number of records analyzed.

View Records	Field: Country/Territory	Record Count	% of 87	Bar Chart	Save Analysis Data to File
<input type="checkbox"/>	SLOVAKIA	45	51.7241 %	<div style="width: 51.7241%;"></div>	
<input type="checkbox"/>	CZECHOSLOVAKIA	36	41.3793 %	<div style="width: 41.3793%;"></div>	
<input type="checkbox"/>	AUSTRIA	4	4.5977 %	<div style="width: 4.5977%;"></div>	
<input type="checkbox"/>	ENGLAND	4	4.5977 %	<div style="width: 4.5977%;"></div>	
<input type="checkbox"/>	RUSSIA	4	4.5977 %	<div style="width: 4.5977%;"></div>	
<input type="checkbox"/>	GERMANY	3	3.4483 %	<div style="width: 3.4483%;"></div>	
<input type="checkbox"/>	FRANCE	2	2.2989 %	<div style="width: 2.2989%;"></div>	
<input type="checkbox"/>	JAPAN	2	2.2989 %	<div style="width: 2.2989%;"></div>	

Analyza záznamov - hodnotenie podľa krajiny, odkiaľ pochádzajú spoluautori

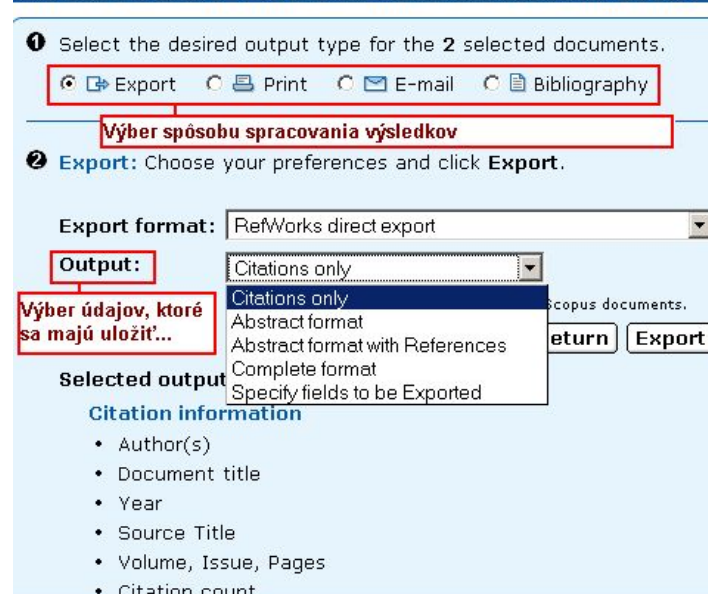
Obr. 3.23 WoK – analýza výsledkov podľa krajiny

V rámci aktuálneho prieskumu si systémy ukladajú jednotlivé vyhľadávania v rámci tzv. histórie vyhľadávania („Search History“). V rámci WoK sa história vyhľadávania zaznamenáva pre každý databázový produkt samostatne. Aktuálnu históriu vyhľadávania si systém pamätá len do ukončenia práce so systémom. Uložiť si určité vyhľadávania aj do budúcnosti je potrebné prostredníctvom „Save History“. História je možné uložiť do profilu používateľa (ak je prihlásený) alebo do súboru na lokálny počítač. WoK ponúka aj kombináciu jednotlivých množín výsledkov prostredníctvom logických operátorov.



Obr. 3.24 WoK – história vyhľadávania

### Output: Export, Print, E-mail or create a Bibliography



Obr. 3.25 SCOPUS – možnosti výstupu



Obr. 3.26 SCOPUS – rozšírené funkcie prístupné registrovaným používateľom



WoK i SCOPUS sú predovšetkým bibliografické a abstraktové informačné zdroje, ale sprostredkujú aj prístup k úplným textom, pokiaľ ich má autorizovaný používateľ predplatené u vydavateľa (WoK: tlačidlo „Full Text“, SCOPUS: „View at Publisher“).

## 3.2 Knižničné katalógy

Knižničné katalógy patria (spolu s bibliografiami, tlačеныmi registrami a pod.) k tradičným nástrojom organizácie poznania, ktorého cieľom je umožniť, aby sa informácia či poznatok dali s odstupom času opätovne vyhľadať.

Knižničné katalógy slúžia v prvom rade na poskytovanie informácií o existencii určitého dokumentu (monografie, zborníka, periodika, článku a pod.) a o jeho výskyte v rámci fondu tej-ktorej knižnice či inštitúcie. Knižničné katalógy spravidla umožňujú používateľom knižnice dané dokumenty vyhľadať a poskytujú aj potrebné informácie o ich sprístupňovaní.

Katalógy možno deliť podľa (Jurčacková 1996):

- spôsobu spracovania na:
  - tradičné lístkové katalógy,
  - strojové (automatizované katalógy),
- použitého súpisného údajja na:
  - menné,
  - predmetové,
  - systematické,
- hľadiska komplexnosti na:
  - úplné (obsahujú záznamy o všetkých dokumentoch knižnice/inštitúcie),
  - čiastkové,
  - súborné (obsahujú záznamy dokumentov z fondov viacerých knižníc).

Vplyvom použitia výpočtovej techniky vznikli tzv. online katalógy. OPAC (Online Public Access Catalog) môže integrovať viacero tradične chápaných katalógov, umožňuje vyhľadávať podľa viacerých kritérií naraz (kombinácia mena autora a predmetového hesla; paralelné multidatabázové vyhľadávanie v katalógoch viacerých knižníc, a pod.).

### 3.2.1 WorldCat

*WorldCat* (<http://www.worldcat.org>) je jednou z najväčších bibliografických databáz na svete. Vytvorili ju a udržiavajú ju knižnice podieľajúce sa na globálnej kooperácii OCLC. Databáza bola sprístupnená v roku 1971, v októbri 2009 obsahovala cca 1,4 miliárd záznamov (o tlačенých i elektronických dokumentoch) vo viac ako 360 jazykoch.

WorldCat ako jedna z najväčších svetových sietí zameriavajúcich sa na knižničný obsah a služby umožňuje:

- prehľadávať fondy rôznych knižníc vo svete a lokalizovať v nich potrebné dokumenty,
- vyhľadávať knihy, hudobniny, videá a iné typy dokumentov,
- hľadať články, elektronické knihy či zvukové dokumenty prístupné na stiahnutie,
- konzultovať prípadné problémy s knihovníkom prostredníctvom služby „Ask librarian“,
- vytvoriť si osobný profil a podeliť sa s ostatnými používateľmi o svoje názory či záujmy,
- vytvárať si zoznamy knižničných materiálov a deliť sa o ne,
- zverejňovať recenzie či hodnotenia na jednotlivé položky, prispievať k nim faktografickými informáciami či inými kontextovými údajmi.

Obr. 3.27 WorldCat – rozšírené vyhľadávanie

### 3.2.2 Projekt KIS3G

Projekt *KIS3G* (Knížnično-informačný systém tretej generácie, <http://www.kis3g.sk>) riešený Slovenskou národnou knižnicou v Martine (SNK) je zameraný na implementáciu spoločného integrovaného systému vo viacerých slovenských knižniciach. Do projektu sú zapojené slovenské vedecké i verejné knižnice ako napr.: SNK, Ústredná knižnica Slovenskej akadémie vied, Univerzitná knižnica v Bratislave, Štátna vedecká knižnica v Banskej Bystrici, knižnica Slovenského národného múzea a iné.

V rámci projektu KIS3G je riešený aj *Súborný katalóg periodík SR* (<http://skp.ulib.sk>), ktorý obsahuje záznamy periodík (odborných, vedeckých i záujmových časopisov, ročníkov, správ z konferencií, zborníkov a iných periodicky vydávaných dokumentov v tlačenej i inej forme) nachádzajúcich sa vo fondoch knižníc a iných inštitúcií Slovenskej republiky. Na tvorbe súborného katalógu periodík spolupracuje vyše 400 slovenských inštitúcií. V elektronickej podobe je katalóg budovaný od roku 1982 a obsahuje vyše 50 tisíc bibliografických záznamov periodík a viac ako 110 000 lokačných záznamov, tzv. holdingov.

### 3.2.3 Katalógy Univerzity Komenského v Bratislave

Univerzita Komenského v Bratislave je producentom dvoch vlastných knižnično-informačných databáz, sú to:

- *súborný online katalóg fakultných knižníc* (prístup k informáciám o dokumentoch, ktoré sa nachádzajú v jednotlivých knižniciach súčasti UK),
- *databáza publikačnej činnosti zamestnancov UK*.

Obe databázy sú prístupné z adresy <http://vili.uniba.sk>.

*Bibliografický záznam* (informácia o jednom vydaní daného titulu) je v súbornom katalógu len raz, k nemu sa pridávajú informácie o tom, ktoré knižnice daný dokument majú, v koľkých exemplároch a či je možné dokument požičať (status).

Registrovaní používatelia môžu:

- poslať žiadanku na dokumenty, ktoré sú prístupné na absenčnú alebo prezenčnú výpožičku,
- skontrolovať svoje používateľské kontá (čo, kde a dokedy majú požičané, rezervované,...),
- predĺžiť lehotu požičaných dokumentov.

System ponúka niekoľko rôznych spôsobov vyhľadávania (výber z hornej navigačnej lišty):

- *Prezeranie* – listovanie v abecedných zoznamoch autorov, názvov, predmetových hesiel, miest vydania a vydavateľstiev, signatúr (výber požadovanej položky z ponuky „Hľadať v“). Stačí zadať niekoľko prvých písmen a kliknúť na tlačidlo „Hľadať“ – systém zobrazí abecedný zoznam termínov (zvyčajne 10), ktoré sa najviac blížia k danému reťazcu.
- *Slovo zo záhlavia* – listovanie v zoznamoch názvov, autorov, predmetov (výber z položky „Typ záhlavia“). Je potrebné uviesť hľadané slovo v úplnosti alebo použiť zástupný znak, nestačí zadať niekoľko prvých písmen hľadaného termínu:
  - \* (hviezdička) – zástupný znak za ľubovoľný počet písmen/znakov,
  - ? (otáznik) – zástupný znak za presne jedno písmeno/znak.

System ponúkne ako výsledok všetky termíny, ktoré obsahujú hľadaný výraz, bez ohľadu na jeho pozíciu.

K dispozícii je možnosť kombinovať termíny pomocou logických operátorov (AND/A, OR/ALEBO, NOT/OKREM). Vyhľadanie možno bližšie špecifikovať voľbou: „Všetky slová“, „Fráza“ alebo „Presná zhoda“.

- *Kľúčové slová* – vyhľadanie konkrétnych záznamov. System vyhľadáva na presnú zhodu, takže je potrebné zadať hľadaný termín v úplnosti alebo použiť znak rozšírenia ? alebo\* podobne ako pri vyhľadávaní „Slovo zo záhlavia“. Termíny možno kombinovať pomocou logických operátorov. Hľadaný termín je potrebné vyhľadávať v rámci autor(ov), názov(u), predmet(u), ISBN, ISSN, žánru či vo všetkých indexovaných poliach (všade) – položka „Hľadať v“.

Zotriediť podľa: Netriediť Zotriediť

Vzostupne  Zostupne

Vyhľadávali ste - Katalog - Autor: Kimička AND Názov: cit\*

Záznamy 1 až 2 z 2

Uložiť do rešerše Vybrať 1-2 Všetko vymazať

<p><b>1</b></p> <p>úplný záznam <b>exempláre</b> (31) marc</p> <p>Vybrať do rešerše <input type="checkbox"/></p>	<p>Počet dostupných: 7</p> <p>Dostupné v: UKOFIUK (500), UKOPAKN (400), UKOPDKN (900)</p> <p><b>Autor</b> Kimička, Štefan</p> <p><b>Názov</b> Publikačná činnosť : Evidencia, vykazovanie a informačné systémy o publikačnej činnosti a citáciách podľa medzinárodných štandardov / Štefan Kimička</p> <p><b>Vyd. údaje</b> Bratislava : Stimul, 1997</p> <p><b>Vydanie</b> 1. vyd.</p>
<p><b>2</b></p> <p>úplný záznam <b>exempláre</b> (45) marc</p> <p>Vybrať do rešerše <input type="checkbox"/></p>	<p>Počet dostupných: 7</p> <p>Dostupné v: UKOEBKN (2900), UKOFIUK (500), UKOTVST</p> <p><b>Autor</b> Kimička, Štefan</p> <p><b>Názov</b> Ako citovať a vytvárať zoznamy bibliografických odkazov podľa noriem ISO 690 pre "Klasické" aj elektronické zdroje / Štefan Kimička</p> <p><b>Vyd. údaje</b> Bratislava : Stimul, 2002</p>

Obr. 3.28 Súborný online katalóg UK – výsledky vyhľadávania

- *Expertné vyhľadávanie* – priamy zápis selekčného výrazu, je potrebné poznať kódy oblastí vyhľadávania, tento spôsob vyhľadávania je určený najmä skúsenejším rešeršérom. Napr. selekčný výraz `au:Kimlička & s:citac*` predstavuje vyhľadávanie všetkých dokumentov, ktoré sa zaoberajú citáciami (s – predmetové heslo/subject) a ich autorom (au) je Kimlička.

Zoznam najpoužívanejších vyhľadávacích oblastí/kategórií:

a – autor, napr.: a:Shakespeare  
t – názov, napr.: t:"Ako sa Vám páči", t:Hamlet  
s – predmet, napr.: s:literatúra  
aw – všade, napr.: aw:matemati\*  
ud02 – rok vydania, napr.: ud02:200?  
ud05 – kód role autora, napr.: ud05:440  
ud10 – jazyk, napr.: ud10:eng  
ud11 – podnázov, napr.: ud11:pohľad  
ud19 – WWW dokument, napr.: ud19:www  
ud25 – žánery, napr.: ud25:príručky, ud25:učebnice  
p – vydavateľ, napr.: p:VEDA  
bl – bibliografická úroveň, napr.: bl:s

Príklady použitia znakov rozšírenia:

\*ková, \*ana\* – vyhľadá Haľková, Hana či Becková, Jana či Hrabková, Anabela  
\*cit\* – vyhľadá cit, pocit, citovať, ocitnúť, citácia atď.  
H?rský – vyhľadá Hůrský, Horský,...

Vyhľadávanie je možné taktiež upresniť použitím tzv. *filtr*ov – napr. na určitý jazyk, rok vydania/rozsah rokov, či lokáciu/knižnicu.

V rámci súborného online katalógu fakultných knižníc UK sa spracúvajú aj *záverečné práce študentov*. Od akademického roku 2005/2006 študenti odovzdávajú svoje záverečné práce aj v elektronickej podobe (formát PDF), ktoré sú následne v súlade s licenčnou zmluvou sprístupnené. Elektronická verzia práce predstavuje samostatný exemplár. Záverečné práce možno vyhľadávať podľa typu (bakalárska, diplomová, dizertačná a pod.) v rámci kľúčových slov – položka „žánery“. Toto vyhľadávanie je vhodné skombinovať s filtrom na fakultu (výsledok = všetky práce daného typu vybranej fakulty), rokom vydania, predmetom či katedrou, na ktorej boli obhájené (názov alebo časť názvu katedry je potrebné uviesť do položky „autor“). Ak má práca prístupnú elektronickú verziu, záznam exemplára obsahuje položku URL, kliknutím na ňu je možné zobrazit' úplný text práce. Podľa licenčnej zmluvy je práca sprístupnená bez obmedzenia – status „Dostupné“ – z ľubovoľného počítača pripojeného k internetu alebo len z počítača zapojeného do siete UK (kontrolujú sa IP adresy počítačov) – status „Prístup len v rámci UK“.

*Databáza evidencie publikačnej činnosti* (EPC), rovnako ako Súborný on-line katalóg UK, využíva rovnaký knižnično-informačný systém (VTLS/Virtua) a z tohto dôvodu ponúka takmer totožné funkcie a možnosti vyhľadávania. Najväčší rozdiel je v tom, že EPC je bibliografická a citačná databáza, dokumenty, ktoré sú v nej spracované sa často vo fakultných knižniciach nenachádzajú a teda ani nie je možné na ne poslať žiadanky.

### 3.3 Zhrnutie

Vývoj výpočtovej techniky spolu s rozšírením internetu, najmä služby World Wide Web, priniesol možnosť efektívneho uchovávanía a vyhľadávania informácií sprístupňovaných prakticky celému

svetu. K významným príkladom efektívnej distribúcie údajov patrí najmä vývoj vedeckých databáz a databázových centier, ktoré preukázali svoju hodnotu a dôležitosť pre širokú vedeckú komunitu. V ostatných rokoch sa vyhľadávanie informácií z elektronických informačných zdrojov stalo de facto primárnym spôsobom informačného prieskumu, keďže tieto nie sú limitované na určité miesto či čas, takže používateľ sa k nim môže dostať pohodlne a prakticky kedykoľvek. Používatelia môžu vyhľadávať vo voľne dostupných zdrojoch alebo môžu uprednostniť dôveryhodnejšie zdroje databázových centier, ktoré sú však často sprístupňované komerčne. Jednotlivé databázy poskytujú sofistikovaný systém vyhľadávania, tezaury, fazetové triedenie či rôzne možnosti na ďalšiu prácu s vyhladanými informáciami.

## Literatúra

BARTOŠEK, Miroslav. 2000. *Web of Science* [online] Jún 2000. Dostupné na internete: <<http://www.ics.muni.cz/mba/wos/ppframe.htm>>.

BELLÉROVÁ, Beáta. 2009. Databáza SCOPUS – jej obsah a nástroje alebo Predstavujeme databázu SCOPUS. In *Itlib*. ISSN 1335-793X, 2009, roč. 13, č. 2, s. 22 – 29.

JURČACKOVÁ, Zora. 1996. *Organizácia poznania v dokumentovej komunikácii*. Bratislava : Centrum VTI, 1996. ISBN 80-85165-61-9.

MAKULOVÁ, Soňa. 1992. *Bázy dát v sústave VTEI*. Bratislava : Univerzita Komenského, 1996. ISBN 80-223-0514-6.

MAKULOVÁ, Soňa. 2002. *Vyhľadávanie informácií v internete*. Bratislava : EL&T, 2002. ISBN 80-88812-16-X.

ROWLEY, J. E. *Organising Knowledge : An Introduction to Information Retrieval*. Aldershot : Gower, 1987.

SINGHAL, Amit. 2001. Modern Information retrieval: A Brief Overview. In *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*. 2001, roč. 24, č. 4, s. 35 – 43.

SCHWARZ, Josef. 2003. Před 40 lety začala zlatá éra citačních rejstříků vědecké literatury. In *Ikaros* [online]. 2003, roč. 7, č. 11 [cit. 2009-09-25]. Dostupné na internete: <<http://www.ikaros.cz/node/1520>>. ISSN 1212-5075.

ŠUŠOL, Jaroslav. 2005. Elektronické informačné zdroje – vzťah používateľských a autorských preferencií. In *Ikaros* [online]. 2005, roč. 9, č. 9 [cit. 2009-08-30]. Dostupné na internete: <<http://www.ikaros.cz/node/2004>>. ISSN 1212-5075.

TOMPSON, S. R. 2007. Electronic Resources Reviews : Scirus – for Scientific Information [online]. In *Issues in Science and Technology Librarianship*. Winter 2007 [cit. 12.09.09]. Dostupné na internete: <<http://www.istl.org/07-winter/electronic3.html>>.

## Zoznam webových sídel

Bio-base  
<http://www.biobase-international.com/>

Compendex  
<http://www.ei.org/compendex>

Current Contents Connect  
[http://thomsonreuters.com/products\\_services/science/science\\_products/a-z/current\\_contents\\_connect](http://thomsonreuters.com/products_services/science/science_products/a-z/current_contents_connect)

Databáza evidencie publikačnej činnosti zamestnancov UK

[http://alis.uniba.sk:8000/cgi-bin/gw\\_49\\_3\\_8/chameleon?skin=epc](http://alis.uniba.sk:8000/cgi-bin/gw_49_3_8/chameleon?skin=epc)

Elsevier

<http://www.elsevier.com>

Embase

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

Essential Science Indicators

[http://thomsonreuters.com/products\\_services/science/science\\_products/a-z/essential\\_science\\_indicators](http://thomsonreuters.com/products_services/science/science_products/a-z/essential_science_indicators)

Geobase

[http://www.elsevier.com/wps/find/bibliographicdatabasedescription.cws\\_home/422597/description](http://www.elsevier.com/wps/find/bibliographicdatabasedescription.cws_home/422597/description)

ISI Web of Knowledge

<http://isiknowledge.com>

Journal Citation Reports

[http://thomsonreuters.com/products\\_services/](http://thomsonreuters.com/products_services/)

[science/science\\_products/a-z/journal\\_citation\\_reports](http://science/science_products/a-z/journal_citation_reports)

KIS3G

<https://www.kis3g.sk/>

Medline

<http://medline.cos.com/>

NISPEZ

<http://nisppez.cvtisr.sk>

OCLC

<http://www.oclc.org/global/default.htm>

Scopus

<http://www.scopus.com>

Súborný online katalóg Univerzity Komenského v Bratislave

<http://www.uniba.sk/?id=1884>

Web Of Science

[http://thomsonreuters.com/products\\_services/science/science\\_products/a-z/web\\_of\\_science](http://thomsonreuters.com/products_services/science/science_products/a-z/web_of_science)

WorldCat

<http://www.worldcat.org>

---

[1] Autorizovaní používatelia v rámci SR majú v súčasnosti prístupné produkty vyznačené hrubým písmom.

[2] Prístup k produktom spoločnosti Thomson Reuters (WoS, CCC, JCR a ESI) a ďalším licencovaným EIZ (Scopus, ProQuest, ScienceDirect, ACM, IEEE a iné) bol pre verejné vysoké školy SR, SAV a vybrané knižnice zabezpečovaný prostredníctvom akademických knižníc, ktoré uzatvárali konzorciá a každoročne pripravovali a podávali na Ministerstvo školstva rozvojové projekty na získanie finančných prostriedkov na zapltenie tohto prístupu. Od roku 2008 sa prístup k vybraným EIZ rieši v rámci projektu Národný informačný systém podpory výskumu a vývoja na Slovensku – prístup k elektronickým informačným zdrojom (NISPEZ) realizovaného v rámci Operačného programu Výskum a vývoj s finančným príspevkom z Európskeho fondu regionálneho rozvoja. Riešiteľom projektu je CVTI SR (viac pozri <http://nisppez.cvtisr.sk>).

# 4

## Vyhľadávanie v internete

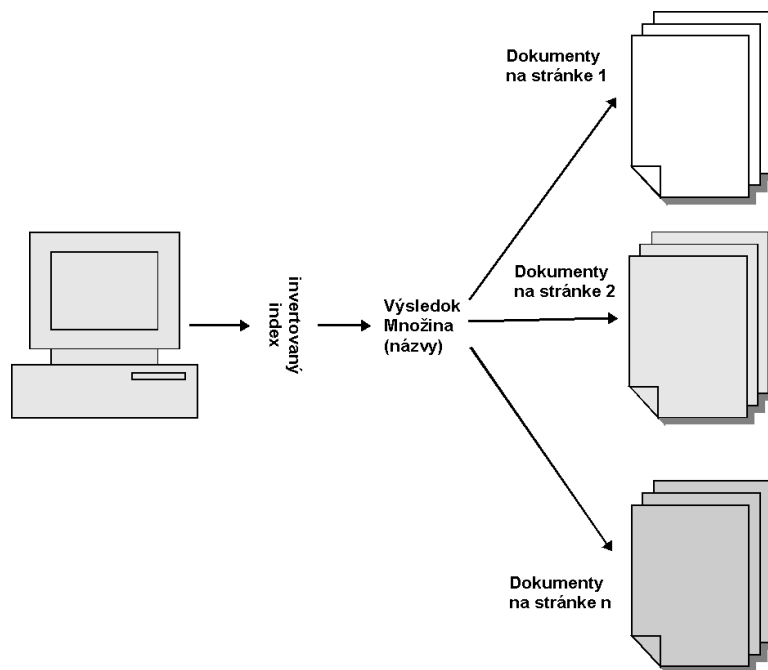
### 4.1

#### Modely webových systémov

Webové prieskumové nástroje sú príkladom distribuovaných prieskumových systémov. Oproti tradičným databázovým systémom sa odlišujú tým, že vyhľadávajú v rozptýlených, fyzicky vzdialených dokumentoch na rôznych rozptýlených miestach (serveroch).

**Centralizovaný model** informačného prieskumového systému pracuje na princípe porovnávania opisu dokumentov a množiny dotazov prostredníctvom invertovaných slovníkov a tezaurov. Základné modely tu predstavujú Booleov model, vektorový priestorový model, pravdepodobnostný, zhukový model. Predpokladá sa, že ku každému dokumentu možno priradiť množinu jedinečných deskriptorov. Pri snahách o vylúčenie synonym sa vytvára tezaurus obsahujúci všeobecnejšie synonymum k jednotlivému deskriptoru (askriptor) tak, aby vznikla báza znalostí s prirodzenou hierarchiou významov (napríklad mačka-zviera). To zjednodušuje vyhľadávanie z hľadiska sémantiky (napríklad pri dotaze zameranom na hľadanie všetkých „zvierat“). Centralizovaný model sa využíva najmä pri vyhľadávaní v jednej databáze a v rámci jedného systému.

**Distribuovaný model** vznikol začiatkom 80. rokov. Všeobecne ide o vyhľadávanie na viacerých miestach a vo viacerých systémoch. Deskriptory môžu byť usporiadané hierarchicky. Tezaurus sa vytvára z rôznych lokálnych slovníkov na rôznych miestach. Odpovede na dotaz z rôznych lokálnych miest (stránok) sa spracujú osobitne a napokon sa kombinujú do konečnej množiny výsledkov. Model distribuovaného dokumentového prieskumu je znázornený na obrázku 4.1 (Grossman, Frieder 2004).



Obr. 4.1 Model distribuovaného dokumentového prieskumu (podľa Grossman, Frieder 2004)

Model ukazuje princípy vyhľadávania vo webových nástrojoch, pričom každý z nástrojov ako Google, Yahoo! alebo aj novšie ako Vivisimo (v súčasnosti už nie je voľne dostupný), Dogpile, Kahzam, Evri, FactBites, SearchCube a i. je jedinečný a obsahuje vlastné a stále sa vyvíjajúce algoritmy. Zdokonaľuje sa kategorizácia, spôsoby triedenia podľa poradia (ranking), sémantika a obrazové znázornenie (vizualizácia). Novší vyhľadávací nástroj WolframAlpha ([www.wolframalpha.com](http://www.wolframalpha.com)) ponúka nové spôsoby vyhľadávania prostredníctvom bázy znalostí a špecifickej sémantickej analýzy. Medzi ďalšie vyhľadávacie nástroje okrem Google patria napríklad Collarity (v súčasnosti už nie je voľne dostupný), Healia (pre zdravotnícke informácie). Napríklad Collarity umožňuje vyhľadávanie v kontexte, podporuje kolaboratívnu personalizáciu (obrázok 4.2).



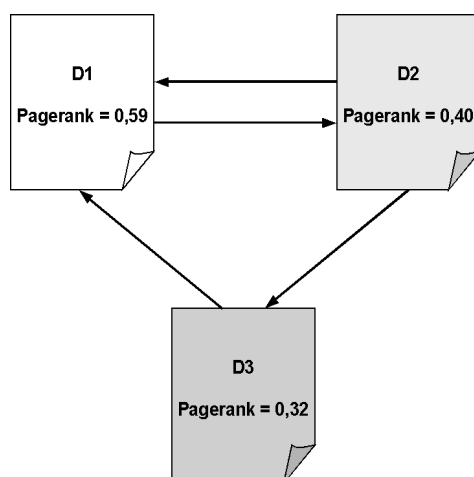
Obr. 4.2 Vyhľadávač Collarity



Vyhľadávanie vo webových nástrojoch využíva práve princípy distribuovaného modelu, pretože umožňujú prístup k webovým serverom a ich dokumentom. Väčšina z týchto systémov má centralizovaný index a väčšina z nich uchováva aj smerovače vo forme hypertextových odkazov na ďalšie servery. Množstvo meniacich sa vyhľadávacích nástrojov je zaznamenaných napríklad na stránkach [www.pandia.com](http://www.pandia.com) či [www.searchengineguide.com](http://www.searchengineguide.com) (špecializované nástroje).

Rozdielom pri vyhľadávaní v internete je aj interpretácia presnosti a odozvy. *Čas a rýchla odozva* sú pre vyhľadávanie vo webe mimoriadne dôležité. Preto návrhári kladú veľký dôraz na prvú vyhľadávanú stránku a prvých desať odkazov. Na zvyšovanie účinnosti vyhľadávania sa využívajú aj analýzy množstva dotazov (query log analysis). V posledných rokoch vo vyhľadávaní dominuje Google, ktorý sa stal z jedného vyhľadávacieho nástroja celým vzorom na vyhľadávanie. Niekedy sa hovorí o tzv. „Google paradigme“. Pre vyhľadávanie vedeckých dokumentov sa často používa Google Scholar.

Vyhľadávací nástroj Google využíva na vyhľadávanie tzv. PageRank algoritmus (pomenovaný podľa autora Larryho Pagea, prvýkrát bol opísaný v roku 2000). Využíva vstupujúce a vystupujúce prepojenia na prispôsobovanie vypočítaného skóre webovej stránky podľa jej obľúbenosti nezávisle od dotazu používateľa. Jednoduché grafické znázornenie algoritmu PageRank je na obrázku 4.3. (Grossman, Frieder 2004). Oproti tradičnému prieskumu sa tu objavuje *princíp populárnosti dokumentu*, ktorý je definovaný ako počet odkazov ďalších webových stránok na dokument.



Obr. 4.3 Znázornenie algoritmu PageRank

## 4.2 Web 2.0

Web 2.0 sa ako pojem rozšíril približne v rokoch 2004 – 2005. Existuje mnoho podobných pojmov, ako napr. sémantický web, poznatkový web alebo digitálne knižnice nových generácií. Tieto pojmy sa prelínajú, web 2.0 sa ujal kvôli jednoduchosti a zrozumiteľnosti. Pojem web 2.0 naznačuje, že vzniká nová kvalita doteraz známych služieb webu a skrýva sa v aktivite človeka v elektronickom prostredí.

Aktivita človeka sa na jednej strane prejavuje myšlienkovými procesmi a interakciou internetovej komunikácie. Preto možno spracovať hlbšie významy informácií a účinnejšie posudzovať relevanciu. Na druhej strane sú komunikačné procesy a spolupráca ľudí pri tvorbe a využívaní webu. Komunikácia a spolupráca umožňujú určovať hlbšie kontexty informácií a zlepšiť efektívnosť vyhľadávania.

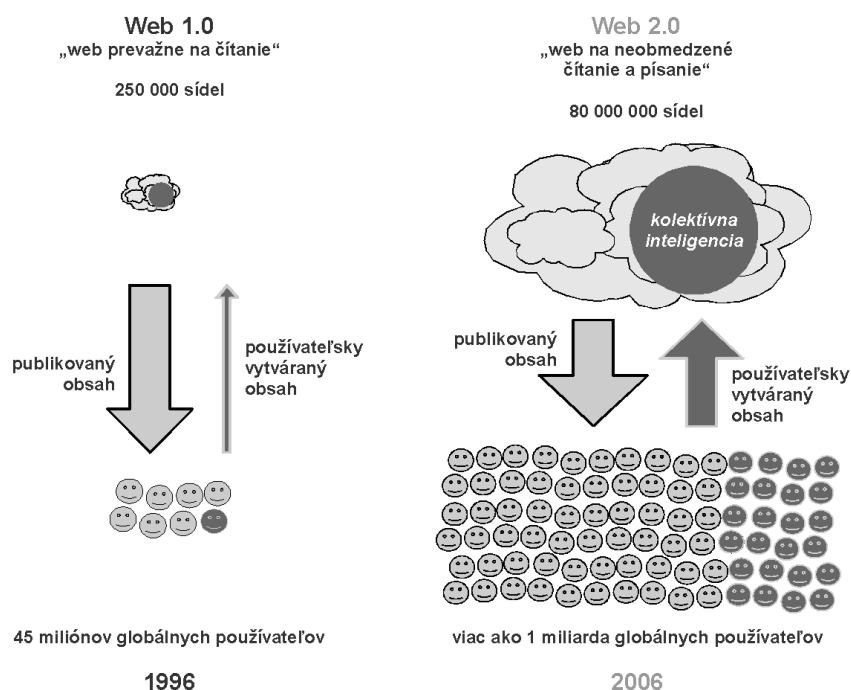
Web 2.0 však vyvolal aj polemiku. Často sa používa intuitívne, bez jasnej definície. Najmä pri podnikaní na internete a získavaní prostriedkov na nové projekty a služby vznikla pojmová „bublina“.

Web 2.0 možno vidieť v kontexte vývoja technológií od začiatku 70. rokov až po súčasnosť. V ére tradičných počítačových systémov sa výpočtové postupy uplatňovali najmä v organizačných súvislostiach na spracovanie transakčných údajov. Neskôr sa posunuli funkcie počítačov smerom k personálnym nástrojom spracovania a využívania informácií. Nová éra nastala pri prepájaní počítačových systémov do celkov a sietí. Prepájajú sa na úrovni fyzickej, logickej aj integráčnej.

Web 2.0 stelesňuje najmä logické prepojenie. Orientuje sa na službu, sémantiku či udalosť, ktoré prinášajú pridanú hodnotu k informáciám. Smeruje však k inteligentným metódam spracovania informácií. Technológia a služba sa nenásilne včleňujú do situácií a udalostí pri riešení každodenných aj odborných informačných problémov. Web 2.0 využíva súbor technológií nazývaných Ajax obsahujúcich špecifické nástroje na prezentáciu, zobrazovanie a interakciu, výmenu a manipuláciu s informáciami, informačný prieskum a integráciu týchto nástrojov.

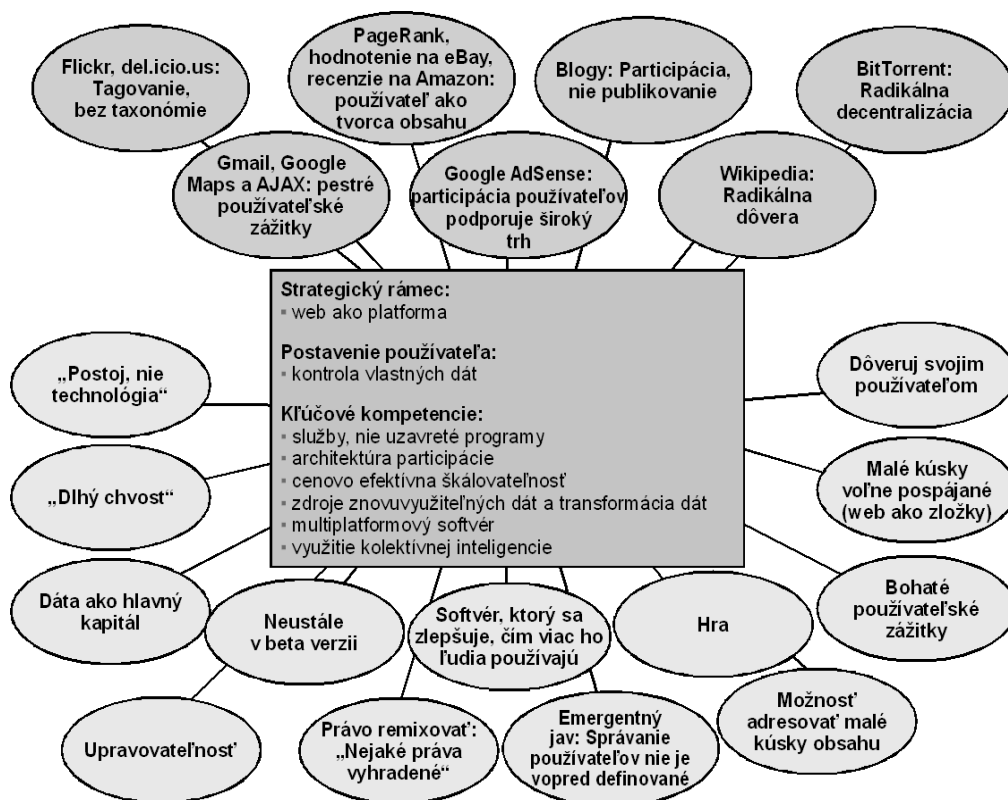
V praxi sa web 2.0 prejavuje ako spoločná tvorba, organizácia a využívanie informácií. Služby sa organizujú okolo komunit. Typická je vznikajúca spoločná kolektívna skúsenosť spolupracujúcich ľudí spojených spoločným záujmom či problémom. Web 2.0 dostáva prívlastky ako interaktívny, inteligentný či mobilný. Dominuje však sociálny aspekt spracovania a využívania informácií – participácia, spolupráca, navigačná štruktúra vytvorená komunitou, dialóg. Vo vedeckej komunikácii sa využívajú možnosti vzájomného hodnotenia produktov – anotácie, recenzie, tvorba v skupinách, diskusia. „Cesta“ v spleti informácií v elektronickom prostredí je „adaptívna“.

Rozdiely medzi web 1.0 a web 2.0 sú znázornené na obrázku 4.4.



Obr. 4.4 Rozdiely medzi web 1.0 a web 2.0 podľa O' Reilly (<http://www.oreillynet.com/>)

Za základné vlastnosti webu 2.0 sa považuje škálovanie služieb a ich prispôsobovanie používateľom na základe poznania ich informačného správania. V tomto zmysle má web 2.0 marketingový aj komerčný rozmer. Používatelia pridávajú hodnotu informáciám svojimi špecifickými spôsobmi využívania informácií a stretávajú sa v sieťovom prostredí. Organizácia informácií a ich využívanie sú založené na decentralizácii. To vedie k možnosti spoločnej tvorby a viacnásobnému využívaniu existujúcich zdrojov, poznatkov a skúseností (napr. wikipedia). Mapa web 2.0 je znázornená na obrázku 4.5.



Obr. 4.5 Mapa funkcií a služieb web 2.0 podľa O'Reilly (<http://www.oreillynet.com/>)

Pri využívaní zdrojov sa v produktoch a službách webu 2.0 objavuje možnosť vytvárať špeciálne „zmesi“ informácií a ich „mixovanie“ podľa aktuálnej potreby. Tým, že si používatelia prispôbujú služby a systémy, vytvárajú relevantné informácie. Prispôbovaním, komunikovaním a participáciou vzniká nová pridaná hodnota. V spoločnej skúsenosti sa vynárajú nové systémy založené na kolektívnom zážitku.

Príkladom sú folksonómie – osobitné ľudové systémy triedenia informácií podľa aktuálnej potreby.

Sociálne aspekty a kolaborácia sa však v elektronickom prostredí prejavujú nielen v pozitívnom zmysle. Mnohé atribúty služieb webu 2.0 môžu priniesť aj negatívne či nechcené efekty. Napríklad informačné preťaženie, problém konsenzu v skupinách, problémy dvojznačnosti pojmov v otvorenom prostredí. Do elektronického prostredia sa prenášajú aj tradičné problémy fungovania skupín ako normy komunikácie, spoločný záujem, motivácia, problémy etiky a kultúry správania, časová záťaž. V elektronickom prostredí vzniká aj problém bezpečnostného rizika a dôvery. Prepojenie funkcií, služieb, produktov a ľudí v rámci web 2.0 je znázornené na obrázku 4.6.



ných knižníc. Kľúčom pre nové služby je samostatná aktivita človeka v elektronickom prostredí. Ukazuje sa, že ak chcú byť akademické knižnice prijateľné pre študentov, musia sa im prispôbiť.

Využívanie informácií nie je podmienené inštitucionálne, ale tematicky. Dôležitá je relevancia informácií, ktorú možno podporiť práve vyššou úrovňou kontextualizácie. Digitálne knižnice však musia podporovať aj kolaboráciu študentov, učiteľov a výskumníkov. Spoločné hodnotenie, anotácie, recenzie, diskusie začínajú vytvárať špeciálnu kultúru informačného správania v elektronickom prostredí.

Príkladmi takýchto knižníc či systémov sú digitálne knižnice s obohatenou funkčnosťou, ako napr. americká digitálna knižnica pre vzdelávanie v oblasti prírodných vied a matematiky, NSDL, (<http://nsdl.org/>) (obrázok 4.7).

Obr. 4.7 NSDL – Národná vedecká digitálna knižnica

Ďalším stelesnením princípov novších systémov a knižníc sú napríklad koncepcie vedeckej spolupráce v elektronickom prostredí („e-science“) a kolaboratória. (sociotechnické komplexné systémy podporované novou kyberinfraštruktúrou). Knižnice sa v nich môžu pretvárať na miesta manažmentu poznatkov.

Pri spracovaní a využívaní informácií môžeme rozlíšiť niekoľko základných vzorcov. **Prvý vzorec** naznačuje, že používatelia preferujú *samostatný pohyb v elektronickom prostredí*. To je aj jedna z príčin úspešnosti jednoduchých rozhraní ako Google alebo Amazon. Aj ďalšie výskumy dokazujú, že používatelia málokedy využívajú pomoc knihovníka. Preto v službách webu 2.0 má byť pomoc používateľovi čo najbližšie k jeho informačnému prostrediu.

S tým súvisí **druhý vzorec** rozloženia obsahu a inštitúcií na menšie časti, tzv. dezagregácia. V elektronickom prostredí vzniká tzv. „mikroobsah“, ktorý sa môže rôznymi spôsobmi mozaikovo skladať a znovu využívať. Dôsledkom je aj *dezintegrácia zdrojov a procesov učenia či výskumu*. To

vplýva aj na tradičné inštitúcie ako knižnice a vydavateľstvá. Zároveň vzrastá úloha kontextu, v ktorom sú informácie spracúvané a využívané.

**Tretím vzorcom** je *spolupráca a skupinové kolaboratívne nástroje* („sociálny softvér“). Z dialógu a spoločného záujmu vzniká viac kontextu. Nová „infosféra“ si vyžaduje ľudský vklad – kreativitu, kogníciu, komunikáciu. Služby webu 2.0 sú len sprostredkovateľmi komunikovania a využívania informácií v spoločnosti.

## 4.4 Webové vyhľadávacie nástroje

### 4.4.1 Východiská a typológia tradičných vyhľadávačov

Vyhľadávacie nástroje (tiež vyhľadávače) slúžia na získavanie informácií v heterogénnom prostredí webu. Pod vplyvom rýchleho vývoja týchto technológií sa mení aj pohľad na typológiu vyhľadávacích nástrojov. Členia sa na dve základné skupiny: 1. *tradičné vyhľadávacie nástroje*, 2. *nástroje vyhľadávania 2.0*. Na to, aby sme mohli lepšie pochopiť nástroje vyhľadávania 2.0, treba si predstaviť niektoré východiská. Prehľad *typológie vyhľadávacích nástrojov*, ktorú predstavila Makulová (2002), je v tabuľke 4.1. Pozostáva z nasledujúcich nástrojov:

1. prieskumové stroje,
2. metaprieskumové stroje,
3. predmetové adresáre,
4. hybridné vyhľadávacie nástroje.

Hlavný rozdiel spočíva v *technológii spracovania dát*. V prípade *prieskumových strojov* (napr. Google, Ask, Bing) sa dáta o webových zdrojoch do dátového skladu dostávajú pomocou vyhľadávacích robotov. Ide o špecializované softvérové agenty, ktoré autonómne prehľadávajú a analyzujú obsah webových stránok. Na druhej strane *predmetové adresáre* (známe aj ako katalógy, napr. Dmoz, IPL2, Bubl) pozostávajú z hierarchicky organizovaných tematických štruktúr, ktoré sú vytvárané kategorizáciou, klasifikáciou alebo priradovaním tagov editormi. Takýto manuálny prístup a automatizované výpočty prieskumových strojov kombinujú hybridné nástroje (ako Yahoo!).

Ďalším rozdiel je v *prístupe pri vyhľadávaní na vstupe rozhrania vyhľadávača*. V zásade môžeme odlišiť vyhľadávanie pomocou kľúčových slov (typické pre prieskumové stroje, hybridné nástroje), prehľadávanie kategórií (v predmetových adresároch) a tagov (v mraku tagov nástroja, ktorý ho sprostredkúva).

Vo vyhľadávačoch sa tiež odlišuje spôsob, akým sa *zobrazujú výsledky vyhľadávania*. Môže to byť *textový zoznam výsledkov* s anotáciami a odkazmi na zdroj (pre ktorý je známy Google). Iný prístup umožňuje *vizualizácia*, ktorá sprostredkúva zobrazovanie výsledkov v zhlukoch (Clusty), mapách (Kartoo) a náhľadoch stránok (Viewzi, RedZ). *Sémantické vyhľadávače* namiesto zoznamu zdrojov prinášajú zmysluplnú odpoveď s podpornými zdrojmi (Lexxe, Hakia, Powerset, WolframAlpha). *Agregátory vertikálnych strojov* umožňujú zobrazovanie výsledkov zo špecializovaných vyhľadávačov na jednej stránke (napr. Hakia, ktorá je aj sémantickým vyhľadávačom). Spoločne je tak možné nájsť nielen webové stránky, ale aj videá, obrázky, aktuality a produkty. Spolu s vývojom techník s podporou vizualizácie a umelej inteligencie sa bude škála vymenovaných typov ešte rozširovať.

Tab. 4.1 Prehľad základných typov vyhľadávacích nástrojov

Typ vyhľadávacieho nástroja	Stručná charakteristika
<b>Prieskumové stroje</b>	Patria medzi najfrekvencovanejší a najviac využívaný typ vyhľadávacích nástrojov. Prístup pri vyhľadávaní je založený na využití kľúčových slov. Fungujú na báze robotov, ktoré prehľadávajú obsah webu. Databázy prieskumových strojov uchovávajú kópie webového obsahu, ktorý je neskôr indexovaný za účelom jednoduchšieho a rýchleho vyhľadania pomocou rozhrania.
<b>Metaprieskumové stroje</b>	Vyhľadávajú a sumarizujú dáta z viacerých databáz vyhľadávacích nástrojov. Paralelne prehľadávajú a syntetizujú najrelevantnejšie výsledky a zároveň eliminujú duplicitné výsledky.
<b>Predmetové adresáre</b>	Na rozdiel od prieskumových strojov je organizácia predmetových adresárov založená na manuálnom triedení zdrojov editormi. Vyhľadávanie je sprostredkované prehľadávaním kategórií. Predmetové adresáre sú triedené pomocou rôznych prístupov – kategorizáciou (priradením zdrojov do kategórií), klasifikáciou (triedením zdrojov pomocou klasifikačných schém, napr. DDC a pod.), ale aj tagovaním (určovaním tagov konkrétnym zdrojom a následné prehľadávanie mraku tagov).
<b>Hybridné vyhľadávacie nástroje</b>	Kombinujú predmetové adresáre s prieskumovými strojmi. Tieto nástroje vznikli kvôli nutnosti rozšírenia výsledkov vyhľadávania alebo jeho zjednodušenia.

Novšia generácia nástrojov sa označuje aj ako *3. generácia vyhľadávačov* (Ezzy, 2006, Zimmer, 2008). Zatiaľ čo relevantné modely predchádzajúcich generácií boli založené na štatistických analýzach frekvencie výskytu pojmov, neskôr rozšírené o analýzu prepojení jednotlivých stránok (napr. PageRank v Google), vyhľadávanie 2.0 prichádza s novšími algoritmi. Vychádzajú z princípov webu 2.0, teda preferencií a skúseností používateľov, prejavov kolaborácie a kolektívnej inteligencie. Vďaka kolaborácii môžu ľudia participovať na tvorbe, ukladaní, spoločnom využívaní, hodnotení a organizovaní rôzneho obsahu (napr. stránky, blogy, podcasty, videá, fotografie, prezentácie, profily a pod.). Tento jav sa tiež označuje pojmom „masová kolaborácia“ (Tenenbaum, 2005). Dôraz sa z pohľadu vyhľadávania kladie nielen na ľudí, ale aj na pokročilé techniky podpory interakcie, napr. pomocou vizualizácie.

Vývojové trendy v oblasti vyhľadávania sa dajú určiť vo viacerých línách. V prvom rade je to vplyv sociálneho softvéru, ktorý má za následok vznik *sociálneho vyhľadávania*. Aktivity ľudí pri vyhľadávaní môžu rozšíriť vyhľadávacie *techniky s podporou umelej inteligencie*. Ďalším výrazným trendom je podpora *interaktivity* pomocou pokročilých techník *vizualizácie*. Veľký vplyv na vývoj v oblasti vyhľadávania majú aj *novšie médiá* ako blogy, RSS, online záložky a pod..

## 4.4.2 Sociálne vyhľadávanie

### 4.4.2.1 Sociálne siete

Sociálne siete sú fenoménom, ktorý ovplyvnil nielen off-line správanie komunit, ale formuje aj vyhľadávanie. Systémy s podporou sociálnych sietí umožňujú vyhľadávanie ľudí, textu (komentárov, správ, príspevkov) a rôznych typov médií ako obrázkov, videí, a pod. Okrem toho komunita prispieva aj k triedeniu tohto obsahu pomocou popisiek a tagov.



Za úplne prvú stránku s podporou sociálnych sietí sa považuje Well, ktorá vznikla už v roku 1985 (Malinowski, 2007). Desať rokov po nej bola založená známa služba Classmates. Začali vznikať rôzne aplikácie s podporou sociálnych sietí na webe ako Geocities alebo dodnes aktívne MySpace, Facebook a pod. Spolu s rastúcim počtom rôznych služieb sa zväčšujú aj on-line komunity. Napríklad zatiaľ čo Facebook mal v októbri 2008 110 miliónov aktívnych členov, o rok to už bolo o 100 miliónov viac. Celkový nárast komunity za rok teda predstavuje až o 90%.

Sociálne siete podporujú nielen aplikácie na tvorbu a prepájanie profilov, ale aj služby s primárnym cieľom zdieľania a organizovania obsahu ako napr. Twitter, Delicious, Flickr, YouTube, či Technorati. V súčasnosti existuje množstvo sociálnych sietí zameraných na rôzne cieľové skupiny. Vznikajú špecializované sociálne siete na základe konkrétnych záujmov, tém, typov vzťahov, pozícií, krajín a pod.

Sociálne siete so sebou prinášajú problémy súvisiace s ochranou a otáznym využívaním osobných údajov. Systémy na podporu sociálnych sietí sa vyznačujú vysokou mierou heterogénnosti, preto by snahy vývojárov mali byť ešte viac orientované na efektívnejšiu integráciu. V súvislosti so sociálnymi sieťami vznikajú otázky spojené so závislosťou na komunikácii vo virtuálnom prostredí a s dopadmi ich využívania na reálne sociálne vzťahy.

Počet aplikácií, ktoré prispievajú k rozrastaniu sociálnych sietí v prostredí internetu stále stúpa. Z hľadiska návštevnosti sa nachádzajú v rebríčkoch spolu so svetovými vyhľadávačmi. Dokazuje to prítomnosť Facebooku, YouTube, MySpace, či Twitter medzi stránkami s najvyššou návštevnosťou podľa štatistík Alexa.com (Alexa, 2009).

#### 4.4.2.2

#### **Nástroje s podporou sociálneho vyhľadávania**

Jednotlivé prieskumové stroje určujú najrelevantnejšie odkazy na prvých pozíciách v zozname výsledkov na základe svojho algoritmu. Základným kritériom je frekvencia výskytu a pozícia hľadaného slova v indexovanom texte, ale aj prepojenosť stránky s ostatnými zdrojmi a pod. Kvantitatívne kritériá určujúce objektívnu relevanciu sa môžu líšiť od toho, čo konkrétny človek v určitej situácii považuje za využiteľné. Práve z toho dôvodu vznikol priestor pre sociálne vyhľadávacie nástroje, ktoré ľudia nielen využívajú, ale môžu pomocou nich aj priamo ovplyvňovať radenie výsledkov alebo vyhľadávať informácie za pomoci editorov a sprievodcov. Tieto nástroje sa označujú aj ako „human-powered“. Toto pomenovanie nie je najvhodnejšie, pretože za tvorbou algoritmov a správou iných typov vyhľadávačov stoja takisto ľudia. Napriek tomu sa tento pojem zaužíval. Trend sociálneho vyhľadávania ovplyvnil aj existujúce vyhľadávače ako Google, ktorý umožňuje v personalizovanom rozhraní presúvanie a aj odstránenie výsledkov v zozname podľa rozhodnutia používateľa, čiže subjektívnej relevancie.

Za predchodcov takýchto služieb sa dajú považovať predmetové adresáre. Tieto organizujú a aktualizujú editori, ktorí sa špecializujú na konkrétnu oblasť. Prvý predmetový adresár s názvom WWW Virtual Library (2008) založil v roku 1991 Tim Bernes-Lee. Ďalším výnimočným a dodnes fungujúcim projektom je Dmoz (2009). Dnes už predmetové adresáre fungujú ako hybridné nástroje, ktoré integrujú aj prieskumový stroj.

#### *iRazoo*

V iRazoo môže registrovaný používateľ hodnotiť konkrétne výsledky v zozname vyhľadaných zdrojov. Výsledok je možné ostatným buď odporúčať alebo neodporúčať. Navyše možno pridávať komentáre (obrázok 4.8). Odporúčané odkazy sú neskôr vo výsledkoch zobrazené v osobitnej sekcii (hneď po sponzorovaných výsledkoch). iRazoo motivuje aktivitu členov komunity pridelovaním bodov a možnosťou získať vecné ceny. Počet bodov narastá aj za získanie nových členov do komunity. Po registrácii iRazoo priradí mailové konto a profil, v ktorom možno sledovať bodové hodnotenie aktivít a uložené odporúčané výsledky.

Okrem stránok, vyhľadáva iRazoo obrázky, videá, aktuality a produkty. Podporuje Razoopediu, ktorá preberá obsah z Wikipédie. Obrázky sú zobrazené neprehľadne, nedajú sa ďalej triediť a organi-



zovať a ani hodnotiť. Niektoré odkazy v iRazoo majú komerčnú povahu a nedajú sa nijako hodnotiť. Ďalším problémom je to, že odporúčané odkazy nie sú ďalej organizované. Ich pribúdajúce množstvo znižuje orientáciu v zobrazených odkazoch.



Obr. 4.8 Pomocou iRazoo sa dajú hodnotiť a komentovať vyhľadane výsledky

### Mahalo

Mahalo je novší typ adresára s podporou prieskumového stroja a sociálnych sietí, v ktorom je obsah organizovaný a aktualizovaný editormi. Mahalo podporuje participáciu členov komunity, ktorí môžu zdieľať a organizovať relevantné odkazy. Mahalo sprostredkúva aj službu Mahalo Answers, pomocou ktorej možno získať odpoveď alebo odpovedať na špecifickú otázku. Mahalo vyhľadáva nielen odkazy na webové stránky, ale aj videá, obrázky, aktuality, blogy, produkty a Twitter správy. V kategóriách Mahalo sa dajú nájsť zaujímavé odporúčania na konkrétne témy. Napr. v kategórii „Technology“ (obrázok 4.9) je výber najviac obľúbených podtém, ale aj výber z otázok z Mahalo Answers. Ďalej sú tu známe osobnosti a firmy, ale aj zaujímavé odkazy na relevantné podstránky Mahalo so zdrojmi.



Search

### Most Popular



Nsfreewind



Cerulean Studios



Ceruleanstudios



Microformats



Apple Claims New



Vasanth Govind



Abeon Hosting



Forbes.com



Webs.com



BestBuy.Com



Classmates.com



CISCO

### Technology Categories (3704)

Technology & Internet » Technology

- People In Tech (272)
- Tech Companies (256)

### Technology Questions

Do you believe that a 3 year old should be able to play online learning games. (6 Answers)

Can you record audio on an iTouch via bluetooth, or better yet hands free, while still maintaining good quality? (0 Answers)

Does anyone know if the "sixth sense" device developed by MIT Media Labs is being commercialized? (0 Answers)

Palm Pre setup how to (0 Answers)

Can drobo share be used with other usb drives? (0 Answers)

[View all Technology Questions](#) | [Ask a Question](#)

### People In Tech

Abhishek Parolkar  
Adam Bly  
Adam Curry  
Adam Savage  
Alan Russell  
Alex Albrecht  
Allan Alcorn  
Allison Hunt  
Amanda Congdon  
Amy Novogratz

Andy Beal  
Andy Bechtolsheim  
Angela Beesley Starling  
Anne Wojcicki  
Anthony Feint  
Arianna Huffington  
Arthur Levinson  
Ashley Qualls  
Aubrey de Grey  
Ben Huh

Obr. 4.9 Kategória „technology“ v Mahalo

### Yoogle

Yoogle prináša kolaboratívne vyhľadávanie, v ktorom možno ovplyvniť poradie zobrazených výsledkov (obrázok 4.10). Yoogle umožňuje odkazy v zozname výsledkov presúvať na základe subjektívnej relevancie. Počet a pozície posunov ovplyvnia kolaboratívne poradie výsledkov. Demokratické zasahovanie do výsledkov prináša aj riziká spojené s „pretlačaním“ odkazov na prvé pozície.

Search results for: web 2.0 Results 1 - 10 of about 213000000 for web 2.0

Result Title	Description	URL	Shifts
<a href="#">Web 2.0 - Wikipedia</a>	description of the term that some refer to as a second phase of development of the World Wide Web, including its architecture and its applications.	<a href="http://en.wikipedia.org/wiki/Web_2.0">http://en.wikipedia.org/wiki/Web_2.0</a> - 94k <a href="#">Cache Copy</a>	17
<a href="#">What is Web 2.0? Definition from Whats.com</a>	Web 2.0 (or Web 2) is the popular term for advanced Internet technology and applications including blogs, wikis, RSS and social bookmarking.	<a href="http://whatis.techtarget.com/definition/0,,sid9_gci1169528,00.html">http://whatis.techtarget.com/definition/0,,sid9_gci1169528,00.html</a> - 56k <a href="#">Cache Copy</a>	1
<a href="#">What Is Web 2.0 - O'Reilly Media</a>	Tim O'Reilly attempts to clarify just what is meant by Web 2.0, the term first coined at a conference brainstorming session between O'Reilly Media and MediaLive ...	<a href="http://oreilly.com/web2/archive/what-is-web-20.html">http://oreilly.com/web2/archive/what-is-web-20.html</a> - 44k <a href="#">Cache Copy</a>	1
<a href="#">Web 2.0 Expo San Francisco 2009 - Co-produced by TechWeb &amp; O'Reilly...</a>	Web 2.0 Expo is a global annual gathering of technical, design, marketing, and business professionals who are building the next generation web.	<a href="http://sf.web2expo.com/">http://sf.web2expo.com/</a> - 49k <a href="#">Cache Copy</a>	1
<a href="#">Web 2.0</a>	Does "Web 2.0" mean anything? ... Tim says the phrase "Web 2.0" first arose in "a brainstorming session between ... conference, what "Web 2.0" seemed to mean ...	<a href="http://www.paulgraham.com/web20.html">http://www.paulgraham.com/web20.html</a> - 31k <a href="#">Cache Copy</a>	7
<a href="#">Enterprise Web 2.0   ZDNet.com</a>	Dion Hinchcliffe on leveraging the convergence of IT and the next generation of the Web. ... Unintended Uses, Enterprise Web 2.0, Identity, Innovation marketplace, ...	<a href="http://blogs.zdnet.com/Hinchcliffe/?cat=2">http://blogs.zdnet.com/Hinchcliffe/?cat=2</a> - 205k <a href="#">Cache Copy</a>	0
<a href="#">Web 2.0 Events: Co-produced by TechWeb and O'Reilly Conferences</a>	Web 2.0 Expo, co-produced by TechWeb and O'Reilly Media, is a conference and tradeshow for the rapidly growing ranks of designers and developers, product managers, ...	<a href="http://www.web2expo.com/">http://www.web2expo.com/</a> - 16k <a href="#">Cache Copy</a>	1

Obr. 4.10 Presúvanie výsledkov v Yoogle

### Otázky & odpovede

Medzi ďalšiu špecifickú kategóriu vyhľadávačov patria nástroje, pomocou ktorých sa dajú nájsť odpovede na otázky. Podporujú sociálne siete a rôzne možnosti interakcie. Medzi najznámejšie a najviac využívané Q&A (Questions & Answers, Otázky a odpovede) služby so všeobecným zameraním patrí Yahoo! Answers.

Služba Yahoo! Answers stavia na živej 179 miliónovej komunite. Okrem vyhľadávania a hodnotenia je možné aj prispieť vlastnou otázkou alebo odpoveďou.

Prezeranie otázok s odpoveďami sprostredkúva prehľadne usporiadaný adresár. Yahoo! Answers podporuje aj jednoduché a pokročilé vyhľadávanie pomocou kľúčových slov. Yahoo! motivuje aktivitu komunity systémom pridelovania bodov a úrovní. Body je možné získať tým, že odpoveď bude vybratá ako najlepšia alebo hlasovaním za odpovede.

Medzi ďalšie Q&A služby patria napr. Answer Bag, Ask Me Help Desk, Fluther a pod. Odlišujú sa najmä veľkosťou používateľskej komunity a niektoré umožňujú spoplatnenie odpovedí na otázky.

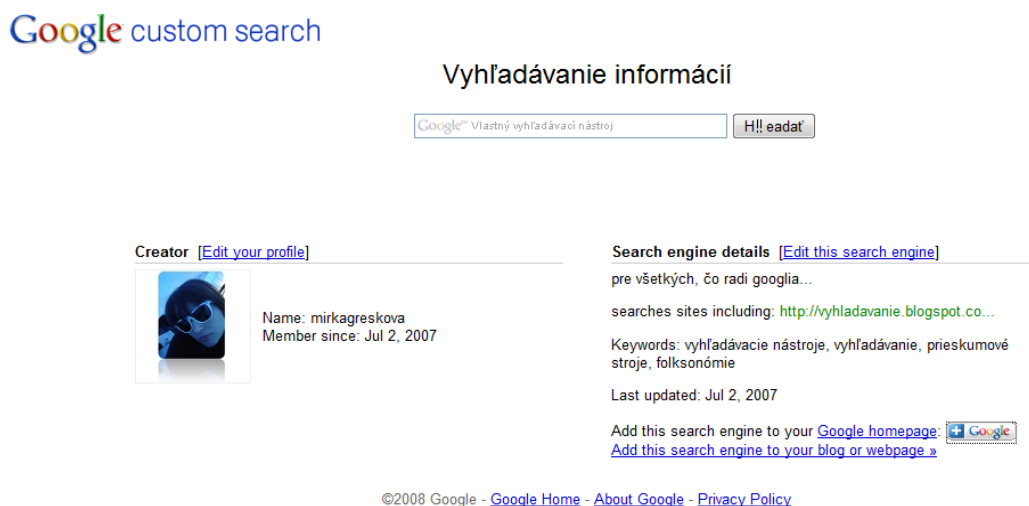
Okrem všeobecných, vznikli aj špecializované Q&A služby napríklad pre oblasť obchodu, vzdelávania a technológií (napr. Fixya). Q&A sú takisto súčasťou služieb na podporu sociálnych sietí (napr. LinkedIn). Mobilné služby Q&A sprostredkujú web aplikácie ako napr. Mosio a Twitter Answers.

Služby otázok a odpovedí majú z hľadiska vyhľadávania aj svoje negatíva. Prvým z nich môže byť zdĺhavosť hľadania otázok a odpovedí, ktorých je veľké množstvo a mnohé z nich nemajú výpočtové názvy ani obsah. Trpezlivosť si tiež vyžaduje čakanie na odpoveď na zadanú otázku. Nie vždy sa dá získať odpoveď, ktorú môžeme skutočne aj využiť. Dôvodom tohto javu je najmä fakt, že „expertom“ v oblasti sa v podstate môže stať ktokoľvek.

### Ďalšie nástroje sociálneho vyhľadávania

Medzi ďalšie služby sociálneho vyhľadávania patria napríklad Chacha alebo Anoox. Tieto nástroje prinášajú určité výhody, ale z hľadiska výkonu ich nemožno porovnávať s bežnými prieskumo-

vými strojmi. Alternatívny prístup k sociálnemu vyhľadávaniu predstavuje tvorba kustomizovaných vyhľadávačov (z angl. „custom search“). Používatelia majú možnosť vytvoriť si vlastné vyhľadávače, ktoré indexujú údaje vo vopred určenom zdroji alebo viacerých zdrojoch a tematických oblastiach. Takúto službu sprostredkávajú napríklad Eurekster Swiki, Rollyo, či Google Custom Search (obrázok 4.11).



Obr. 4.11 Príklad kustomizovaného vyhľadávača vytvoreného pomocou Google Custom Search

So sociálnym vyhľadávaním súvisia aj prístupy *umelej inteligencie*. Sociálne vyhľadavanie obohacuje prístupy umelej inteligencie v týchto trendoch vyhľadávania:

- personalizáciu,
- odporúčanie informácií,
- spracovanie prirodzeného jazyka,
- a inteligentných informačných agentov.

## 4.5 Prístupy umelej inteligencie

### 4.5.1 Personalizácia

Personalizácia z pohľadu vyhľadávania predstavuje uplatnenie techník umelej inteligencie, ktoré zabezpečujú prispôbenie vyhľadávacích nástrojov v prospech človeka na základe jeho profilu (potrieb) a vzorcov správania. Personalizácia si kladie za cieľ zníženie informačného preťaženia, vyhľadavanie relevantnejších výsledkov, zvýšenie efektivity interakcie a využitia času. Prvky personalizácie boli využívané aj v tradičných vyhľadávacích nástrojoch spred tridsiatich rokov. V pravom slova zmysle ju však pre širokú verejnosť nasadil v roku 2004 Eurekster. Využil na to prispôbovanie vyhľadávania na základe sociálnych sietí, správania ich členov a histórie nájdených výsledkov. Hneď po Eurekster nasledovali ďalšie vyhľadávacie služby ako Google, Yahoo!, AskJeeves (dnes Ask) a pod. Collarity predstavil nový prístup škálovania personalizácie od individuálneho, cez komunitný až po sociálny stupeň.

Vo všeobecnosti sa dá zhrnúť, že pri personalizácii sú aplikované mechanizmy založené na modelovaní správania používateľa, progresívnom učení sa v dynamickom prostredí a na základe predchádzajúcich skúseností, kontextuálnej analýze, web miningu (neurónové siete) a kolaboratívnom filtrovaní. Špecifické miesto v tejto oblasti predstavuje adaptívna personalizácia, t.j. profilovanie na základe automatizovaného vytvárania profilu bez nutnosti vstupu používateľa.

## 4.5.2 Odporúčanie informácií

Systémy na odporúčanie umožňujú spoznať nové informácie pomocou filtrovania obsahu, kolaboratívneho filtrovania alebo využitím oboch prístupov naraz (potom hovoríme o hybridnom odporúčaní informácií). Okrem toho však systémy na odporúčanie využívajú aj implicitné metódy, napr. monitorovanie správania používateľa (ukladanie záznamov o kúpe, kliknutiach, náhľadoch na produkty), zisťovanie záujmov ostatných členov v sociálnej sieti a pod. Všetky tieto dáta sú spracované a ovplyvnia, aký obsah bude na výstupe. V súčasnosti sú známe aplikácie s podporou odporúčania informácií na webe zamerané na hudbu, webové stránky, fotografie, videá, RSS a pod. Osobitne významnú úlohu zohráva odporúčanie produktov v súvislosti so stále rozmáhajúcim sa elektronickým obchodom. Príkladmi systémov na odporúčanie informácií sú Amazon, Last.fm (obrázok 4.12), Gnod a pod.

Obr. 4.12 Odporúčania hudby, podujatí a videí na Last.fm

### 4.5.3

## Inteligentní informační agenti

Inteligentní informační agenti sú aplikácie s podporou multiagentových systémov. Ich cieľom je zvýšiť relevanciu vyhľadovaných výsledkov a zjednodušiť prácu s informáciami. Zvyčajne integrujú pokročilé funkcie analýzy zdrojov, hodnotenia relevancie, personalizácie, tvorby informačných produktov a pod. Medzi inteligentných informačných agentov dostupných ako aplikácie na webe patria napríklad SearchBot. Ďalším príkladom je inteligentný informačný agent na vyhľadávanie informácií Copernic. Umožňuje nastavbové možnosti ako napr. upozorňovanie o zmenách vo výsledkoch vyhľadávania a zdrojoch, sumarizáciu textu, organizáciu výsledkov (triedenie, zoskupovanie), vytváranie vlastných anotácií, odstraňovanie duplicitných liniek, vytváranie vlastných kategórií a pod. Jednou z najzaujímavejších funkcií je sumarizácia, pomocou ktorej sú z textu extrahované kľúčové slová a abstrakty. CA dokáže monitorovať zmeny vo výsledkoch vyhľadávania alebo jednotlivých zdrojoch a zasielať notifikácie o zmenách.

### 4.5.3.1

## Spracovanie prirodzeného jazyka

NLP (Natural Language Processing, spracovanie prirodzeného jazyka) nachádza využitie v oblastiach informačného prieskumu a extrakcie, strojového prekladu, sumarizácie a rozpoznávania reči. Jednou z najatraktívnejších aplikačných oblastí sú chatboti. Zaujímavým príkladom je vyhľadávací nástroj AskVox, ktorý integruje prieskumový stroj s chatbotom (chatovacím robotom) a sociálnymi sieťami. Chatbot Vox dokáže odpovedať na otázky používateľa v prirodzenom jazyku pomocou hlasového výstupu.

Cestu využitia NLP, často v kombinácii so sémantickým vyhľadávaním, si vybrali aj viaceré vyhľadávacie služby ako Ask, Hakia, Lexxe, PowerSet a Wolfram Alfa.

NLP z používateľského hľadiska zjednodušuje prístup k informáciám. Neredukuje informačnú požiadavku na dotaz, ale umožňuje komunikovať s vyhľadávacím nástrojom v prirodzenom jazyku. Ďalším krokom je zadávanie dotazu pomocou reči. Takéto vyhľadávanie bude umožnené technológiou rozpoznávania reči, ktoré predstavuje odklon od najčastejšie prítomnej textovej interakcie.

## 4.6

## Vizuálne vyhľadávanie

Vizualizácia umožňuje zobrazenie entít v dvojrozmernom, trojrozmernom priestore alebo vo virtuálnej realite. Dôležitú úlohu pri výskume a vývoji v oblasti vizualizácie zohrala NSF (National Science Foundation, Národná agentúra pre výskum), keď v roku 1987 publikovala správu o vizualizácii vo výpočtovej technike pre vedu.

Vizualizácia výsledkov vyhľadávania umožňuje lepšiu orientáciu v problematike, pochopenie kontextu hľadaných informácií, môže podporiť tvorivé myslenie. Zvyšovanie interaktivity vizualizáciou výsledkov je sprostredkované rôznymi technikami. Podľa Kulesa, et al. (2008) sú tieto techniky triedené do základných kategórií: klasifikácia, organizácia výsledkov a doplnkové funkcie.

### 4.6.1

## Klasifikácia

Do prvej kategórie, klasifikácie, sú zaradené hierarchické klasifikácie, fazetové klasifikácie, zhľukovanie a sociálne „klasifikácie“. Príkladom vizualizácie pomocou viacúrovňovej fazetovej klasifikácie je mSpace (obrázok 4.13). V hornej časti obrázka nájdete fazety, ktorých prehľadávanie a kombinovanie MSpace podporuje. V príklade je výsledok kombinácie „Science & Technology“ a „Nanotechnology“.

The screenshot shows the mSpace website interface. At the top, there is a navigation bar with 'mSpace' logo and 'exploration reinvented' tagline. The main navigation includes 'mSpace', 'Projects', 'Research & Papers', and 'Who & How'. A search bar is located on the right. Below the navigation, there is a breadcrumb trail: '/ 1990s [1 more] / 1995 [2 more] / Science & Technology / Nanotechnology / (Story Title)'. A faceted search interface is displayed with four columns: 'Decade', 'Year', 'Theme', and 'Subject'. The 'Theme' column has 'Science & Technology' selected. The 'Subject' column has 'Nanotechnology' selected. The 'Story Title' column shows three results: 'Nanotechnology: Over-view of Uses:', 'Science: Nanotechnology Advances:', and 'Science: Nanotechnology: Mps Bemo...'. Below the faceted search, there is a message: 'You are currently browsing an online newsfilm archive'. The main heading is '/ Science & Technology / Nanotechnology (3 results)'. Below this, there is a sub-heading: 'Nanotechnology: Manipulating objects at the molecular scale'. A pagination bar shows '(1-3 of 3)' with 'Prev' and 'Next' buttons. The first result is 'Nanotechnology: Over-view of Uses:' with a thumbnail image of a person. The second result is 'Science: Nanotechnology Advances: More Funding Announced' with a thumbnail image of a robot.

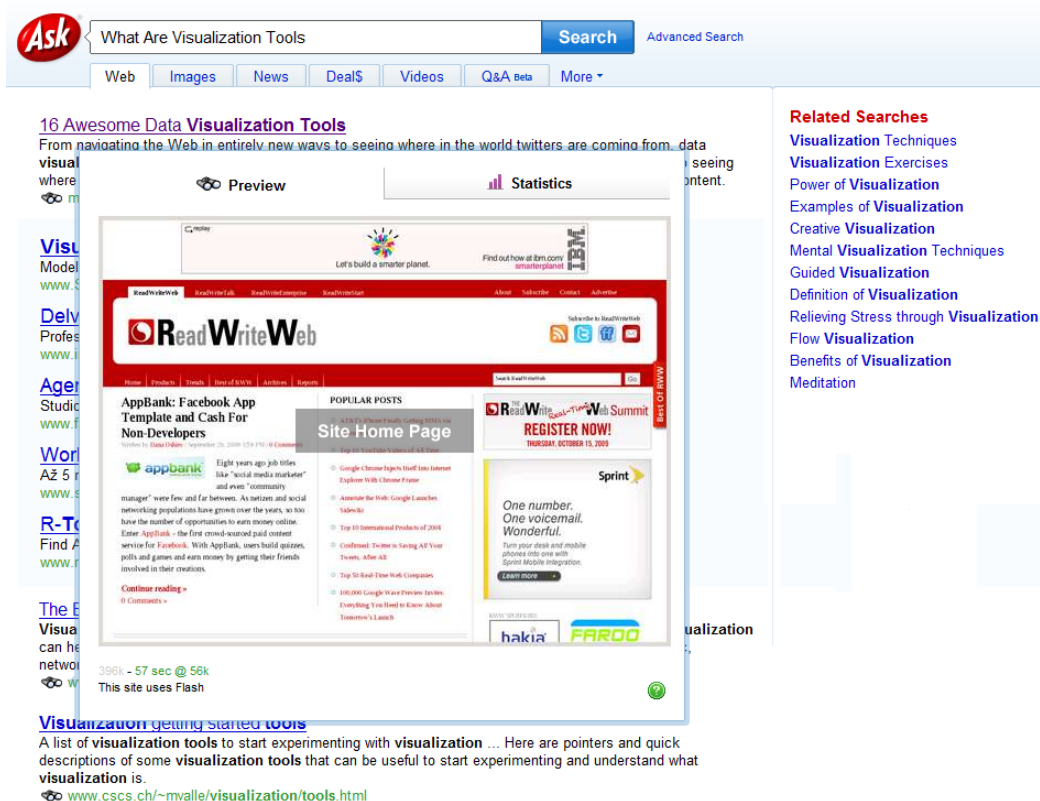
Obr. 4.13 Viacúrovňová fazetová klasifikácia pomocou mSpace

#### 4.6.2 Organizácia výsledkov

Do kategórie nástrojov s podporou vizualizácie pomocou klasifikácie radíme aj známy zhlukovací vyhľadávač Clusty. Výhoda týchto nástrojov spočíva v možnosti efektívneho zužovania výsledkov pomocou tematických zhlukov. Sociálna „klasifikácia“ je aplikovaná v nástrojoch s podporou relevantnej spätnej väzby (napríklad Yoope) a prostredníctvom folksonómii.

Ďalšiu kategóriu vizualizácie predstavujú náhľady, ale aj reprezentácia výsledkov vyhľadávania v 2D/3D priestore (Kules et al., 2008). Náhľady vo forme „momentiek“ (snapshots) poskytuje napríklad Ask.com (obrázok 4.14).





Obr. 4.14 Momentka stránky z výsledkov na Ask.com

Reprezentácia výsledkov v 2D priestore je sprostredkovaná samoorganizujúcimi sa mapami (SOM). Stromové mapy v kombinácii s hyperbolicými stromami využíva napríklad Kartoo. Vizualizáciu výsledkov vo virtuálnej realite umožňovali výstupy výskumných projektov ako Data Mountain, ktorý bol zameraný na organizáciu webových dokumentov. VR-VIBE vizualizoval výsledky pri dotazoch a ich relevanciu. Aktuálnejším príkladom je využitie CAVE na multimodálnu interakciu pri objavovaní zdrojov vo fonde digitálnej knižnice. Niektoré projekty organizácie výsledkov vo virtuálnej realite stroskotávajú buď na špeciálnych technických požiadavkách alebo neproduktívnosti.

### 4.6.3 Doplnkové funkcie

Do kategórie vizualizácie pomocou doplnkových funkcií patria náhľady faziety (napríklad už spomínaný mSpace), vizualizácia prepojení medzi údajmi, animácia, sémantické približovanie. Pri formulovaní požiadavky zvyčajne zadávame textový vstup vo forme dotazu, odkazu na kategóriu alebo tag. Iný prístup poskytuje napr. QBE (query-by-example, zadanie dotazu prostredníctvom príkladu), ktorý je aplikovaný v zaujímavom nástroji Retrievr (2006). Retrievr vyhľadáva vo Flickr na základe jednoduchého obrázku, ktorý je nakreslený namiesto textového dotazu (na obrázku 4.15 vľavo). Výsledkom sú vyhľadané obrázky a fotografie, ktoré obsahujú požadované farby a tvary. Okrem QBE je tiež známa technika query-by-humming, ktorá vyžaduje uloženie konkrétneho obsahu (obrázku, videa) na server služby. Na základe tohto obsahu sú vyhľadané podobné zdroje. Z hľadiska využiteľnosti je však tento prístup menej efektívny.

Ďalšie zaujímavé nástroje s podporou rôznych techník vizualizácie sú charakterizované v zdroji Blog.vyhľadavanie (Grešková, 2009).



## All images

Search by:  
Sketch • [Image](#)

From [javjuice](#)

From [Lou Rouge](#)

From [the nixonator](#)

From [ropeboy / sambr](#)

From [Reinhard Pantke](#)

From [Jean Sol Parre](#)

From [Lou Rouge](#)

From [Jorge Orte Tudela](#)

From [keymistress](#)

From [R.e.a.s.o.n.](#)

From [wmacphail](#)

From [magic\\_eye](#)

From [iamjon](#)

From [Bloggerati Star](#)

From [Coffee Mokka Break](#)

From [Eric Rolph](#)

From [theGolfer](#)

From [Precious Roy](#)

From [Reinhard Pantke](#)

From [Zickie](#)

**New!** Please help us rate sketches for The Art of retrievr! Only a million to go ...

Still new! You can search by uploading images now as well. Also: Your sketches have URLs! Send 'em around.

This is an experimental service. Please treat it nicely and send copious amounts of feedback! For some background, [read here](#).

retrievr relies on parts of the [System One platform](#).

Obr. 4.15 Retrievr vyhľadáva fotografie a obrázky na základe nakresleného obrázku s výberom z ponuky škály farieb a hrúbok čiary.

## 4.7 Zhrnutie

Vyhľadávanie vo webových systémoch a službách je založené na distribuovanom modeli vyhľadávania. Tento model prepája rozptýlené miesta a rozptýlené zdroje. Stále sa vyvíjajú algoritmy vyhľadávania, zlepšuje sa vizualizácia, sémantika a interakcia. V dominantnom vyhľadávači Google sa uplatňuje známy algoritmus PageRank založený na princípe popularnosti dokumentu. Okrem Google a GoogleScholar existujú však aj novšie vyhľadávacie nástroje v internete, napríklad Healia, WolframAlpha a rôzne iné, ktoré sa v interakcii približujú myšlienkovým procesom človeka, pracujú s čoraz bohatším kontextom a inteligentnými metódami spracovania informácií.

Vyhľadávanie informácií v prostredí webu dynamicky mení tvár. Počas mesiaca vznikne a zanikne niekoľko vyhľadávacích služieb. Keďže vyhľadávanie na internete sa stalo pre mnohých každodennou nevyhnutnosťou, rastie aj tlak na výskum a vývoj nových techník zabezpečovania relevantných a kvalitných informácií dôležitých pri bežných, ale aj strategických rozhodnutiach. Spolu s počtom používateľov vyhľadávacích nástrojov narastá aj snaha využiť tento kanál na marke-

tingové účely. Vznikla špecializovaná oblasť zameraná na marketing prostredníctvom prieskumových strojov – SEM, Search Engine Marketing.

Web 2.0 je nová koncepcia webu založená na spolupráci, tvorbe obsahu používateľom a bohatej interakcii a funkčnosti. V službách a produktoch sú zahrnuté poznatky o informačnom správaní používateľov. Obsah sa v produktoch rôzne kombinuje do personalizovaných „zmesí“ (mash-ups). Vzorce vyhľadávania sú tu charakterizované samostatným pohybom, viacnásobným využitím obsahu a spoluprácou. Novšie nástroje vyhľadávania v internete stelesňujú najmä princípy personalizácie, vizualizácie a spolupráce. Na základe toho sa rozlišuje vizuálne a sociálne vyhľadávanie.

V súvislosti so sociálnym vyhľadávaním a monitorovaním správania ľudí pri vyhľadávaní môžu byť prieskumovými strojmi agregované osobné dáta. Tieto môžu byť následne využité pri prispôbovaní služieb a výsledkov vyhľadávania. V súčasnosti je v mnohých prípadoch problematické určiť, ako sa spracúvajú stroje osobné dáta citlivé na ochranu. Týmto problémom je potrebné intenzívne sa zaoberať, aby sa zamedzilo negatívnym dopadom na jednotlivcov v kontexte súkromia.

Z novej úrovne služieb web 2.0 vznikla aj koncepcia knižnice 2.0. Ide o spojenie knižničných služieb s komunitami v elektronickom prostredí a s funkciami web 2.0. Knižnice sú súčasťou komplexných informačných prostredí. Na úrovni akademickej sa spájajú s komunitami študentov, učiteľov, vedcov. Na úrovni verejných knižníc sa v elektronickom prostredí spájajú s komunitnými centrami.

V súčasnosti sa nedá predpovedať, akým smerom sa vyvinú vyhľadávacie nástroje v budúcnosti. Možno vznikne úplne nová koncepcia reprezentácie, organizácie a vyhľadávania informácií, ktorá sa nebude v ničom podobať súčasným trendom. Predpokladáme, že vo sfére vyhľadávania v najbližších rokoch budú ovplyvňovať vývoj najmä sociálne vyhľadávanie, vizualizácia a umelá inteligencia.

## Literatúra

- BAWDEN, D. et al. 2007. Information Retrieval Curricula : contexts and perspectives. In Eds. MacFARLANE, A. et al. *Proceedings of the First International Workshop on Teaching and Learning of Information Retrieval* [online]. London : British Computer Society, January 10th 2007 [cit. 2010-01-12]. S. 55 – 60. Dostupné na internete: <<http://www.bcs.org/ewic/tlir2007>>.
- DEMPSEY, Lorcan. 2009. Always on : Libraries in a World of Permanent Connectivity. In *First Monday* [online]. January 2009, vol. 14, no. 1–5 [cit. 2010-01-11]. Dostupné na internete: <<http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2291/2070>>. ISSN 1396–0466.
- EZZY, E. 2006. Search 2.0 versus traditional search. In *ReadWriteWeb* [online]. 2006 [cit. 2008-03-05]. Dostupné na internete: <[http://www.readwriteweb.com/archives/search\\_20\\_vs\\_tr.php](http://www.readwriteweb.com/archives/search_20_vs_tr.php)>.
- GREŠKOVÁ, M. 2009. *Blog.vyhľadavanie* [online]. Last update: 2009-09-02 [cit. 2009-10-09]. Dostupné na internete: <<http://vyhladavanie.blogspot.com/search/label/Vizualiz%C3%A1cia>>.
- GROSSMAN, David A., FRIEDER, Ophir. 2004. *Information Retrieval : Algorithms and Heuristics*. Sec. ed. Dordrecht : Springer, 2004. 332 s. ISBN 1-4020-3004-5.
- KULES, B. et al. 2008. *From Keyword Search to Exploration : How Result Visualization Aids Discovery on the Web* [online]. University of Maryland; Advanced Computer Studies; HCI Lab, 2008 [cit. 2008-03-09]. Dostupné na internete: <<http://hcil.cs.umd.edu/trs/2008-06/2008-06.pdf>>.
- MAKULOVÁ, Soňa. 2002. *Vyhľadavanie v internete : problémy, východiská, postupy*. Bratislava : EL&T, 2002. 376 s. ISBN 80-88812-16-X.

- MALINOWSKI, Erik. 2007. Proto Social Network 'The Well' Runneth Over. In *Wired* [online]. 12.20.2007, vol. 16, no. 1 [cit. 2009-09-09]. Dostupné na internete: <[http://www.wired.com/techbiz/media/magazine/16-01/st\\_15thewell](http://www.wired.com/techbiz/media/magazine/16-01/st_15thewell)>.
- STEINEROVÁ, Jela. 2007a. Manifest o digitálnych knižniciach a využívanie informácií v informačnej spoločnosti. In *Itilib*. ISSN 1335-793X, 2007, roč. 11, č. 2, s. 15 – 19.
- STEINEROVÁ, Jela. 2007b. Relevancia vo vedeckej komunikácii a nové služby elektronického prostredia. In *CASLIN 2007*. Bratislava : UKB, 2007. 12 s. CD-ROM.
- STEINEROVÁ, Jela. 2007c. Informačná gramotnosť a relevancia : človek v informačnej spoločnosti. In *Infos 2007*. Knižnice – piliere vedomostnej spoločnosti. Prínos pamäťových inštitúcií k budovaniu vedomostnej spoločnosti. Stará Lesná, 16. – 19. apríla 2007. Bratislava : Spolok slovenských knihovníkov; Albertina Icome, 2007. 13 s. CD-ROM.
- STEINEROVÁ, Jela. 2007d. Komunikácia poznatkov v novej paradigme informačnej vedy. In *Itilib*. ISSN 1335-793X, 2007, roč. 11, č. 3, s. 5 – 13.
- STEINEROVÁ, Jela. 2007e. Zmeny vo využívaní informácií a nová paradigma informačnej vedy. In *Nová paradigma využívania informácií*. Zborník príspevkov zo seminára. 21.11.2007. Bratislava : Univerzita Komenského v Bratislave, 2007. ISBN 978-80-223-2415-1. s. 58 – 71. CD-ROM.
- STEINEROVÁ, Jela. 2008a. Informačné systémy v novej paradigme vzdelávania a výskumu. In *Itilib*. ISSN 1335-793X, 2008, roč. 12, č. 2, s. 4 – 12.
- TENENBAUM, J. M. 2005. *AI Meets Web 2.0 : Building the Web of Tommorrow, Today* [online]. CommerceNet Labs Technical Report 05-07, December 2005 [cit. 2007-05-13]. Dostupné na internete: <<http://wiki.commerce.net/images/a/a2/CN-TR-05-07.pdf>>.
- ZIMMER, M. 2008. The Externalities of Search 2.0 : The Emerging Privacy Threats when the Drive for the Perfect Search Engine meets Web 2.0. In *First Monday* [online]. March 2008, vol. 13, no. 3–3 [cit. 2008-03-17]. Dostupné na internete: <<http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2136/194>>. ISSN 1396–0466.

## Zoznam webových sídel

Alexa Internet <a href="http://www.alex.com/topsites">http://www.alex.com/topsites</a>	AskVox <a href="http://www.askvox.com/">http://www.askvox.com/</a>
Amazon <a href="http://www.amazon.com/">http://www.amazon.com/</a>	Bing <a href="http://www.bing.com/">http://www.bing.com/</a>
Anoox <a href="http://www.anoox.com/">http://www.anoox.com/</a>	Bubl <a href="http://bubl.ac.uk/">http://bubl.ac.uk/</a>
Answer Bag <a href="http://www.answerbag.com/">http://www.answerbag.com/</a>	Chacha <a href="http://www.chacha.com/">http://www.chacha.com/</a>
Ask <a href="http://www.ask.com/">http://www.ask.com/</a>	Clusty <a href="http://clusty.com/">http://clusty.com/</a>
Ask Me Help Desk <a href="http://www.askmehelpdesk.com/">http://www.askmehelpdesk.com/</a>	Collarity <a href="http://www.collarity.com">http://www.collarity.com</a>

Copernic  
<http://www.copernic.com/>

Delicious  
<http://delicious.com/>

DMOZ (Open Directory Project)  
<http://www.dmoz.org/>

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

Eurekster Swiki  
<http://www.eurekster.com/>

Evri  
<http://www.evri.com/>

Facebook  
<http://www.facebook.com/>

FactBites  
<http://www.factbites.com/>

Fixya  
<http://www.fixya.com/>

Flickr  
<http://www.flickr.com/>

Fluther  
<http://www.fluther.com/>

Google  
[www.google.com](http://www.google.com)  
[www.google.sk](http://www.google.sk)

Google Custom Search  
<http://www.google.com/cse/>

Gnod  
<http://www.gnod.net/>

Hakia  
<http://www.hakia.com/>

Healia  
<http://www.healia.com/>

IPL2  
[www.ipl.org](http://www.ipl.org)

iRazoo  
<http://www.irazoo.com>

Kahzam  
<http://www.kahzam.com/>

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

Last.fm  
<http://www.last.fm/>

Lexxe  
<http://www.lexxe.com/>

Linkedin  
<http://www.linkedin.com/>

Mahalo  
<http://www.mahalo.com>

Mosio  
<http://www.mosio.com/biz/>

mSpace  
<http://mspace.fm/>

MySpace  
<http://www.myspace.com/>

National Science Foundation  
<http://www.nsf.gov/>

NSDL  
<http://nsdl.org/>

O'Reilly  
<http://oreilly.com/>

Powerset  
<http://www.powerset.com/>

RedZ  
<http://redz.com/>

Retrievr  
<http://labs.systemone.at/retrievr>

Rollyo  
<http://www.rollyo.com/>

SearchBot  
<http://www.searchbot.net/>

SearchCube  
<http://search-cube.com/>

Search Engine Watch  
[www.searchenginewatch.com](http://www.searchenginewatch.com)

Search Engine Guide  
[www.searchengineguide.com](http://www.searchengineguide.com)

Technorati  
<http://technorati.com/>

The WWW Virtual Library  
<http://vlib.org/>

Twitter  
<http://twitter.com/>

Twitter Answers  
<http://twitter.com/answers>

Viewzi  
<http://www.viewzi.com/>

Vivisimo  
<http://vivisimo.com/>

Wikipedia  
[www.wikipedia.org](http://www.wikipedia.org)

WoframAlpha  
[www.wolframalpha.com](http://www.wolframalpha.com)

Yahoo!  
<http://m.www.yahoo.com/>

Yahoo! Answers  
<http://answers.yahoo.com/>

Yahoo! Answers Blog  
<http://yanswersblog.com/>

Yoogle  
<http://www.yoogle.net>

YouTube  
<http://www.youtube.com/>

# 5

## Relevancia

Relevanciu možno skúmať v širokom zmysle slova ako všeobecný pojem, ktorý skúma aj filozofia, komunikácia, logika, psychológia, právo. Relevancia sa často používa intuitívne. Znamená to, že hovoríme o niečom, čo patrí k danej veci, problému. Je to vhodnosť, užitočnosť, vecná príslušnosť. Samotný pojem sa začal používať v 16. storočí v Škótsku v oblasti práva ako niečo zákonne vhodné (Saracevic 1996). Všeobecne sa relevancia sa používa na vyjadrenie vhodnosti, prepojenosti a použiteľnosti.

Najvýznamnejšie teórie, ktoré sa pri skúmaní relevancie v informačnej vede využili, boli filozofia a komunikácia.

Vo filozofii skúmal relevanciu fenomenológ Alfred Schutz ako vlastnosť, ktorá určuje vzťahy v zložitom sociálnom svete (Wilson 2000b, Saracevic 2007). Vychádza z bytia človeka v sociálnom svete, pričom sa najčastejšie koncentruje na tému v určitom kontexte (horizonte) ako je fyzický priestor, sociálny vývoj a skúsenosti. Z toho vyvodzuje tri druhy relevancie: tematickú, interpretačnú a motivačnú. Predpokladá, že existuje celý systém navzájom sa dopĺňujúcich relevancií v každodennom informačnom správaní človeka.

Z hľadiska komunikácie sa v informačnej vede uplatnila najmä komunikačná teória Sperbera a Wilsona (Saracevic 2007). Tu sa zdôrazňuje kognitívny princíp, ktorý znamená, že ľudská myseľ sa organizuje s cieľom maximalizovať relevanciu. Komunikatívny princíp súvisí s poukazovaním na niečo (ostenzia) a zameranosťou (intenciou). Relevancia je včlenená do predvídania správania.

Relevanciou sa v informačnej vede zaoberalo množstvo autorov. Najvýznamnejším z nich je Tefko Saracevic, ktorý už v roku 1976 definoval relevanciu ako **vzťah**. Na základe analýzy mnohých definícií autorov vytvoril všeobecnú definíciu relevancie ako

*Relevancia je A B existujúce medzi C a D a určené E.*

Za premenné A, B, C, D, E možno dosadiť rôzne výrazy: Napríklad A – miera, stupeň, dimenzia, vzťah, B – zhoda, spojenie, užitočnosť, C – dokument, text, informácia, D – dotaz, požiadavka, uhol pohľadu, využitie informácií, E – osoba, posudzovateľ, používateľ, informačný špecialista.

### 5.1

## Všeobecné vlastnosti relevancie

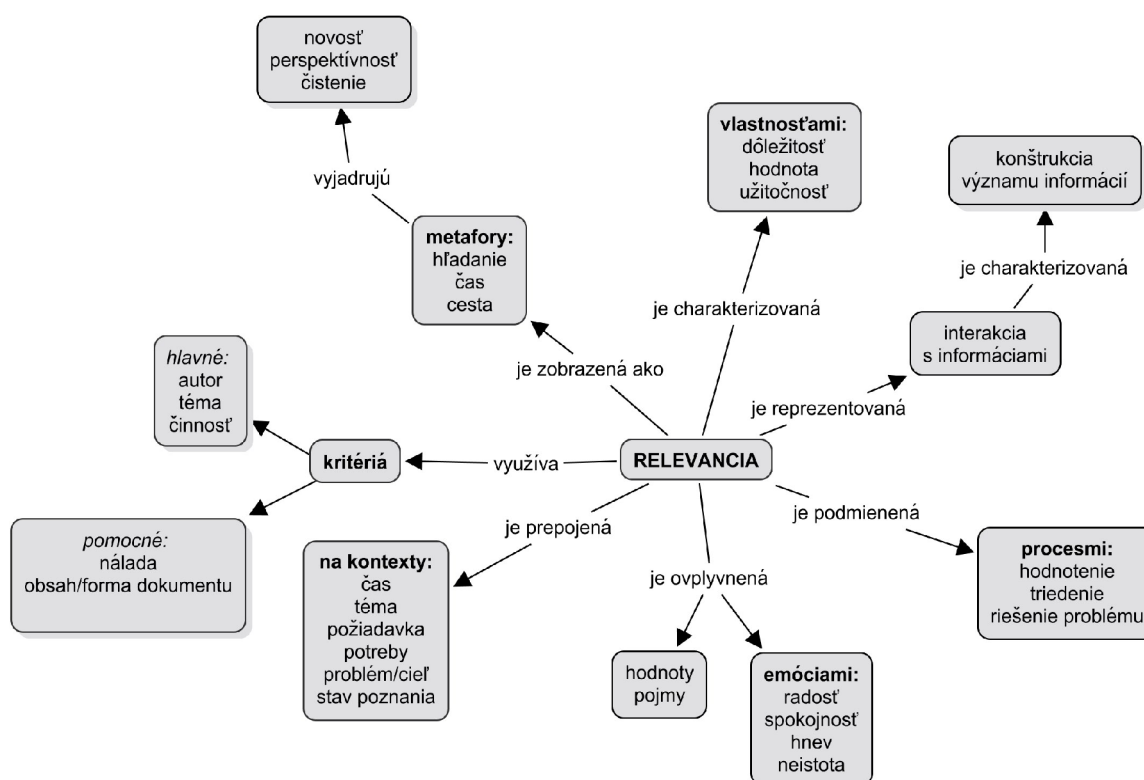
Vlastnosti relevancie možno vysvetliť podľa Saracevica (1996) takto:

1. Relevancia vyjadruje a predpokladá vzťah, ktorý je často komunikačný.
2. Relevancia sa vyznačuje zámerom, súvisí s cieľmi, rolami, očakávaniami, motiváciou.
3. Kľúčovú úlohu pri relevancii zohráva kontext ako východisko aj cieľ hodnotenia informácií.
4. Relevancia je založená na myšlienkovom procese vyvodzovania (inferencie).

5. Relevancia je dynamická a interaktívna. Interpretácie relevancie sa menia s vývojom poznania človeka.

Tieto vlastnosti relevancie komplikujú možnosti jednoduchého vysvetlenia jej podstaty a prejavov. Novšia definícia relevancie ju určuje ako kritérium, ktoré odráža efektívnosť výmeny informácií medzi ľuďmi, prípadne medzi človekom a systémami a zdrojmi informácií (informačnými objektmi) v komunikačnom vzťahu a v kontexte.

Relevanciu môžeme definovať ako vzťah medzi hľadanými a nájdenými informáciami (Šušol 2007) alebo ako súbor viacrozmerných spojení človeka a informácií pri vytváraní významu (Steinerová 2007). Definovanie relevancie je znázornené na pojmovej mape na obrázku 5.1.



Obr. 5.1 Pojmová mapa Relevancia (Steinerová, Grešková, Šušol 2007)

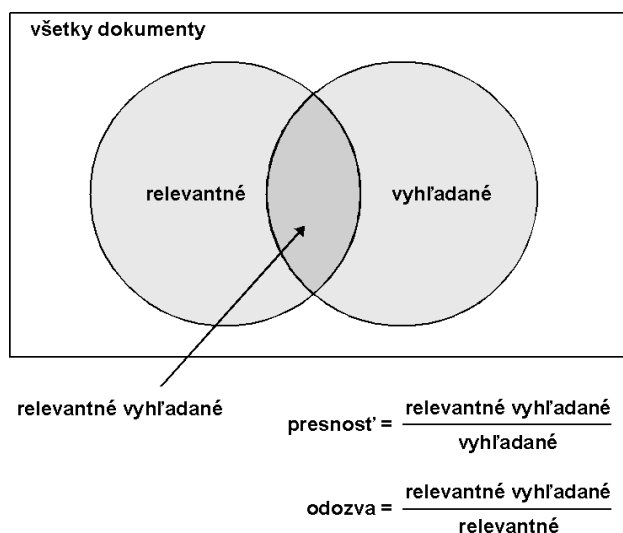
## 5.2 Systémový pohľad na relevanciu

Relevanciu definujeme ako vzťah medzi vyhľadanými informáciami, prípadne dokumentmi a informačnou požiadavkou posudzovaný konkrétnym človekom. Vo vyhľadávaní informácií sa pôvodne využívalo čierno-biele, jednoduché posudzovanie vhodnosti informácií – vyhľadané, nevyhľadané, relevantné, nerelevantné. Takéto posudzovanie relevancie súvisí aj s vývojom informačných prieskumových systémov. V začiatkoch ich vývoja, ktoré môžeme vidieť už v roku 1945 pri predstave Vannevara Busha o systéme Memex, bola relevancia považovaná za efektívnosť samotných informačných systémov. Vedci sa pri budovaní informačných systémov snažili merať účinnosť vyhľadávania prostredníctvom mier odozvy (úplnosti) a presnosti (pôvodne nazvanej relevanciou). **Odozva** (recall) je mierou efektívnosti pri vyhľadávaní informácií, ktorá vyjadruje pomer vyhľadaných relevantných

informácií k relevantným. Vzťahuje sa k celkovému počtu relevantných dokumentov (pomer počtu relevantných vyhladaných dokumentov k celkovému počtu dokumentov v súbore, o ktorých sa predpokladá, že sú relevantné. Počet relevantných dokumentov v súbore sa odhaduje rôznymi metódami. **Presnosť** (precision) je mierou účinnosti pri vylučovaní nerelevantných prvkov (miera čistoty vyhľadávania). Presnosť možno vypočítať v rôznych aspektoch odozvy. Ak sme odhadli, že v súbore existujú dva relevantné dokumenty a našli sme jeden relevantný dokument, odozva bola 50%. Presnosť v tomto aspekte je tiež 50%, pretože sme vyhládali dva dokumenty a jeden z nich je relevantný. Na to, aby sme dosiahli 100% odozvy musím pokračovať vo vyhľadávaní až kým nenájdem obidva relevantné dokumenty. Tieto dve miery relevancie boli výsledkom experimentov v Cranfiede a testov SMART (Steinerová 1996). Pri testovaní systému SMART ako výskumu prieskumového systému Gerarda Saltona bola určená najmä relevantná spätná väzba.

Pri meraní presnosti si kladieme takúto otázku: *Aká časť vyhladaných výsledkov je relevantná pre informačnú potrebu?* Pri meraní odozvy sa pýtame: *Akú časť relevantných výsledkov v súbore systém vyhládaval?*

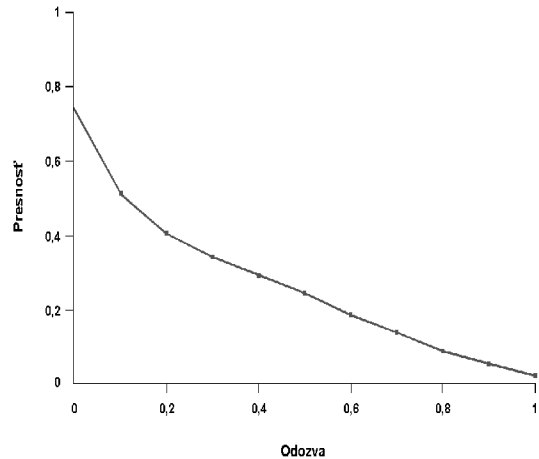
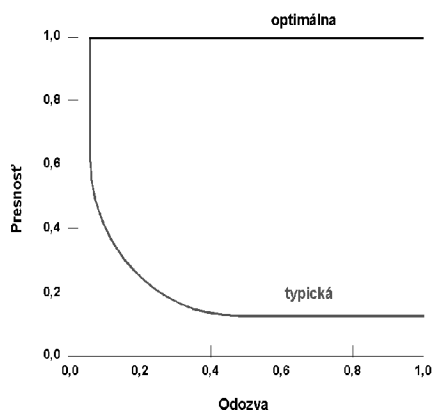
Na obrázku 5.2 je znázornená situácia medzi relevantnými a vyhladanými dokumentmi a miery relevancie prostredníctvom odozvy a presnosti.



Obr. 5.2 Vzťahy medzi relevantnými a vyhladanými dokumentmi, odozva a presnosť (podľa Grossman, Frieder 2004)

V dokonalom systéme by sa vyhľadávali iba relevantné informácie. V reálnych podmienkach je však vzťah medzi odozvou a presnosťou inverzný. Grafické znázornenie presnosti v rôznych momentoch odozvy sa nazýva krivka presnosti a odozvy. To si môžeme zobrazit' na obrázku 5.3





Obr. 5.3a Vzťahy medzi odozvou a presnosťou (optimálny, typický) (podľa Grossman, Frieder 2004)  
 Obr. 5.3b Graf presnosť/odozva (podľa Manning et al. 2009)

V 60. rokoch sa relevancia skúmala ako kritérium na meranie vzťahov, presnosti, vhodnosti, hodnoty (napr. Rees 1967, Cuadra, Katter 1967) (Grossman Frieder 2004). Prvé skutočné rozpoznanie relevancie v informačnej vede sformuloval Kent et al. (1955).

Systémový pohľad zdôrazňuje relevanciu ako vlastnosť systému, ktorú možno ovplyvniť reprezentáciou, organizáciou a internými procesmi systému. Systémový prístup dominoval v začiatkoch informačného prieskumu. Výkon systému ovplyvňuje vnútorný model organizácie informácií a algoritmickeho vyvodzovania. Slabosťou pôvodného systémového princípu je modelovanie používateľa, chýbajúca interakcia a dynamika a kontext.

Relevancia a vyhľadávanie v systémovom kontexte sú však podmienené aj technologickým vývojom, ktorý možno rozdeliť historicky na tri etapy – etapu sálových počítačov (dominancia IBM), etapu osobných počítačov (dominancia Microsoftu) a etapu internetu a rôznych verzií webu (dominancia Google). V novej etape elektronického prostredia sa aj systémová relevancia zmenila na interaktívnu, intuitívnu a viackriteriálnu. Čoraz viac sa približuje prirodzeným princípom posudzovania relevancie človekom.

### 5.3 Používateľský pohľad na relevanciu

Hodnotenie relevancie je súčasťou kognitívnych schopností človeka od začiatkov jeho vývoja. Ľudia ju vlastne vždy používali intuitívne. Podľa Saracevica (1996) je to vstavaný mechanizmus kognície, ktorý umožňuje človeku vycítiť, čo je preňho vhodné.

Z používateľského hľadiska upozornil na situačnú relevanciu Wilson (1973). Relevancia podľa neho závisí od situácie, v ktorej sa používateľ nachádza a od informačnej potreby. Dôležitá je jedinečná situácia jednotlivca. Preto treba skúmať ako ľudia používajú informácie a ako sa pri využívaní informácií mení myslenie a ďalšie spracovanie informácií. Logickú relevanciu určil Cooper (1971) ako tematickú vhodnosť vychádzajúcu z tematického zamerania informačnej potreby.

Prelom pri nazeraní na relevanciu nastal v rámci kognitívnej paradigmy vyhľadávania informácií (Steinerová 1996). Významný bol v 90. rokoch článok Schamberovej, Eisenberga a Nilana (1990) definujúci situačnú relevanciu, dynamiku a multidimenzionálnosť relevancie. Vznikla celá Syrakúzska škola (Syracuse University) skúmania relevancie. Psychologickú koncepciu relevancie predstavil Harter (Harter 1992). Relevanciu považuje za psychologickú kategóriu uskutočňujúcu sa v mysli používateľa a prejavujúcu sa v správaní používateľa.

Maglaughlinová a Sonnenwaldová (2002) v rámci používateľského pohľadu rozlišujú psychologickú, situačnú a subjektívnu relevanciu. Relevanciu autori určujú ako proces v rámci rôznych štádií riešenia problému a definujú rôzne stupne relevancie.

Ingwersen a Järvelin (2005) definujú relevanciu ako odhad užitočnosti témy alebo úžitku informačných zdrojov realizovaný kognitívnym aktérom alebo aj informačným systémom v danom čase a situácii. Môže sa dynamicky meniť v čase a rozlišuje sa objektívna a subjektívna relevancia. Je súčasťou interaktívneho informačného prieskumu.

Novšie koncepcie relevancie dokazujú, že je to proces a že sa stáva súčasťou každodenného informačného správania človeka a predstavuje špeciálnu skúsenosť (Anderson 2006).

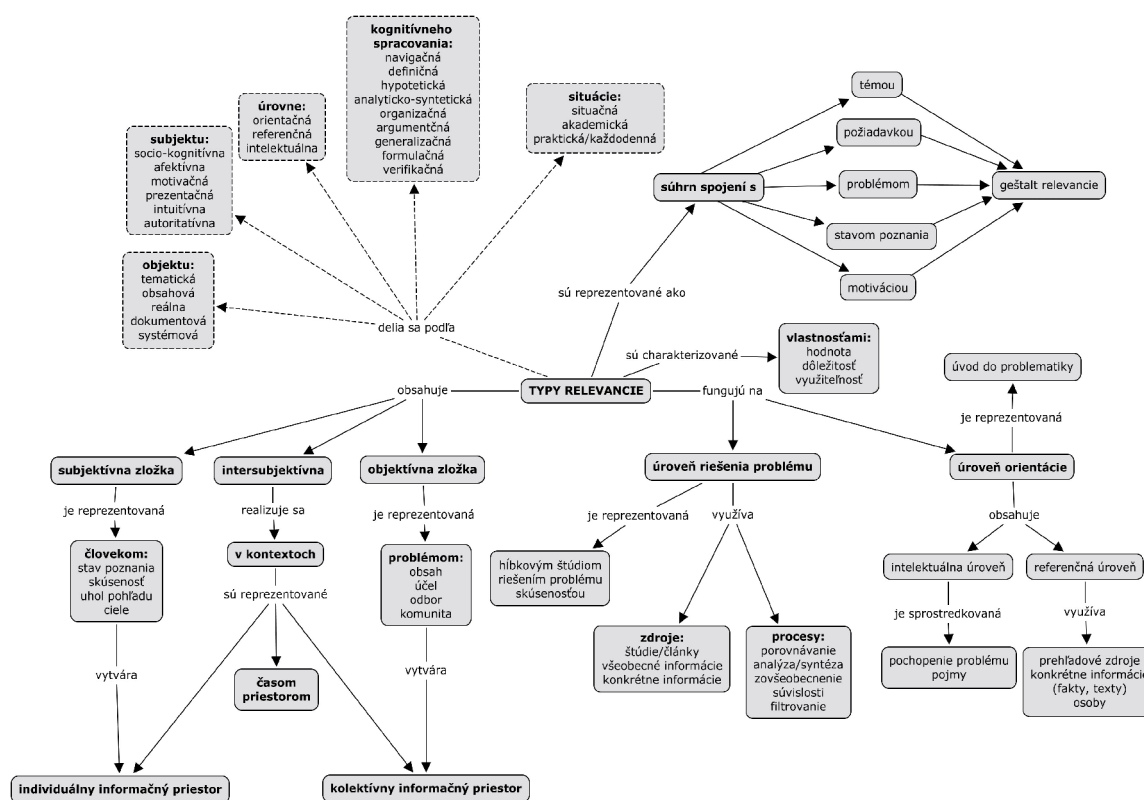
Saracevic sa viackrát v sumarizujúcich interpretáciách relevancie v priebehu takmer 30 rokov (1976 – 2007) (Saracevic 2007a) pokúsil o prepojenie používateľského a systémového pohľadu na relevanciu. P. Borlundová (Borlund 2003) skúmala situačnú relevanciu a dokázala, že pri posudzovaní relevancie funguje dynamika a viacrozmernosť kombinácie rôznych kritérií. Vo svojom modeli zdôrazňuje prepojenie relevancie s informačnou potrebou. Komplexné štúdie o výskumoch relevancie v informačnej vede spracovali Froelich (1994), Mizzaro (1997) a Saracevic (2007b).

## 5.4 Základné koncepcie a typy relevancie v informačnej vede

V priebehu vývoja informačnej vedy a praxe informačného prieskumu sa rozvinuli tieto koncepcie relevancie: systémová, komunikačná, situačná, psychologická, interaktívna a procesuálna. Relevancia je potom komplex vzťahov medzi informačnou potrebou človeka a informáciami získavanými z informačného prostredia. Podľa typu tohto vzťahu sa rozlišujú tieto druhy relevancie (Saracevic 2007a):

1. systémová, algoritmickej relevancia – vyjadruje vzťah medzi dotazom a informačnými objektmi. Dôležitý je spôsob reprezentácie a organizácie informácií. Nezohľadňuje sa používateľ, ale vzťah medzi vyjadrenou požiadavkou a vyhládanými dokumentmi. Prieskumové systémy hodnotia iba systémovú relevanciu. Nie je dôležité, či sa vyhladané objekty zhodujú na kognitívnej úrovni a či budú užitočné.
2. tematická, predmetová relevancia – vyjadruje vzťah medzi predmetom alebo témou v dotaze a témou alebo predmetom vo vyhladávaných textoch. Kritériom vyvodenia relevancie je téma/predmet. Tento typ relevancie je pri vyhladávaní informácií najčastejší. Pri tematickej relevancii chýba kontext potrieb a situácie. Je založená na pevnom predpoklade o vzťahu medzi témou dokumentu a skúmaným problémom.
3. kognitívna relevancia, pertinencia – vyjadruje vzťah medzi stavom poznania, kognitívnu informačnou potrebou a vyhladanými textami/informáciami v systéme. Pri vyvodzovaní kognitívnej relevancie je kritériom kognitívna korešpondencia, novosť a kvalita informácií. Pertinencia je v informačnom prieskume definovaná ako subjektívna relevancia. Oproti formulovanej požiadavke (dotazu) pri systémovej relevancii a informačnej požiadavke pri tematickej relevancii pri pertinencii dominuje informačná potreba.
4. situačná relevancia alebo užitočnosť – vyjadruje vzťah medzi situáciou, úlohou, problémom a vyhladanými textami/informáciami zo systému. Kritériom relevancie je tu užitočnosť pri rozhodovaní a vhodnosť informácií pri riešení problémov. Pri posudzovaní relevancie sa znižuje neurčitnosť v situácii.
5. motivačná alebo afektívna relevancia – vyjadruje vzťahy medzi zámermi, cieľmi, motiváciou používateľa a vyhladanými textami/informáciami. Kritériom vyvodenia relevancie je spokojnosť, úspech, dosiahnutie cieľa.

Typy relevancie sú otvorené. Ako príklad si môžeme uviesť pojmovú mapu (obrázok 5.4), ktorá znázorňuje niekoľko typov relevancie, rôzne úrovne relevancie (objektívna, subjektívna) a dva stupne relevancie (na úrovni orientácie, na úrovni riešenia problému).



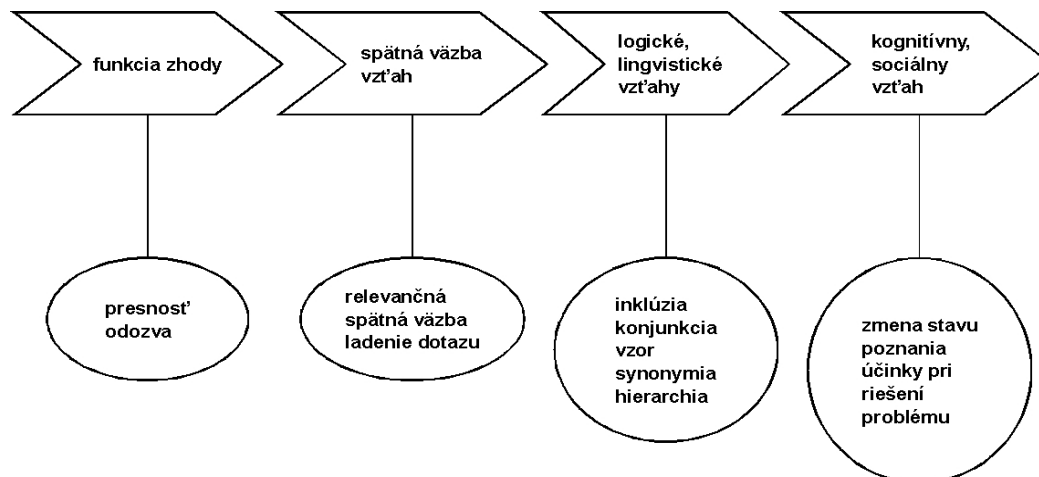
Obr. 5.4 Typy relevancie – pojmová mapa (Steinerová, Grešková, Šušol 2007)

Výskumy relevancie za posledných 30 rokov klasifikuje Tefko Saracevic (2007a) vo svojej štúdií podľa predmetu výskumu.

- Štúdie zamerané na kritériá relevancie. Základná otázka je *Čo si ľudia všimajú na informáciách, aby vedeli hodnotiť relevanciu?*
- Štúdie sledujúce dynamiku a stabilitu relevancie. Otázky sú zamerané na *odlišnosti posudzovania relevancie v závislosti od času u jedného používateľa a pri jednej úlohe*
- Štúdie zamerané na relevantnú spätnú väzbu. Otázky výskumu riešia, *ktoré faktory ovplyvňujú proces relevantnej spätnej väzby a ktoré faktory ovplyvňujú posudzovanie relevancie*
- Štúdie skúmajúce individuálne rozdiely v posudzovaní relevancie. Otázky výskumu sú: *Aké veľké sú individuálne rozdiely v posudzovaní relevancie, aký vplyv majú na proces posudzovania*
- Štúdie skúmajúce tematickú relevanciu. Skúmajú sa otázky, či ľudia posudzujú relevanciu iba podľa témy
- Štúdie skúmajúce binárnosť relevancie. Otázky riešia problémy, či je posudzovanie relevancie binárne (relevantné, nerelevantné) alebo sa využíva viac stupňov relevancie
- Štúdie skúmajúce „nezávislosť“ relevancie. Otázky výskumov sú: *Sú informácie hodnotené nezávisle od seba a má poradie hodnotenia jednotlivých informácií vplyv na relevanciu?*
- „Jednotnosť“ a „nejednotnosť“ relevancie. Výskumy odpovedajú na otázky: *Je posudzovanie relevancie jednotné pri jednotlivých hodnoteniach? Ako ovplyvňuje nejednotnosť posudzovania relevancie výsledky informačného prieskumu.*

## 5.5 Vývoj názorov na relevanciu v informačnej vede

Zjednodušene možno znázorniť interpretovanie koncepcie relevancie v informačnej vede na obrázku 5.5. Vývoj prešiel od výlučne systémovej interpretácie k zložitejším predstavám v rámci používateľskej relevancie.



Obr. 5.5 Vývoj interpretácie relevancie v informačnej vede

## 5.6 Relevancia v sieťovom elektronickom prostredí

Novšie výskumy informačnej vedy a digitálnych knižníc prepájajú systémove a používateľské prúdy relevancie.

Posudzovanie relevancie v elektronickom prostredí je podmienené *rozptýlenosťou a premenlivosťou kritérií*. Základom je postupné mozaikovitité skladanie významu informácií z rozptýlených zdrojov. Relevancia je v elektronickom prostredí podporená pokročilými technologickými funkciami rozhraní a prieskumových nástrojov. Pridaná hodnota relevantných informácií vzniká väzbami v kontextoch, pokročilým vyhľadávaním a inteligentnými funkciami systémov.

Na základe výsledkov projektu VEGA 1/2481/05 na KKVI sme charakterizovali nový model relevancie 2.0 (Steinerová, Grešková, Šušol 2007). Cieľom prieskumu relevancie bolo odhaliť subjektívne predstavy doktorandov o relevancii, vnímanie kritérií relevancie, postupy pri posudzovaní relevancie a triedení informácií a určiť odlišnosti v posudzovaní relevancie v elektronickom prostredí.

Respondenti potvrdili časté využívanie elektronických zdrojov, pričom najvýznamnejšie faktory, ktoré na to vplyvajú, sú *téma, disciplína a úloha*. Tradičné zdroje preferujú kvôli intelektuálnej hĺbke, stabilite, predvídateľnosti. V elektronickom prostredí oceňujú dostupnosť, rýchlosť, prepojitelnosť, vyhľadateľnosť, multimediálnosť a informácií. Mnohí tiež zdôraznili *odlišný spôsob vnímania elektronického textu*. Pritom je konštrukcia významov z elektronických textov komplikovanejšia. Spoločným menovateľom v oboch prostrediach je rozhodovanie o prijatí alebo odmietnutí informácie založené na overovaní ich spoľahlivosti a kvality. Elektronické prostredie podporuje toto rozhodovanie o relevancii prepojeniami a tematickými kontextmi, interakciami a komunikáciou.

Novší model relevancie v sieťovom informačnom prostredí vychádza z takých funkcií spracovania informácií ako *flexibilita, navigácia, vizualizácia, nelineárnosť*. Dôležitá je aj možnosť kolektív-

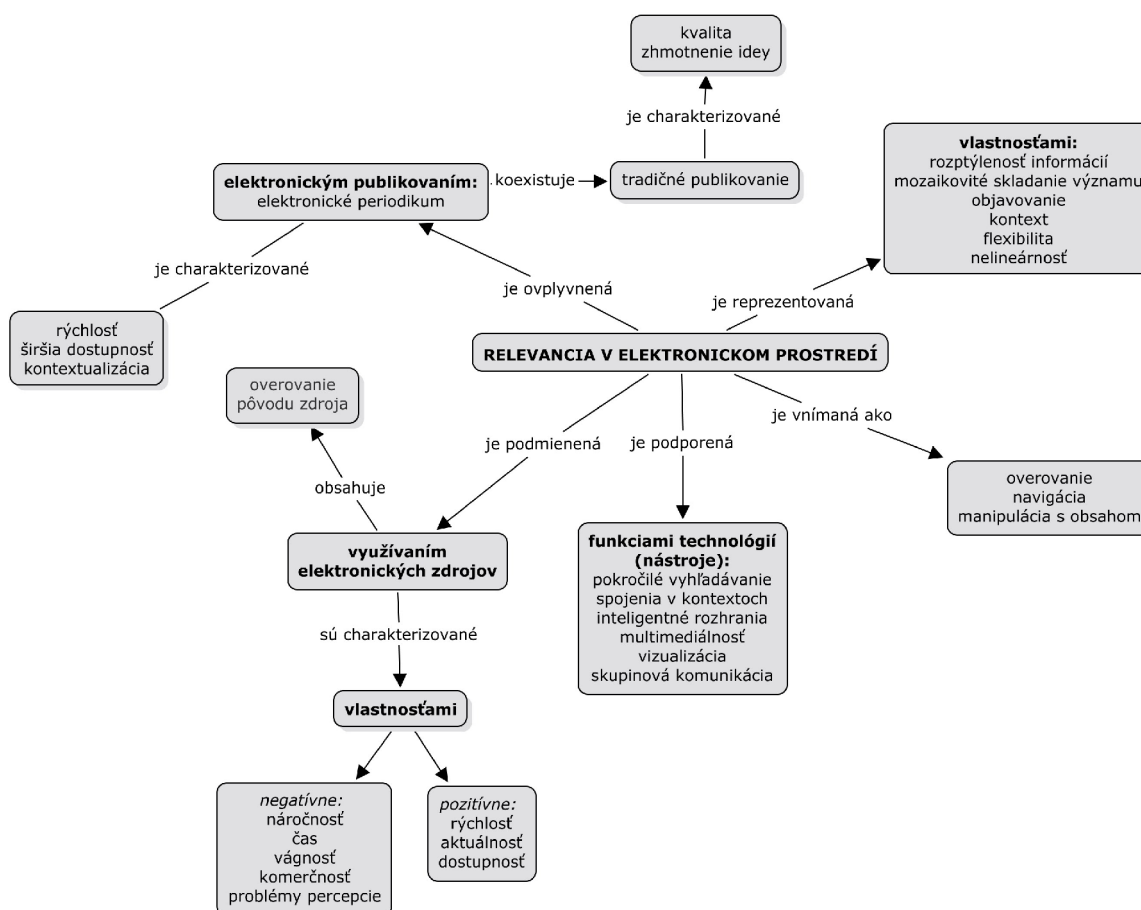
neho spracovania informácií a spoločného rozhodovania o relevancii. Relevanciu obohacujú pocity radosti a spokojnosti a hnevu a neistoty. Elektronické prostredie umožňuje uplatniť tvorivosť a objaviť nové informácie aj pri posudzovaní relevancie.

Pri organizácii informácií sa najčastejšie používajú *témy, typy informácií a kontexty*. Princípy organizácie informácií sú hierarchia, asociácie a aspekty (fazety). Cieľom kategorizácie informácií pri posudzovaní relevancie je určiť kontext a znížiť kognitívnu záťaž. Respondenti zdôraznili *viac kritérií* pri posudzovaní relevancie a intuitívne rozhodovanie.

Nový model relevancie využíva všetky možnosti bohatej kontextualizácie elektronického prostredia. Relevancia tu spája *prvotné hodnotenie s hypertextovými prepojeniami a odporúčaniami v pokročilých systémoch*. Relevancia v elektronickom prostredí približuje procesy hodnotenia a organizovania informácií k samotnému kognitívnemu procesu riešenia problému. Za explicitné zložky relevancie možno považovať pojmy a kategórie prejavujúce sa v kritériách hodnotenia informácií. Implicitné zložky obsahujú hodnoty, kontexty a informačné štruktúry.

Výsledky výskumu tiež ukazujú, že relevancia je *stratifikovaná (rozvrstvená), intuitívna aj interaktívna*. Pri kritériách dominujú témy, typy problémov a dostupnosť informácií.

Relevancia je aj v elektronickom prostredí vnorená v pracovných úlohách a kontextoch. V praktickom správaní pri posudzovaní relevancie zaznamenávame aj neformálne a spontánne odhalenie významu informácií. Na obrázku 5.6 znázorňujeme relevanciu v elektronickom prostredí.



Obr. 5.6 Relevancia v elektronickom prostredí – pojmová mapa (Steinerová, Šušol, Grešková 2007)

*Relevancia 2.0* vychádza z nových možností webu označovaných ako web 2.0, keď technológia podporuje kolaboráciu, participáciu a interakciu. Je založená na procese vytvárania zmyslu z informá-

cií. Odzrkadľuje hlbšie poznanie kognitívnych procesov pri rozhodovaní o relevancii. Základným procesom je postupné, intuitívne a adaptívne skladanie kritérií. Najčastejšie kritériá sú autor, téma, činnosti a kontexty. V elektronickom prostredí sa študenti sústreďujú najmä na zisťovanie pôvodu zdrojov. Relevancia je rozptýlená a vyžaduje si mozaikovitú skladanie informácií. *Kontext* pridávajú prepojenia na úrovni témy, spolupráce (kolektívny diskurz, sociálne tagovanie) a personalizácie. Rozhodovanie pri posudzovaní relevancie podporujú také vlastnosti elektronických zdrojov ako je aktuálnosť, rýchlosť, dostupnosť, multimediálnosť a prepojitelnosť. Dôležitá je aj *vizualizácia*, ktorá pridáva hodnotu prostredníctvom obrazov, foriem, pojmových máp (pre expertov, pre nováčikov), tematických máp, intelektuálnych priestorových máp typu „metra“ a pod.

*Relevancia 2.0* sa vyznačuje pridanou hodnotou vznikajúcou z kolektívnej skúsenosti. Základom je *delenie sa o informácie*. V tomto zmysle relevanciu ovplyvňuje aj *kolaborácia*. Môže ísť o spoluprácu ľudí pri vyhľadávaní a využívaní informácií, pri *formulácii dotazu* a pri spoločnom *filtrovaní* informácií. Z hľadiska relevancie sú dôležité rôzne *odporúčacie systémy*, ktoré môžu používateľovi uľahčiť hodnotenie veľkého množstva informácií.

*Kolaboratívne filtrovanie* má odporúčať informácie používateľom na základe poznatkov o predchádzajúcich preferenciách. V konkrétnom systéme môže ísť o formu aktívneho filtrovania (prostredníctvom informačného profesionála alebo samotným používateľom), systémového kolaboratívneho filtrovania, alebo aj o kolaboratívne filtrovanie citačných informácií (Foster, 2006). Nové systémy by mali v novom modeli podporiť nelineárnosť, neurčitnosť, náhodné objavovanie relevancie. Dôležitá je aj podpora pozitívnych emócií, najmä radosti z objavu.

Výskumy tiež dokazujú, že členovia výskumných skupín sa delia o informácie najmä v takých situáciách ako je informovanie o nových trendov, poznanie kontextov, riešenie úloh a procesy výskumu. Rôzne druhy praxe pri delení sa o informácie sa môžu týkať vytvorenia stratégie výskumu, určenia nového smeru výskumu, vzájomnej komunikácie medzi mladšími a staršími výskumníkmi alebo udržiavania vzťahov v komunite.

Zistenia zahraničných štúdií napríklad dokazujú, že najväčšiu váhu pri posudzovaní relevancie webových stránok (Tombros, Ruthven, Jose 2003) má *téma (obsah) a usporiadanie stránky*. V kritériách dominuje poradie text, štruktúra, kvalita, netextové prvky. Iné štúdie dokazujú, že žiaci základných a stredných škôl málokedy hodnotia relevanciu informácií na internete, skôr predpokladajú, že všetky informácie sú správne. Intuitívne hodnotenie je založené na základe názvu stránky, množstva informácií, obrázkoch alebo pútavosti.

Interaktívna, intuitívna a exploratívna relevancia obsahuje viacnásobné kritériá hodnotenia a konštruovanie zmyslu v kontextoch integrované pozitívnymi aj negatívnymi emóciami. Ďalšie zložky relevancie obsahujú prepájanie individuálneho a kolektívneho, subjektívneho a objektívneho v limitoch času a priestoru. Nová paradigma informačnej vedy prináša práve nové modely spracovania a využívania informácií. Tieto modely sa prejavujú v novej organizácii informácií, vývoji digitálnych knižníc a v podpore informačnej gramotnosti.

Rozdiely medzi chápaním relevancie v tradičnom a elektronickom prostredí sumarizuje tabuľka 1.

Tab. 5.1 Rozdiely v podpore posudzovania relevancie medzi tradičným knižničným prostredím a elektronickým prostredím

	Tradičné knižničné prostredie	Elektronické prostredie (internet, web, digitálna knižnica)
<b>Systémová relevancia</b>	Bibliografická úroveň Formálna zhoda	Interaktívnosť Zložitosť vzťahov
<b>Tematická relevancia</b>	Orientačná, navigácia k zdrojom – informácie	Obsahová, intelektuálna – znalosti
<b>Kritériá</b>	Jednorozmerné	Viackritériálny prístup
<b>Organizácia informácií</b>	Lineárnosť	Nelineárnosť
<b>Pridaná hodnota</b>		Kontext, vizualizácia, spolupráca
<b>Kognitívna relevancia</b>	Predpokladaný kognitívny stav, štýl, kategórie používateľov	Personalizácia, inteligentné rozhrania
<b>Informačné štruktúry</b>	Deterministické	Flexibilné, prepojitelnosť pojmov Skladanie mozaiky
<b>Komunikačná relevancia</b>	Referenčné interview, štýly komunikácie	Skupinové triedenie, odporúčania, dialóg
<b>Situačná relevancia</b>	Referenčné interview, úlohy, problémy, situácie	Redukcia neistoty kontextom, prepojeniami, organizáciou
<b>Motivačná relevancia</b>	Preferencie, úlohy	Lahký prístup, rôzne spôsoby služby, možnosť objavovania

Rozdiely v prostrediach poukazujú na nové koncepcie relevancie informácií, ktoré modelujú kontext používateľa ( *kto, kde, kedy, načo potrebuje informácie*). Oproti „orientačnému“ typu relevancie pri vyhľadávaní v tradičných prieskumových nástrojoch (napr. Google) sa tak získavajú aktuálnejšie a personalizované výstupy z digitálnych knižníc. Na strane systémov vznikajú nové sofistikované algoritmy relevantného rangovania a kontextového prepájania.

Poznatky o informačnom správaní, informačných štýloch a hodnotení relevancie v elektronickom prostredí integrujeme v novej koncepcii informačnej ekológie pre digitálne knižnice.

## 5.7 Využitie výsledkov výskumov relevancie v praxi

### *Relevantné správanie*

Relevancia je súčasť informačného správania človeka. Otvorenosť elektronického prostredia uľahčuje používateľovi intelektuálnu činnosť. Jej súčasťou pri spracovaní informácií je aj posudzovanie relevancie. Kvalitatívne metódy prieskumov ako pozorovania a rozhovory sú kľúčové pri pochopení spôsobov informačného správania, pri výbere stratégie vyhľadávania a pri výbere kritérií a rozhodovaní.

### *Digitálne knižnice*

Integrované prostredia na analýzy a navigovanie používateľa začleňujú do svojich funkcií určité prvky relevantného správania človeka. Praktické aplikácie sú v pokročilých digitálnych knižniciach s obohatenou funkčnosťou. Viac kontextu, anotovanie, indexovanie, participácia spolupracujúcich odborníkov sú už reálne v takých systémoch ako NSDL (americká Národná vedecká digitálna knižnica) v oblasti prírodných vied a matematiky, systém BRICKS zameraný na budovanie zdrojov pre integrované služby v oblasti kultúrneho dedičstva, systém COLLATE alebo projekt ARTIST, integrované prostredie na analýzu a navigovanie v oblasti umenia (Steinerová 2007, Steinerová 2009).

### *Informačná gramotnosť*

Relevancia je súčasťou informačnej gramotnosti pri kritickom hodnotení informácií. Pri informačnom preťažení a množstve dostupných elektronických zdrojov už nejde len o nájdenie informácií, ale skôr o ich hodnotenie a efektívne využitie. Používatelia však stále nevedia posudzovať relevanciu. Najmä mladí ľudia v prostredí webu majú tendenciu uprednostňovať kvantitu pred kvalitou, rýchlosť pred postupným budovaním zmyslu. Informačná veda tu môže pomôcť pri navrhovaní programov vzdelávania a informačnej gramotnosti.

### *Vyhľadávanie v elektronickom prostredí*

Internet a elektronické prostredie sú základom na riešenie informačných problémov človeka. Napriek množstvu zdrojov je vyhľadávanie stále komplikované, najmä z hľadiska voľby vhodnej informačnej stratégie a efektívnosti vyhľadávania. Aktuálne je hodnotenie relevancie informácií na úrovni orientácie (predbežnej relevancie) a analýzy a hodnotenia informácií. Na úrovni orientácie pôsobia najmä také faktory ako voľba vyhľadávacieho nástroja a formulovanie dotazu. Dôležité sú referenčné informácie a metainformácie, ktoré navigujú k plnému textu. Na úrovni analýzy ide o hodnotenie nájdených zdrojov z hľadiska hodnoty, užitočnosti a dôležitosti. Prechodom cez prepojené zdroje si používateľ skladá vlastnú mozaiku relevancie. Posudzovanie relevancie je podporené funkciami systémov (najmä pokročilým vyhľadávaním, inteligentnými rozhraniami) a komunikáciou, interakciou a spoluprácou. V elektronickom prostredí sa overuje najmä pôvod, spoľahlivosť a dôveryhodnosť zdroja. Pri pôvode sa najčastejšie overujú charakteristiky inštitúcie (organizácia, prevádzkovateľ) a autora.

### *Sémantický web a efektívnosť prieskumových nástrojov*

Prieskumové stroje fungujú na princípe automatizovaného zberu dát z webových sídiel. Ukladajú sa do databázy, ktorú prehľadáva robot a na základe algoritmu oznámi výsledok prieskumu. Algoritmus neberie do úvahy kontext. Ide vlastne o systémovú relevanciu, ktorá rozlišuje najmä formálnu zhodu a zoradovanie výsledkov podľa vlastného hodnotenia. Preto sa pracuje na novších algoritmoch. Google aplikuje patentovanú technológiu PageRank na hodnotenie dôležitosti webových stránok a hodnotí aj odkazujúce stránky. Problém efektívnosti prieskumových strojov sa často interpretuje ako to, že dáta nie sú zrozumiteľné pre stroje z hľadiska významu (Makulová 2005). Preto vznikla koncepcia sémantického webu založená konzorciom W3C. Cieľom je definovať významy dostupných informácií a spracovať ich prostredníctvom inteligentných agentov. Inteligentné nástroje by mali umožniť zachytiť sémantiku informácií s priblížením k ľudskému mysleniu. Úspešnosť vyhľadávania by mala závisieť od účinnejšej spolupráce človeka a systému. Okrem toho v sémantickom webe vzniká aj používateľsky generovaný obsah. Metadáta a folksonómie vytvárajú kontext a umožňujú lepšie chápať význam informácií aj v službách webu. Vyhľadávací nástroj Google založený na odkazoch a stránkach, ktoré iní používatelia označili za relevantné, môže preto poskytovať lepšie výsledky ako nástroje vyhľadávajúce prostredníctvom kľúčových slov. Podobne aj spoločné poznanie skupín ľudí pri tvorbe napríklad Wikipédie (odbornej encyklopédie) môže pomôcť zvyšovať relevanciu obsiahnutých informácií.

## 5.8

### **Ako posudzovať relevanciu odborných/vedeckých informácií**

Aj keď je posudzovanie relevancie komplexným problémom, môžeme na záver zovšeobecniť empirické skúsenosti aj analýzy a výskumy informačnej vedy do niekoľkých odporúčaní.

1. K posudzovaniu relevancie treba pristupovať ako primárne k intelektuálnej, myšlienkovvej činnosti, ktorá si vyžaduje určitý stupeň koncentrácie.



2. Posudzovanie relevancie súvisí s premyslenou a predbežne formulovanou informačnou požiadavkou, ktorá vychádza z informačnej potreby. Informačné potreby sa zvyčajne vynárajú z kontextov štúdia, práce, záujmu, riešenia každodenného problému.
3. Predbežné stopy relevancie, prvotnú predstavu o možnej užitočnosti informácie si intuitívne formulujeme ako „náčrt“ relevancie, ktorý sa môže meniť v priebehu vyhľadávania.
4. Relevanciu predvídame už pri výbere informačného zdroja a vyhľadávacieho nástroja. Svoju pozornosť zameriame na také zdroje, v ktorých predpokladáme čo najvyššiu odozvu. Oddelením zdrojov a nástrojov z informačného horizontu, ktorý vidíme vo svojom informačnom prostredí, si zabezpečujeme predbežnú, tzv. orientačnú relevanciu.
5. Na úrovni orientácie sa pri posudzovaní relevancie zvyčajne snažíme pochopiť základné pojmy tematiky a pochopiť problém prostredníctvom slovníkov, encyklopédií. Často sa obraciame na osoby, ktoré môžu vysvetliť problém (učiteľ, príbuzný, spolužiak).
6. Na úrovni riešenia problému hlbšie analyzujeme, hodnotíme, filtrujeme, porovnávame, triedime a rozhodujeme sa o tom, či informáciu prijať alebo neprijať. Najčastejšie používame kritériá ako je téma, autor, názov.
7. V elektronickom prostredí často overujeme pôvod zdroja (prevádzkovateľ stránky, inštitúcia, autorita).
8. Veľkú úlohu zohrávajú aj čas a dostupnosť informácií. Často akceptujeme aj menej relevantné informácie, ak nemáme prístup k relevantnejším, prípadne ak nemáme čas hlbšie informácie hodnotiť.
9. Na hodnotenie relevancie vplývajú prepojenia s inými informáciami, odporúčania informácií inými používateľmi a vizuálne vyjadrenie informácií.
10. Najčastejšou informačnou stratégiou aj pri posudzovaní relevancie je „hodenie siete“ kvôli získaniu čo najvyššieho počtu možných relevantných informácií, a následné zužovanie výberu, filtrovanie informácií.
11. V elektronickom prostredí je významná aj informačná stratégia prezerania, ktorá podporuje aj náhodné získavanie nových informácií a objav nových súvislostí.
12. Konečné rozhodnutie o relevancii prijímame ako výsledok skladačky viacerých kritérií súvisiacich s obsahom a formou informácií, stavom poznania, náladou, cieľmi a kontext problémov a záujmov.

## 5.9 Zhrnutie

Relevancia je základná kategória informačnej vedy. Relevancia je vzťah medzi vyhladanými informáciami a formulovanou požiadavkou. Informačná veda skúma relevanciu informácií už od 50. a 60. rokov 20. stor. Uplatňovala pritom v teórii najmä komunikačné prístupy a filozofické prístupy. Najvýznamnejším autorom zaoberajúcim sa relevanciou v informačnej vede je Tefko Saracevic. Medzi základné vlastnosti relevancie patria vyjadrenie vzťahu, zameranosť, kontext, dynamika a premenlivosť. Množstvo teórií a prístupov k relevancii možno rozdeliť na systémové prístupy a používateľské prístupy. Systémová relevancia je kritériom na meranie účinnosti systému a spája sa s vnútorným algoritmom (modelom) prieskumového systému. Dôležitú úlohu tu zohrala koncepcia relevantnej spätnej väzby. Používateľská relevancia sa spája s kognitívnymi procesmi človeka pri posudzovaní hodnoty informácií. Historický vývoj názorov na relevanciu naznačuje prechod od algoritmickej, formálnej zhody k rozširovaniu kategórie relevancie smerom ku kognitívnym procesom a sociálnym vzťahom.

Základné typy relevancie v informačnej vede sú systémová (algoritmická), tematická (predmetová), kognitívna (pertinencia), situačná (užitočnosť), motivačná (afektívna). Výskumy tiež dokazujú, že relevancia je dynamická, založená na viacnásobných prepojeniach, procesuálna. Rozlišujú sa aj rôzne stupne relevancie a objavujú sa aj ďalšie nové typy relevancie.

Relevancia v sieťovom elektronickom prostredí vytvára nový model založený na interaktívnosti, nelineárnosti, spolupráci, vizualizácii. Nový model možno označiť aj ako relevanciu 2.0 založená na tvorbe zmyslu. Rozdiely pri posudzovaní relevancie v tradičnom a elektronickom prostredí poukazujú na odlišnosti pri hľadaní pôvodu informácií a pri práci s kontextom. Využívajú sa algoritmy kontextového prepájania na strane systémov a poznatky o informačnom (relevantnom) správaní používateľov. Výsledky výskumov relevancie možno využiť ako pri formovaní služieb a produktov digitálnych knižníc, tak aj pri podpore informačnej gramotnosti človeka. Relevancia je vždy súčasťou úspechu vyhľadávania informácií. Posudzovanie relevancie je primárne intelektuálny proces. Nájdenie relevantných informácií závisí od mnohých faktorov, najmä od formulovania požiadavky, výberu vhodného zdroja a určenia vhodných kritérií hodnotenia informácií.

## Literatúra

- ANDERSON, Theresa Dirmdorfer. 2006. Uncertainty in action : observing information seeking within the creative processes of scholarly research. In *Information Research* [online]. 2006, vol. 12, no.1, paper 283 [cit. 2010-01-11]. Dostupné na internete: <<http://InformationR.net/ir/12-1/paper283.html>>.
- BORLUND, Pia. 2003. The concept of relevance in IR. In *JASIS*. ISSN 0002-8231, 2003, vol. 54, no. 10, s. 913 – 925.
- FOSTER, Alan. 2005. Nonlinear Information Seeking. In *Theories of Information Behavior*. Medford : ASIST – Information Today, 2005. s. 154 – 258.
- FROELICH, T. J. 1994. Relevance Reconsidered – Towards an Agenda for the 21st Century : Introduction to Special Topic Issue on Relevance Research. In *JASIS*. ISSN 0002-8231, April 1994, vol. 45, no. 3, s. 124 – 134.
- HARTER, S. P. 1992. Psychological Relevance and Information Science. In *JASIS*. ISSN 0002-8231, October 1992, vol. 43, no. 9, s. 602 – 615.
- INGWERSEN, Peter – JÄRVELIN, Kalervo. 2005. *The Turn : Integration of Information Seeking and Retrieval in Context*. Dordrecht : Springer, 2005. 448 s. ISBN 1-4020-3850-X.
- MAGLAUGHLIN, K. L., SONNENWALD, D. H. 2002. User perspective on relevance criteria : a comparison among relevant, partially relevant, and not-relevant judgements. In *JASIS*. ISSN 0002-8231, 2002, vol. 53, no. 5, s. 327 – 342.
- MAKULOVÁ, Soňa. 2005. Návrh riešenia problémov pri vyhľadávaní informácií v internete alebo od kvantity ku kvalite. In *Knihovna* [online]. 2005, roč. 16, č. 1 [cit. 2009-08-21]. S. 9 – 22. Dostupné na internete: <<http://knihovna.nkp.cz/knihovna51/5123makul.htm>>. ISSN 1801-3252.
- MIZZARO, Stefano. 1997. Relevance : The whole history. In *JASIS*. ISSN 0002-8231, 1997, vol. 48, no. 9, s. 810 – 832.
- SARACEVIC, Tefko. 1996. Relevance reconsidered. In *Information science : integration in perspectives* [online]. Proc. of the Second Conference on Conceptions of Library and Information Science (CoLIS2). Copenhagen, 14-17 Oct. 1996, s. 201 – 218 [cit. 2010-01-11]. Dostupné na internete: <<http://www.scils.rutgers.edu/~tefko/>>.
- SARACEVIC, Tefko. 1997. The stratified model of information retrieval interaction : Extension and application. In *Proc. of the ASIS* [online]. 1997, vol. 34, s. 313 – 327 [cit. 2010-01-11]. Dostupné na internete: <<http://comminfo.rutgers.edu/~tefko/ProcASIS1997.doc>>.

- SARACEVIC, Tefko. 2007a. Relevance : A review of the literature and a framework for thinking on the notion in information science. Part II : Nature and manifestations of relevance. In *JASIS*. ISSN 0002-8231, 2007, vol. 58, č. 13, s. 1915 – 1933.
- SARACEVIC, Tefko. 2007b. Relevance : A review of the literature and a framework for thinking on the notion in information science. Part III : Behavior and effects of relevance. In *JASIS*. ISSN 0002-8231, 2007, vol. 58, č. 13, s. 2126 – 2144.
- SCHAMBER, L., EISENBERG, M. B., NILAN, M. S. 1990. A re-examination of relevance : toward a dynamic, situational definition. In *Inf. Process. Mgmt.* ISSN 0306-4573, 1990, vol. 26, č. 6, s. 755 – 776.
- SOMPTEL, H. van de et al. 2004. Rethinking Scholarly Communication : Building the System that Scholars Deserve. In *D-Lib Magazine* [online]. 2004, vol. 10, no. 9 [cit. 2009-02-08]. Dostupné na internete: <<http://www.dlib.org/dlib/september04/vandesompel/09vandesompel.html>>. ISSN 1082-9873.
- SPINK, A., GREISDORF, H., BATEMAN, J. 1998. From highly relevant to not relevant : Examining different regions of relevance. In *Inf. Process. Mgmt.* ISSN 0306-4573, 1998, vol. 34, s. 599 – 621.
- STEINEROVÁ, J. 1996. *Teória informačného prieskumu*. Bratislava : CVTI SR, 1996. 262 s. ISBN 80-65165-58-9.
- STEINEROVÁ, Jela. 2007f. Relevance assessment for digital libraries. In *Mousaion*. ISSN 0027-2639, 2007, vol. 25, no. 2, s. 37 – 57.
- STEINEROVÁ, J., GREŠKOVÁ, M., ŠUŠOL, J. 2007. *Prieskum relevancie informácií*. Výsledky rozhovorov s doktorandmi FiFUK. Bratislava : CVTI SR, 2007. 150 s. ISBN 978-80-85165-93-7.
- STEINEROVÁ, Jela. 2009. Digitálne knižnice ako inovatívne produkty vo vzdelávaní, vede a výskume. In *Infos 2009*. Pamäťové inštitúcie v digitálnom prostredí. Zborník z 35. medzinárodného sympózia, 27. – 30. apríl, Stará Lesná. Bratislava : SSK, 2009. ISBN 978-80-969674-3-8, s. 166 – 177.
- ŠUŠOL, J. 2007. Posudzovanie relevancie v tradičných a elektronických informačných zdrojoch. In *Ikaros* [online]. 2007, roč. 11, č. 9 [cit. 2009-09-11]. Dostupné na internete: <<http://www.ikaros.cz/node/4294>>. ISSN 1212-5075.
- TALJA, Sanna. 2005. The Domain Analytic Approach to Scholars' Information Practices. In *Theories of Information Behavior*. Medford : ASIST – Information Today, 2005, s. 123 – 127.
- TAYLOR, Robert. 1986. *Value-added processes in information systems*. Norwood : Ablex Publ., 1986. 257 s.
- TODD, R. J. 2005. Information Intent. In *Theories of Information Behavior*. Eds. FISHER, Karen D., ERDELEZ, Sanda., McKECHNIE, Lynne. Medford : Information Today, 2005, s. 198 – 203.
- TOMBROS, A., RUTHVEN, I., JOSE, J. 2003. *Searchres' criteria for assessing web pages* [online]. 2003 [cit. 2009-09-11]. Dostupné na internete: <<http://www.dcs.gla.ac.uk/~jj/publications/postersdemo/tombros-sigir03.pdf>>
- TUOMINEN, K., TALJA, S., SAVOLAINEN, R. 2005. Social constructionist viewpoint on information practices. In Eds. FISHER, Karen D., ERDELEZ, Sanda., McKECHNIE, Lynne. *Theories of information behavior*. Medford : ASIST – Information Today, s. 328 – 333.
- VAKKARI, P., HAKALA, N. 2000. Changes in relevance criteria and problem stages in task performance. In *Journal of Documentation*. ISSN 0022-0418, 2000, vol. 56, issue 5, s. 540 – 562.
- VAKKARI, Pertti – PENNANEN, Mikko. 2001. Sources, relevance and contributory information of documents in writing a research proposal : a longitudinal case study. In *The New Review of Information Behaviour Research*. 2001, vol. 2, s. 217 – 232.

WILSON, T. D. 2000a. Human Information Behavior. In *Informing Science*. Special Issue on Information Science Research. 2000, vol. 3, no. 2, s. 49 – 55.

WILSON, T. D. 2000b. Alfred Schutz, phenomenology and research methodology for information behaviour research. In *Information Research* [online]. 2000 [cit. 2010-01-11]. Dostupné na internete: <<http://informationr.net/tdw/publ/papers/schutz02.html>>.

# 6

## Informačné správanie v elektronickom prostredí

V posledných desiatich rokoch odhaľujú výskumy informačného správania človeka v elektronickom prostredí zmeny v práci s informáciami. Vznikajú nové modely vedeckej komunikácie v elektronickom prostredí. Do takýchto výskumov sa zapájajú aj akademické knižnice a rôzne výskumné organizácie, najmä v USA a Veľkej Británii. Výsledky sú zosumarizované v správach OCLC (Sharing 2007), správe Pew Internet and American Life Project (Horrigan 2007), správach CIBER (Centre for Information Behaviour and the Evaluation Research, University College of London) (Information Behaviour 2008) a v projekte IBEC (Information Behavior in Everyday Contexts) na University of Seattle, Washington, USA.

Nové funkcie informačných systémov sa vynárajú zo *vzorcov informačného správania mladých ľudí v elektronickom prostredí*. V teórii sumarizuje výskumy monografia D. Casea (Case 2007) a zborník *Theories of Information Behavior* (2005).

Mladá generácia sa pri využívaní informácií niekedy označuje ako tzv. Google generácia, „net“ generácia alebo „screenagers“ – obrazovková generácia. Odlišuje sa svojím informačným správaním od starších generácií. Uprednostňuje personalizáciu, pohodlie, rýchlosť. V informačnom správaní sú mladí ľudia netrpezliví, praktickí, orientovaní na výsledky. V elektronickom prostredí prepájajú spoluprácu a sociálny kontakt so samostatným vyhľadávaním. Uspokoja sa aj s menej relevantnými informáciami a preferujú alternatívne formy vzdelávania.

Čo sa vlastne mení pri komunikácii v elektronickom prostredí? Najvýraznejšia je interaktivita a spolupráca. Preto vznikli interaktívne modely vyhľadávania informácií (Saracevic, Spinková, Ingwersen). Z individuálneho komunikačného aktu sa stáva kolektívna komunikácia a spolupráca viacerých aktérov.

Informačné správanie v elektronickom prostredí vyplýva zo **vzťahov k informačným a komunikačným technológiám** (IKT). Skúmaním týchto vzťahov vznikla jedna z novších typológií používateľov (Horrigan 2007). Používatelia boli rozdelení do troch kategórií.

*Prvá kategória* sú elitní používatelia IKT (31%). Obsahuje štyri skupiny veľmi intenzívnych používateľov internetu a mobilných telefónov zapojených do používateľskej tvorby obsahu. Odlišujú sa najmä postojom pri hodnotení pozitív technologických funkcií. Tieto skupiny sú označené ako „všezravci“, *prepájajúci*, *bezduchí veteráni*, *podporovatelia produktivity*. Intenzívni používatelia využívajú technológie na konzumáciu služieb a produktov rýchlym tempom. Technológie sú pre nich platformou na účasť (participáciu) a sebavyjadrenie.

*Druhá kategória* sú používatelia strednej cesty (20%). Obsahuje dve skupiny používateľov zameraných na využívanie technológií pri riešení úloh a komunikovaní. Jedna skupina je orientovaná na mobilnú komunikáciu a plne využíva všetky funkcie mobilných prístrojov. Druhá skupina predstavuje pripojených používateľov, ktorí však považujú neustále pripojenie za obťažovanie.

*Tretia kategória* sú menej technologicky aktívni používatelia (49%), ktorí využívajú moderné technológie skôr okrajovo. Štyri skupiny v tretej kategórii sú neskúsení experimentátori, používatelia spokojní s málo funkcií, ľahostajní, a ľudia mimo siete (15% starších ľudí). Táto kategorizácia dokazuje, že pri informačnom správaní v elektronickom prostredí nefunguje jednotný model. Problémom je včleňovanie technológií do života človeka.

V elektronickom prostredí sa používatelia stávajú „konzumentmi“, rýchlo preskakujú medzi zdrojmi. Často im chýba postupné budovanie poznatkov a vytváranie zmyslu z ich štúdiá. Používatelia tu zamieňajú dostupnosť a vlastníctvo informácií za cieľ – ich premenu na poznatky. Preto je potrebné hlbšie skúmať vzťahy človeka a elektronického informačného prostredia.

## 6.1

### **Problémy informačného správania mladých ľudí v elektronickom prostredí**

Výskumy dokazujú, že pri informačnom správaní „digitálnych domorodcov“ vznikajú problémy so sústredenou reflexiou a kritickým myslením. Niektorí hovoria o „digitálnej DNA“ mladých ľudí, ktorá sa vyznačuje paralelným spracovaním informácií, náhodným prístupom k zdrojom a obrazným myslením. Sledovanie týchto zvláštností a rozdielov medzi informačnými štýlmi môže pomôcť pri navrhovaní digitálnych knižníc a pri koncepciách vzdelávania.

Perspektívne je prepojenie inteligentných technológií s ľudským prístupom. Mnohé štúdie potvrdzujú spoločné vzorce informačného správania tzv. „Google“ generácie alebo „sieťovej generácie“ (Connaway 2008, Prabha 2007, Information Behaviour 2008, Sharing 2007). Je to uprednostňovanie okamžitých riešení, selektívnosť a spolupráca s kolegami. Predstavy o tzv. „Google“ generácii (narodená po roku 1993) upresňuje najnovší výskum CIBER (Information Behaviour 2008).

Prirodzenou vlastnosťou informačného správania je rozmanitosť. V internete dominuje horizontálne vyhľadávanie informácií. Vyznačuje sa rýchlym prezeraním niekoľkých zdrojov, preskakovaním medzi nimi a online prezeraním. Napriek tomu, že Google generácia je technologicky mimoriadne zručná, jej informačná gramotnosť sa nezlepšuje.

Mladí ľudia majú pri informačnom správaní v elektronickom prostredí tieto problémy:

- pochopenie vlastných informačných potrieb
- vytváranie informačnú stratégiu a hodnotenie informácie z hľadiska relevancie a dôveryhodnosti.
- nedostatok hlbších informácií o fungovaní internetu (mentálna mapa).

Všeobecne podliehajú rýchlym, intuitívnym a zjednodušeným riešeniam. Preto vo svojom informačnom správaní obchádzajú knižnice. Čoraz populárnejší je *obrazový, vizuálny prieskum*. Problémom býva aj online čítanie elektronického textu.

V sociálnom sieťovaní v elektronickom prostredí sa prelínajú roly tvorcu a príjemcu. Mladšia generácia je skôr konzumentom ako tvorcom informácií. Typickými črtami tejto generácie je mozaikovitité skladanie častí textov (generácia „cut and paste“), samostatné učenie a preferencia interaktívnych médií. Takéto správanie porušuje tradičný rešpekt k autorským právam a intelektuálnemu vlastníctvu informácií.

Často sa stáva, že informačné zručnosti mladých predbiehajú informačnú gramotnosť. Chýbajú im poznatky o funkciách prieskumových systémov, reprezentácii informácií a sémantike pri vyhľadávaní informácií. Pri rýchlym prezeraní elektronických zdrojov nehodnotia relevanciu a nevenujú sa hlbšiemu analytickému spracovaniu informácií.

## 6.2

### Typy vyhľadávania informácií v elektronickom prostredí

V súčasnosti rozlišujeme v elektronickom prostredí dva typy vyhľadávania. Prvý typ je tzv. **intuitívne vyhľadávanie informácií**, druhý typ je **analytické (profesionálne) vyhľadávanie informácií**.

**Intuitívnemu vyhľadávaniu** v podstate zodpovedá strategický typ informačného správania. Je to voľný pohyb vo webe s dôrazom na rýchlosť prístupu k zdrojom. Dotazy sú krátke, vyhľadávanie je zamerané na okamžité použitie informácií. Intuitívne vyhľadávanie funguje bez komplikovanej prípravy a mimoriadnych nárokov na kvalitu a relevanciu informácií.

Oproti tomu **profesionálne vyhľadávanie** súvisí s informačným správaním analytického typu, ktoré vzniká najmä z informačných požiadaviek vedcov a expertov. Profesionálne vyhľadávanie obsahuje aj vyhľadávanie v špecializovaných databázach a hĺbkovom webe. Zabezpečujú ho školení informační profesionáli (rešeršéri).

Intuitívne vyhľadávanie vyplýva z pohybu študentov v elektronickom prostredí. Študenti tu hľadajú „priestor na učenie“ v skupinách a udržiavanie sociálnych kontaktov. To sa stáva základom na plánovanie služieb akademických knižníc, tzv. „learning commons“. Knižnice sa pretvárajú na priestory spoločného (kolaboratívneho) učenia a miesta stretnutí ľudí pri učení a výskume. Tu je dôležitá spolupráca študentov, učiteľov a knižníc.

Dotazníkový prieskum akademických knižníc v SR z rokov 2002 – 2004 (Steinerová a kol. 2004) priniesol zistenia o problémoch používateľov s kooperáciou pri vyhľadávaní informácií, o potrebe podporiť študentov a učiteľov v prístupe k informačným zdrojom, o preferenciách elektronických zdrojov, neochote platiť či optimistických očakávaniach používateľov. Rozlíšili sme dva základné typy informačného správania, strategický a analytický typ.

Prvý typ je *strategický, pragmatický* typ. Vyznačuje sa plytkým, povrchovým spracovaním informácií a snahou získať informácie rýchlo a hneď ich použiť.

Druhý typ je *analytický, hĺbkový* typ. Preferuje hlbšie spracovanie informácií.

Podobné výskumy rozlišujú ďalšie typické spôsoby informačného správania, napríklad tzv. „široké prezeranie“ (broad scanning) a „hĺbkový ponor“ (deep diving) pri vyhľadávaní informácií v elektronickom prostredí.

## 6.3

### Štádiá vyhľadávania informácií

Na základe výskumov aj skúseností možno rozlišovať dva základné štádiá vyhľadávania informácií: **orientačné a intelektuálne**.

**Orientačné štádium** umožňuje zaznamenať vznik problému, formulovať problém a vybrať zdroje a informácie. V orientačnom štádiu používatelia vyžadujú podporu pri porozumení a formulovaní problému a pri orientácii v možných zdrojoch. Tu sa uprednostňuje ľahká dostupnosť zdrojov, možnosť okamžitého použitia informácií a prehľadná organizácia témy. V orientačnom štádiu sa stretávame s neistotou, pretože nevieme formulovať svoju informačnú potrebu do informačnej požiadavky. Používatelia tu však majú aj pomerne optimistické očakávania.

V **intelektuálnom štádiu** sa vyžaduje hlbšie štúdium a analýza informácií. Prebieha neustála interpretácia a reformulácia problému, ale aj reorganizácia a reštrukturalizácia informácií až po syntézu a prezentáciu. Systémy a sprostredkovatelia v tomto štádiu by mali potvrdiť formulovaný problém. Používatelia sa tu inšpirujú poznatkami z dokumentov pri hľadaní riešenia informačného problému. V intelektuálnom štádiu sa zapájajú najmä myšlienkové procesy a kreativita. V elektronickom prostredí tak vzniká exploratívny prieskum. Systémy umožňujú „objavovať“ zdroje. Digitálne knižnice

poskytujú nové funkcie na obohacovanie práve intelektuálneho štádia vyhľadávania informácií (anotácie, odporúčania, analýzy).

## 6.4 Informačné štýly

Informačný štýl je zmesou kognitívnych štýlov, štýlov učenia a riešenia problémov pri spracovaní a vyhľadávaní informácií. V nadväznosti na výsledky výskumov rozlišujeme dva informačné štýly pri vyhľadávaní a spracovaní informácií.

**Pragmatický štýl** preferuje jednoduchý prístup k informáciám, prehľadnú organizáciu poznatkov, nízku cenu a rýchlu dostupnosť elektronických zdrojov. Dominuje kvantita a čas (rýchlosť). **Analytický štýl** sa vyznačuje hlbším intelektuálnym spracovaním informácií. Kladie dôraz na spoľahlivosť a overiteľnosť informácií a proces recenzovania. Dominuje kvalita a relevancia. V prieskumoch sme zistili prevahu pragmatického (strategického) štýlu spracovania informácií u používateľov knižníc. Na prieskumy sme vyvinuli originálnu metodológiu integrujúcu poznatky zo sociálnych a kognitívnych vied do oblasti informačného správania.

Tieto poznatky sme využili pri návrhoch modelov informačného správania. Poznatky o informačných štýloch sme spracovali aj pri tvorbe odporúčaní na podporu informačnej gramotnosti a tvorivosti človeka v informačnom prostredí.

V tab. 6.1 sa sumarizujú rozdiely medzi pragmatickým a analytickým štýlom.

Tab. 6.1 Rozdiely medzi pragmatickým a analytickým informačným štýlom

	Pragmatický	Analytický
Vyhľadávanie	horizontálne	exploratívne
Terminológia	jasnosť	multidisciplinárnosť
Hodnotenie	náhodné, povrchové	skúsenosť pri hodnotení relevancie
Organizácia	povrchová závislosť od poľa	integrácia na základe poznania a skúsenosti
Plánovanie	intuitívne, jednoduché dotazy	komplexnejšie dotazy
Účel	orientácia	intelektuálne spracovanie
Pocity	dôvera, optimizmus	pochybnosti
Motivácia	rýchle riešenie	pochopenie súvislostí
Pristup k informáciám	navigácia	interpretácia

## 6.5 Informačné správanie a sociálne siete

V informačnom správaní mladých ľudí sú v súčasnosti aktuálne **sociálne siete** (Sharing 2007, Information Behaviour 2008). Až 56% študentov využíva sociálne siete ako MySpace (prevažne stredoškôláci), Facebook (prevažne vysokoškôláci) a Mixi (japonskí študenti). Sociálna sieť je sociálna štruktúra prepojená vzťahmi založenými na delení sa o informácie, postoje, záujmy. Cieľom je interakcia na základe spoločných hodnôt, vízie, presvedčení, finančných transakcií alebo priateľstva. Tak vzniká „nová webová kultúra“. Internet už nie je nová technológia, ale samozrejme infraštruktúra ekonomických a sociálnych aktivít. Integráciou do domácností nadobúda rozmer každodennosti. Globálna dominancia služieb webu sa vyvažuje personalizáciou, intuitívnosťou a mobilitou.



Sociálne siete v elektronickom prostredí sa niekedy delia *na miesta sociálneho sieťovania* vyznačujúce sa interakciou (sociálne portály ako MySpace, Mixi, Facebook) a *sociálne médiá* umožňujúce výmenu rôzneho obsahu (napr. YouTube, flickr, Snapfish, delicious). Kultúra „zdieľania“ je založená na výmene osobných informácií, záujmov, fotografií a pod. Výskum OCLC dokazuje, že hlavnou motiváciou týchto interakcií je spojenie s priateľmi, zábava a začlenenie do komunity. V tomto svete „trhoviska“ sa knižnice presadzujú ťažšie, lebo predpokladajú záujem o intelektuálne štúdiá a systematicky organizované poznanie.

## 6.6 Emócie pri informačnom správaní

**Emócie** sú komplexné psychické javy s funkciami hodnotenia, regulácie, adaptácie a organizácie v informačnom správaní. Emócie sa kombinujú do nálad, rozpoloženia, afektov. Emócie sa členia na jednoduchšie, biologicky podmienené (hnev, radosť, strach, smútok) a vyššie, sociálne podmienené (intelektuálne, etické, estetické city). Dôležitú úlohu emócií v informačnej vede skúmali v rámci svojich výskumov N. Belkin, B. Dervinová, P. Ingwersen, T. Wilson. Ako jedna z prvých priniesla opis emócií vo svojom modeli procesu vyhľadávania informácií C. Kuhlthauová (1993). Dokázala postupné odstraňovanie neistoty v priebehu vyhľadávania a využívania informácií. V informačnom správaní je opísaný model ACS – postupnosť afektívneho, kognitívneho a senzomotorického (Steinerová 2005). To znamená, že emócie predchádzajú myšlienkové procesy a aktivity.

V súčasnosti sa emóciami v informačnom správaní zaoberajú Diane Nahlová a Dania Bilalová (2007). Emócie sú súčasťou informačného prostredia. Prežívajú sa aj na sociálnej úrovni podľa zámerov a motivácie. Príjem a využitie informácií sú podmienené emocionálnou hodnotou informácií v komunite.

Emócie ovplyvňujú to, či sa informácia príjme alebo odmietne a čo človek bude s informáciami robiť. Napríklad či ich odovzdá ďalej alebo inak využije na dosiahnutie cieľa. Autori skúmajú tzv. afektívnu realitu, ktorá súvisí s emocionálnou reakciou na technológie.

Diane Nahlová rozpracovala model informačného správania obsahujúci afektívnu, kognitívnu a senzomotorickú zložku (Steinerová 2005). V afektívnej zložke navrhuje metódy na prekonávanie odporu k vyhľadávaniu informácií, posilňovanie informačnej motivácie a vytváranie celoživotných návykov informačne gramotného človeka.

Z hľadiska emócií je pri využívaní informácií najdôležitejšia **motivácia**, pocit „chcenia“ robiť niečo s informáciami. Emócie sú pri spracovaní informácií rôznorodé, dynamické, situačné, závislé od kontextu. Fungujú ako virtuálne brány, ktoré umožňujú všimnúť si informácie, ignorovať alebo akceptovať ich, protirečiť alebo vykonať aktivity. Rôzne skupiny používateľov sa vo vzťahu k informáciám a technológiám líšia práve emóciami. Často sa skúmajú deti, študenti, ale aj špeciálne skupiny, napríklad zrakovo postihnutí.

Emocionálne rozdiely v spracovaní informácií sú viditeľné pri rodovom rozdelení. Naše výskumy potvrdili, že pri vyhľadávaní informácií sú ženy opatrnejšie, úzkostlivejšie, muži asertívnejší a netrepezlivejší (Steinerová, Šušol 2007). Afektívne informačné správanie sa skúma aj pri čítaní, emocionálnej pamäti, informačnej gramotnosti a riadení efektívnosti vlastnej informačnej činnosti.

Zložitosť emócií vo vzťahoch medzi črtami osobnosti a informačným správaním skúmala aj B. Dervinová (Dervin, Reinhard 2007). Výskum dokazuje, že socio-emocionálne črty osobnosti sú variabilné v rôznych informačných prostrediach. Dôvera, kooperácia, kontrola, vyrovnanosť podporujú spracovanie informácií. Na druhej strane využívanie informácií a technológií môže často sprevádzať frustrácia spôsobená časovým tlakom, neistotou a pochybnosťami.

Potvrdilo sa, že v priebehu vyhľadávania sa začiatočná neistota mení na pocity spokojnosti. Po ukončení riešenia informačného problému používateľa pociťujú najčastejšie spokojnosť, úľavu a zvedavosť.

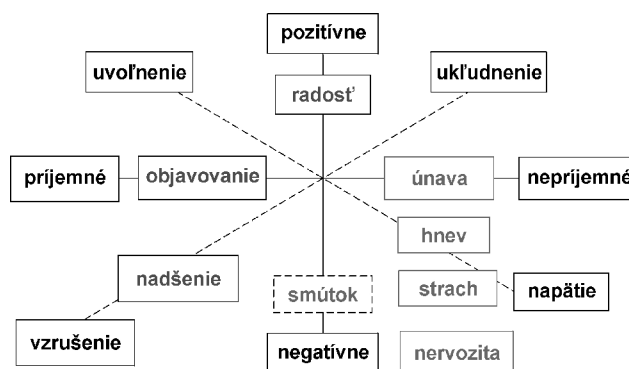
V našich výskumoch sme zistili aj **emocionálne charakteristiky jednotlivých informačných štýlov**.

**Pragmatický štýl** sa vyznačuje skôr externou motiváciou. Často sa tu objavuje stres z vyhľadávania informácií, strach z neúspechu a neistota. To používateľom spôsobuje ťažkosti pri riešení nepredvídateľných situácií využívania informácií. Pragmatický štýl vykazuje tendencie optimistického postoja k využívaniu informácií, dominovala uvoľnenosť, dôvera a optimizmus. Potvrdila sa vyššia miera dôvery vo funkcie systémov a služieb, na druhej strane aj vyššia miera sklamaní. **Analytický štýl** sa vyznačuje skôr stabilitou a istotou osobnosti. Motivácia je tu vnútorná, konštruktívny postoj k informáciám je založený na prirodzenej zvedavosti a otvorenosti novým poznatkom. Analytický štýl sa spája s úzkostlivejším prístupom k vyhľadávaniu s vyššou mierou pochybností. Hlbší záujem o informácie podnecuje aj analytický a kritický prístup k informáciám (Steinerová, Šušol 2005).

Aj vo výskume relevancie sme sa venovali emóciám ako dôležitej súčasť procesov hodnotenia a triedenia informácií (Steinerová, Grešková, Šušol 2007). Emócie sa tu prejavujú v rôznej miere intenzity. Základné pozitívne pocity príjemného predstavujú najmä radosť, nadšenie a objavovanie. Základné negatívne pocity sprevádzajúce posudzovanie relevancie informácií sú hnev, strach, nervozita a únava. Funkciami pozitívnych pocitov pri posudzovaní relevancie je najmä aktivácia, vitalita, seba-dôvera. Na základe hodnotového systému sa získava želaná informácia ako výnimočná hodnota. Objavuje sa niečo skryté a nové a postupuje sa smerom k vedeckej pravde. Funkciami negatívnych pocitov pri posudzovaní relevancie sú reakcie na prekážky a obmedzenia. Často vyplývajú z nedostatku alebo prebytku informácií. Problémom býva najmä kontrola situácie podmienená novosťou, nezvyčajnosťou a momentom prekvapenia pri interakcii s informáciami. Nerovnováha medzi očakávaniami a výsledkami a nedostatok času sú tiež príčinami negatívnych pocitov.

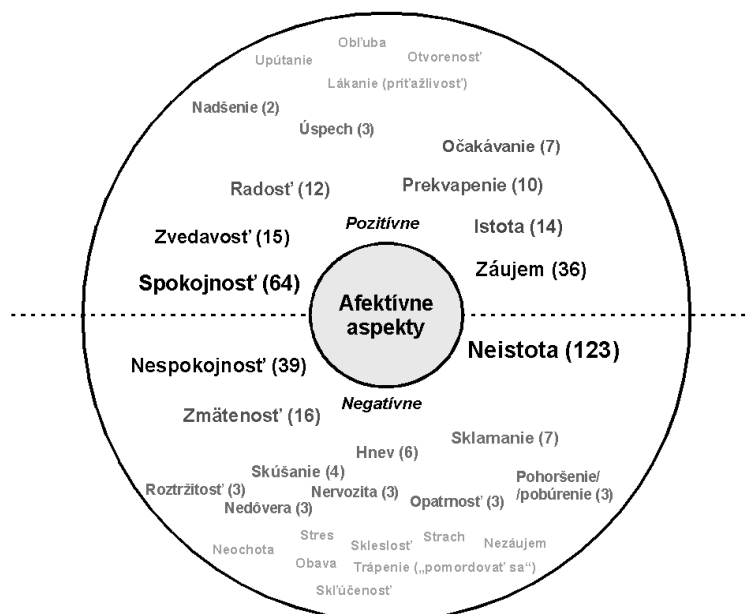
Analýzy potvrdzujú, že emócie integrujú procesy hodnotenia, objavovania vzťahov a riešenia problémov pri posudzovaní relevancie. Od nepríjemných pocitov neistoty a napätia sa postupne prechádza k pocitom uvoľnenia. Tento postup však nie je priamočiary. Keďže je vysoko individuálny a premenlivý, ťažko sa zovšeobecňuje.

Príklad kategorizácie emócií z výskumu relevancie (Steinerová, Grešková Šušol 2007) je znázornený na obrázku 6.1.



Obr. 6.1 Kategorizácia emócií pri posudzovaní relevancie

Ďalší príklad kategorizácie emócií vyplynul z výskumu interakcie človeka a inteligentného agenta (Grešková 2008). Grafické znázornenie je naznačené na obrázku 6.2.



Obr. 6.2 Kategorizácia emócií pri interakcii človeka so systémom

## 6.7 Zhrnutie

Informačné správanie v elektronickom prostredí vyplýva zo vzťahu človeka k informačným a komunikačným technológiám. Mladá generácia sa v elektronickom prostredí správa ako konzumenti. Preferujú rýchlosť, paralelné spracovanie informácií, obrazné formy informácií. Informácie spracúvajú povrchnejšie, preskakujú medzi zdrojmi. Mnohé štúdie dokazujú, že najvýraznejšie vlastnosti informačného správania v elektronickom prostredí sú interaktívnosť a spolupráca. Problémami je hlbšie analytické myslenie a chápanie informačných potrieb. V elektronickom prostredí rozlišujeme intuitívne, laické vyhľadávanie a profesionálne vyhľadávanie informácií.

Na základe výskumov rozdeľujeme vyhľadávanie na dve štádiá: orientačné (pochopenie a formulácia problému) a intelektuálne (analýza a hlbší prieskum).

Informačný štýl je súbor preferencií pri vyhľadávaní a využívaní informácií. Podľa výsledkov výskumu rozlišujeme pragmatický a analytický štýl vyhľadávania informácií. Pragmatický štýl preferuje rýchlosť a množstvo informácií. Analytický štýl uprednostňuje kvalitu, hodnotu a recenzované informácie.

Vo vyhľadávaní a spracovaní informácií v elektronickom prostredí v súčasnosti dominuje sprostredkovanie prostredníctvom sociálnych sietí. V internete sa delia na sociálne médiá a sociálne portály.

Pri informačnom správaní zohrávajú dôležitú úlohu aj emócie ako komplexné psychické celky, ktoré riadia myšlienkové procesy aj konkrétne aktivity. Pragmatický informačný štýl sa vyznačuje optimistickými pocitmi, analytický informačný štýl obsahuje emócie spojené s úzkostlivým prístupom k využívaniu informácií. Výskumy dokazujú, že emócie sa prejavujú aj pri posudzovaní relevancie a pri interakcii človeka a systému. Emócie podporujú alebo aj spomaľujú informačné správanie človeka.

## Literatúra

- BERMAN, F., BRADY, H. 2005. *Final Report : NSF SBE-CISE Workshop on Cyberinfrastructure and the Social Sciences* [online]. 2005 [cit. 2008-04-07]. Dostupné na internete: <[http://ucdata.berkeley.edu/pubs/CyberInfrastructure\\_FINAL.pdf](http://ucdata.berkeley.edu/pubs/CyberInfrastructure_FINAL.pdf)>.
- CASE, Donald O. 2007. *Looking for Information : A Survey of Research on Information Seeking, Needs and Behavior*. Sec. ed. Amsterdam : Elsevier, 2007. 423 s. ISBN 13 978-0-12-369430-0.
- CHOO, Chun Wei. 2006. *The Knowing Organization : How Organizations Use Information to Construct Meaning, Create Knowledge, and Make Decisions*. Sec. ed. New York : Oxford University Press, 2006. 354 s. ISBN 0-19-517678-2.
- CONNAWAY, L. S. 2008. Make Room for the Millennials : Changes to Systems and Services Will Appeal to This Generation. In *NextSpace* [online]. October 2008, no. 10 [cit. 2010-01-11]. s. 18 – 19. Dostupné na internete: <<http://www.oclc.org/nextspace/010/research.htm>>. ISSN 1559-0011.
- DERVIN, B., REINHARD, C. D. 2007. How Emotional Dimensions of Situated Information Seeking Relate to User Evaluations of Help from Sources : An Exemplar Study Informed by Sense-Making Methodology. In *Information and Emotion : the Emergent Affective Paradigm in Information Behavior Research and Theory*. Medford : Information Today, 2007, s. 51 – 84.
- FISHER, K. E., ERDELEZ, S., McKECHNIE, L. 2005. *Theories of Information Behavior*. Medford : Information Today, 2005. 431 s. ISBN 1-57387-230-X.
- FOSTER, Jonathan. 2006. Collaborative Information Seeking and Retrieval. Chapter 8. In Ed. CRONIN, Blaise. *ANNUAL Review of Information Science and Technology*. Medford : Information Today, 2006. ISSN 0066-4200, vol. 40, s. 329 – 356.
- GREŠKOVÁ, Mirka. 2008. *Kognitívne základy informačnej vedy : interakcia človeka s agentom na vyhľadávanie informácií* [Dizertačná práca]. Univerzita Komenského v Bratislave; Filozofická fakulta; Katedra knižničnej a informačnej vedy. Bratislava : FiFUK, 2008. 139 s.
- HORRIGAN, J. B. 2007. *A Typology of Information and Communication Technology Users : Pew Internet & American Life Project* [online]. 2007 [cit. 2008-03-28]. 65 s. Dostupné na internete: <<http://www.pewinternet.org/>>.
- KUHLTHAU, Carol C. 1993. *Seeking Meaning : A Process Approach To Library And Information Services*. Norwood : Ablex, 1993. 199 s.
- MURSU, Á. et al. 2007. Activity theory in information systems research and practice – theoretical underpinnings for an information systems development model. In *Information Research* [online]. April 2007, vol. 12, no. 3 [cit. 2008-04-29]. Dostupné na internete: <<http://InformationR.net/ir/12-3/paper311.html>>.
- NAHL, Diane – BILAL, Dania. 2007. *Information and Emotion : the emergent affective paradigm in information behavior research and theory*. Medford : Information Today, 2007. 359 s. ISBN 978-1-57387-310-9.
- NARDI, B. A., O'Day, V. L. 1999. *Information Ecologies : Using Technology with Heart*. Cambridge : MIT Press, 1999. 231 s. ISBN 0-262-14066-7.
- NICHOLAS, David et al. 2005. The Digital Information Consumer. Chapter 11. In Eds. SPINK, A., COLE. Ch. *New Directions in Human Information Behavior*. Dordrecht : Springer, 2005, s. 203–228. ISBN 1-4020-3667-1.

PRABHA, Ch. et al. 2007. What is enough? : Satisficing information needs. In *Journal of Documentation* [online]. 2007, vol. 63, no. 1 [cit. 2009-02-08]. s. 74 – 89. Pre-print. Dostupné na internete: <<http://www.oclc.org/research/publications/archive/2007/prabha-satisficing.pdf>>. ISSN 0022-0418.

RADFORD, Marie L., CONNAWAY, Lynn Silipigni. 2005. *Seeking Synchronicity : Evaluating Virtual Reference Services from user, Non-user, and Librarians perspectives*. A research project funded by the National Leadership Grants for Libraries program of the Institute of Museum and Library Services (IMLS) [online]. Rutgers, The State University of New Jersey; OCLC Online Computer Library Centre, Inc [cit. 2009-02-08]. Dostupné na internete: <<http://www.oclc.org/research/projects/synchronicity>>.

ROWLANDS, I., FIELDHOUSE, M. 2008. *Information behaviour of the researcher of the future : a ciber briefing paper* [online]. University College of London; Centre for Information Behaviour Research, January 2008 [cit. 2008-03-28]. 35 s. Dostupné na internete: <<http://www.ucl.ac.uk/slais/research/ciber/downloads/ggexecutive.pdf>>.

SOMPEL, H. van de et al. 2004. Rethinking Scholarly Communication : Building the System that Scholars Deserve. In *D-Lib Magazine* [online]. 2004, vol. 10, no. 9 [cit. 2009-02-08]. Dostupné na internete: <<http://www.dlib.org/dlib/september04/vandesompel/09vandesompel.html>>. ISSN 1082-9873.

STEINEROVÁ, Jela. 1998a. *Tvorba informačných produktov : nové prístupy informačnej vedy*. Bratislava : CVTI SR, 1998. 130 s. ISBN 80-85165-73-2.

STEINEROVÁ, Jela a kol. 2004. *Správa o empirickom prieskume používateľov knižníc ako súčasť grantovej úlohy VEGA 1/9236/02 Interakcia človeka s informačným prostredím v informačnej spoločnosti*. Bratislava : Filozofická fakulta UK, KKIV, 2004. 113 s.

STEINEROVÁ, Jela. 2005a. *Informačné správanie : Pohľady informačnej vedy*. Bratislava : CVTI SR, 2005. 189 s. ISBN 80-85165-90-2.

STEINEROVÁ, Jela. 2007a. Manifest o digitálnych knižniciach a využívanie informácií v informačnej spoločnosti. In *Itilib*. ISSN 1335-793X, 2007, roč. 11, č. 2, s. 15 – 19.

STEINEROVÁ, Jela. 2007f. Relevance assessment for digital libraries. In *Mousaion*. ISSN 0027-2639, 2007, vol. 25, no. 2, s. 37 – 57.

STEINEROVÁ, J., GREŠKOVÁ, M., ŠUŠOL, J. 2007. *Prieskum relevancie informácií. : Výsledky rozhovorov s doktorandmi FiFUK*. Bratislava : CVTI SR, 2007. 150 s. ISBN 978-80-85165-93-7.

STEINEROVÁ, J., ŠUŠOL, J. 2006. *Využívanie informácií v informačnej spoločnosti*. Zborník z medzinárodnej konferencie. Bratislava SR, 10. – 11. októbra 2006. Information Use in Information Society. Proceedings of the International conference. Bratislava, Slovakia, October 10-11, 2006. Bratislava : CVTI SR, 2006. 228 s. ISBN 80-85165-92-9.

STEINEROVÁ, J., ŠUŠOL, J. 2007. Users' information behaviour – a gender perspective. In *Information Research* [online]. 2007, vol. 11, no. 3, paper 251 [cit. 2010-01-11]. Dostupné na internete: <<http://InformationR.net/ir/11-3/paper251.html>>.

## Zoznam webových sídel

e-Horizons Project

<http://www.e-horizons.ox.ac.uk/>

Open Archives Initiative

<http://www.openarchives.org>

OCLC

<http://www.oclc.org/reports/sharing/default.htm>

Oxford e-Research Centre

<http://www.oerc.ox.ac.uk/>

# 7

## Kategorizácia a klasifikácia ako základ vyhľadávania informácií

### 7.1

#### Kategorizácia a klasifikácia ako myšlienkové procesy

Kategorizácia a klasifikácia sú myšlienkové procesy, ktoré umožňujú deliť a organizovať informácie a poznatky. Je to vlastne kognitívny mechanizmus, t.j. funkcia mozgu vyznačujúca sa zložitými procesmi prispôsobovania sa prostrediu, pružnosťou (flexibilitou) a plasticitou (zjednodušene zvláštna schopnosť meniť vnútorné štruktúry neurónových sietí pod vplyvom individuálnej skúsenosti jednotlivca). Zložitosť procesov kategorizácie a klasifikácie vyplýva zo zložitosti fungovania máp skúseností, s ktorými pracuje ľudský mozog. Tieto mapy nám umožňujú spracúvať zrakové, sluchové, dotykové a iné skúsenosti. Navyše sa ešte pod vplyvom skúsenosti menia a vyvíjajú (Beňušková 2002). Podľa kognitívnych vedcov sa v individuálnom vývine človeka schopnosť organizovať poznatky, a teda aj kategorizovať a klasifikovať, objavuje až približne vo veku 7 rokov.

**Kategorizácia** je voľnejší myšlienkový proces, ktorý pracuje s pojmami. Základné funkcie kategorizácie sú:

- Uľahčenie interakcie človeka s prostredím
- Odstraňovanie kognitívnej záťaže pri spracovaní informácií
- Vyhľadávanie a využívanie informácií

Medzi základné vlastnosti kategorizácie patrí

- Zavedenie hraníc do významu pojmov (čo patrí do kategórie a čo je mimo významu)
- Určenie spoločných vlastností objektov prostredníctvom kategórie
- Určenie kontextu (nakoľko je kategória závislá alebo nezávislá od súvisiacich pojmov)

Napríklad objekty sa rozčlenia a hľadajú sa ostrovy významov a určia sa hranice. Tak môžeme odlišovať napr. 15. a 16. storočia.

Klasická teória kategórií podľa Roschovej (1973, 1975) predpokladala, že existuje množina definovaných kritérií na vytváranie kategórie. Príslušnosť členov danej kategórie bola absolútna (Jacob 2004).

Napríklad kategória *vták* je definovaná týmito kritériami: *má krídla, kladie vajcia, lieta, stavia hniezda*.

Príslušnosť členov v tejto kategórii je absolútna, žiadny člen nie je uprednostnený. Napríklad *vták*: *holub, papagáj, pštros, vrabec*.

Kategórie sa tu vyznačujú svojou *intenziou* – čo je definovanie kategórie prostredníctvom vlastností/kritérií a *extenziou* – čo je vlastne príslušnosť členov v danej kategórii.

Novšia teória kategórií (Barsalou 1987) počíta aj s možnosťou stupňa príslušnosti jednotlivého prvku do kategórie podľa reprezentatívnosti (Jacob 2004). Napríklad v kategórii *vták* je *vrabec* *reprezentatívnejší pojem* ako *pštros* (*menej reprezentatívny*).

Pri kategorizácii a klasifikácii sa teda spracúvajú podnety z prostredia a premieňajú sa na využiteľné pojmy. Na tom sa podieľajú také kognitívne procesy ako vnímanie, myslenie, pamäť a riešenie problémov. Niektoré experimenty dokazujú, že pri vnímaní a spracovaní informácií hrá veľkú úlohu aj intuícia (napr. numerické a fyzikálne intuície, Rybár 2002). Intuítnosť kategorizácie a klasifikácie sa prejavuje aj v praxi pri intuitívnom vyhľadávaní, najmä v elektronickom prostredí.

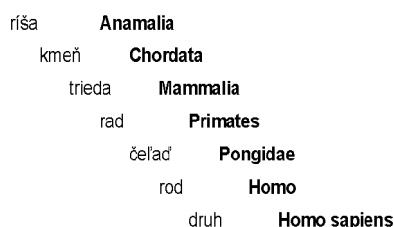
**Klasifikácia** je systematické zaradovanie entity do jednej triedy v rámci systému navzájom sa vylučujúcich a neprekrývajúcich sa tried. Vlastnosťami klasifikácie sú:

- Zákonnosť (je založená na súbore pravidiel)
- Systematickosť (pravidlá sa používajú dôsledne a konzistentne)
- Arbitrárnosť (pri vytváraní tried sa zámerne vyberá jeden aspekt objektu, pojmu)

Najznámejším príkladom klasifikácie je *taxonomická klasifikácia*. Vznikla pri určovaní druhov v biológii, botanike, medicíne. Vyznačuje sa nomenklatúrou – určitým formalizovaným jazykom. Názvy jednotlivých tried v tomto jazyku sú usporiadané hierarchicky a hniezdovým princípom. Jednotlivé pojmy majú presne určenú intenziu (definície) aj extenziu (príslušnosť). Taxonómia prísne organizuje pojmy v hierarchickom kontexte a tým uľahčuje poznávanie objektov (napríklad rastlinných druhov, zvierat a pod.). Obmedzením taxonómie je to, že je postavená na výbere jednej vlastnosti objektu, aby sa zjednodušil systém spracovania a využívania informácií. Nadradené a podradené triedy umožňujú úsporné uchovávanie a vyhľadávanie informácií.

Napríklad veta *Benny je retriever* poskytuje informácie nielen o asociácii s triedou retriever, ale zároveň kontext poskytuje ďalšie informácie z hierarchickej štruktúry, najmä že retriever je prepojený s nadradenými triedami pes, cicavec, stavovec a pod.

Príklad taxonómie je na obrázku 7.1.



Obr. 7.1. Taxonómia delenia druhov organizmov

## 7.2 Systémy klasifikácie a kategorizácie

Klasifikačné systémy pracujú so vzájomne sa vylučujúcimi triedami. Triedy sa nikdy neprekrývajú a sú usporiadané v hierarchii. Jednotlivé triedy majú určené jedinečné mená a pravidlá priradovania pojmov do tried. V tradičnom klasifikačnom systéme je stelesnená jednota poznania, nepočíta sa s jeho premenlivosťou a vývojom. Minulé, súčasné aj budúce poznatky sú spracované rovnako, jednotne. Napríklad ťažko by sme klasifikovali výrok *včera bolo pekne*, pretože ide o intuitívne a prirodzené, každodenné poznanie.

Kategorizácia a klasifikácia sú základom organizácie informácií a ich vyhľadávania. Keďže sú založené na myšlienkových procesoch človeka a používaní jazyka, zápasia s približnosťou významov,

metaforami, analógiami, tvorivou konštrukciou významov. Pre tvorbu významov je dôležité aj odvodzovanie (inferencia), ale aj určovanie dôležitého (relevancie), objavovanie nového a učenie.

Bibliografické klasifikačné systémy možno považovať za príklad klasifikácie. Používajú sa už stáročia na organizovanie dokumentov vo fondoch knižníc a iných pamäťových inštitúcií. Tradičné systémy, ako napríklad DDT (Deweyho desatinné triedenie), MDT (Medzinárodné desatinné triedenie, UDC, Universal Decimal Classification), klasifikačný systém Kongresovej knižnice (Library of Congress Classification System, LCCS) sú deduktívne. Delia poznanie logicky od všeobecného k špeciickému (zhora dolu). Výsledkom je prísna hierarchia založená najmä na rodovo druhových vzťahoch. (napríklad *rodíč-dieťa*). Okrem všeobecných, medzinárodných systémov existujú aj národné (napr. nemecké, švédske, holandské) klasifikácie.

V tab. 7.1. sú naznačené rozdiely medzi kategorizáciou a klasifikáciou podľa Jacobovej (2004).

Tab. 7.1 Rozdiely medzi kategorizáciou a klasifikáciou (podľa Jacob 2004)

Klasifikácia	Kategorizácia
<b>Proces</b>	
Systematické rozdelenie entít založené na analýze charakteristík nevyhnutných a postačujúcich na zaradenie do tried	Tvorivá syntéza entít založená na kontexte alebo na odporovanej podobnosti
<b>Hranice/Ohraničenie</b>	
Triedy sú jedinečné a neprekrývajú sa, hranice sú pevné	Členstvo v kategórii nie je záväzné, hranice nie sú jasné
<b>Členstvo</b>	
Prísne: entita buď je alebo nie je členom nejakej triedy	Flexibilné: príslušnosť ku kategórii je založená na zovšeobecnených poznatkoch alebo na bezprostrednom kontexte
<b>Kritériá priradenia</b>	
Kritériá sú vopred určené princípy	Kontextovo závislé aj kontextovo nezávislé kritériá
<b>Reprezentatívnosť</b>	
Všetky entity jednej triedy sú rovnako reprezentatívne	Členovia kategórie môžu byť pre ňu viac alebo menej typickí
<b>Štruktúra</b>	
Hierarchická štruktúra pevných tried	Zhluky entít; môžu vytvárať hierarchickú štruktúru

Taxonomická klasifikácia je obmedzená tým, že určuje kontext len z hierarchických spojení medzi triedami a triedy sú definované len na jednej vlastnosti.

Klasifikačné systémy sa vyznačujú lineárnosťou. Do istej miery reprezentujú usporiadanie v regáloch. Keď sa takýto klasifikačný systém „číta“, posúva sa v hierarchii od najvyššieho k nižšiemu prvku. Prítom záleží na bezprostrednej blízkosti a postupnosti prvkov, ktoré za sebou nasledujú. To určil už Ranganathan, významný indický knihovník, pri vytvorení svojej klasifikácie. Definoval lineárnu štruktúru APUPA (Alien – Penumbra – Umbra – Penumbra – Alien), U je prítom ohnisková, najrelevantnejšia téma, P sú prvky najbližšie spojené s témou, S sú vzdialené, nespojené s témou. V klasifikácii tak má dôležitú úlohu princíp **blízkosti**.

*Napríklad môžeme si predstaviť, že na polici v knižnici sú knihy nejakým spôsobom vedľa seba zoradené. Systém ich radenia ukazuje na dôležitú súvislosť, vzťah medzi nimi – najčítanejšie romány, súvisiace anglické slovníky (výkladové, prekladové, synonymické a i.). Hierarchická štruktúra umiestňuje do blízkosti tie triedy, ktoré sú tematicky najbližšie spojené.*



*Fazetové systémy* sú analyticko-syntetické klasifikačné systémy. Oproti dedukcii v tradičnej klasifikácii sú fazetové systémy induktívne. Tvoria sa zdola hore. Na základe identifikácie individuálnych prvkov a ich vlastností sa formujú fazety ako aspekty, uhly pohľadu. Fazety tvoria klasifikačné triedy a sú usporiadané v hierarchii. Výsledkom je riadený slovník pojmov a vzťahov. Napríklad na reprezentáciu pojmu *auto* použijeme *fazety: farba, typ, tvar, riadenie*. To sa využíva ja pri rámcovej reprezentácii poznania.

*Klasifikácia* všeobecne určuje hranice vybratej témy alebo domény. Umožňuje aj ustáliť používané výrazy, zabezpečuje normalizáciu jazyka a stabilný slovník. Hierarchické vzťahy sú založené na zmysluplnej odlišnosti pojmov. Vzťahy sú stabilné v tom zmysle, že sa nemenia triedy ani kontext. To je na druhej strane nevýhodou, pretože štruktúra tried je pevná a ťažšie sa zatriedujú nové poznatky mimo existujúceho systému.

*Kategorizácia* je z hľadiska kontextu pružnejšia, kreatívna. Medzi prvkami sú nezáväzné asociácie, pričom sa rozpoznáva podobnosť. Kontext je spojený s okamžitou podobnosťou v čase a priestore. Umožňuje spracovať premenlivosť významu, individuálne skúsenosti aj osobné ciele. Členstvo v kategóriách sa môže prekrývať, definície sa môžu meniť podľa reakcie. Nie každý systém zoskupovania je systémom organizácie. *Napríklad zoznam vecí na nákup triedený podľa typu obchodu nie je systémom organizácie*. Systém kategorizácie však oproti tomu stanovuje zmysluplné vzťahy medzi kategóriami.

### 7.3

## Vplyv informačných štruktúr na vyhľadávanie informácií

Informačné štruktúry vznikajú ako výsledok kategorizácie a klasifikácie. Možno ich charakterizovať ako súbor pojmov a vzťahov medzi nimi. Posilňujú kontexty informácií a vyhľadávanie informácií. Veľký vplyv má jazyk a formulovanie termínov a kľúčových slov na vyhľadávanie. V praxi vyhľadávania sa často používa spojenie dvoch pojmov. Kognitívni vedci dokazujú, že existuje skratkovitá, vnútorná reč, ktorá funguje podobne – *prepája dva pojmy*. Aj keď vyhľadávame vo voľnom texte, istý typ vzťahu medzi pojmi je myslený, hoci nie vyjadrený. Vo vyhľadávaní v internete mnohé štúdie dokázali dominanté používanie jedného alebo maximálne dvoch termínov na formulovanie dotazu (Makulová, 2002, 2005)

Vyhľadávanie informácií na základe vplyvu informačných štruktúr možno rozdeliť na:

1. vyhľadávanie vo voľnom texte,
2. postkoordinované indexovanie,
3. prekoordinované indexovanie,
4. klasifikáciu.

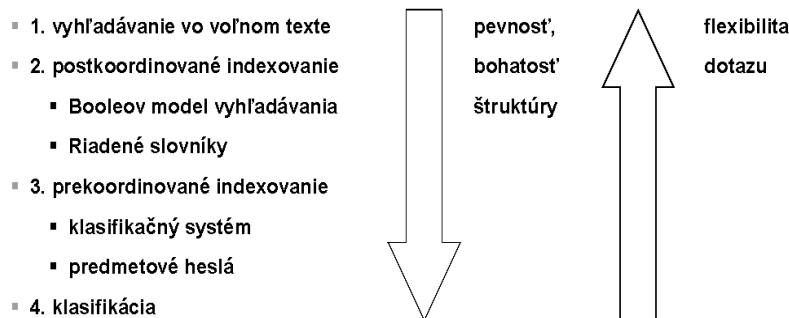
Informačné kontexty vznikajú na základe kognitívnej kategorizácie. V prípade indexovania a klasifikácie sa využívajú indexačné systémy ako špeciálne jazyky na vyjadrenie obsahu informácií. Známý je napríklad systém predmetových hesiel Kongresovej knižnice (*Library of Congress Subject Headings, LCSH*) alebo tezaurus v medicíne MeSH (*Medical Subject Headings*).

- *Pri vyhľadávaní vo voľnom texte* sa využíva voľná kategorizácia založená na kognitívnom procese. Chýba určený súbor princípov na riadenie štruktúry tried a ich vzťahov. S inými systémami má spoločné to, že sa vytvárajú vzájomne sa neprekrývajúce triedy a ich vzťahy. Na dotaz sa využíva napríklad alfanumerický reťazec znakov. Pri analýze textu aj pri vyhľadávaní sa berie do úvahy spoločná povrchová podobnosť, bez hlbšej sémantiky. Nedostatkom je a) výlučne syntaktický prístup zoskupovania a b) binárnosť vyhľadávania (delenie textu na skupiny ktoré sú buď zhodné s hľadaným reťazcom, alebo nie). Vzniknú dve protikladné

triedy podľa toho, či sú zhodné alebo nezhodné s reťazcom. Do úvahy sa neberú významové rozdiely medzi týmito triedami.

- *Postkoordinované systémy* sú založené na využívaní termínov indexačného jazyka nazývaných deskriptory. Deskriptor je znak informačného jazyka, ktorý má za cieľ reprezentovať intelektuálny obsah dokumentu. Deskriptor je teda súčasť informačného jazyka a normalizuje slovník používaný na opis a vyhľadávanie dokumentov. Deskriptor môže byť napríklad označenie triedy v klasifikačnom systéme, predmetové heslo, termín, alebo fráza. Riadený slovník umožňuje vyberať termíny z textov a určuje vzťahy medzi termínmi. Výhodou týchto systémov je interaktívnosť – spojenie termínov pri formulovaní dotazu (nie sú pevne spojené v systéme). Mechanizmus zoskupovania je založený na zhode alebo nezhode (zdroje, ktorých deskriptory zodpovedajú dotazu a tie, ktorých deskriptory nezodpovedajú dotazu). Postkoordinovaný systém je flexibilný, pretože používateľ si môže kombinovať deskriptory do vlastných dotazov a môže ich meniť. Zachytáva sa tak okamžitá informačná potreba. Kategórie sa vytvárajú flexibilne, podobne ako pri kognitívnej kategorizácii. Keďže indexátor priradzuje deskriptory z indexačného jazyka, existuje tu určitý stupeň kontroly pojmov (sémantická základňa). Príkladom vyhľadávania prostredníctvom postkoordinovaných systémov je vyhľadávanie prostredníctvom kľúčových slov a ich kombinácií (*Booleov model, riadené slovníky, tezaury*).
- *Prekoordinované systémy* sú založené na tom, že indexátori priradzujú dokumentom deskriptory vopred. Príkladom sú klasifikačné systémy, ale aj systémy predmetových hesiel. Predmetové heslá sú vlastne systémom kategorizácie, pričom vzťahy nemusia byť absolútne dobre definované, kategórie sa často prekrývajú. Viaceré deskriptory sa môžu priradiť jednému zdroju, čo umožňuje viaceré prístupové body pri vyhľadávaní. Interakcia so systémom je obmedzená podľa toho, ako vytvoril indexátor reťazce predmetových hesiel. Štruktúra vzťahov v predmetových heslách však nie je systematická, vzťahy musia byť dodané človekom. Napríklad „*potkany ako nositelia chorôb*“ alebo sú často opisné, v kruhu, napr. *filozofia – humanitné vedy – humanizmus – filozofia*. Komunikácia so systémom predmetových hesiel je jednosmerná, od systému k človeku. Komplikuje sa nedostatkom pevného rámca pri interpretovaní vzťahov medzi pojmami. Pri klasifikácii existuje jednoznačný vzťah medzi triedou a pojmom. Pri predmetových heslách je možno pričleniť viac deskriptorov k jednému zdroju, systém je kreatívny, flexibilný. Pri klasifikácii sú vzťahy trvalejšie a preto aj informačne bohatšie. Pri vyhľadávaní určuje možnosť formulovať dotazy konečný súbor hesiel alebo názvov tried. Sú napokon úplnou množinou možných prieskumových dotazov a vopred určujú zloženie vyhľadávaných množín. Príkladmi prekoordinovaného indexovania sú *klasifikačné systémy a predmetové heslá*.
- Najvyšší stupeň štruktúracie predstavuje *klasifikácia*. Štruktúra tried je vytvorená podľa pravidiel, triedy sú dobre definované. Systém je pevný, založený na hierarchických rodovo druhových vzťahov. Každý zdroj možno zaradiť len do jednej triedy v štruktúre. Existuje tu aj pojmový rámec, ktorý umožňuje vytvoriť zmysluplné štruktúrne vzťahy. Napriek tomu, že vzťahy sú tu bohaté, vyhľadaný súbor ako odpoveď na dotaz je obmedzený členstvom v jedinej triede. Na druhej strane klasifikačné systémy sú nezávislé od kontextu. Vzťahy sú nemenné, pretrvávajú v čase a priestore, preto sú odolné voči kontextu použitia. Komunikácia pri vyhľadávaní je jednosmerná, od systému k človeku. Súbor možných odpovedí je vopred predpísaný hierarchiou systému pojmov. Človek sa musí spoľahnúť na pochopenie štruktúrnych vzťahov medzi triedami. Príkladom je prísna *taxonómia druhov (biológia a pod.)*

Vplyv informačných štruktúr na vyhľadávanie informácií je naznačený na obrázku 7.2. Od voľného textu smerom ku klasifikácii stúpa pevnosť a bohatosť štruktúry, na druhej strane klesá flexibilita pri formulovaní dotazu vo vyhľadávaní.



Obr. 7. 2 Vplyv informačných štruktúr na vyhľadávanie informácií

## 7.4 Sémantické vzťahy

Informačné štruktúry vo vyhľadávaní informácií posilňujú kontexty informácií a podporujú prístup k informáciám a zdrojom. Prejavujú sa ako v kognitívnej vrstve (myšlienkové manipulácie s významom), tak aj vo fyzickej vrstve (uloženie dokumentov v knižnici, uloženie súborov v pamäti systému, triedenie informácií v hierarchii súborov a v osobnom informačnom systéme).

Základom tvorby informačných štruktúr sú sémantické vzťahy medzi pojmami. Základnou schémou je: **pojmem – vzťah – pojem**. Sémantické vzťahy vznikajú prostredníctvom vzťahu jazyka a myslenia. Spojenia medzi pojmami sú nositeľmi významu a hrajú najdôležitejšiu úlohu pri budovaní tezaurov a slovníkov ako nástrojov na reprezentáciu poznania.

Medzi vlastnosti sémantických vzťahov patria:

- Otvorenosť
- Produktívnosť
- Binárnosť (protirečivé pojmy, antonymá)
- Spojenie s udalosťou (*napríklad udalosť: nákup, vzťahy: predajca, kupujúci, objekt, cena*).

Druhy sémantických vzťahov možno rozdeliť na:

1. Vzťahy ekvivalencie
2. Vzťahy hierarchie
3. Vzťahy asociácie
4. Iné vzťahy

*Vzťahy ekvivalencie* obsahujú najmä homonymiu (znak obsahuje viac konceptov) a polysémiu (sémanticky príbuzné koncepty), synonymiu (jeden koncept označujú viaceré znaky a polylexiu (viac slov s jedným významom)).

*Vzťahy hierarchie* predstavujú najmä hypernymia (nadradené pojmy) a hyponymia (podradené pojmy). Rodovo-druhovú vzťahy vyjadrujú vzťah všeobecného a špecifického, prípadne triedy a výskytu (rodič – dieťa). Najčastejšie sa objavujú v klasifikácii a v slovníkoch (tezauroch).

*Vzťahy asociácie* môžu byť rôznorodé, najčastejšie však časť – celok, predmet – použitie, predmet – vlastnosť, príčina – účinok (kauzalita), proces – prostriedok.

Ďalšie dôležité vzťahy predstavujú antonymiu (protiklady), nulový znak (nevyjadrený znak má svoj význam v systéme pojmov) a autonymiu (znak zastupuje seba samého).

## 7.5 Reprezentácie poznania

S informačnými štruktúrami súvisí aj problém **reprezentácie poznania**. Reprezentácia poznania je dohovorená štruktúra informácií, zároveň však aj proces opisu objektov. V psychológii a kognitívnej vede hovoríme o reprezentáciách poznania. Sú to rôzne druhy myšlienkových modelov vznikajúcich pri vnímaní objektov a spracovaní informácií.

V oblasti informatiky a umelej inteligencie sa reprezentácia definuje ako uloženie, usporiadanie znalostí (informácií) s vyjadrením väzieb (sémantických vzťahov) medzi nimi s cieľom odvodzovať ďalšie informácie. Pri reprezentácii poznania ide o modelové zachytenie objektu (najmä jedného vybraného aspektu objektu) so zachytením sémantických vzťahov. Problémom reprezentácie poznania sú otázky o tom, čo reprezentovať (predmet reprezentácie), prostredníctvom čoho (prostriedok reprezentácie) a čím reprezentovať (reprezentant objektu). Základné metódy reprezentácie poznania v inteligentných informačných systémoch možno deliť takto (Steinerová, 1996):

1. Deklaratívne metódy
  - predikátová logika
  - produkčné pravidlá (ak – potom)
2. Asociatívna reprezentácia
  - sémantické siete
  - kauzálne siete
3. Procedurálna reprezentácia
4. Rámcová reprezentácia
5. Iné (hybridné).

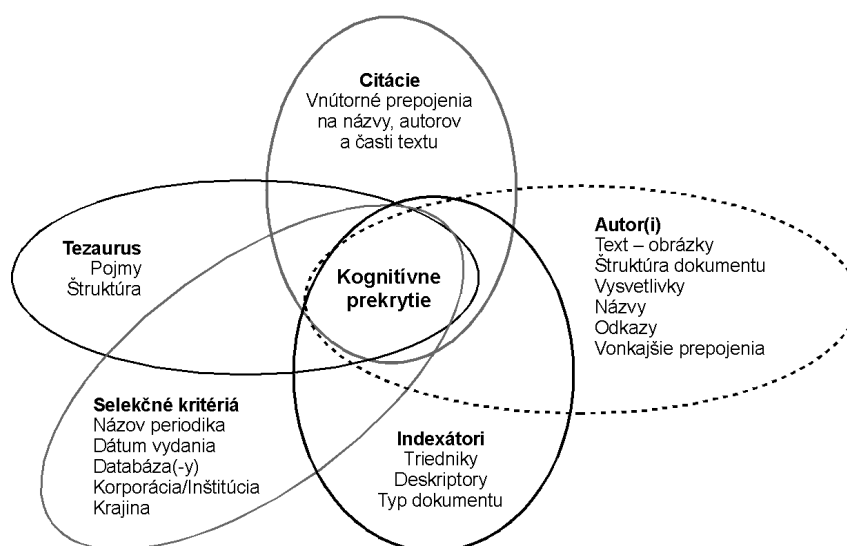
Metódy a nástroje reprezentácie poznania riešia problémy **vyjadrenia poznatkov** (deklarácia), **začleňovania poznatkov** (asociácia) do vzťahu k iným poznatkom a **použitia poznatkov** (procedúry). Podľa toho sa deklaratívne metódy zameriavajú na riešenie vyjadrovania poznatkov. Môže ísť napríklad o nástroje formálnej logiky (matematický algebraický jazyk na zápis viet s vyjadrením logických vzťahov, najmä logický súčin, logický súčet, negácia, implikácia – čo z niečoho vyplýva, ekvivalenciu, zovšeobecnenie a i.). Produkčné pravidlá sa vyvinuli z tzv. situačného programovania. Umožňujú vyjadriť poznatok tak, že sa zapíše pravidlo, ktoré vyjadruje určitú situáciu obsahujúcu východiská a dôsledky, ktoré platia pre danú situáciu. Takéto pravidlá sa nazývajú produkčné pravidlá alebo pravidlá typu AK (podmienka) – POTOM (dôsledok). Pravidlá môžu byť aj štruktúrované a prepájať sa do zložitejších celkov v bázach znalostí. Asociatívna reprezentácia kladie dôraz na prepojenia, súvislosti. Zvyčajne sa graficky znázorňuje ako orientovaný graf s uzlami (napríklad pojmy) a hranami (vzťahmi medzi pojmi). Sémantická sieť je podkladom aj webových hypertextových prepojení. Sémantické siete môžu byť usporiadané hierarchicky, ale aj prepojené do širších celkov. Iným typom sú napríklad kauzálne siete, ktoré znázorňujú príčinu a účinok (kauzalitu), prípadne iné kauzálne vzťahy. Procedurálna reprezentácia označuje postupnosť procedúr, ktoré sa spracúvajú v systéme (napríklad začiatok, vykonanie operácie, ukončenie). Možno ju dobre vyjadriť v programovacích jazykoch, napríklad už klasickým jazykom pri tvorbe báz dát a expertných systémoch bol jazyk Prolog.

Rámcová reprezentácia poznania integruje všetky tri aspekty reprezentácie – deklaráciu, asociáciu aj procedúru. Prvýkrát ju v 70. rokoch 20. stor. definoval významný odborník z oblasti umelej inteligencie M. Minsky. Rámec je štruktúra, ktorá môže nadobúdať rôzne hodnoty podľa konkrétnej situácie. Možno si ju predstaviť ako tabuľku, v ktorej sa vyplňa definícia pojmu, jeho vlastnosti, asociácie s inými pojmi aj procedúry, ktoré sa s objektom/pojmom spájajú. Príklad rámca je v tabuľke 7.2.

Tab. 7.2 Príklad rámcovej reprezentácie poznania (rámec „Auto“)

AUTO	vlastnosti	typy	vzťahy
	4 kolesá	Osobné – 5 osôb	Dopravné prostriedky
	riadenie	mikrobus	Značky, typy
	Preprava osôb	nákladné	Športové jazdenie, preteky
	Rýchlosť, pohonné hmoty	kabriolet	Vodičský preukaz

V reálnych systémoch sa vyskytuje tzv. polyreprezentácia poznania, ktorú skúmal Peter Ingwersen. Vychádza z kognitívnej teórie informačného prieskumu, ktorou sa zaoberajú aj Amanda Spinková, Charles Cole, Nigel Ford. Počíta s tým, že paralelne človek spracúva rôzne druhy reprezentácií. Majú rôzne sémantické obsahy. Kognitívne priestory používateľa sú charakteristické variabilitou požiadaviek, problémov, pracovných úloh a záujmov. V praxi môže ísť o rôzne reprezentácie dokumentu ako citácie, tezaury, autorské texty, znaky indexátorov. Princípy polyreprezentácie sú naznačené na obrázku 7.3 (podľa Ingwersen, Järvelin 2005).



Obr. 7.3 Polyreprezentácia poznania (podľa Ingwersen, Järvelin 2005)

Polyreprezentačný charakter vyhľadávania informácií a informačného prieskumu dokazujú aj experimenty s veľkými súbormi textov v rámci testov TREC (Text Retrieval Conference). Tieto testy sú pokračovaním Cranfieldských testov. Začali sa v roku 1992 a pokračujú každý rok testovaním spracovania novinových článkov v rôznych jazykoch. Podobne sa rozvíja aj viacjazyčný informačný prieskum (CLIR – Cross-language information retrieval). Znamená to vyhľadávanie informácií vo viacerých prirodzených jazykoch.

Pri polyreprezentácii poznania sa rozlišuje štruktúrovaný a neštruktúrovaný prieskum. Štruktúrovaný prieskum je spojený s vyhľadávaním v databázových systémoch, neštruktúrovaný je založený na vyhľadávaní v plnom texte. Každý text má svoju vnútornú štruktúru, avšak ide o to, ako je vyjadrená v reprezentácii poznania (napríklad aj v značkovacích jazykoch ako HTML a XML). Rozdiely naz-

načuje obrázku 7.4. Sú založené aj na vývoji technológií, od tradičného prieskumu k inteligentnému prieskumu (Steinerová 2005).

- **Štruktúrovaný prieskum**
  - Presná zhoda
  - Množiny
  - Prekrytia
  - Pseudo-rangovanie
- **Neštruktúrovaný prieskum**
  - Najlepšia zhoda
  - Prahy
  - Prepojenie rangov
  - Kontinuálne rangovanie

Obr. 7.4 Rozdiely medzi štruktúrovaným a neštruktúrovaným prieskumom

Význam reprezentácií poznania a sémantických vzťahov možno charakterizovať najmä prostredníctvom vyjadrovacej účinnosti, vyvodzovania v prieskumových systémoch a efektívnosti vyhľadávania. Pri vyjadrovaní prostredníctvom jazyka reprezentácie poznania je cieľom zachytiť dôležité vlastnosti, funkcie a vzťahy objektov. Pre fungovanie prieskumového systému je dôležité aj to, ako je algoritmicky zabezpečené vyvodzovanie záverov prípadne nových informácií zo zaznamenaných informácií v pamäti. V algoritmoch je dôležitá manipulácia s informáciami a ich hodnotenie. Efektívnosť vyhľadávania s najčastejšie vyjadruje prostredníctvom hodnotenia relevancie informácií, ktorá je čiastočne včlenená v algoritmoch a funkciách porovnávania. V používateľskom zmysle je hodnotenie relevancie myšlienkovým, kognitívnym procesom. Význam reprezentácií poznania je schematicky znázornený na obrázku 7.5.

- **Vyjadrovacia účinnosť poznania**
  - Zachytenie podstatných vlastností, funkcií a vzťahov objektov
- **Vyvodzovacia účinnosť informačného prieskumového systému**
  - Procesy – manipulácia, spracovanie informácií, hodnotenie informácií
- **Účinnosť, efektívnosť vyhľadávania**
  - Relevancia

Obr. 7.5 Význam reprezentácií poznania pre vyhľadávanie

## 7.6

### **Kultúrne odlišnosti v reprezentáciách**

Klasifikácia a kategorizácia sú procesy, ktoré sú postavené na mentálnych reprezentáciách. Ide o vyjadrenie myšlienkových procesov pri vytváraní a manipulovaní s významom na základe faktov (sémantika a definície pojmov, deklaratívna reprezentácia) a na základe hodnôt (axiologické významy, skupinové, kultúrne, transkultúrne).

Hodnotová zložka reprezentácií sa prejavuje v rozdieloch pri kategorizácii a klasifikácii v rôznych kultúrach. Kultúrne odlišnosti pri tvorbe kategórií súvisia s kultúrou ako spôsobom života a prispôso-

bovaním sa skupiny svojmu prostrediu (adaptácia). Rozdiely možno vidieť už medzi tzv. západnou (americko-západoeurópskou) civilizáciou a inými spoločenstvami (africké, ázijské). Prejavuje sa to aj v konkrétnom a abstraktnom myslení, v participácii na živote komunit. Napríklad nie všetci využívajú matematickú západnú desiatkovú sústavu, niektoré kmene v Austrálii či Afrike používajú na počítanie dvojkovú sústavu. Rozdiely vidieť aj pri vyjadrovaní emócií (prekvapenie, súhlas, nesúhlas) a pri označovaní a hodnotení prostredníctvom pojmov. Napríklad v Británii existuje veľa výrazov na lietanie, africký kmeň pozná iba jeden. Psychológovia experimentálne skúmali aj kategorizáciu pojmov u rôznych izolovaných kmeňov (Ruisel 2004). Niektorí poznajú iba dva pojmy na označenie farby (tmavé, svetlé, tma/svetlo). Už A. Lurija dokázal, že jazykové označovanie odtieňov závisí od praxe určitej kultúrnej skupiny.

Kategorizácia však závisí aj od stupňa vývinu jednotlivca, pričom u mladších detí často splývajú vlastnosti nápadnosti, kvality. Stupňovité triedenie objektov sa uskutočňuje podľa schémy farba – forma – funkcia. Rozdiely vyplývajú nielen z vnímania objektov, ale aj z riešenia problémov. V západnom spoločenstve sa zdôrazňuje manipulácia s predmetmi, analytické myslenie, ovládanie objektov, vo východných kultúrach sa uprednostňujú sociálne vzťahy, spolupráca, bytie.

Pri klasifikácii sa pojmy tvoria tiež na základe skúsenosti. Často sa pre nás klasifikované predmety spájajú s určitými dôsledkami (napríklad *dobrý pes/zlý pes*). Pojmy sa všeobecne delia na formálne pojmy (vyjadrujúce podstatné charakteristiky objektov, sú definované a možno ich hierarchicky organizovať v doménach) a prirodzené pojmy (subjektívne vnímanie vonkajšej reality každodenného života, opis bežných predmetov).

## 7.7 Zhrnutie

Kategorizácia a klasifikácia sú základom interakcie človeka s informačným prostredím. Umožňujú ho organizovať. Odlišujú sa prísnyim delením tried (klasifikácia) a pružnejším a kreatívnejším narábaním s kategóriami (kategorizácia).

V informačnej vede sú výsledkom profesionálnej klasifikácie a kategorizácie informačné štruktúry, ktoré ovplyvňujú možnosti vyhľadávania informácií. Podľa pevnosti týchto štruktúr sa rozlišujú štyri druhy vyhľadávania: vyhľadávanie vo voľnom texte, vyhľadávanie na základe postkoordinovaného indexovania (napríklad Booleov model, riadené slovníky), vyhľadávanie na základe prekoordinovaného indexovania (napríklad predmetové heslá, klasifikačný systém) a vyhľadávanie na základe prísnej klasifikácie (taxonómia). Informačné štruktúry sú založené na sémantických vzťahoch ako hierarchia, ekvivalencia, asociácia. Pri vyhľadávaní informácií je dôležitá aj reprezentácia poznania ako typ informačnej štruktúry naznačujúcej usporiadanie poznatkov v bázach. Medzi základné spôsoby reprezentácie poznania patria deklaratívne metódy (predikátová logika, produkčné pravidlá), asociatívne metódy (sémantické siete, kauzálne siete) a procedúrne metódy.

Význam kategorizácie a klasifikácie a reprezentácií poznania je v podpore efektívnosti vyhľadávania, vyvodzovania a vyjadrovania významu informácií. Procesy kategorizácie a klasifikácie sú kultúrne podmienené a závisia od stupňa vývoja jednotlivca. Zložitosť reprezentácií, kategorizácie a klasifikácie vplyvajú na možnosti vyjadrovania informačnej požiadavky a určujú aj efektívnosť vyhľadávania informácií.

## Literatúra

BEŇUŠKOVÁ, Ľubica. 2002. Kognitívna neuroveda. In Eds. RYBÁR, J., BEŇUŠKOVÁ, Ľ., KVASNIČKA, V. *Kognitívne vedy*. Bratislava : Kalligram, 2002. ISBN 80-7149-515-8, s. 47 – 101.

- INGWERSEN, Peter – JÄRVELIN, Kalervo. 2005. *The Turn : Integration of Information Seeking and Retrieval in Context*. Dordrecht : Springer, 2005. 448 s. ISBN 1-4020-3850-X.
- JACOB, Elin K. 2004. Classification and categorization : a difference that makes a difference. In *Library Trends* [online]. Wntr, 2004 [cit. 2009-02-07]. Dostupné na internete: <[http://findarticles.com/p/articles/mi\\_m1387/is\\_3\\_52/ai\\_n6080402/?tag=content:col1](http://findarticles.com/p/articles/mi_m1387/is_3_52/ai_n6080402/?tag=content:col1)>.
- MAKULOVÁ, Soňa. 2002. *Vyhľadavanie v internete : problémy, východiská, postupy*. Bratislava : EL&T, 2002. 376 s. ISBN 80-88812-16-X.
- MAKULOVÁ, Soňa. 2005. Návrh riešenia problémov pri vyhľadávaní informácií v internete alebo od kvantity ku kvalite. In *Knihovna* [online]. 2005, roč. 16, č. 1 [cit. 2009-08-21]. s. 9 – 22. Dostupné na internete: <<http://knihovna.nkp.cz/knihovna51/5123makul.htm>>. ISSN 1801-3252.
- PLICHTOVÁ, Jana. 2001. Reprezentácie : mentálne alebo sociálne? In Eds. VIŠŇOVSKÝ, E., POPPER, M., PLICHTOVÁ, J. *Príbehy o hľadani mysle*. Bratislava : Veda SAV, 2001. ISBN 80-224-0679-1, s. 106 – 137.
- RUISEL, Imrich. 2004. *Inteligencia a myslenie*. Bratislava : Ikar, 2004. 432 s. ISBN 80-551-0766-1.
- RYBÁR, Ján. 2002. Kognitívna psychológia. In Eds. RYBÁR, J., BEŇUŠKOVÁ, Ľ., KVASNIČKA, V. *Kognitívne vedy*. Bratislava : Kalligram, 2002. ISBN 80-7149-515-8, s. 105 – 136.
- SEDLÁKOVÁ, Miluše. 2004. *Vybrané kapitoly z kognitívni psychologie : Mentální reprezentace a mentální modely*. Praha : Grada Publishing, 2004. 252 s. ISBN 80-247-0375-0.
- STEINEROVÁ, Jela. 1996. *Teória informačného prieskumu*. Bratislava : CVTI SR, 1996. 262 s. ISBN 80-65165-58-9.
- STEINEROVÁ, Jela. 2005a. *Informačné správanie : Pohľady informačnej vedy*. Bratislava : CVTI SR, 2005. 189 s. ISBN 80-85165-90-2.



# 8

## Organizácia poznania v elektronickom prostredí

Snahu o organizáciu dát, informácií a poznatkov zaznamenávame už od nepamäti. V súčasnej ére sa musia ľudia vyrovnávať s dopadmi informačnej explózie. Organizácia informácií zjednodušuje neskoršie vyhľadávanie a sprostredkúva efektívnejšiu prácu s informáciami.

### 8.1 Pohľad do histórie

Korene organizácie poznania môžeme hľadať v Aristotelovej filozofii, ktorý približne v polovici 4. storočia p.n.l. predstavil teóriu kategórií a tiež rozlíšil druhy a rody. Následne položil základy vied o klasifikácii a taxonómii.

Okolo roku 300 p. n. l. vznikla výnimočná knižnica v Alexandrii, ktorá bola preslávená svojim najväčším repozitárom antickej literatúry. Po asi 50 rokoch od jej založenia obsahovala už 750 000 zvitkov. O tento nárast sa postarali Ptolemaiovcí, ktorí zvitky získavali a zbierali niekedy aj kurióznym spôsobom. Návštevníkom Alexandrie napríklad skonfiškovali knihy, ktoré si so sebou priniesli a ešte neboli uložené v knižnici. Ptolemaios III. si požičiaval najvzácnejšie rukopisy priamo z Atén. Tie sa však do starovekého Grécka nikdy nevrátili, pretože Ptolemaios III. sa postaral o to, aby sa ich originály dostali do Alexandrijskej knižnice. Do Atén sa tak vrátili už len kópie rukopisov napr. Sofokla, či Euripida, spolu s úplatkami vo forme striebra. Alexandrijská knižnica čoraz viac narastala a preto ju bolo potrebné nejako organizovať. Riešenie priniesol historicky prvý katalóg, ktorý vznikol v roku 331 p.n.l. a radil zdroje v knižnici podľa abecedy.

V histórii organizácie poznania nastalo ešte mnoho zlomových bodov, ako napr. vznik Linného taxonómie, či prvé tezaury a klasifikácie. Taxonómia triedi pojmy do hierarchicky usporiadaných kategórií. Taxonómie v biológii klasifikujú organizmy do taxonomických úrovní (riša – kmeň – trieda – rad – čeľaď – rod – druh). Z hľadiska nasadenia počítačov na pomoc pri organizovaní informácií je dôležité spomenúť vznik špecializovanej oblasti informačného prieskumu. Tento pojem zaviedol v roku 1951 C. N. Mooers a vníma ho dosť široko. Podľa neho zahŕňa aspekty popisu informácií a špecifikácie vyhľadávania. Tieto operácie mohli byť realizované systémami, technikami alebo strojmami. Stroje nám pri organizácii a vyhľadávaní informácií pomáhajú už viac ako polstoročie. Najznámejšími príkladmi sú v súčasnosti prieskumové stroje Google, Ask, Bing, a pod. Cesta k ich vývoju ako ich poznáme dnes viedla cez dlhodobé experimentovanie, skúmanie a programovanie (napr. Cranfieldske testy, TREC). Na začiatku boli dávkové, neskôr dialógové systémy a napokon systémy vyhľadávania na internete.

## 8.2 Od tradícií k metadátam

Organizácia poznania v elektronickom prostredí súvisí s usporiadaním informácií a poznatkov v pamäti prieskumových systémov. Nové nástroje organizácie poznania sa vyvinuli aj pri využívaní **meta-dátových štruktúr** v digitálnych knižniciach. Metadáta sú informácie, ktorých funkciou je uľahčiť prístup k dokumentu a jeho použitie. Sú súčasťou informačných zdrojov (dokumentov) v elektronickom prostredí. Informujú o obsahu a forme dokumentov a bývajú zakódované v značkovacích jazykoch (HTML, XML, SGML). Tieto štruktúry výrazne uľahčujú vyhľadávanie informácií v elektronickom prostredí. Najznámejšou iniciatívou v tejto oblasti je Dublin Core Metadata Initiative (DCMI).

Slovníky, tezaury a terminologické databázy v elektronickom prostredí majú cieľ mapovať lingvistické prostriedky v používateľskom prístupovom rozhraní.

Informačné správanie v elektronickom prostredí je voľné, dynamické a interaktívne. S tým súvisí aj individuálna a komunitná tvorba vlastných nástrojov organizácie poznania. To umožňujú napríklad systémy manažmentu obsahu (Content Management Systems, CMS), ktoré ponúkajú flexibilnú a oproti tradičným systémom organizácie poznania jednoduchú a nezávislú „samoobslužnú“ štruktúru informácií. Jedným z prejavov sú napríklad aj systémy manažmentu obsahu predmetov (Course Management Systems) v akademickom prostredí, ktoré sa budujú v rámci digitálnych knižníc alebo vzdelávacích portálov. V digitálnych knižniciach sa uplatňujú najmä také princípy organizácie poznania ako odkazovanie (linkovanie), anotované bibliografie, recenzovanie, fazetový prieskum a riadené slovníky.

Neformálnou elektronickou komunikáciou na internete vznikajú aj osobitné ľudové „triediace systémy špeciálnych komunit (tzv. „folksonómie“). Sú výsledkom komunikácie a spolupráce ľudí pri špecifických problémoch.

Novšie nástroje organizácie poznania v elektronickom prostredí kombinujú rozličné princípy tradičných schém s otvorenosťou a nelineárnosťou webu. Nové štruktúry vyplývajú z jazyka komunikácie medzi členmi komunit. Podobne ako cesty informačného správania sú tieto nové štruktúry kľukatými prepletenými štruktúrami prepojení. Možno ich metaforicky opísať ako „sémantické mapy“ alebo „rizómy“. Rizóm“ je prepletený koreňový systém bez začiatku a konca pripomínajúci pavučinu hypertextových prepojení. Organizácia poznania je potom nástrojom, ktorý umožňuje riešiť problémy a kooperovať v pracovných a iných doménach.

V praxi sa potvrdzuje prechod od neutrálnych a univerzálnych systémov organizácie poznania ku kontextovo bohatým štruktúram. Napríklad poznatky z nášho výskumu správania používateľov dokazujú rozdielnu mieru závislosti používateľov od kontextu pri využívaní informácií. Analytický typ kladie dôraz na súvislosti, spoľahlivosť a viac kontextu. Pragmatický typ kladie dôraz na rýchlosť, nízku cenu a menej menej kontextu.

Budovanie sémantického webu a reprezentácií poznania v inteligentných systémoch sa opiera o hlbšie poznanie modelov sémantiky a syntaxe prirodzených jazykov. Do používateľských rozhraní sa včleňujú slovníky alebo fazetové klasifikácie. Rôzne systémy a slovníky sa navzájom kombinujú a vznikajú nové nástroje organizácie poznania. Využíva sa aj možnosť oddeľovať obsah (poznatkov) od služby (prezentácií). Prezentácia sa pod vplyvom informačného správania môže pretvoriť do personalizácie, rozhrania, špecifického slovníka, rôznych štádií využívania informácií a rôznych prostredí (odlišných kultúrne, lokálne, historicky alebo skupinovo). Sémanticky bohaté kontexty umožňujú aj filtrovať informácie a riešiť informačné preťaženie.

Nové nástroje organizácie poznania spájajú predmetový prístup s hypertextovými prepojeniami, špecializované slovníky a hierarchie. Často umožňujú objavovať nové poznatky.

Za tradičné systémy organizácie poznania v knižničnom a informačnom prostredí považujeme najmä klasifikačné systémy, fazetové systémy a systémy predmetovej organizácie poznania. V týchto

tradičných systémoch sa vytvárali také štruktúry poznania, ktoré znižovali kognitívnu záťaž pri chápaní zložitých významov objektov.

Tradičný hierarchický klasifikačný systém pracoval so „zjednodušeným“ obrazom univerza poznania a s vopred definovanými a navzájom sa vylučujúcimi triedami. Predpokladal, že tento obraz sveta je pravdivý a relatívne stabilný. Usporiadáva chaotický svet informácií. Systémy organizácie poznania vznikali najmä vo vedeckých disciplínach a predpokladali systematický spôsob informačného správania pri využívaní informácií.

V týchto systémoch hrajú dôležitú úlohu zložky semiotiky: *sémantika*, t.j. význam informácií sprostredkovaný pojmami a kategóriami, *syntax* – štruktúra vytvorená bohatosťou vzťahov medzi pojmami ako významný kontext využívania informácií a *pragmatika* ako širší kontext použitia kategorizovaných informácií v disciplíne, doméne a komunite.

Vyhľadávanie informácií a organizácia informácií v internete sa v zásade zakladajú na dvoch princípoch. Na jednej strane sú to populárne **klúčové slová** umožňujúce špecifické vyhľadávanie prostredníctvom vyhľadávačov s indexovanými súbormi. Na druhej strane sú **adresáre a katalógy**, ktoré poskytujú určitú štruktúraciu či triedenie tematickej oblasti. Niektoré systémy kombinujú tieto dva spôsoby štruktúracie. Málokto však využíva tradičné knižničné a bibliografické systémy organizácie poznania. Problémom je tradícia a mechanizmy vyhľadávania, ktoré vychádzajú z výskumov informačného prieskumu v 50. rokoch 20. stor. Ukazuje sa, že ich mechanický prenos do interaktívneho elektronického prostredia spôsobuje množstvo problémov. Ide najmä o prácu s významom informácií, riadenie dvojznačnosti prirodzeného jazyka a problém konceptualizácie a kategorizácie informácií.

## 8.3

### Zoznamy termínov a klasifikačné systémy

Informačné správanie používateľov v elektronickom prostredí dokazuje potrebu kognitívnej podpory, ktorú poskytovali tradičné systémy organizácie poznania. Na jednej strane sa tradičné systémy prenášajú do elektronického prostredia. Na druhej strane sa vytvárajú nové nástroje na organizáciu poznania. Sú to napríklad riadené slovníky, tezaury, fazetové triedenia, klasifikácie, taxonómie, miestopisné slovníky, ontológie. V elektronickom prostredí fungujú nielen ako nástroje organizovania, vyhľadávania a objavovania skrytých poznatkov, ale z používateľského hľadiska najmä ako nástroje orientácie, navigácie a konštrukcie významu.

Z hľadiska organizácie informácií môžeme rozdeliť nástroje v elektronickom prostredí na tieto druhy:

- zoznamy termínov: súbory autorít, registre, slovníky
- klasifikačné systémy a systémy kategorizácie: systémy predmetových hesiel, klasifikačné systémy, taxonómie
- relačné zoznamy: tezaury, sémantické siete, ontológie
- pojmové mapy, tematické mapy
- používateľsky tvorené systémy: folksonómie

*Zoznamy termínov* sú najjednoduchšími nástrojmi, ktorých funkciou je poskytnúť ustálenú formu termínov a určiť ich správne alebo preferované formy. V knižničných automatizovaných systémoch sú to najmä *súbory autorít* rôznych druhov. Obsahujú mená autorov, krajín, organizácií a určujú ich správny tvar, prípadne vylučujú nesprávne. Zväčša sú usporiadané abecedne a predstavujú lingvistikú povrchovú štruktúru. Hierarchia aj asociácie sú obmedzené. Každý väčší, prípadne národný knižnično-informačný systém pracuje so súborom autorít.

*Registre* obsahujú rôzne súbory termínov, niekedy obohatené definíciami. Termíny môžu byť vybraté z dokumentu (napríklad tradičný register k odbornej publikácii) alebo spracované ako termi-

nológia špecifickej oblasti (napr. environmentalistika). Tvorbu registrov v knižničnej a informačnej vede určujú aj tradičné normy (Steinerová 1998).

*Slovníky* sú všeobecnejšie ako registre, reprezentujú ich ustálené súbory jazykových jednotiek – slová, slovné spojenia, mená, termíny skratky, znaky. Môžu obsahovať rôzne výklady, varianty slova, odkazy, synonymá informácie o pôvode slova (etymologické slovníky) a pod.

*Miestopisné slovníky* (gazetteers) obsahujú zoznamy názvov geografických miest. Spájajú sa najmä s geografickými informačnými systémami. Obsahujú aj jednoduchú klasifikáciu alebo kategorizáciu (mestá, rieky, budovy a pod.).

*Klasifikačné systémy* vytvárajú prísne triedy pri informačnej štruktúre. Delia univerzum poznania do určitých tematických okruhov, stanovujú terminológiu pri pomenovaní tried. Sú enumeratívne, často obsahujú označenia tried prostredníctvom špecifických značení (notácií). Predmetové heslá sú výrazy, ktoré vyjadrujú tematiku dokumentu. Vyberajú termíny z heslára a dodržiavajú pravidlá ich spájania. Využívajú aj odkazy na vylúčovanie alebo pridružovanie termínov. Najznámejšie systémy predmetových hesiel sú LCSH (Library of Congress Subject Headings) a MeSH (Medical Subject Headings).

*Relačné zoznamy* sú súbory termínov s určitou špecifickou štruktúrou. Najznámejším je tezaurus ako riadený slovník termínov v jednej tematickej oblasti s určenými vzťahmi medzi termínmi, najmä hierarchickými, synonymiou, asociáciou a i.

Sémantické siete určujú pojmy a vzťahy medzi nimi a reprezentujú tak asociatívne poznatky. Graficky sa znázorňujú ako orientovaný graf obsahujúci uzly (pojmy) a hrany (vzťahy, prepojenia). Pojmy možno zoskupovať do ďalších skupín (širšie-užšie, hierarchie, asociácie). Fungovanie WWW a hypertextov podmieňuje vždy sémantická sieť prepletených pojmov. Sledovanie spojení pri vyhľadávaní je vlastne pohybom po sémantickej sieti.

Ontológie sú špeciálne pojmové modely, ktoré modelujú komplexné vzťahy medzi pojmi. Podrobnejšie sú opísané v ďalších častiach textu.

## 8.4

### **Delenie nástrojov podľa zložitosti štruktúr**

Nové nástroje organizácie poznania v elektronickom prostredí by mali podporovať proces konštrukcie významov pri vyhľadávaní informácií. Vyznačujú sa nielen otvorenosťou a dynamikou, ale aj možnosťou zasahovať do informačných štruktúr aj v priebehu prieskumu. Tradičné systémy neboli takýmto spôsobom otvorené a flexibilné.

Z hľadiska zložitosti možno tieto novšie nástroje organizácie poznania v elektronickom prostredí rozdeliť na:

1. Metadátové štruktúry (Dublin Core Metadata Initiative, DCMI)
2. Taxonómie
3. Tematické mapy (pojmové modely disciplín)
4. Pojmové mapy
5. Myšlienkové mapy
6. Systémy CMS (content management systems)
7. Folksonómie
8. Ontológie

#### *Metadátové štruktúry*

Ide o súbor definovaných prvkov na opis zdrojov (informačných objektov) s cieľom zabezpečiť interoperabilitu a jazykovú kompatibilitu systémov a služieb, jazykov a jednotlivcov (sémantický, syntaktický, pragmatický aspekt).

Príklad metadátovej štruktúry Dublin Core v prvej verzii je na obrázku 8.1. Funkciou metadát je opisovať dokumenty a informačné objekty, ale aj zabezpečiť rôzne administratívne funkcie (napr.

riadenie prístupu k zdrojom v digitálnej knižnici, riadenie autorských práv a pod.). Podľa toho sa metadáta ešte klasifikujú, napríklad na formálno-opisné, lokačné, obsahové, administratívne, používateľské a pod.

názov	autor	predmet
opis	vydavateľ	prispievateľ
dátum	typ	formát
identifikátor	zdroj	jazyk
vzťah	pokrytie	práva

Obr. 8.1. Metadáta podľa Dublin core metadata initiative (DCMI), 2003

### *Taxonómie*

Taxonómie sú klasifikačné hierarchické systémy. Etymologicky je taxonómia zložené slovo z gréčtiny, pričom taxis je usporiadanie, poriadok a nomos je zákon. Pôvodne bola taxonómia v biológii klasifikácia živých organizmov (rastliny, živočchy). Známa je Linného taxonómia rastlín. Neskôr sa však začala uplatňovať aj v širšom zmysle na klasifikáciu rôznych objektov. Taxonómia je teda usporiadaný klasifikačný systém. V najužšom zmysle slova je prísne hierarchický a obsahuje rodovo druhové vzťahy (rodič/dieťa). Informácie sú v taxonómii zoskupené podľa predpokladaných prirodzených vzťahov. Ako príklady taxonómie možno uviesť periodickú tabuľku chemických prvkov, Deweho desatinnú klasifikáciu, triedenie kategórií v Yahoo!

### *Ontológie*

Ontológia je všeobecne filozofická disciplína, ktorá skúma bytie a objekty reality. Ontológie ako systémy organizácie poznania sú pojmové modely, ktoré opisujú poznanie v špecifickej oblasti alebo v organizácii. Oproti tradičným prístupom spájajú logickú a sociálnu štruktúru poznatkov. Reprezentujú zložité vzťahy medzi objektmi a využívajú systém pravidiel na riadenie sémantiky a syntaxe.

Ontológiu možno definovať ako špecifický pojmový model vyjadrujúci logickú a sociálnu štruktúru poznatkov vo vysoko formalizovanom jazyku.

Z pohľadu knižničnej a informačnej vedy je definovanie ontológie problematické. Nie sú to enumeratívne schémy ani slovníky, sú príkladom vynorenia sa novej štruktúry na základe poznania informačného správania ľudí. Ontológie umožňujú vyjadriť dynamiku poznatkov (neustály vývoj poznania v organizáciách), zložitost' poznatkových štruktúr a ich sociálnu rozptýlenosť. Z hľadiska manažmentu poznatkov má takýto systém pomôcť riešiť problémy informačného správania. Tieto problémy vyplývajú z nejednoznačnosti pojmov, informačného preťaženia a delenia sa o poznatky (komunikovanie v skupinách).

Význam ontológií dokazuje aj výskum sémantického webu a vytváranie jazyka ontológie webu (OWL) a mnohé ďalšie výskumy smerujúce k poznatkovému či inteligentnému webu.

Ontológie sa opierajú o špeciálne pojmové modely a môžu využívať rôzne moderné modelovacie nástroje či jazyky (napr. UML, Unified Modeling Language).

V ontológii sa definujú pojmy, termíny a reťazce. Oproti tezaurum majú formálnejšiu hierarchiu alebo detailnú štruktúru vzťahov s prísne definovanými pravidlami kódovania. Týmto smerom sa vyvíja napríklad tezaurus FAO AGROVOC.

Ontológie sa začali v znalostnom inžinierstve uplatňovať najmä v 90. rokoch 20. stor. Vzniklo aj ontologické inžinierstvo, ktoré prepája sémantický web a ontológiami. Jeho cieľom je zdokonaľovať reprezentáciu významu informácií a interakciu človeka so systémom.

Sociálny rozmer ontológie predstavuje reprezentácia pravidiel komunikovania v určitej komunite. Niekedy je ťažké rozlíšiť, či ide o inteligentný tezaurus, sémantickú sieť alebo ontológiu. Príkladmi môžu byť Tezaurus FAO Agrovoc, WordNet, Cyc, OntoWeb, KnowledgeWeb. Z projektov EÚ sú

zaujímavé napríklad aj CoMMA – Corporate Memory Management through agents, On-To-Knowledge, „inteligentný web“ a i.

Na obrázku 8.2 je naznačený príklad génovej ontológie. Je založená na hierarchii troch zložiek, molekulárnej funkcie, biologických procesoch a bunkových zložkách. Pritom každý prvok je možnou vlastnosťou (atribútom) génu, každý prvok môže nadobúdať rôzne hodnoty a pojmy sú prepojené rôznymi sémantickými vzťahmi.

je	<b>bunka</b>
má časť	<b>vnútrobunková</b>
má časť	<b>cytoplazma</b>
má časť	<b>plastid</b>
je	<b>chloroplast</b>
má časť	<b>tylakoid</b>
má časť	<b>membrána tylakoidu</b>

Obr. 8.2. Génová ontológia Gene ontology consortium, 2003.  
<http://www.geneontology.org/>

Medzi prvky ontológie patria:

1. Triedy
2. Koncepty, kategórie
3. Rámce
4. Individuá
5. Príklady
6. Relácie
7. Funkcie, vlastnosti, roly.

Príklad definície v jazyku ontológie môže byť:

X je inštanciou triedy *anglicky-hovoriaci-človek* práve vtedy, ak preňho slot *hovori jazykom* nadobúda hodnotu (jednu z hodnôt) *angličtina*, pričom individuum *angličtina* je inštanciou triedy *jazyk*.

Rozlišujeme tieto druhy ontológií:

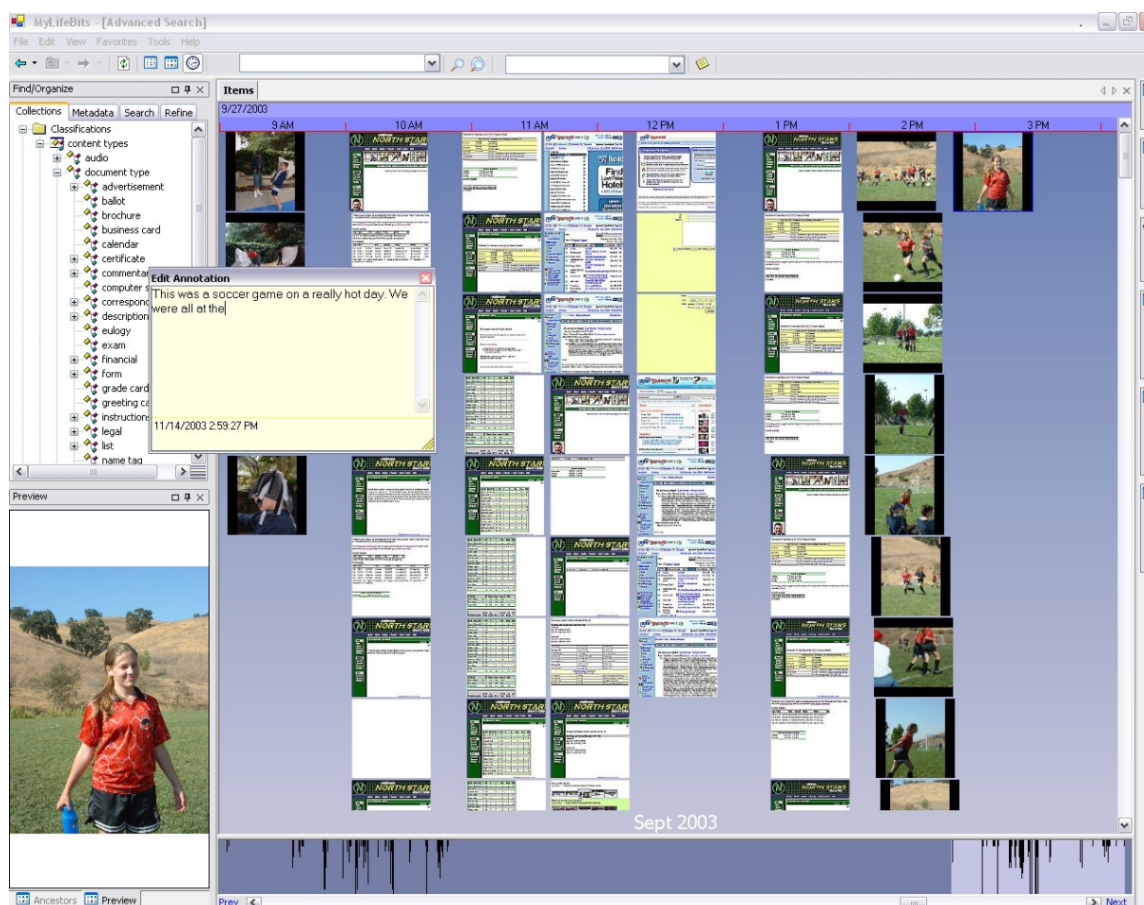
1. Lexikálne a terminologické
2. Informačné (databázové systémy)
3. Znalostné (umelá inteligencia)
4. Doménové (širšie oblasti: medicína, firma, užšie: choroba, úver, napr. Enterprise Ontology, On9, MeSH)
5. Generické (Cyc)
6. Špecifické (task-based, diagnostika a i.)
7. Aplikačné

Príklady aplikačných ontológií možno nájsť vo vytváraní podnikovej infraštruktúry. napr. KAON – ontologická infraštruktúra pre podnikové informačné systémy, alebo aj pri budovaní univerzitných systémov (Karlsruhe Ontology infrastructure (OpenSource)).

## 8.5 Organizácia poznania v praxi

Rôznym aspektom organizácie poznania sa venujú informační špecialisti, informační architekti, knihovníci, jazykovedci, programátori a pod. Procesy a systémy organizácie poznania majú teda interdisciplinárny povahu. Uplatňujú sa nielen v prostredí webu, ale aj v tradičných kultúrnych a pamäťových inštitúciách ako sú knižnice, múzeá, galérie, či archívy.

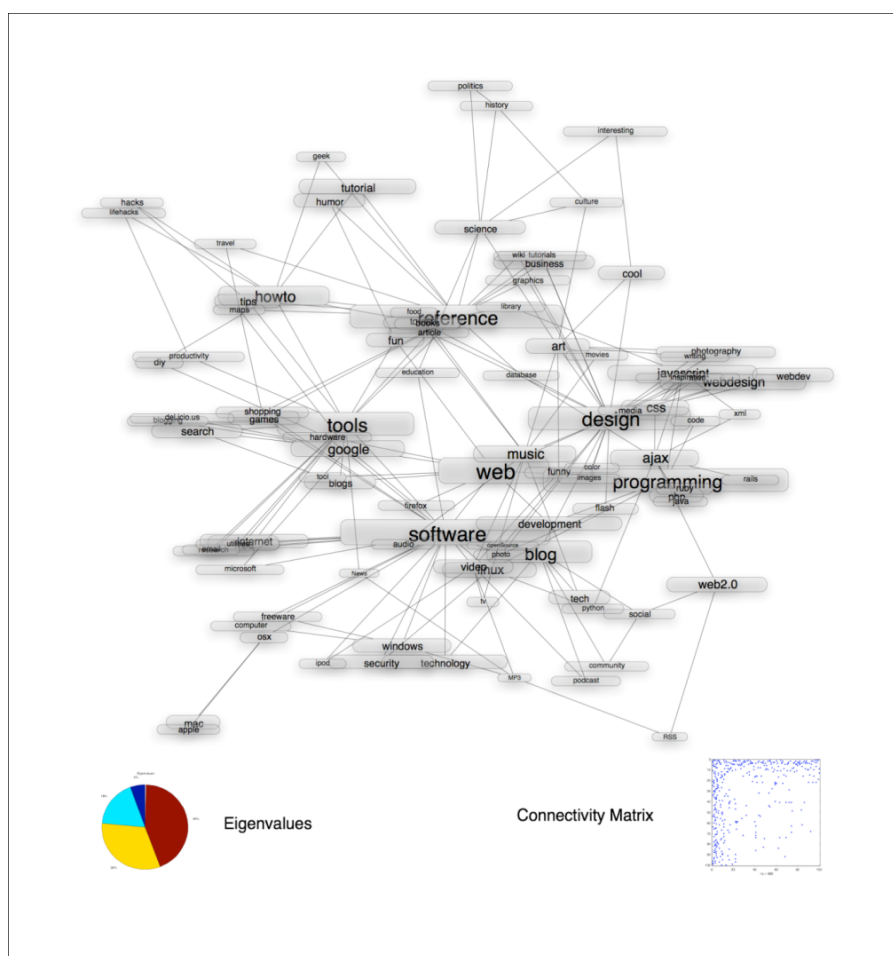
Organizácia poznania sa viaže aj na našu každodennú prácu s informáciami, tzv. osobný informačný manažment. Na tieto účely slúži množstvo užitočných nástrojov, ktoré človeka podporia napr. pri organizovaní poznámok a anotácií, kontaktov, dokumentov, elektronickej pošty, plánov, úloh, webového obsahu a pod. Medzi najzaujímavejšie projekty patria systémy na archiváciu všetkých aktivít, ktoré sme realizovali pomocou svojho počítača počas svojho života. Znáмым príkladom je MyLifeBits (2009), ktorý slúži ako umelá externá pamäť, v ktorej sa dajú vyhľadať informácie v digitálnej forme. Ide napr. o záznamy telefonátov, história rozhovorov (chatov), e-maily, dokumenty, videá, fotky, linky atď. Všetky sú organizované do prehľadných kategórií a možno si ich prezerať aj usporiadané podľa časového hľadiska (obrázok 8.3). V súčasnosti sú aktuálne otázky súvisiace s kolaboratívnou organizáciou poznatkov komunitami v rámci sociálnych sietí.



Obr. 8.3 MyLifeBits – časový prehľad aktivít

## 8.6 Organizujeme sami

Organizácia poznania nadobúda za posledných päť rokov nový rozmer. Spôsob triedenia informácií na internete je čoraz viac v rukách používateľov. Ľudia participujú nielen na tvorbe a publikovaní obsahu, ale aj na jeho organizovaní. Nové aplikácie umožňujú vytváranie vlastných knižníc kategórií, ktoré sprehľadňujú nielen webové linky (záložky), ale aj prezentácie, videá, obrázky, záujmy, stanovené ciele a pod. Aktivity používateľov pri organizovaní obsahu do kategórií vedú ku vzniku nových štruktúr organizácie poznania, akými sú folksonómie (obrázok 8.4). Spoločným znakom folksonómii je to, že sú vytvárané a neskôr zdieľané komunitou. Medzi ďalšie aktivity používateľov, súvisiacich s organizáciou poznania, patrí vytváranie vlastných komentárov k zdrojom, ale aj anotácií a hodnotení.



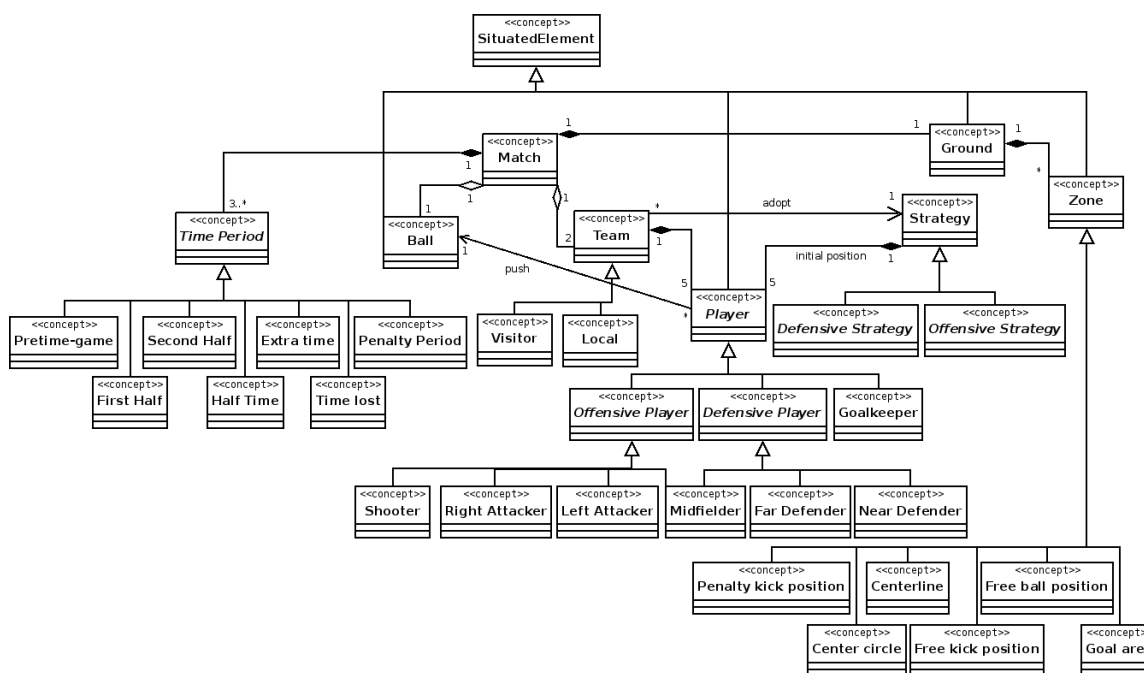
Obr. 8.4 Vizualizácia folksonómie z Del.icio.us (Shaw, 2006)

Všetky tieto prístupy majú dramatický vplyv na vyhľadávanie na internete. Žiaľ, nielen pozitívny. V mnohých prípadoch totiž môžu prispieť ešte k vyššej heterogénnosti a chaosu, ktorý vládne internetu. K riešeniu týchto problémov prispievajú aktivity na podporu sémantického webu, ktoré boli iniciované W3C (Herman, 2009).



## 8.7 Hľadanie významu/sémantický web

Sémantický web má byť „*rámcom, ktorý umožňuje zdieľanie a viacnásobné využitie dát*“ z viacerých zdrojov (Herman, 2009) v rôznych systémoch a kontextoch. Sémantický web má tiež opisovať dáta tak, aby sa čo najviac približovali objektom reálneho sveta. Rozhodujúcu úlohu pri tvorbe sémantického webu majú XML, RDF a ontológie (príklad ontológie pozri obrázok 8.5). Tieto štruktúry sú zároveň aj systémami organizácie poznania. Pomocou nich je možné nielen definovať význam, obsah a štruktúru webových objektov, ale aj opisné metadáta a formalizovanú štruktúru poznatkov.



Obr. 8.5 Ontológia riešenia problémov v robotickom futbale (Cossentino et al., rok neznámy)

Nástroje sémantického webu majú zabezpečiť efektívnejšiu prácu strojov vo webovom prostredí. Avšak ani sémantický web nie je akceptovaný bez kritiky. Možnosť porozumenia významu webu strojmi vyznieva pre niektorých odborníkov prinajmenšom absurdne. Napriek tomu sa do objavovania nových možností investuje množstvo financií a aj roky výskumu. Samozrejme, web 3.0 nie je konečnou stanicou. Dnes sa už totiž polemizuje o webe 4.0 ako o inteligentom alebo znalostnom webe.

Systémy organizácie poznania sa odlišujú svojou štruktúrou a funkciou. Spája ich však schopnosť sprostredkovať informácie v organizujúcej schéme, podporovať manažment poznatkov a vyhľadávanie informácií.

Systémy organizácie poznania vyjadrujú vzťahy medzi poznatkami. Medzi tradičné systémy patria napríklad katalógy, bibliografie a registre. K novším systémom organizácie poznania v elektronickom prostredí radíme napr. hypertext, folksonómie, tematické mapy, ontológie

## 8.8 Sémantické siete

Špecialisti z oblasti umelej inteligencie sa už viac ako polstoročie snažia o simuláciu a modelovanie myslenia a inteligencie človeka. V 60. rokoch podnietili záujem o sémantické siete, ktoré mali odhaliť niektoré aspekty fungovania poznávacích procesov ľudí a zvýšiť efektívnosť práce počítačových systémov.

Sémantické siete sa využívajú a skúmajú v informatike, konkrétne v umelej inteligencii. Medzi oblasti ich využitia patria napr. porozumenie prirodzeného jazyka, strojové videnie, objektovo-orientovaná analýza, dynamická kontrola bojových lietadiel a informačný prieskum. Ich aplikácia sa neobmedzuje len na technické disciplíny, veľa pozornosti sa im venuje aj v sociálnych vedách.

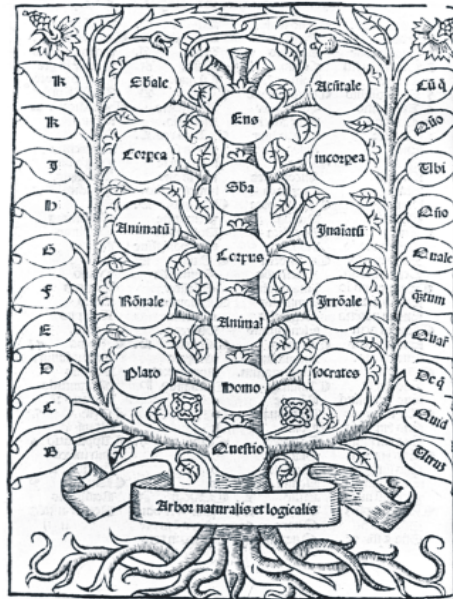
Sémantické siete sa dajú chápať nielen ako komplexné systémy organizácie, ale aj reprezentácie poznania. Sémantická sieť pozostáva z uzlov, ktoré predstavujú pojmy. Hrany siete reprezentujú významové vzťahy medzi pojmami. Pojem môže byť v rámci siete zastúpený slovom alebo skupinou slov (kategóriou) (McCann). Vzťahy medzi pojmami môžu mať rôznu povahu od najjednoduchších po komplexnejšie. Príkladmi sú vzťahy ako napr. definuje, je typom, je vyrobený, pozostáva, reprezentuje a pod.

V rámci sémantických sietí funguje princíp dedičnosti. Konkrétne vlastnosti sa dedia na pojmy (uzly) v hierarchicky nižšej úrovni. Napr. trieda zelenina je nadtriedou triedy mrkva. Mrkva bude mať špecifické vlastnosti a zdedí aj vlastnosti zeleniny. Sowa (2006) sprehľadnil hlavné typy sémantických sietí, ktoré sa triedia na definičné, rozhodovacie, implikačné, vykonávacie, učiace sa a hybridné.

### 8.8.1 Vznik sémantických sietí

Ross M. Quillian, ktorý sa zaoberal možnosťami simulácie ľudskej asociatívnej pamäte a prirodzeného jazyka, predstavil sémantické siete vo svojej dizertačnej práci, keď hovoril o organizácii ľudskej sémantickej pamäti alebo pamäti slovných pojmov (Sharples et al., 1996). Idea sémantických sietí je však omnoho staršia.

Vznik sémantických sietí sa datuje do čias Aristotelovej starovekej filozofie. Porfýrios z Tyru opierajúc sa o Aristotelovu teóriu kategórií navrhol stromovú štruktúru, ktorá sa považuje za prvú (definičnú) sémantickú sieť (Sowa, 2006). Túto sieť využil napr. malorský filozof Raymond Lull v 13. storočí vo svojom logickom diagrame (obrázok 8.6). Sémantické siete sa stali predmetom záujmu viacerých disciplín. Ešte pred ich rozšírením do umelej inteligencie boli využívané nielen vo filozofii, ale aj v psychológii a jazykovede.



Obr. 8.6 Porfýriov strom ako prvá sémantická sieť

## 8.8.2 Gigantická sémantická sieť

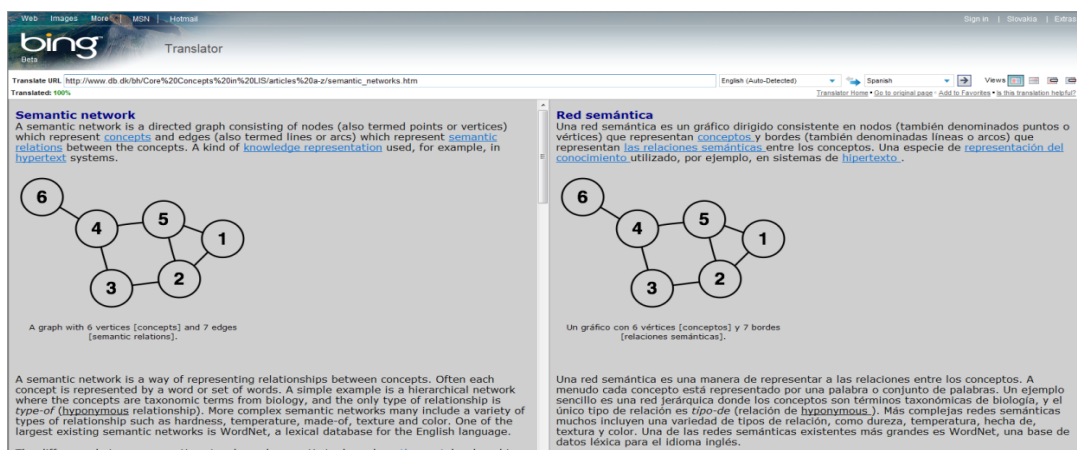
Jednou z najznámejších a najrozsiahlejších sémantických sietí je WordNet (obrázok 8.7).

Za jeho myšlienkou stojí profesor George A. Miller, známy kognitívny psychológ z Princetonskej univerzity. Od roku 1985 až dodnes na ňom pracuje tím ľudí a je podporený grantovými agentúrami, fondmi a firmami ako DARPA, NSF, Google a pod. WordNet (2009) obsahuje viac ako 155 tisíc unikátnych slov a spolu so synonymami je ich celkový počet takmer 207 tisíc. Jedinečnosť tohto systému spočíva v opise vzťahov medzi pojmami. Práve preto je WordNet využívaný napr. pri tvorbe ontológií, v informačných systémoch.

Obr. 8.7 WordNet – záznam slova „technology“.

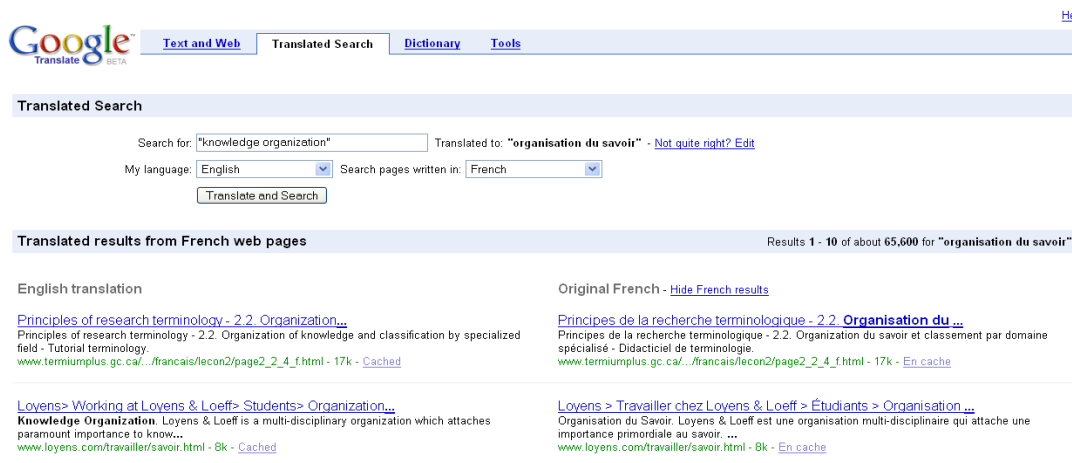
## 8.8.3 Strojový preklad

V novodobej informatike sa sémantické siete začali úspešne využívať koncom 50. rokov 20. stor., kedy boli aplikované v oblasti strojového prekladu. Ide o počítačový preklad z jedného prirodzeného jazyka do druhého. Medzi najznámejšie voľne dostupné aplikácie, ktoré využívajú strojový preklad patrí napr. Bing Translator (2009) od Microsoftu. Umožňuje preklad z 12 svetových jazykov napr. z kórejštiny do angličtiny, z francúzštiny do nemčiny a pod. Okrem výsekov textov s maximálnym rozsahom 500 slov dokáže Bing Translator preložiť aj celú webovú stránku. Výsledný preklad je dostupný vo viacerých zobrazeniach. Veľmi prehľadné je napr. stĺpcové, ktoré umožní porovnať preložený text s originálom (obrázok 8.8).



Obr. 8.8 Bing Translator – strojový tlmočník na webe

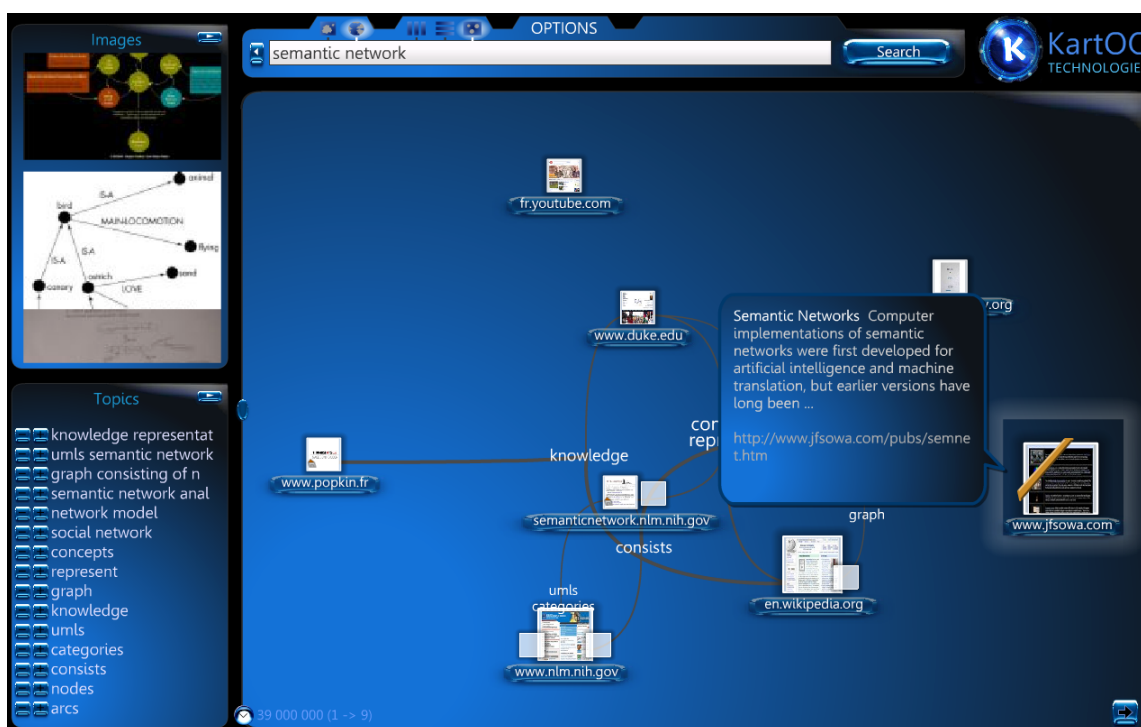
Na strojový preklad nezanevrelí ani ďalšie firmy, ktoré v minulosti sprostredkúvali primárne vyhľadávacie služby. Patria medzi ne Google aj Yahoo!. Yahoo! Babel Fish (2008) poskytuje v porovnaní s Bing Translator veľmi podobnú funkcionálnosť. Na rozdiel od predchádzajúcich systémov, Google Language Tools (2009) umožňuje okrem prekladu časti textu alebo stránky aj preklad vyhľadaných výsledkov (obrázok 8.9).



Obr. 8.9 Výsledky z Google Translate po zadaní anglického dotazu „knowledge organization“ (organizácia poznania) preložené do francúzštiny.

## 8.8.4 Sémantické vyhľadávanie informácií

Sémantické siete sú využívané aj vo vyhľadávacích nástrojoch napr. prieskumových strojoch, predmetových adresároch, digitálnych knižniciach, inteligentných informačných agentoch a pod. V takýchto systémoch môžu sémantické siete slúžiť na reprezentáciu znalostí, dotazov a zdrojov. Niektoré vyhľadávače prezentujú výsledky vo forme sémantických sietí. Znáмым príkladom je pokročilý metaprieskumový stroj Kartoo (2009) s podporou vizualizácie. Kartoo (obrázok 8.10) zobrazuje výsledky vo forme máp. V jednotlivých mapách sú uzly zastupujúce výsledky a hrany, ktoré prepájajú súvisiace zdroje.



Obr. 8.10 Výsledky v Kartoo ako sémantická sieť.

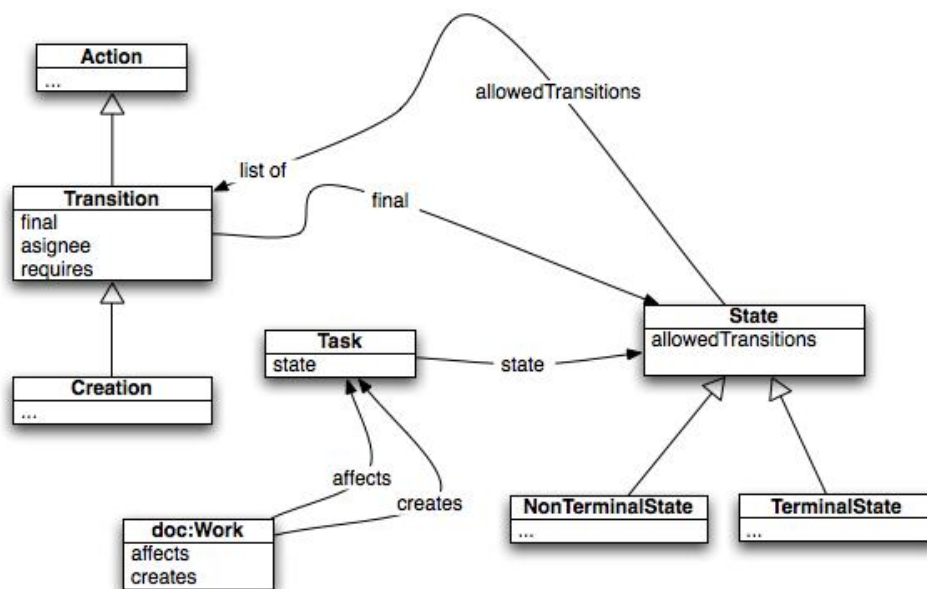
V súvislosti so sémantickým webom sa často spomínajú pojmy ako ontológie, XML a RDF. Práve ontológie sú v určitom zmysle blízke sémantickým sieťam.

## 8.9 Ontológie

Ontológie, ako kľúčový pilier sémantického webu, umožňujú organizáciu heterogénneho prostredia databáz, systémov a služieb založených na znalostiach.

Ontológia je v inom význame filozofickou kategóriou. Predstavuje vedu o bytí ako takom a v stredoveku sa dávala do opozície metafyzike. Metafyzikou, vedou o hmote, sa stredovekým učencom darilo vysvetliť takmer všetky problémy. Dnes sa oba tieto pojmy, ontológia a metafyzika, nevnímajú tak jednoznačne. Pojem ontológia našiel nový význam v informatike. Ontológia sa v tejto oblasti považuje za účelový produkt, ktorý slúži na opis špecifických oblastí poznania. Je to komplikovaná sieť definujúca triedy, vlastnosti, vzťahy a ďalšie objekty.

Existujú rôzne typy ontológií, ktoré sa líšia cieľom, funkciou, či oblasťou. Na základe toho sa ontológie typologicky triedia na generické, znalostné, úlohové (obrázok 8.11), aplikačné, doménové a pod. Medzi známe generické ontológie patrí Cyc, ktorá vznikla ako ambiciózný projekt znalostnej bázy obsahujúcej všetky poznatky z rôznych domén, ktoré sú prepojené na fakty a dáta.



Obr. 8.11 Model odvodený z ontológie priebehu práce – „Workflow ontology“, ktorú vytvoril Tim Bernes-Lee pod hlavičkou W3C (Story, 2007)

Ďalším typom sú doménové ontológie. Medzi širšiu doménovú radíme ontológiu medicíny a užšiu ontológiu konkrétnej choroby. Typickým príkladom širšej doménovej ontológie z oblasti medicíny je MeSH. Využíva ju aj prieskumový stroj GoPubMed (2009) na obrázku 8.12. Tento špecializovaný vyhľadávač po zadaní priezviska „John Wilson“ zobrazuje okrem záznamov aj hierarchické kategórie (v ľavej časti obrázku – napr. autor s týmto menom publikoval o chorobách: alkoholizmus, Alzheimerova choroba, cirhóza pečene atď.). Medzi kategóriami sú konkrétne oblasti, ktorým sa autor venuje spolu s počtom záznamov. V jednotlivých hierarchických kategóriách sa dá dobre orientovať a rýchlejšie vyhľadávať.

Obr. 8.12 Vyhľadanie autora „John Wilson“ pomocou GoPubMed, ktorý využíva ontológiu MeSH.

Ontológie nachádzajú svoje využitie nielen v oblasti organizácie a vyhľadávania informácií, ale aj v B2B elektronickom obchode, manažmente poznatkov, inteligentných databázach a inteligentných softvérových agentoch a pod.

### Úloha ontológií

Ontológie integrujú odlišné databázy, systémy a služby. Zohrávajú nezastupiteľnú úlohu v oblasti sémantického webu. Význam ontológií sémantického webu spočíva v integrácii a prelínaní dát z rôznych zdrojov. Pre lepšiu zrozumiteľnosť uvedieme konkrétny príklad. V prostredí webu sú dostupné aplikácie, ktoré využívame nezávisle, napr. zdieľanie fotografií (Flickr, Picasa), zoznamy úloh (RememberTheMilk, Backpack), kalendár (Google Calendar), zdieľanie a hodnotenie záložiek (Del.icio.us, Digg), a pod. Sémantický web umožní prepájanie týchto služieb. Napr. v kalendári budú zjednotené fotky a záložky, ktoré boli pridané v konkrétny deň, ale aj rozpis úloh.

### Jazyky na tvorbu ontológií na webe

W3C ako iniciatíva na vývoj štandardov webu, poskytuje rozsiahlu knižnicu definícií a odporúčaní. V oblasti ontológií sprostredkúva odporúčanie, ktoré podrobne opisuje štruktúru a postupy využívania jazyka OWL (Bechhofer et al., 2004). OWL, čiže Web Ontology Language, je jazyk definujúci ontológie na webe. Samozrejme, nie je jediným pokročilým jazykom na tvorbu webových ontológií. Okrem neho sú známe napr. DAML, RDFS, DAML+OIL (predchodca OWL), XOL atď. Všetky spomínané jazyky sú založené na dátovom modeli RDF (Herman, Swick, 2009). RDF, ktoré je podobne ako OWL W3C odporúčaním, slúži na opis metadát objektov vo webovom prostredí.

### Príklady vyhľadávania ontológií

Príklady webových ontológií definovaných v rôznych jazykoch sa dajú vyhľadať veľmi jednoducho. Prístup k ontológiám poskytujú adresáre, ale aj prieskumové stroje, akými sú Swoogle a OntoSearch2. Swoogle (2007) je výsledkom práce výskumného tímu na Marylandskej univerzite. Názov Swoogle vznikol spojením slov Semantic Web Ontology a Google. Tento nástroj vyhľadáva dokumenty sémantického webu v RDF, medzi ktoré patria aj ontológie. Umožňuje využitie pokročilej syntaxe, pomocou ktorej sa dajú vyhľadávať kľúčové slová v konkrétnych poliach (definícia, URL,



kódovanie, dátum poslednej aktualizácie, a pod.). V prípade, že chceme nájsť ontológiu, v ktorej sa nachádza definícia RDF, zadáme dotaz def:RDF. Výsledky vyhľadávania môžeme vidieť na nasledujúcom obrázku (obrázok 8.13).

Obr. 8.13 Zoznam výsledkov po zadaní „def:RDF“ do vyhľadávača Swoogle

ONTOSEARCH2 (2009) je živý projekt Aberdeenskej univerzity. Práve z dôvodu neustáleho vývoja tento prieskumový stroj priebežne mení svoju podobu. Podobne ako v predchádzajúcom príklade, aj ONTOSEARCH2 umožňuje do databázy pridať ontológiu manuálne. V porovnaní so Swoogle navyše pri vyhľadávaní umožňuje aj využitie špecializovaného dotazovacieho jazyka SPARQL. V prípade, že napr. chceme vyhľadať podtriedy kľúčového slova „person“, dotaz bude mať nasledujúcu formu (tento príklad je uvedený priamo na stránke ONTOSEARCH2).

```
#TQ#
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX os: <http://www.ontosearch.org/NS/>

SELECT ?X ?Y WHERE {
  ?Y os:hasKeyword "person" . #TH# 0.6
  ?X rdfs:subClassOf ?Y .
}
```

Výsledok po zadaní dotazu formulovaného pomocou SPARQL bude vyzeráť nasledovne (obrázok 8.14):



?x	?y
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#PostDoc">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#PostDoc</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Dean">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Dean</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Chair">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Chair</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#VisitingProfessor">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#VisitingProfessor</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#AssistantProfessor">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#AssistantProfessor</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#FullProfessor">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#FullProfessor</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#AssociateProfessor">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#AssociateProfessor</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Professor">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Professor</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Lecturer">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Lecturer</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Faculty">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Faculty</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#ClericalStaff">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#ClericalStaff</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#SystemsStaff">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#SystemsStaff</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#AdministrativeStaff">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#AdministrativeStaff</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Director">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Director</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Employee">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Employee</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#TeachingAssistant">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#TeachingAssistant</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#ResearchAssistant">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#ResearchAssistant</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#UndergraduateStudent">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#UndergraduateStudent</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#GraduateStudent">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#GraduateStudent</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>
<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Student">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Student</a>	<a href="http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person">http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl#Person</a>

Obr. 8.14 Zoznam podtried kľúčového slova „person“ vyhľadaný ONTOSEARCH2 s využitím protokolu SPARQL

Zoznam nástrojov na vyhľadávanie ontológií sa neobmedzuje len na spomínané prieskumové stroje. Sú užitočné najmä pre špecialistov, ktorí vytvárajú ontológie alebo systémy s ich podporou.

Okrem ontológií sú ďalšou známou technológiou sémantického webu aj tematické mapy. Tematické mapy určujú základné témy, ktoré reprezentujú konkrétny pojem. Sú užitočným nástrojom na organizáciu a vyhľadávanie informácií vo webových stránkach, portáloch, katalógoch, indexoch sídiel, systémoch riadenia obsahu a majú aj ďalšie dôležité úlohy.

## 8.10 Tematické mapy

Ideálne predstavy o informačnej spoločnosti, v ktorej budú sprostredkované vhodné informácie pravej osobe v pravom čase, sa zatiaľ nepodarilo úplne naplniť. Namiesto toho žijeme v spoločnosti, pre ktorú je typické informačné preťaženie. Každodenne sme nútení triediť, analyzovať, ukladať a neskôr využiť enormné množstvo informácií. Ďalším nástrojom, ktorý umožňuje lepší manažment a organizáciu informácií v kontexte, sú tematické mapy.

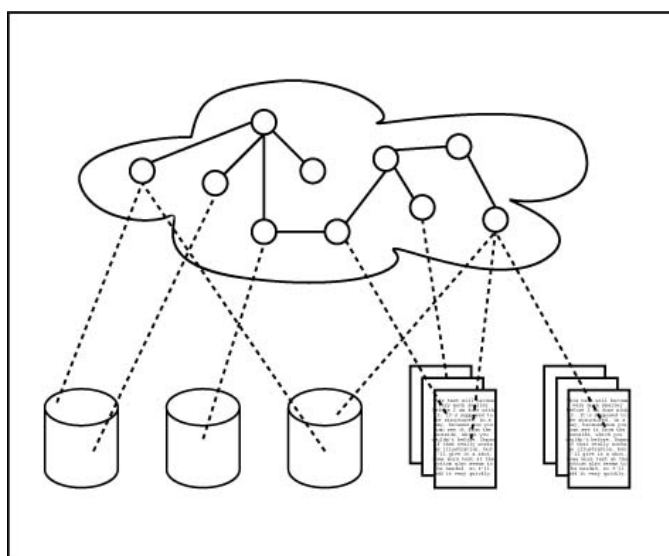
### Vývoj tematických máp

Tematické mapy boli prvýkrát vytvorené pri snahe o zlúčenie elektronických indexov. O kontinuálny vývoj tematických sa zaslúžil najmä Steven R. Newcomb. Začiatkom 90. rokov minulého storočia inicioval vznik aktivity, ktorá niesla názov SOFABED (Standard Open Formal Architecture for Browsable Electronic Documents, štandardná otvorená formálna architektúra na prezeranie elektronických dokumentov). Iniciatíva bola zameraná na možnosti využitia štandardu HyTime pri reprezentácii interoperabilných indexov, slovníkov, tezaurov a obsahov dokumentov. HyTime definuje štruktúru hypermediálnych dokumentov a je založený na štandarde SGML. Základný model tematic-

kých máp bol navrhnutý S. R. Newcombom za spolupráce s M. Biezunskim v roku 1995. V nasledujúcich rokoch bol vypracovaný návrh štandardu, ktorý bol v roku 2000 spracovaný a zverejnený ISO. Ďalší vývoj tematických máp ovplyvnili zmeny a nové štandardy v oblasti webu. Spolu so vznikom XML boli iniciované aktivity na jeho špecifikáciu pre tematické mapy. Výsledkom týchto snáh bolo vytvorenie štandardu XTM (XML Topic Maps). Najaktuálnejšou je špecifikácia XTM vo verzii 1.0 (Pepper, Moore, 2001).

#### *Možnosti využitia tematických máp*

Tematické mapy pozostávajú z viacerých prvkov. Témy v mape slúžia na opis konkrétnej reality pomocou vyjadrenia názvov, výskytu a asociácií. Tematické mapy umožňujú vytváranie indexu, ktorý je uložený mimo opisovaných informácií (obrázok 8.15). V prostredí internetu môžu byť tematické mapy vytvárané z dokumentov alebo databáz a prepojené s pôvodnými objektmi pomocou URI. Tematické mapy poskytujú flexibilný prístup k opisu informácií v kontexte, v ktorom nadobúdajú význam.

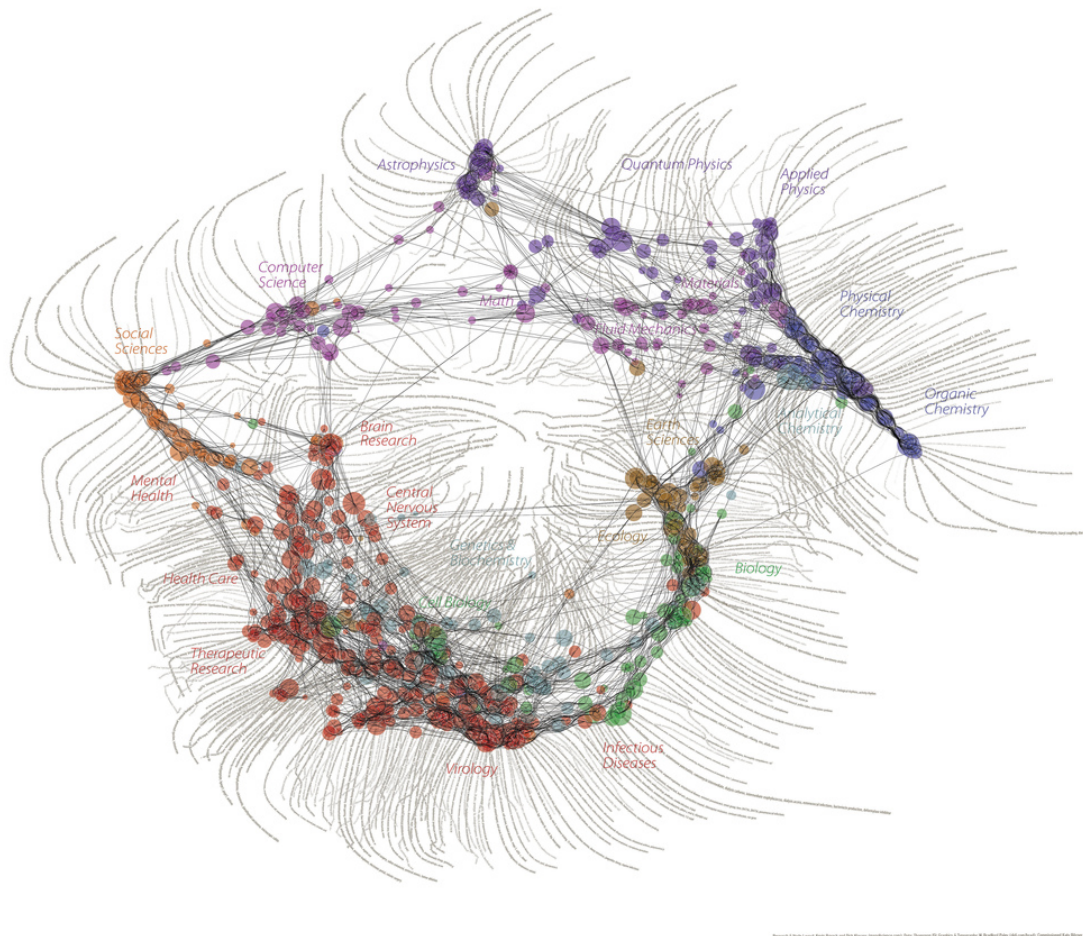


Obr. 8.15 Zjednodušený model tvorby tematickej mapy z dokumentov a databáz (Garshol, 2002)

Tematické mapy sú využívané v rôznych oblastiach. V širšom pohľade sú aplikované vo vyhľadávaní, filtrovaní a segmentácii, organizácii a klasifikácii, navigácii, vizualizácii a zlučovaní. S tematickými mapami sa môžeme stretnúť v priemysle, verejnom sektore, či vede. Často sa využívajú v manažmente poznatkov, informačnej architektúre, sémantickom indexovaní, e-learningu a pri integrácii informácií. Aplikčné možnosti tematických máp v ďalších rôznych sférach zjednodušuje ich vysoká miera štandardizácie. ISO štandardy detailne opisujú tematické mapy z odlišných hľadísk. Napr. ISO/IEC 13250 a jej jednotlivé časti špecifikujú dátový model, XML syntax, referenčný model, dotazovací jazyk atď.

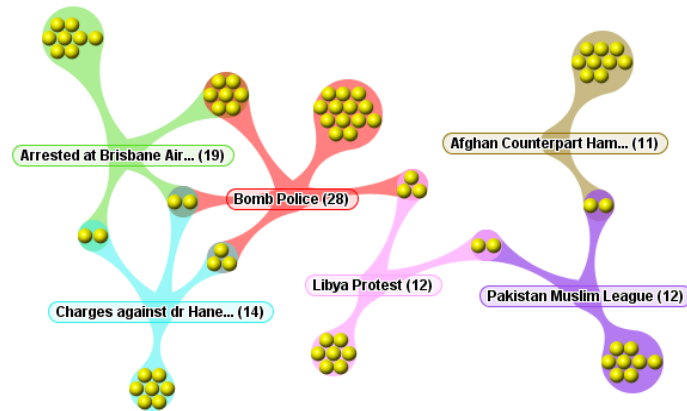
#### *Vzťahy medzi článkami*

Tematické mapy môžu slúžiť na sprehľadnenie tém z rôznych dokumentov. Nasledujúci príklad je zameraný na vedecké články z roku 2003 (Klavans, Boyack, 2006). Mapa na obrázku 8.16 vizualizuje výsledky kocitačnej analýzy 800 tisíc článkov z rôznych odborov. Zjednodušene povedané, mapa zobrazuje disciplíny a vzťahy medzi nimi sú vyjadrené na základe toho, ako jednotliví autori na seba odkazovali.



Obr. 8.16 Príklad tematickej mapy (Payley, 2006) – vzťahy medzi vedeckými paradigmami

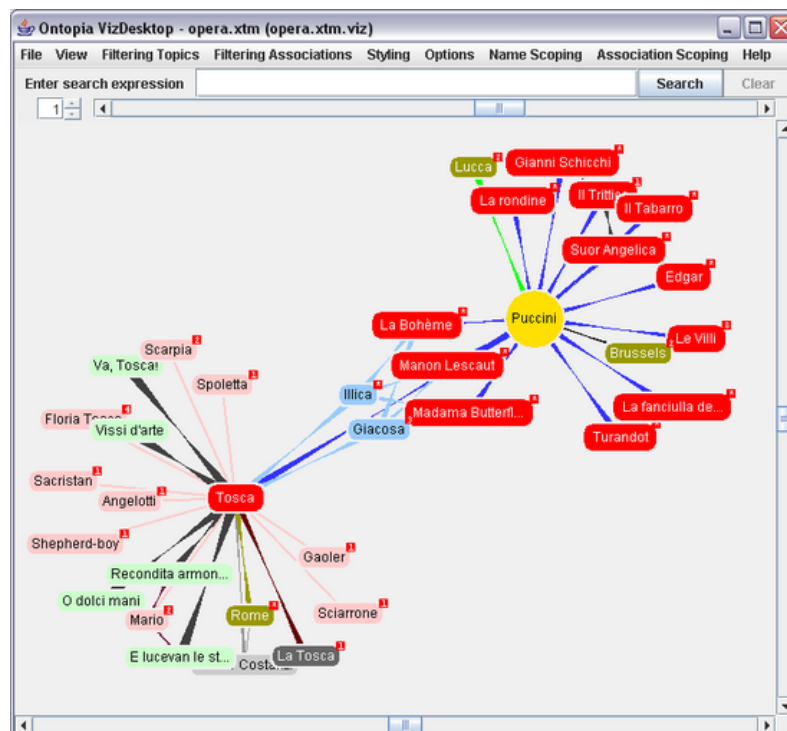
Na ďalšej tematickej mape (obrázok 8.17) môžeme vidieť vzťahy medzi jednotlivými témami, ktoré boli súčasťou viacerých článkov BBC zameraných na terorizmus (BBC, 2007). Konkrétne články sú v mape zobrazené žltými guľôčkami. Nielen vizualizáciu, ale aj pokročilé vyhľadávanie umožňuje špecializovaný nástroj BBC News Browser (Prehliadač BBC News). Tento nástroj je čiastkovým výsledkom pilotného projektu, ktorý sa realizoval v NaCTem (National Centre for Text Mining, Národné centrum pre dolovanie textu) na Manchesterskej univerzite. Prístupy na vizualizáciu tematických máp bola uplatnená vo viacerých projektoch NaCTem, napr. vo vyhľadávacom nástroji INTUTE (2009).



Obr. 8.17 Príklad tematickej mapy 2 – zobrazenie súvislostí medzi témami v článkoch na BBC (2007)

### 8.10.1 Príklady aplikácie tematických máp

Príkladom nástroja na zobrazovanie a navigáciu v tematických mapách je Omnigator (2009). Umožňuje prezeranie tematických máp a ich súčasti. Ide napr. o nasledujúce charakteristiky tematických máp – typy tém, asociácií, výskytov, typy hierarchií, metadát a pod. Omnigator aplikuje sadu nástrojov (nielen) na tvorbu tematických máp – Ontopia Knowledge Suite. Omnigator okrem prezerania podporuje aj ďalšie funkcie. Patria medzi ne napr. možnosti zlučovania tematických máp, vyhľadávania, sémantickej validácie, exportu a pod. Súčasťou tohto nástroja je aj funkcia vizualizácie, pomocou ktorej sa zobrazujú témy mapy v súvislostiach (obrázok 8.18). Na obrázku je grafická vizualizácia tematickej mapy talianskej opery.



Obr. 8.18 Vizualizácia tematickej mapy talianskej opery pomocou funkcie Vizigator (2009)

Slovník skratiek z oblasti webu a XML (Web, 2009) (obrázok 8.19) využíva na organizáciu a vyhľadávanie informácií tematické mapy. Na úvodnej stránke nájdeme zoznam skratiek. Po výbere konkrétnej z nich sa nájdú ďalšie špecifikácie, ktoré vyplývajú z aplikácie tematických máp. Vo výsledku je navyše okrem definície aj typ asociácie spolu s konkrétnymi súvisiacimi témami. Napr. ak si zo zoznamu vyberieme „topic maps“, zobrazia sa aj asociácie na ďalšie pojmy a konkretizácie typov vzťahov. Zo záznamu sa teda dozvieme nielen, že tematické mapy sú založené na HyTime (štruktúrovaný jazyk) a TMDM ( dátový model na tvorbu tematických máp), ale aj to, že sú štandardizované organizáciami IEC a ISO a pod.

The screenshot shows the 'Web and XML Glossary' website. At the top, there is a navigation bar with the title 'Web and XML Glossary' and an alphabetical index '1 2 3 4 5 6 7 A B C D E F G H I J K L M N O P Q R S T U V W X Y Z'. The main content area is titled 'Topic Maps' and includes a 'Description' section, 'Type Associations', 'Associations', and 'Bibliographic References'. The 'Description' section explains that Topic Maps enable vast information resources to be classified and navigated in a consistent manner. The 'Type Associations' section lists 'Meta Data' as a source. The 'Associations' section lists 'HyTime', 'TMDM', and 'HyTM' as related concepts. The 'Bibliographic References' section lists several international standards and research papers.

Obr. 8.19 Záznam zo slovníka Web and XML Glossary (Web, 2009), ktorý využíva tematické mapy

## 8.10.2 Výhody využitia tematických máp

Tematické mapy majú široké možnosti praktického využitia v rôznych sférach. Ich šíreniu pomáhajú aj mnohé výhody, ktoré prinášajú. Tematické mapy zjednodušene reprezentujú informácie a štruktúrujú ich v prirodzenej forme. Takýto prístup modelovania informácií je pre používateľov jednoduchšie pochopiteľný a umožňuje efektívnejšie vyhľadávanie a navigáciu. Ďalšou nepopierateľnou výhodou tematických máp patrí ich schopnosť efektívne integrovať informačné systémy. Tematické mapy sú flexibilné a ich tvorbu sprostredkúva relatívne jednoduchý formát XTM.

V predchádzajúcich častiach sme sa snažili o sprehľadnenie niektorých aktuálnych prístupov organizácie poznania. Okrem tematických máp, ktorým sme sa venovali v tejto časti, sme už predtým charakterizovali aj sémantické siete a ontológie. Ďalšia časť sa bude zaoberať konceptuálnymi a myšlienkovými mapami.

## 8.11 Konceptuálne a myšlienkové mapy

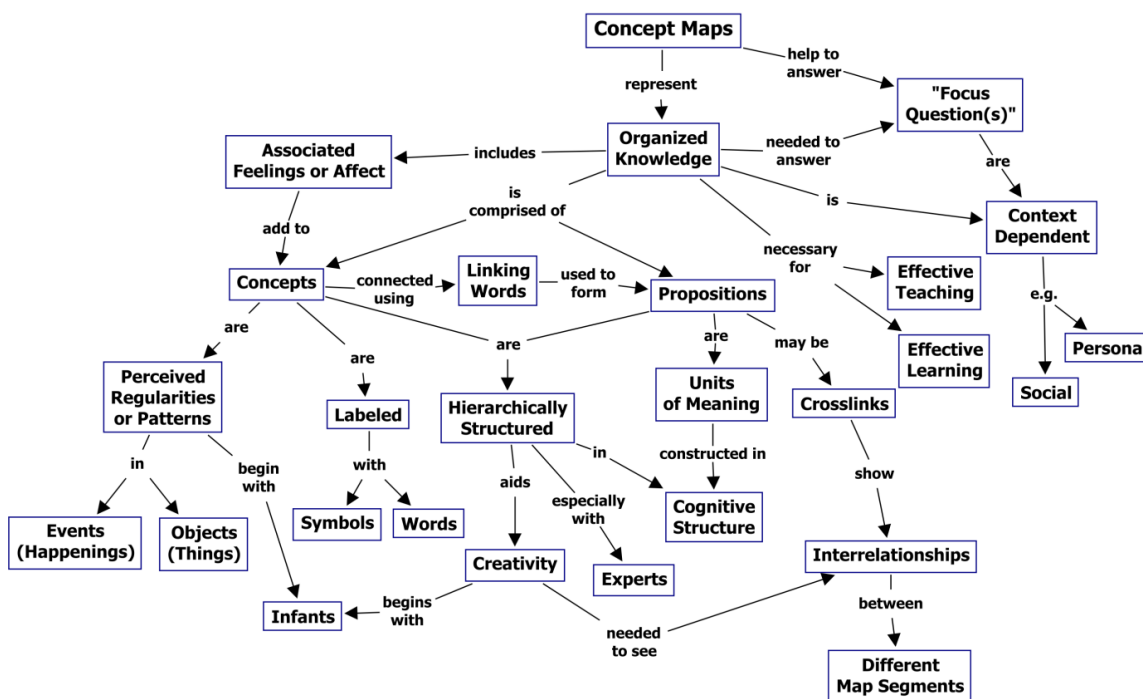
### Výhody mapovania

Pozitívna vizualizácia informácií pomocou máp sú nesporné. Nástroje na mapovanie myšlienok a pojmov pomáhajú zefektívniť prácu a podporiť tvorivosť. Mapy umožňujú lepšie pochopiť súvislosti, ale aj sprehľadniť vzťahy medzi pojmami. Informácie zobrazené v mape sú jednoduchšie zapamätateľné. Pri tvorbe máp môžeme objavovať nové vlastnosti pojmov.

Vytváranie máp podporuje nielen kreativitu, ale aj predstavivosť. Vyhľadávanie máp, zobrazujúcich konkrétnu oblasť, môže byť efektívne z časového hľadiska, pretože pomáha rýchlejšie pochopiť súvislosti medzi jednotlivými pojmami.

### Konceptuálne versus myšlienkové mapy

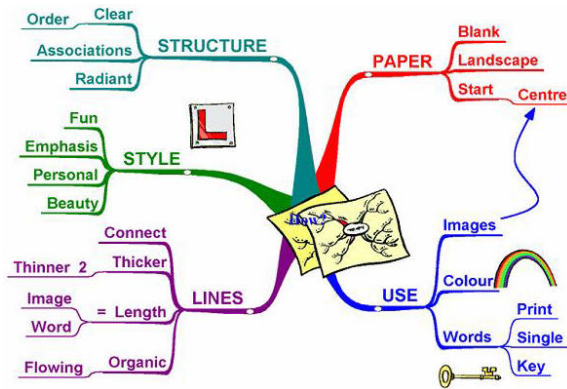
Tieto dva pojmy sa často zamieňajú, preto ich vysvetlíme a odlíšime. Pomocou *konceptuálnej (tiež pojmovej) mapy* sa dá zobraziť konkrétna oblasť alebo problém ako súbor pojmov, ktoré sú poprepájané vzťahmi. Na nasledujúcom obrázku (8.20) je príklad pojmovej mapy, v ktorej sú pomenované vzťahy medzi pojmami.



Obr. 8.20 Konceptuálna mapa zobrazuje vlastnosti pojmových máp (Novak, Cañas, 2008)

V porovnaní s konceptuálnou mapou, je *mapa mysle* (obrázok 8.21) nástrojom na vizualizáciu myšlienok a súvislostí medzi nimi, či postupov pri konkrétnych aktivitách a činnostiach. Mapy mysle sa často využívajú vo firemnom prostredí a sú produktom alebo pomôckou na brainstorming. Tieto mapy nachádzajú svoje využitie aj v oblasti manažmentu poznatkov, systémového inžinierstva, vzdelávaní, vede a pod. Často majú formu stromovej štruktúry vytváranej od jadra (centrálneho pojmu).





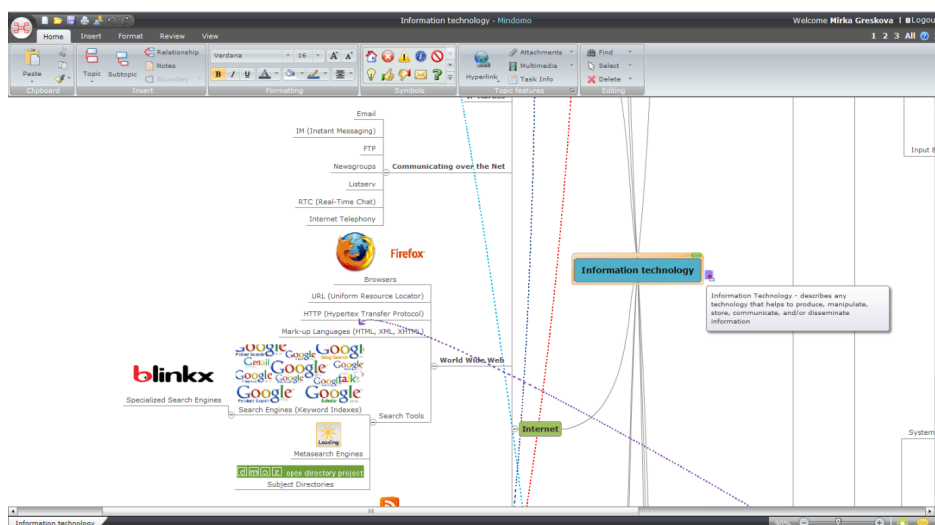
Obr. 8.21 Mapa mysle – proces vytvárania mapy mysle (van Halen, 2009)

V súvislosti s konceptuálnymi a myšlienkovými mapami sa dá stretnúť aj s pojmom kognitívne (mentálne) mapy. Kognitívna mapa reprezentuje priestorové súvislosti ako sú uložené v pamäti. Vo svojej myslí môžeme mať napr. vytvorenú kognitívnu mapu cesty z centra mesta na železničnú stanicu. Pomocou tejto mapy sa nemusíme cestu učiť vždy odznova a vieme, ako sa úspešne dostať do cieľa. Skúmaniu kognitívnych máp sa venuje pozornosť najmä v kognitívnej psychológii a neurovede.

### 8.11.1 Príklady online nástrojov na tvorbu máp

#### *Mindomo*

Mindomo patrí medzi najlepšie nástroje na tvorbu myšlienkových máp. Okrem základnej funkcionality tvorby hierarchických máp pozostávajúcich z tém a podtém ponúka aj možnosti pridávania obrázkov, ikon, poznámok, odkazov a pod. Takisto podporuje prepájanie jednotlivých objektov pomocou vzťahov (obrázok 8.22). Výsledné mapy sa dajú exportovať do rôznych formátov, dokonca aj spracovať do formy textu, ktorá sa môže stať podkladom na správu. Mindomo poskytuje aj viaceré možnosti zdieľania. Celkovo sa dá tento nástroj hodnotiť ako veľmi vhodný nielen na tvorbu štruktúr mapujúcich nejakú doménu, ale aj na podporu manažmentu projektov.



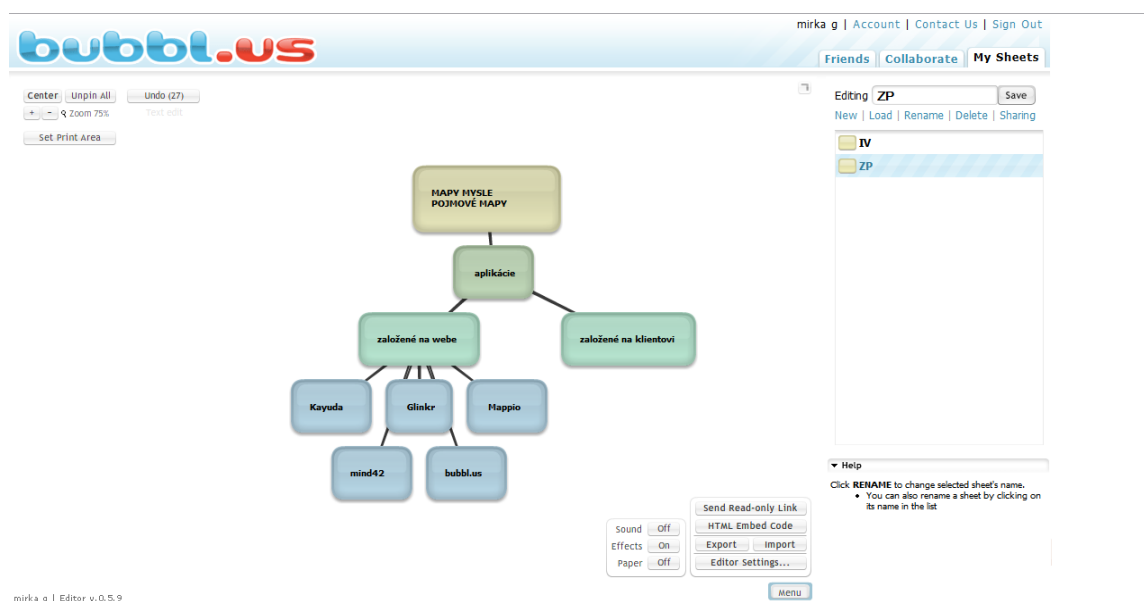
Obr. 8.22 Mindomo – mapa sprehľadňujúca oblasť informačných technológií. Farebné čiary predstavujú vzťahy prepájajúce jednotlivé pojmy alebo oblasti

## Bubbl.us

Bubbl.us je jednoducho použiteľný nástroj na tvorbu máp. Umožňuje efektívne zorganizovať informácie a procesy do hierarchickej štruktúry. Po otvorení pracovného listu možno začať editovať jednotlivé pojmy. Od rodičovského pojmu sa môžu postupne dopĺňať ďalšie pojmy. Jednotlivé hierarchické úrovne mapy sú farebne odlišené (obrázok 8.23). Jednotlivé mapy sú uložené v prehľadných adresároch.

Hotovú mapu môžeme zdieľať s ostatnými používateľmi Bubbl.us alebo pozvať svojich priateľov, aby ju editovali. Na mapu tiež možno odkazovať alebo ju priamo ukotviť na stránku alebo blog.

Bubbl.us je nielen jednoduchý, ale aj rýchly a interaktívny. Na druhej strane práve jeho jednoduchosť môže byť nevýhodou. Tento nástroj nie je vhodný na modelovanie komplexných štruktúr, skôr ho odporúčame na vytváranie jednoduchších máp bez určenia zložitých vzťahov medzi pojmami.



Obr. 8.23 Mapa vytvorená pomocou Bubbl.us– aplikácie na vytváranie pojmových máp a máp mysle

## Mappio

Mappio je služba, pomocou ktorej môžeme ukladať, zdieľať a vyhľadávať mapy. Mapy, ktoré chceme uložiť na Mappio môžu byť vo formátoch známych nástrojov na tvorbu máp FreeMind a MindManager, ale aj v čistom texte alebo JPG. Mapy určené pre vlastnú potrebu je možné uložiť ako súkromné.

V Mappio, kde je uložených viac ako tisíc máp, sa dá vyhľadávať pomocou kľúčových slov (obrázok 8.24). Prezeranie máp je možné na základe popularity a aktuálnosti. V položke „featured“ nájdeme najzaujímavejšie mapy vybrané administrátormi. Spolu s nárastom počtu máp sa bude musieť zdokonaľiť aj vyhľadávanie, ktoré má zatiaľ veľmi obmedzené možnosti. Používatelia okrem názvu opisujú mapy kategóriami. Mrak tagov a ani pokročilé vyhľadávanie však nie sú k dispozícii. Vývojári pridávajú nové funkcie aj na základe podnetov z používateľskej komunity.



Popular My MindMaps Sign in

13 results found matching: goals

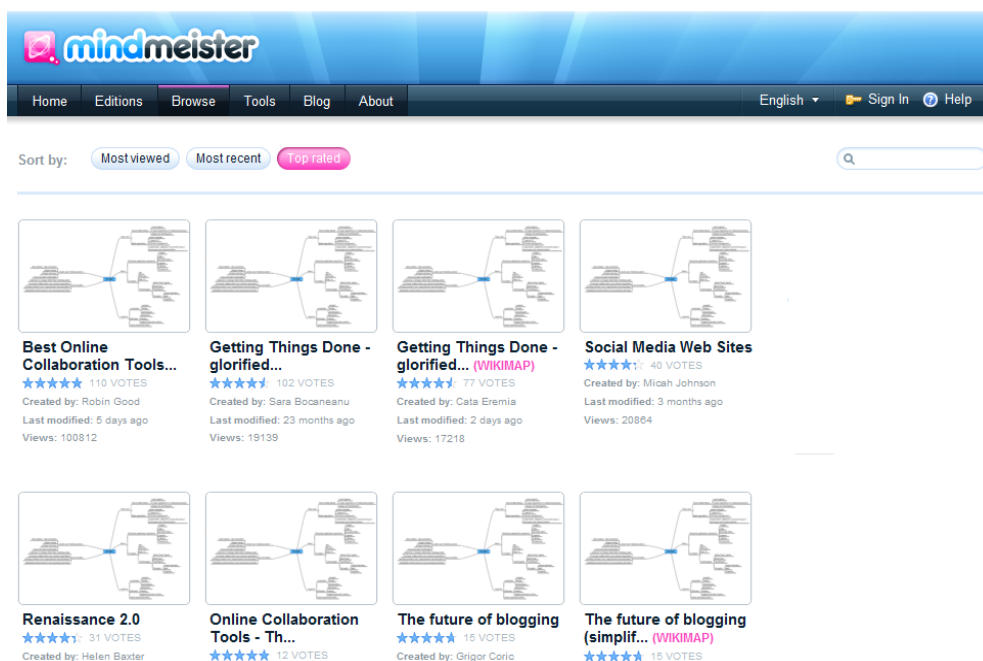
 <p><a href="#">Goals Mind Map Template</a> 1360 views submitted by <a href="#">erichlue</a> 12/19/2007</p>	 <p><a href="#">My Life - Lifetime Partner - TPAssist Sample</a> 420 views submitted by <a href="#">TPAssist</a> 2/25/2008</p>	 <p><a href="#">My Life - Fun &amp; Friends - TPAssist Sample</a> 180 views submitted by <a href="#">TPAssist</a> 2/24/2008</p>	 <p><a href="#">My Life - TPAssist Sample</a> 210 views submitted by <a href="#">TPAssist</a> 2/25/2008</p>
 <p><a href="#">My Life - Family &amp; Home - TPAssist Sample</a> 230 views submitted by <a href="#">TPAssist</a> 2/24/2008</p>	 <p><a href="#">Goals+*</a> submitted by <a href="#">dlakmg</a> 11/25/2007</p>	 <p><a href="#">Life Compass</a> <a href="#">How To Achieve Your Personal Goals - Life Compass 1/24/2008 11:31:11 AM</a> 210 views submitted by <a href="#">yeksoon</a> 1/24/2008</p>	 <p><a href="#">Money Management</a> submitted by <a href="#">himaytinence</a> 12/15/2007</p>
 <p><a href="#">My Life - Finance - TPAssist Sample</a> submitted by <a href="#">TPAssist</a> 2/24/2008</p>	 <p><a href="#">My Life - Health &amp; Fitness - TPAssist Sample</a></p>	 <p><a href="#">2009 goals</a> submitted by <a href="#">Yasuvuki Kishi</a></p>	 <p><a href="#">Money Management</a> submitted by <a href="#">quest</a> 12/15/2007</p>

Obr. 8.24 Mappio – odporúčané mapy a jednoduché vyhľadavanie

### MindMeister

Pri tvorbe a zdieľaní máp pomocou MindMeister sa dajú využiť nastavbové funkcie. Príkladom je možnosť viacerých používateľov spolupracovať na jednom projekte v reálnom čase. Efektívnejšiu kolaboráciu pri tvorbe máp podporuje integrácia Skype. Jednou z najväčších výhod MindMeister je možnosť pracovať na projektoch aj v offline móde.

Zverejnené mapy sú dostupné aj bez registrácie (obrázok 8.25). Jednotlivé mapy možno vyhľadávať podľa kategórií a kľúčových slov. Prezeranie máp zjednodušuje zoradovanie podľa konkrétnych kritérií, kde sú mapy triedené podľa aktuálnosti, návštevnosti a popularity.



Obr. 8.25 Vyhľadavanie verejne dostupných máp pomocou služby MindMeister

Zoznam existujúcich aplikácií na vytváranie máp je široký. Medzi najznámejšie offline aplikácie na vytváranie máp patria FreeMind, Cmap, MindManager atď. Ďalšie služby a aplikácie na vytváranie, zdieľanie a vyhľadávanie máp je možné nájsť napríklad na stránkach MindMapping (2009) alebo MindMap Search (2009).

## 8.12 Folksonómie

Organizácia poznania sa čoraz viac prepája na aktivity komúní na internete. Ľudia triedia a kategorizujú videá, fotografie, záložky, podcasty, blogové príspevky a pod. V nadväznosti na to vznikajú nové štruktúry organizácie poznania, medzi ktoré patria aj folksonómie. Folksonómie môžu pomôcť nielen pri prezeraní obsahu na internete, ale aj pri jeho vyhľadávaní. Veľké vyhľadávače ako napr. Google a Yahoo! pristúpili k integrácii folksonómii na podporu zvýšenia relevancie výsledkov vyhľadávania.

Pojem „folksonómie“ spája výrazy „folk“ (ľudia) a „taxonómie“. Okrem folksonómii sa môžeme stretnúť aj s pojmami ľudové klasifikácie, etnoklasifikácie, distribuované klasifikácie a pod. Tieto pomenovania sú z hľadiska organizácie poznania sporné. Dôvodom je to, že v skutočnosti nejde o klasifikácie, ani taxonómie, kategorizácie. Folksonómie sú často lineárne a vzťahy medzi jednotlivými pojmami väčšinou nie sú presne určené. Používatelia priradujú konkrétnym objektom kategórie – „tagy“. Tagy sa dajú v rôznych systémoch aj organizovať, vyhľadávať, či zdieľať. V rámci komúní ľudia vytvárajú „mraky tagov“ (tag clouds). Mraky tagov, ktoré prepájajú folksonómie jednotlivých používateľov, prezentujú témy, ktoré sa najčastejšie vyskytujú.

Folksonómie môžeme chápať vo viacerých významoch. Okrem procesuálneho hľadiska, kedy folksonómie označujú priradovanie kategórií, sa dajú vnímať aj ako produkty – knižnice tagov, mraky tagov. V abstraktnej rovine ide o pojmové štruktúry, ktoré reflektujú záujmy komúní.

### *Slabé stránky folksonómii*

S folksonómiami súvisí viac problémov. V prvom rade ide o ich lineárnu povahu (výnimkou je napr. možnosť zoskupovania tagov do tém v Delicious). Kategorizácia je pri priradovaní tagov subjektívne podmienená. Takisto vytváranie folksonómii môže byť motivované rôznymi cieľmi. Buď ide o organizáciu informácií pre vlastné účely alebo potreby širšej komunity, či dokonca marketingové ciele. S priradovaním kategórií súvisia aj problémy synonymie a polysémie, ktoré môžu prispievať k nejednoznačnosti. Vnímanie konkrétnych tém a domén je vo vysokej miere ovplyvnené poznatkovou úrovňou používateľov. Napriek týmto nedostatkom sú možnosti vytvárania folksonómii integrované v mnohých systémoch, ktoré budeme charakterizovať ďalej.

### 8.12.1 Podpora folksonómii v aplikáciách založených na webe

#### *CiteULike*

CiteULike, ktorý je alternatívnou známeho Del.icio.us, je určená pre vedeckú komunitu. Pomocou CiteULike možno organizovať a vyhľadávať odborné zdroje, alebo odkazy na vedecké publikácie. Tento nástroj naplno podporuje sociálne siete. Vo vlastnom profile je možnosť vytvorenia knižnice zaujímavých odkazov a sledovanie diania v komunite. Vedci tak majú prehľad o najnovších zdrojoch v záujmových oblastiach a môžu ich zdieľať so svojimi priateľmi. CiteULike podporuje aj ďalšie funkcie, ktoré z neho robia účinný nástroj nevyhnutný pre každého odborníka, ktorý chce ostať v ohnisku diania. Zoznam kategórií, ktoré priradujú jednotliví používatelia k odkazom nájdete v pravej časti nasledujúceho obrázka (8.26).

CiteULike
MyCiteULike

---

**2,975,572 articles - 4,620 added today.**

Latest News - Tuesday 1 September ([More News](#))

**BiomedExperts**

There's a new plugin for posting papers from [BiomedExperts](#). Also, one can do author lookups on BiomedExperts from any author page, e.g.,

<http://www.citeulike.org/author/Greenfield:D>

BiomedExperts deal exclusively with biomedicine so it'll be a waste of time to look up authors in other fields.

**Everyone's Tags**

Most active tags on CiteULike

Filter:

[\[Display as Cloud\]](#)

<a href="#">review</a>	306
<a href="#">evolution</a>	131
<a href="#">model</a>	115
<a href="#">statistics</a>	109
<a href="#">cancer</a>	107
<a href="#">network</a>	104
<a href="#">theory</a>	99
<a href="#">simulation</a>	90
<a href="#">networks</a>	85
<a href="#">structure</a>	83
<a href="#">learning</a>	79
<a href="#">development</a>	77
<a href="#">software</a>	76
<a href="#">bioinformatics</a>	70
<a href="#">analysis</a>	68
<a href="#">methods</a>	67
<a href="#">social</a>	67
<a href="#">human</a>	65
<a href="#">design</a>	62
<a href="#">ontology</a>	60
<a href="#">protein</a>	60
<a href="#">genome</a>	57
<a href="#">clustering</a>	56
<a href="#">microarray</a>	54
<a href="#">research</a>	53

---

**Everyone's library**

Some recent papers posted to CiteULike - all mixed together. You can also see [just your library](#).

[RIS](#) [BibTex](#) [RTF/PDF](#)

[Hide details](#)

**Efficacy of Pneumococcal Vaccination in Children Younger Than 24 Months: A Meta-Analysis**  
*Pediatrics*, Vol. 123, No. 6, (1 June 2009), pp. e1103-1110.  
 by [Maria Pavia](#), [Aida Bianco](#), [Carmelo G. A. Nobile](#), [Paolo Marinelli](#), [Italo F. Angelillo](#)  
 posted to [streptococcus pneumoniae pneumococcus pneumococcal conjugate vaccine paediatric invasive pneumococcal disease bacterial vaccines](#) by [bensal](#) on 2009-09-03 10:25:37 as ★★ [along with 1 person](#)

**The promise and reality of personal genomics.**  
*Genome biology*, Vol. 10, No. 9, (4 September 2009)  
 posted to [review human genomics](#) by [daforeroq](#) on 2009-09-03 10:24:22 as ★★

**Secretory Phospholipase A2 Receptor-Mediated Activation of Cytosolic Phospholipase A2 in Murine Bone Marrow-Derived Mast Cells**  
*J Immunol*, Vol. 165, No. 5, (1 September 2000), pp. 2773-2782.  
 by [Alfred N. Fonteh](#), [Gen-Ichi Atsumi](#), [Tiffany Laporte](#), [Floyd H. Chilton](#)  
 posted to [spla2 pla2](#) by [guhiv](#) on 2009-09-03 10:23:51 as ★★

**Phospholipase A2 receptor: a regulator of biological functions of secretory phospholipase A2**  
*Prostaglandins & Other Lipid Mediators*, Vol. 68-69 (August 2002), pp. 71-82.  
 by [K. Hanasaka](#), [H. Arita](#)  
 posted to [spla2 pla2r pla2](#) by [guhiv](#) on 2009-09-03 10:22:48 as ★★

Obr. 8.26 CiteULike – zoznam kategórií priradených článkom používateľmi

*Picasa*

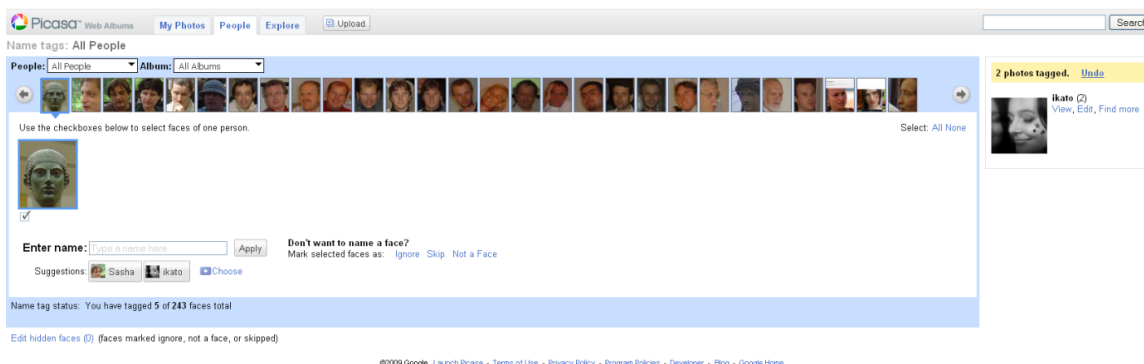
Picasa patrí medzi najznámejšie aplikácie na organizáciu a zdieľanie digitálnych fotiek, ale aj obrázkov a videí. Ide nielen o veľmi intuitívny, ale aj jednoduchý systém. Pri organizácii vašich digitálnych fotiek si vytvárame jednotlivé albumy, ku ktorým možno okrem názvu priradiť aj podrobnejšiu charakteristiku. Vo webových albumoch sa dajú jednotlivé fotky označiť popiskami a zoradiť ich v závislosti od aktuálnosti alebo názvu.

Picasa podporuje aj geotagging – pridávanie zemepisných značiek, ktoré sa dajú neskôr zobrazit' na Google maps (obrázok 8.27).

Obr. 8.27 Fotografie ulíc Barcelony od používateľov Picasa na Google maps

151

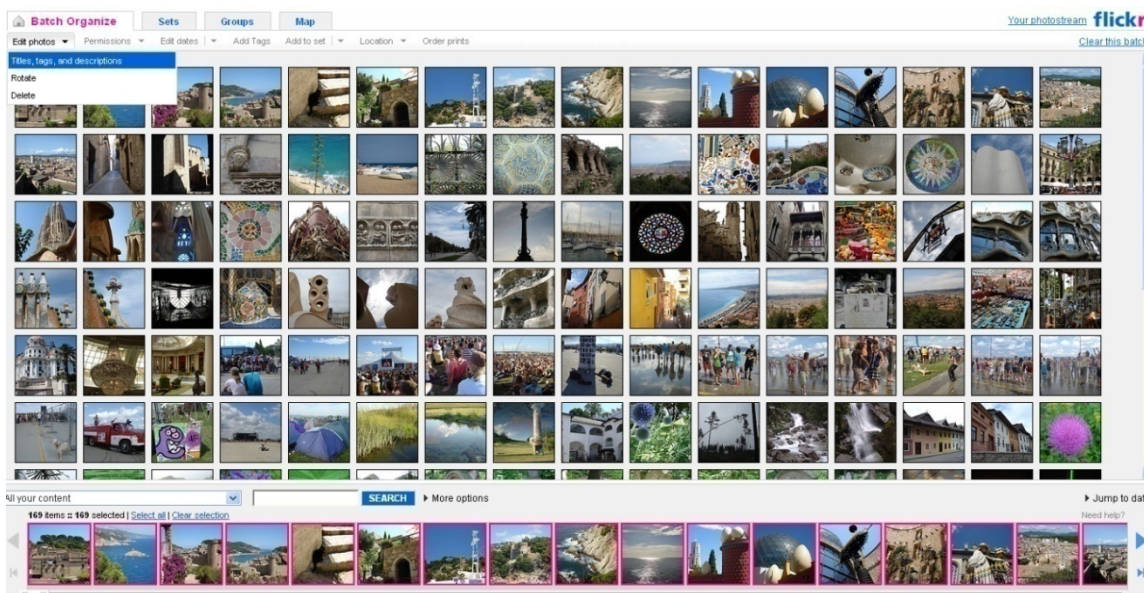
Okrem geotagovania Picasa ponúka zaujímavú novinku – automaticky vyhľadá všetky fotky, na ktorých sú tváre ľudí. K nájdeným tváram sa dajú priradiť značky s menami, ale aj e-mailový kontakt. Táto funkcia zjednodušuje spätné vyhľadávanie fotiek, na ktorých sú konkrétni ľudia. Picasa umožňuje aj odstraňovanie nesprávne rozpoznávaných objektov, ktoré majú charakteristiky ľudskej tváre, napríklad sôch (obrázok 8.28), či ornamentov.



Obr. 8.28 Pridávanie značiek k tváram a odstraňovanie nesprávne rozpoznávaných objektov

### Flickr

Flickr, podobne ako Picasa, podporuje sociálne siete. Vďaka tejto funkcii máme prehľad o najaktuálnejších fotkách, ktoré uložili naši známi. Pre niektorých používateľov môže byť Flickr na prvý pohľad až príliš komplikovaný. Preto zjednodušene vysvetlíme, aké možnosti ponúka z hľadiska organizácie digitálnych fotiek. Fotografie si možno triediť do albumov (tzv. „sety“) a každú fotku označiť názvom, opisom a tagmi (obrázok 8.29).

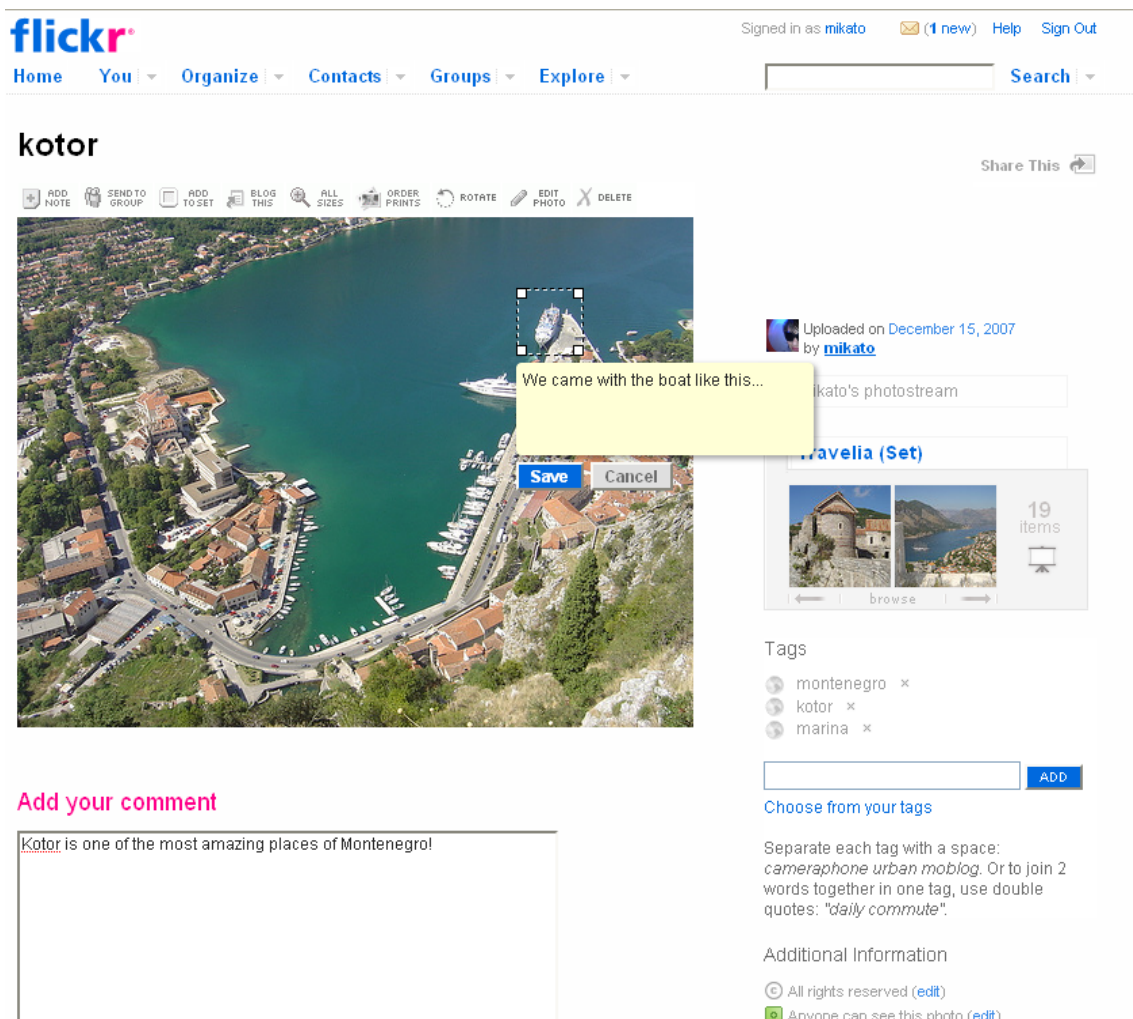


Obr. 8.29 Flickr – triedenie fotografií do setov a ďalšie možnosti organizácie



Ďalej sa dá nastaviť čas a lokácia fotky na mape. Flickr podporuje pridávanie fotiek na mapu. Tejto funkcii sa dá vytknúť to, že mapa nie je natoľko podrobná a pri zväčšovaní môžeme naraziť na problém absencie náhľadov na lokalitu.

Priamo na fotografiu sa dajú pridať aj poznámky, ktoré sú viditeľné pri pohybe myšou nad obrázkom (obrázok 8.30). Flickr je známy najmä pre svoju silnú a rozsiahlu komunitu profesionálnych fotografov. Pre používateľov, ktorí si nerozšíria a nezaplatia Pro konto, sú možnosti limitované určitými obmedzeniami. Napríklad množstvo fotiek, ktoré môžu uchovať na Flickri alebo aj počet albumov, ktoré si môžu vytvoriť.



The screenshot shows the Flickr interface for a photo of Kotor. At the top, the Flickr logo is on the left, and the user is signed in as 'mikato'. Navigation links include Home, You, Organize, Contacts, Groups, and Explore. A search bar is on the right. The photo title is 'kotor'. Below the photo, there are various editing tools: ADD NOTE, SEND TO GROUP, ADD TO SET, BLOG THIS, ALL SIZES, ORDER PRINTS, ROTATE, EDIT PHOTO, and DELETE. A yellow comment box is overlaid on the photo with the text 'We came with the boat like this...'. Below the photo, there is a 'Save' button and a 'Cancel' button. To the right of the photo, there is a 'Share This' button. Below the photo, there is a section for 'Add your comment' with a text input field containing the text 'Kotor is one of the most amazing places of Montenegro!'. To the right of the photo, there is a section for 'Tags' with a list of tags: montenegro, kotor, and marina. Below the tags, there is a text input field and an 'ADD' button. Below the tags, there is a section for 'Additional Information' with two options: 'All rights reserved (edit)' and 'Anyone can see this photo (edit)'.

Obr. 8.30 Flickr – pridávanie poznámok na fotografiu a možnosť komentovania.

### Technorati

Technorati monitoruje a triedi objekty blogosféry. Umožňuje organizovať pridané blogy podľa tagov, popisu a kľúčových slov. Tieto údaje, spolu s obsahom priebežne aktualizovaných príspevkov sú neskôr využité ako vyhľadávacie kritériá. Blogy sú zoradené do kategórií prehľadného tematického adresára. Najviac priradené tagy sú zobrazené pomocou mraku, ktorý nájdete na nasledujúcom obrázku (8.31).

The screenshot shows the Technorati website interface. At the top, there is a navigation bar with categories like Technology, Business, Entertainment, etc. Below that is a search bar and a user profile. The main content area is titled 'Tags by Popularity' and includes a description of tags. A list of 'Top 100 Tags from A to Z (all languages)' is displayed, with 'blogging' and 'weblog' being prominent. To the right, there is a 'Top Tags' section featuring a line graph comparing the popularity of tags like 'Apple vs. Art vs. Blogging vs. Blogroll' from August 5th to September 4th. Below the graph is a 'Compare' button and a section titled 'What bloggers are writing about now:' listing various topics.

Obr. 8.31 Blog registrovaný v adresári Technorati. Sú tu dostupné opisné informácie o blogu, príspevky z blogu, vyhľadávanie, ale aj štatistiky a ďalšie nástroje.

Technorati je účinný nástroj, ktorý monitoruje štatistiky jednotlivých blogov a blogosféry, ale poskytuje aj rôzne iné nástroje, ktoré možno využívať.

Foksonómie sú podporované aj ďalšími známymi web 2.0 nástrojmi na tvorbu a organizáciu textu. Napr. systémy na editovanie blogov (Blogger, Wordpress) tiež umožňujú triedenie blogových príspevkov pomocou pridávania tagov. Podobne je to na stránkach na uchovávanie, zdieľanie a vyhľadávanie videí, podcastov, RSS a pod.

Spolu s rozšírením a obrovskou popularitou sociálnych sietí sa zvyšuje aj potreba organizovať obsah pridávaný používateľmi (napr. fotografie na Facebook-u sa dajú označiť popiskou, ale aj tagom). Aj napriek svojim slabým stránkam bude tagovanie naďalej podporované ako zásadný nástroj na organizáciu obsahu webu. Ľudia tak sami spoločne rozhodujú o tom, ako sa tento obsah bude dať neskôr vyhľadať.

## 8.13 Zhrnutie

Nástroje organizácie poznania v elektronickom prostredí sú rozmanité. Vychádzajú z rôznych typov metadát a smerujú k reprezentácii hlbšieho významu informácií (sémantický web). Nástroje na organizáciu poznania možno deliť podľa obsahu (zoznamy termínov, klasifikačné a kategorizačné schémy ai.) a podľa zložitosti štruktúr (metadáta, taxonómie, ontológie a i.). Smermi vývoja nástrojov v elektronickom prostredí sú sémantické siete a ontológie, tematické mapy, konceptuálne mapy, myšlienkové mapy a folksonómie. Tieto nástroje sa veľmi dynamicky menia, zanikajú a vznikajú.

Sémantický web je rámec vývoja webu, ktorý poskytuje model dát aj reprezentáciu a organizáciu informácií smerom k hlbšiemu významu pojmov. Sémantické siete sú komplexné systémy organizácie a reprezentácie poznania prostredníctvom uzlov (pojmov) a vzťahov medzi nimi.

Ontológie sú špecifické pojmové modely vyjadrujúce logické a sociálne vzťahy vo formalizovaných jazykoch. Tematické mapy opisujú informácie a dokumenty v určitých témach pomocou názvov a asociácií. Konceptuálne mapy zobrazujú pojmy a vzťahy medzi nimi, myšlienkové mapy vizualizujú

myšlienky, súvislosti a činnosti. Folksonómie sú prirodzené komunitné systémy kategorizácie rôzneho obsahu (citácie, záložky, fotografie, blogy a i.).

Najvýraznejšími trendmi v tejto oblasti je sociálna organizácia poznania sprostredkovaná komunitami a webová inteligencia založená na sémantickom webe. Organizácia poznania v elektronickom prostredí je poznačená prechodom k hlbšiemu významu informácií a k začleneniu používateľov do ich tvorby.

## Literatúra

BECHHOFFER, S. et al. 2004. OWL Web Ontology Language Reference : W3C Recommendation 10 February 2004. In Eds. DEAN, Mike – SCHREIBER, Guus. *W3C* [online]. Document Status Update, 12 November 2009 [cit. 2009-09-10]. Dostupné na internete: <<http://www.w3.org/TR/owl-ref/>>.

BERNERS-LEE, Tim. 2009. WorkflowOntology. In *Baetle* [online]. Last update 25-04-2007 [cit. 2009-09-03]. Dostupné na internete: <[http://baetle.googlecode.com/svn/trunk/img/W3C\\_WorkFlowOntology.jpg](http://baetle.googlecode.com/svn/trunk/img/W3C_WorkFlowOntology.jpg)>.

BERNERS-LEE, T., HENDLER, J., LASSILA, O. 2001. The Semantic Web : a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. In *Scientific American*, vol. 284, no. 5 (May), s. 34 – 45.

BRICKLEY, D., GUHA, R. V. 2004. RDF vocabulary description language 1.0 : RDF schema : W3C Recommendation 10 February 2004. In Ed. McBRIDE, Brian. *W3C* [online]. 2004 [cit. 2010-01-10]. Dostupné na internete: <<http://www.w3.org/TR/rdf-schema/>>.

BUSH, V. 1945. As we May Think. In *The Atlantic* [online]. July 1945 [cit. 2009-10-12]. Dostupné na internete: <<http://www.theatlantic.com/doc/194507/bush>>.

COSENTINO, Massimo et al. 2008. *ASPECS Methodology : Problem Ontology Description* [online]. ASPECS : Agent-oriented Software Process for Engineering Complex Systems, 13 May 2008 [cit. 2009-09-08]. Dostupné na internete: <[http://set.utbm.fr/upload/gestionFichiers/SoccerOntology\\_809.png](http://set.utbm.fr/upload/gestionFichiers/SoccerOntology_809.png)>.

CUADRADO, John L. – CEGLOWSKI, Maciej. 2009. What is a Memex? In *Peer-to-Peer Semantic Search Engines : Building a Memex* [online]. NITLE, © 2002-2009 [cit. 2009-08-10]. Dostupné na internete: <<http://www.knowledgesearch.org/presentations/etcon/slide0002.htm>>.

GARSHOL, Lars Marius. 2002. *What are topic maps?* [online]. September 11, 2002 [cit. 2009-08-12]. Dostupné na internete: <<http://www.xml.com/pub/a/2002/09/11/topicmaps.html>>.

GRUBER, T. R. 1993. *What is an ontology? : Ontologies as a specification mechanism* [online]. Accessed March 2003 [cit. 2010-01-10]. Dostupné na internete: <<http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>>.

HALEN, Cees van. 2009. *Showing the linkages from an issue, concept or problem with other concepts or issues : Explanation of Mind Mapping* [online]. Bilthoven (Holandsko) : 12 Manage the Executive Fast Track, 2009 [cit. 2009-10-17]. Dostupné na internete: <[http://www.12manage.com/methods\\_mind\\_mapping.html](http://www.12manage.com/methods_mind_mapping.html)>.

HERMAN, I. 2009. W3C Semantic Web Activity. In *W3C* [online]. Last update 2009/12/09 [cit. 2009-08-20]. Dostupné na internete: <<http://www.w3.org/2001/sw/>>.

HERMAN, Ivan – SWICK, Ralph – BRICKLEY, Dan. 2009. Resource Definition Framework (RDF). In *W3C* [online]. Last update 2009/08/29 [cit. 2009-08-12]. Dostupné na internete: <<http://www.w3.org/RDF/>>.

- HJØRLAND, Birger. 2004. *Theory of knowledge organization and the feasibility of universal solutions* [online]. Eighth International ISKO Conference, London, (Friday) July 16th 2004, Session 9B 2.15-3.45. Dostupné na internete: <<http://www.ucl.ac.uk/isko2004/sysweb/9bHjorland.ppt>>.
- JACOB, Elin – LOEHRLEIN, Aaron. 2003. *What ontologies are not : [draft] theoretical framework for the analysis of representational systems* [online]. Updated 11 April 2003.
- KLAVANS, R., BOYACK, K. W. 2006. Quantitative evaluation of large maps of science. In *Scientometrics*. ISSN 0138-9130, December 2006, vol. 68, no. 3, s. 475 – 499.
- LANE, Johnathan. 2008. Information Architecture – Planning out a web site. In *Dev.Opera* [online]. 8 Jul, 2008 [cit. 2009-09-12]. Dostupné na internete: <<http://dev.opera.com/articles/view/6-information-architecture-planning-o/>>.
- McCANN, John M. 1997. *Generation of Marketing Insights : Semantic Networks* [online]. Durham : Duke University, 1997 [cit. 2009-09-15]. Dostupné na internete: <<http://www.duke.edu/~mccann/mwb/15semnet.htm>>.
- NOVAK, Joseph D. – CANAS, Alberto J. 2008. *The Theory Underlying Concept maps and How to Construct and Use Them* [online]. Pensacola : Florida Institute for Human and Machine Cognition, Rev. 2008-01 [cit. 2009-10-03]. Dostupné na internete: <<http://cmap.ihmc.us/Publications/ResearchPapers/TheoryCmaps/TheoryUnderlyingConceptMaps.htm>>.
- PALEY, W. Bradford. 2006. *Topic Map : How Scientific Paradigms Relate* [online]. 2006 [cit. 2009-09-03]. Dostupné na internete: <[http://wbpaley.com/images/finalTopicMap5+1\\_.jpg](http://wbpaley.com/images/finalTopicMap5+1_.jpg)>.
- PEPPER, Steve – MOORE, Graham. 2001. *XML Topic Maps (XTM) 1.0 : TopicMaps.Org Specification* [online]. TopicMaps.org, Last update 2001/08/06 [cit. 2010-01-11]. Dostupné na internete: <<http://topicmaps.org/xtm/1.0/index.html>>.
- SHARPLES, Mike et al. 1996. *Computers and Thought : A practical Introduction to Artificial Intelligence* [online]. Last update Oct 22, 1996 [cit. 2010-01-10]. Dostupné na internete: <<http://www.cs.bham.ac.uk/research/projects/poplog/computers-and-thought/>>.
- SHAW, Blake. 2006. Visualizing Folksonomies using Machine Learning Algorithms. In *Metablake : A collection of projects and ideas* [online]. Last update 2006-03-01 [cit. 2009-08-20]. Dostupné na internete: <<http://www.metablake.com/advml/map1.png>>.
- SOWA, John F. 2006. *Semantic Networks* [online]. Last Modified: 06/02/2006 [cit. 2009-10-12]. Dostupné na internete: <<http://www.jfsowa.com/pubs/semnet.htm>>.
- STEINEROVÁ, Jela. 1998b. Register. In Eds. KATUŠČÁK, D., MATTHAEIDESOVÁ, M., NOVÁKOVÁ, M. *Informačná výchova : Terminologický a výkladový slovník : Odbor Knižničná a informačná veda*. Bratislava : SPN, 1998. 375 s. ISBN 80-08-02818-1.
- STEINEROVÁ, Jela a kol. 2004. *Správa o empirickom prieskume používateľov knižníc ako súčasť grantovej úlohy VEGA 1/9236/02 Interakcia človeka s informačným prostredím v informačnej spoločnosti*. Bratislava : Filozofická fakulta UK, KKIV, 2004. 113 s.
- STEINEROVÁ, Jela. 2005b. Informačné správanie človeka v knižniciach a v elektronickom prostredí – dôsledky pre organizáciu poznatkov v digitálnych knižniciach. In *Infos 2005*. Zborník príspevkov. Medzinárodné informatické sympóziu. Bratislava : CVTI SR, 2005, s. 149 – 160.
- SVÁTEK, Vojtech. 2002. Ontologie a WWW. In *DATAKON 2002*. Brno, 19. – 22.10.2002, s. 1 – 35.
- WILDE, Erik. 2009. *WEB and XML Glossary* [online]. Last modification: 13-May-2009 [cit. 2009-08-29]. Dostupné na internete: <<http://dret.net/glossary/>>.



## Zoznam webových sídel

- Ask  
[www.ask.com](http://www.ask.com)
- Backpack  
<http://backpackit.com/>
- BBC  
<http://www.bbc.co.uk/>
- BBC News Browser Pilot Project  
<http://www.nactem.ac.uk/bbc/>
- Bing Translator  
<http://www.microsofttranslator.com/>
- Blogger  
<https://www.blogger.com/start>
- Bubbl.us  
<http://www.bubbl.us>
- CiteULike  
<http://www.citeulike.org>
- Cyc  
[www.opencyc.org](http://www.opencyc.org)  
[www.cyc.com](http://www.cyc.com)
- DAML Ontology Library  
<http://www.daml.org/ontologies/>
- Del.icio.us  
<http://delicious.com/>
- Digg  
<http://digg.com/>
- Dublin Core Metadata Initiative  
<http://dublincore.org/>
- Flickr  
<http://www.flickr.com>
- Gene ontology consortium  
<http://www.geneontology.org/>
- Google  
[www.google.com](http://www.google.com)  
[www.google.sk](http://www.google.sk)
- Google Calendar  
[www.google.com/calendar/](http://www.google.com/calendar/)
- Google Language Tools  
[www.google.com.au/language\\_tools](http://www.google.com.au/language_tools)
- GoPubMed  
<http://www.pubmed.org/>
- Hypertextopia  
<http://www.hypertextopia.com>
- INTUTE  
<http://www.intute.ac.uk/>
- KAON  
<http://kaon.semanticweb.org>
- Karlsruhe On-tology infrastructure  
<http://kaon.semanticweb.org/>
- Kartoo  
<http://www.kartoo.com/>
- KnowledgeWeb  
<http://www.knowledgeweb.org/>
- LCSH (Library of Congress Subject Headings)  
<http://id.loc.gov/authorities/search/>
- Mappio  
<http://mappio.com>
- Memex and beyond Web Site  
<http://www.cs.brown.edu/memex/>  
[ACM\\_HypertextTestbed/](http://www.cs.brown.edu/memex/ACM_HypertextTestbed/)  
[http://www.cs.brown.edu/memex/ACM\\_HypertextTestbed/papers/HARTadj5in.jpg](http://www.cs.brown.edu/memex/ACM_HypertextTestbed/papers/HARTadj5in.jpg)
- MeSH (Medical Subject Headings)  
<http://www.nlm.nih.gov/mesh/>
- Mind-mapping.org  
<http://www.mind-mapping.org/>
- MindMap Search  
<http://www.mindmapsearch.org>
- MindMeister  
<http://www.mindmeister.com>
- Mindomo  
<http://www.mindomo.com>
- MyLifeBits  
<http://research.microsoft.com/en-us/projects/mylifebits/>
- NACTEM  
<http://www.nactem.ac.uk/>

NaCTem Intute  
[http://www.nactem.ac.uk/intute/intute\\_tutorial.htm](http://www.nactem.ac.uk/intute/intute_tutorial.htm)

Omnigator  
<http://www.ontopia.net/omnigator/>

On-To-Knowledge  
<http://www.ontoknowledge.org/>

Ontosearch2  
<http://www.ontosearch.org/>

OntoWeb  
<http://www.ontoweb.org/>

Picasa  
<http://picasaweb.google.com>

RememberTheMilk  
<http://www.rememberthemilk.com/>

SnapShots  
<http://www.snap.com/>

StorySpace  
<http://www.eastgate.com/storyspace>

Swoogle: semantic web search  
<http://swoogle.umbc.edu/>

Technorati  
<http://www.technorati.com>

TechTarget, Inc. (definicia Taxonomy)  
[http://whatis.techtarget.com/definition/0,,sid9\\_gci331416,00.html](http://whatis.techtarget.com/definition/0,,sid9_gci331416,00.html)

Agrovoc Thesaurus  
<http://aims.fao.org/website/AGROVOC-Thesaurus/sub>

Vizigator  
<http://www.ontopia.net/solutions/vizigator.html>

Web and XML Glossary  
<http://dret.net/glossary/>

Wiki  
<http://en.wikipedia.org/wiki/Wiki>

WordNet  
<http://wordnet.princeton.edu/>

Wordpress  
<http://wordpress.org/>

W3C  
<http://www.w3.org/>

Yahoo! Babel Fish  
<http://babelfish.yahoo.com>

# 9

## Digitálne knižnice

### 9.1

#### Stručná história

Históriu digitálnych knižníc možno odvodiť už od predstavy Vannevara Busha o Memex z roku 1945. Mala to byť mechanizovaná knižnica obsahujúca osobný systém organizácie informácií. V 50. a 60. rokoch 20. stor. sa začali budovať online katalógy, ktoré vznikli z počítačového indexovania mikrofilmov 20 000 vedeckých článkov na Massachusetts Institute of Technology (MIT).

V 60. rokoch prvýkrát predstavil ucelenejšiu koncepciu digitálnych knižníc J. C. R. Licklider v monografii o knižniciach budúcnosti. Išlo o oslobodenie človeka v interakcii s poznaním. Koncom 60. rokov sa objavili aj prvé aplikácie na indexovanie právnych a vedeckých dokumentov, napríklad the Ohio Bar Automated Research (OBAR) System alebo Legal Information Through Electronics (LITE), systém amerických leteckých síl.

Najznámejším systémom zo 70. rokov bol systém IBM nazvaný STAIRS (IBM Storage and Retrieval System). Pôvodne bol určený pre veľké počítače a zabezpečoval prístup k bibliografickým informáciám a abstraktom stoviek knižníc. Systém bol obmedzený možnosťami technológií a orientáciou na štruktúrovaný text. S vývojom technológií sa postupne zdokonaľoval nielen informačný prieskum, ale aj možnosti spracovania hypertextu, obrazov, fotografií a zvuku.

V roku 1994 vypracovala Kongresová knižnica koncepciu národnej digitálnej knižnice ako základ národnej informačnej infraštruktúry. V roku 1995 americké nadácie na podporu vedy a výskumu ako NSF, NASA, DARPA podporili projekt výskumu digitálnych knižníc (Digital Library Initiative (DLI) v dvoch fázach (DLI-1, 1994-1999, DLI-2, 2000 – 2004). Tak vzniklo celé odvetvie výskumu digitálnych knižníc a množstvo aplikácií v oblasti vedy, vzdelávania, kultúry a pamäťových inštitúcií (knižnice, galérie, múzeá, archívy).

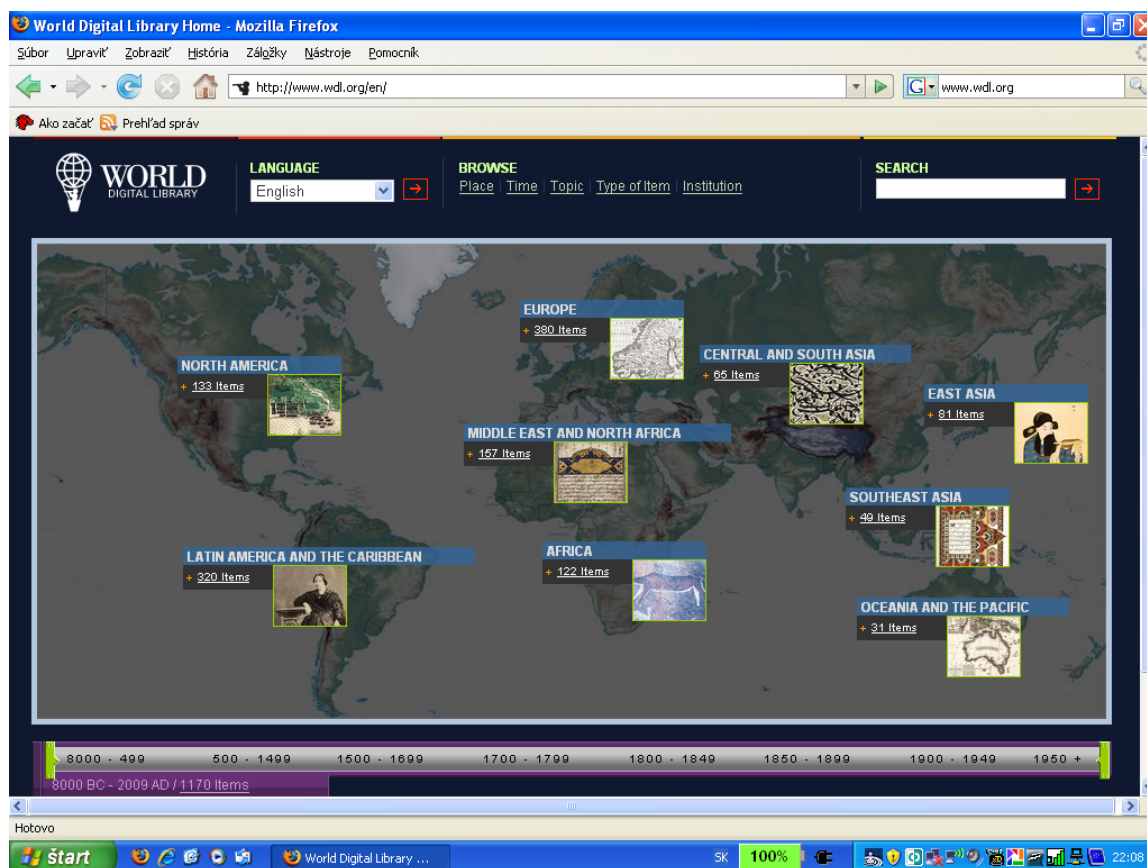
Termín digitálna knižnica prvýkrát použil Gorman v roku 1974. Všeobecne prijatý a využívaný je od začiatku 90. rokov 20. storočia. Súvisiace pojmy sú virtuálna knižnica a elektronická knižnica. Niektorí autori ich rozlišujú v tom, že virtuálna knižnica obsahuje odkazy na dokumenty (bibliografické záznamy) a digitálna knižnica obsahuje skutočné digitálne objekty. Krajné významy pojmu digitálna knižnica možno definovať z hľadiska počítačovej vedy (informatiky) a z hľadiska knižničnej a informačnej vedy. Z hľadiska informatiky je digitálna knižnica distribuovaný informačný systém. Z hľadiska knihovníckeho je predstava digitálnej knižnice spojená s prenosom tradičnej knižnice do elektronického prostredia.

Vo výskumoch digitálnych knižníc možno rozlíšiť dve etapy. Prvá etapa sa zameriavala na technické hľadisko, technológie, štandardy a protokoly. Druhá etapa sa neskôr začala orientovať na sociálne kontexty služieb, prepojenie s komunitami a hodnotenie ich efektívnosti. V prvej etape sa začali

objavovať aj prvé publikácie o digitálnych knižniciach. Napríklad v rokoch 1993 a 1994 vyšla Digital Library Resource Book od Edwarda Foxa, v roku 1994 vyšla monografia Karen Drabenstottovej o knižnici budúcnosti. Prvá konferencia o digitálnych knižniciach sa konala v Texase, College Station v roku 1994. Vznikli aj špecializované tematické čísla časopisov JASIS (Journal of the American Society of Information Science) v roku 1993, Communications of the ACM v roku 1995. V roku 1995 začal vychádzať aj špecializovaný časopis D-Lib Magazine.

Významné monografie spracovali títo autori, profesori informačnej vedy: Gary Marchionini v roku 1995, Michael Lesk (1997), Christine L. Borgmanová (2000, 2007). W. Y Arms (2000), P. Lyman (1999).

Paralelne sa rozvíjali výskumy aj na úrovni Európy. Najznámejším je projekt DELOS (<http://www.delos.info>) predstavujúci sieť excelentnosti výskumov digitálnych knižníc ako súčasť programu zameraného na informačnú spoločnosť a technológie (IST). Praktickým výsledkom sú také projekty ako The European Library (<http://www.theeuropeanlibrary.org>), the UNESCO World Digital Library (obrázok 9.1) a iné.



Obr. 9.1 World Digital Library

Významným kontextom fungovania systémov a projektov na európskej úrovni je stratégia rozvoja informačnej spoločnosti do roku 2010: EC – IST, i2010-DL ([http://ec.europa.eu/information\\_society/](http://ec.europa.eu/information_society/)). V najnovších výskumoch sa zdôrazňuje využívanie digitálneho obsahu v kultúrnych inštitúciách, najmä tri aspekty: 1. digitalizácia a budovanie centier digitalizácie a konzervácie, 2. riešenie otázok autorského práva pre „bezhraničné digitálne objekty“ a 3. podpora experimentov s „open access“ softvérom pri podpore poznatkov a tvorivosti v elektronickom prostredí (creative commons, knowledge commons).

V roku 2006 vyšiel Manifest o digitálnych knižniciach kolektívu autorov – vedcov z USA aj Európy, ktorý sumarizuje doterajšie výskumy a predvída trendy vývoja (The Digital Library Manifesto 2006). Z množstva konferencií o digitálnych knižniciach sú známe najmä ECDL (European Conference on Digital Libraries), JCDL, LIDA a iné.

## 9.2 Definovanie pojmu

Pojem digitálna knižnica v súčasnosti definujeme v širšom zmysle ako digitálne prostredie vhodné na sprístupňovanie digitálneho obsahu a vytváranie služieb a produktov s pridanou hodnotou. Základní autori (E. Fox, Ch. Borgmanová, N. Belkin, D. Soergel, M. Lesk, G. Marchionini, D. Bawden, I. Rowlands, M. Goncalves a i.) sa zhodujú v tom, že digitálna knižnica predstavuje nielen *priestor na ukladanie* informačných objektov, ale aj *priestor na komunikovanie*.

Digitálna knižnica je interaktívnym, flexibilným systémom spojeným s potrebami a komunikáciou komunity. Pridanou hodnotou digitálnych knižníc je aj špeciálna organizácia poznatkov a konsolidované služby a produkty. Napríklad v koncepcii G. Marchioniniho (1999) sú digitálne knižnice „sharium“ – interaktívne digitálne prostredia navrhnuté na podporu spolupráce medzi používateľmi aj knihovníkmi.

Spoločný priestor na komunikovanie poznatkov môže uľahčiť komunikáciu a spoluprácu medzi vedcami, študentmi a inými komunitami. V spoločnom priestore existujú „informačné objekty“ (digitálny obsah), ktoré sa rôzne prepájajú, rozkladajú a znovu integrujú podľa konkrétnych situácií a potrieb. Sociálni aktéri, napríklad učitelia, výskumníci, študenti a ľudia z praxe, tu využívajú technológie na efektívny prístup k zdrojom a kolegom, expertom, službám, nástrojom, publikáciám, empirickým údajom a prekračujú geografické, časové a inštitucionálne bariéry.

Existuje už mnoho definícií digitálnych knižníc. Digital Library Federation (1999) definuje digitálne knižnice ako organizácie, ktoré poskytujú zdroje aj špecializovaných pracovníkov na výber, štruktúrovanie a poskytovanie intelektuálneho prístupu, interpretáciu, rozširovanie, zabezpečenie konzistentnosti a pretrvania v čase súborov digitálnych diel tak, aby boli pripravené a prístupné na využívanie definovanou komunitou alebo súborom komunit za ekonomicky výhodných podmienok. Iná definícia určuje, že digitálna knižnica je knižnica, ktorá udržiava celý alebo podstatnú časť svojho fondu vo forme prístupnej prostredníctvom počítačového systému ako alternatíva konvenčných tlačených a mikrofilmových dokumentov, ktoré v súčasnosti dominujú vo fondoch knižníc. Termín fond/súbor označuje dokumenty, ktoré knižnica nadobúda alebo udržiava (Saffady 1995).

Podľa Borgmanovej (2007) sú digitálne knižnice súbory elektronických zdrojov a zodpovedajúcich technických funkcií na vytváranie, vyhľadávanie a využívanie informácií. Sú v tomto zmysle rozšírením systémov informačného prieskumu, ktoré manipulujú s digitálnymi údajmi v akomkoľvek médiu (text, obrazy, zvuky, statické alebo dynamické obrazy) a existujú v distribuovaných sieťach. Obsahom digitálnych knižníc sú údaje, metaúdaje, ktoré opisujú rôzne aspekty údajov, linky a vzťahy k iným dátam a metadátam, či už interné alebo externé pre digitálnu knižnicu.

Spoločné črty definícií digitálnych knižníc možno zosumarizovať takto:

1. ide o súbor digitálnych zdrojov
2. sú selektované podľa určitých kritérií,
3. sú určitým spôsobom organizované,
4. sú sprístupnené pre vyhľadávanie v počítačových sieťach,
5. sú určené pre definovanú komunitu používateľov.

Digitálna knižnica je ukľadací priestor pre poznatky a tiež komunikačný kanál, nástroj aj prostriedok na poskytovanie služieb. Druhá definícia Borgmanovej (2007) zdôrazňuje, že digitálne knižnice sú vytvorené (zhromaždené a organizované) komunitou používateľov. Ich funkcie majú podporovať informačné potreby a využívanie informácií touto komunitou. Sú zložkou týchto komunit,

v ktorých sú vo vzájomnej interakcii jednotlivci a skupiny pri využívaní údajov, informácií a zdrojov znalostí a systémov. Digitálne knižnice sú rozšírením informačných inštitúcií ako fyzických miest podporujúcich využívanie organizovaných zdrojov určitou komunitou. Príkladmi týchto informačných inštitúcií nie sú len knižnice, múzeá, archívy, ale aj školy, kancelárie, triedy, laboratória, domácnosti a verejné priestory. Distribuované súbory digitálnych knižníc obsahujú rôzne druhy médií a informačných objektov. Pritom primárne a sekundárne zdroje sa prelínajú a tvoria kontinuum.

Informačný objekt definujeme ako hlavnú jednotku informácie, ktorá je riadená digitálnou knižnicou. Informačný objekt je zdrojom digitálnej knižnice. Identifikuje ho identifikátor informačného objektu a obsahuje metadáta pre rôzne účely manažmentu. Navyše informačný objekt môže mať rôzne vydania, prejavy a môže obsahovať anotácie. Informačný objekt je abstraktný pojem. Môže nadobúdať rôzne prejavy na rôznych úrovniach konkrétnosti (napríklad edícia, artefakt a pod.).

Príkladom práce s informačným objektom je aj objektovo orientovaný konceptuálny referenčný model (Conceptual Reference Model, CRM, norma ISO 21127 CIDOC CRM). Využíva sa na zabezpečenie kompatibility pri technologickom spracovaní údajov o objektoch v pamäťových inštitúciách (múzeách, galériách, knižniciach, archívoch). Informácie o rôznych kultúrnych objektoch a artefaktoch sa môžu nachádzať v rôznych pamäťových inštitúciách a nový model ich umožňuje prepájať vo významových vzťahoch. Tieto vzťahy určujú vzťahy osôb (aktérov), času, miesta, udalosti a i. Napríklad artefakty uložené v rôznych inštitúciách sa môžu viazať na udalosť podpisu určitej politickej dohody. Model umožňuje prepojiť informácie týkajúce sa tejto udalosti – článok z novin, zvukový záznam prejavu politikov, biografické informácie aktérov, miesto podpisu (mesto), historickú situáciu, rukopis dokumentu, informácie o majiteľovi hotela, v ktorom sa udalosť odohrala či o pere, ktorým bola dohoda podpísaná atď. Model reprezentuje vzťahy v pravidlách a ukladá ich v báze znalostí. V súčasnosti sa pracuje aj na integrácii modelov IFLA-FRBR a CIDOC CRM. Cez koncepciu „informačného objektu“ sa v praxi prechádza od dát a informácií k znalostiam integrovaným v spoločných pokročilých inteligentných (znalostných) systémoch.

Význam pojmu digitálna knižnica sa vyznačuje tromi posunmi. Prvým posunom je prechod od orientácie na systém (prístup k zdrojom) a na obsah (jeho organizácia, informačná architektúra) k *orientácii na rôzne činnosti človeka v elektronickom prostredí*. Druhým významovým posunom je zmena tradičnej predstavy o uchovávaní a následnom vyhľadávaní informácií smerom k *podpore komunikácie, kolaborácie a interakcie členov komunit v elektronickom prostredí*. Tretí posun umožňuje vidieť digitálnu knižnicu ako pojem smerujúci od uloženia textov a multimédií k obsiahnutiu senzorických údajov, mobilnej komunikácii a nadväzným službám.

### 9.3

## Vlastnosti a problémy digitálnych knižníc

Digitálne knižnice obsahujú digitálne dokumenty bez fyzikálnych alebo administratívnych obmedzení. Poskytujú ucelený pohľad na obsiahnuté digitálne objekty. Dokumenty tu už nekopírujú tradičné tlačeneé zdroje, ale sú rozložené do tzv. digitálnych objektov (text, obraz, zvuk, trojrozmerný objekt a pod.). Tie sa uchovávajú v menších častiach a na výstupe sa pre používateľov spájajú do špeciálnych personalizovaných služieb a produktov.

Digitálna knižnica poskytuje zdroje organizovaným spôsobom a využíva špeciálne nástroje na štruktúraciu informácií. Zabezpečuje inteligentný prístup k zdrojom a slúži konkrétnym komunitám. Obsahuje všetky procesy a služby typické pre tradičné knižnice.

Problémy digitálnych knižníc vyplývajú z nestability elektronických informácií. Ide najmä o ľahké kopírovanie a možnosť viacnásobného použitia. Problémy vznikajú na technickej úrovni, na úrovni budovania súborov/fondov, na úrovni ich spracovania (metaúdaje) a na úrovni ich využívania (zabezpečenie prístupu, autorské právo).

Problém technickej architektúry súvisí s tým, ako navzájom spolupracujú rôzne systémy. Spájajú sa bibliografické bázy, adresáre, obrazy, nástroje informačného prieskumu, elektronické časopisy. Tieto rôznorodé zdroje majú byť v digitálnej knižnici sprístupnené cez jedno rozhranie či bránu (portál), pričom používateľ má dojem, že komunikuje len s jedným systémom. Príkladom môže byť JIB – Jednotní informačná brána českej národnej knižnice (<http://www.jib.nkp.cz>). Často sa budujú aj odborovo orientované rozhrania a brány.

Problém budovania digitálnych súborov súvisí s tým, ako získať pôvodné digitalizované dokumenty (digital-born materials) a ako určiť hranicu medzi internými a externými elektronickými zdrojmi. Pri digitalizácii tlačенých dokumentov vznikajú problémy pri výbere kritérií na výber dokumentov. Medzi metódami digitalizácie sa najčastejšie používa konverzia tlačенých zdrojov do digitálnej formy (skenovanie). Používajú sa však aj metódy vzorkovania (vzorka reprezentatívnych dokumentov) alebo prepisovania. Koordinované budovanie digitálnych fondov si vyžaduje prepracované akvizičné stratégie prepojené s dlhodobým archivovaním, zabezpečovaním prístupnosti a viacnásobným využívaním digitálnych objektov. Najčastejšie kritériá na výber dokumentov na digitalizáciu sú potenciál na dlhodobé využitie, intelektuálna a kultúrna hodnota, možnosť lepšieho elektronického prístupu podľa konkrétnej požiadavky.

Z hľadiska spracovania sa pri digitálnych knižniciach uplatňujú najmä metadáta (metaúdaje). Ide o údaje, ktoré opisujú obsah a vlastnosti konkrétneho dokumentu. Predstavujú paralelu k tradičným katalogizačným a bibliografickým záznamom. Ich hlavnou funkciou je identifikovanie zdrojov a ich využitie. Keďže organizácia poznania v elektronickom prostredí by bola pri množstve údajov tradičným spôsobom príliš drahá, vznikla pôvodne jednoduchšia schéma metadát známa ako Dublin Core. Obsahuje 15 prvkov („jadro“) potrebných na opis elektronickým dokumentov. Metadáta sa delia na sémantické, syntaktické, deskriptívne, administratívne a iné. Spracovanie elektronických informácií určujú aj ďalšie schémy ako PURL, MARC, TEI (Text Encoding Initiative), RDF (Resource Description Framework) alebo DOI (Digital Object Identifier). Podrobnejšie informácie o týchto schémach sú napríklad v publikácii Šušol 2003.

Pri spracovaní sa často riešia aj problémy pomenovania, názvov (identifikátorov) digitálnych objektov. Tieto identifikátory sú reťazce, ktoré jedinečne identifikujú digitálny objekt a sú súčasťou metaúdajov dokumentu. V premenlivom elektronickom prostredí je dôležité zabezpečiť určitú stálosť týchto pomenovaní. Jedným z prvých riešení bol pokus oddeliť názov dokumentu od jeho lokácie v systéme PURL (perzistentný systém jednotnej lokácie zdrojov), ktorý bol vyvinutý v OCLC. URN (Uniform Resource Name) obsahuje identifikátor pomenovania a identifikátor objektu. DOI (Digital Object Identifier) systém rieši problém rozdelením informácií na lokačné a na informácie prístupu k digitálnemu objektu.

Problémom digitálnych knižníc je aj otázka manažmentu autorských práv a prístupových práv. Tradičné modely autorských práv sa v elektronickom prostredí rozpadajú. V spolupráci s vydavateľmi a knižnicami preto vznikajú nové systémy na manažment práv v digitálnych knižniciach, ich funkciou je aj sledovanie využívania zdrojov, identifikovania a ochrana používateľov, informovanie o statuse digitálneho objekt z hľadiska obmedzeného prístupu, ale aj spracovanie transakcií (poplatky za prístup, kópie, požiadavky a i.).

Z hľadiska uchovávania a ochrany digitálnych objektov sa v digitálnych knižniciach riešia problémy spojené s technickým zastarávaním pamäťových médií. Najmä v pamäťových inštitúciách je často potrebné presúvať digitálne objekty z jedného média na aktuálnejšie. Problémy tvoria aj rôzne formáty, verzie a štandardy digitálnych objektov a generácie informačných systémov. Knižnice tu môžu zohrať významnú úlohu pri vytváraní stratégií na dlhodobé a systematické uchovávanie digitálnych objektov.

Oproti pôvodným ideálnym predstavám o jednotnej digitálnej knižnici ako univerzálnom digitalizovanom systéme sa vytvárajú v súčasnosti digitálne knižnice spojené s určitým odborom alebo komunitou (napríklad veda, vzdelávanie, kultúrne inštitúcie). Pri budovaní digitálnych knižníc sú aktuálne postupné kroky, evolúcia po menších, zvládnuteľných častiach, ktoré sa napokon môžu prepojiť do širších systémov (napríklad systém európskych národných knižníc Europeana).

Pri realizácii projektov digitálnych knižníc treba rozlišovať rôzne typy a funkcie tradičných knižníc. Napríklad národné, akademické a vedecké knižnice, verejné a školské knižnice, špeciálne a firemné knižnice. Úlohou národných knižníc je najmä uchovanie a ochrana národného kultúrneho dedičstva či kultúrnej pamäti. V tomto zmysle majú veľkú úlohu aj pri digitalizácii a budovaní národnej digitálnej knižnice (napríklad Slovenská digitálna knižnica Slovenskej národnej knižnice). Akademické a verejné knižnice sa zas v digitálnom prostredí užšie spájajú s inštitúciami a komunitami. Akademické digitálne knižnice sa včleňujú do procesov vzdelávania, vedy a výskumu, učenia. Verejné knižnice sa v digitálnom priestore spájajú so životom komunit, sú nástrojom na prístup občanov k informáciám (portály e-vlády, e-zdravia, e-kultúry a i.). Digitálne knižnice verejných knižníc môžu byť nástrojom regionálnej informačnej politiky a umožňujú účasť ľudí na demokracii. Firemné a špeciálne knižnice sa včleňujú do podpory informačného správania zamestnancov a samotných firiem. Digitálne knižnice sú súčasťou ich informačných systémov.

Najvýznamnejšie projekty digitálnych knižníc sú v oblasti vzdelávania, vedy a výskumu. Známe sú také projekty ako digitálna knižnica Alexandria (Alexandria Digital Library, <http://www.alexandria.ucsb.edu>) zameraná na geopriestorové multimediálne dokumenty, California Digital Library, digitálna knižnica Perseus Digital Library, National Science Digital Library (NSDL). Vo vede sa menia tradičné funkcie uznania výsledkov výskumu, rozširovania výsledkov, prístupu k nim a dlhodobej archivácie výstupov. Napríklad nové formy manipulácie s dátami sa objavujú vo výskumoch ľudského genómu ([http://www.ornl.gov/sci/techresources/Human\\_Genome/home.shtml](http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml)) alebo v pozorovaniach digitálnej oblohy v astronómii (US National Virtual Observatory Project, <http://www.us-vo.org>). Informácie a dáta sú vnorené vo vedeckých sociálnych procesoch, pričom medzi disciplínami môžu existovať významné rozdiely.

Hľadajú sa stále nové nástroje na vyhľadávanie a objavovanie informácií, analýzy a interpretácie. Videozáznamy, organizačné plánovanie a sociálne siete sú výhodami digitálnych knižníc. V digitálnych knižniciach sa integrujú senzorické údaje, správa a analýza údajov, vizualizácia analýz a pojmových štruktúr, simulácia a modelovanie javov. Ako príklady možno spomenúť Britské národné centrum e-Science (<http://www.nesc.ac.uk>), americké centrum údajov pre interakcie človeka s prostredím (CIESIN) (<http://sedac.ciesin.columbia.edu/wdc/#>), informačné infraštruktúry univerzít v Oxforde (<http://www.oerc.ox.ac.uk>) alebo Indiane (<http://iuni.slis.indiana.edu/>).

## 9.4 Štruktúra digitálnej knižnice podľa Manifestu digitálnych knižníc

### 9.4.1 Univerzum digitálnych knižníc – základné pojmy

Tri základné vrstvy univerza digitálnych knižníc predstavujú tieto pojmy: digitálna knižnica (DL), systém digitálnej knižnice (DLS) a systém manažmentu digitálnej knižnice (DLMS). Digitálna knižnica je organizácia, ktorá zhromažďuje, riadi a dlhodobo uchováva bohatý digitálny obsah a poskytuje používateľským komunitám špeciálnu funkčnosť v merateľnej kvalite a podľa predpísaných pravidiel. Systém digitálnej knižnice je programový systém založený na architektúre a poskytovaní funkcií, ktoré sa vyžadujú pre konkrétnu digitálnu knižnicu. Zabezpečuje interakciu používateľov s digitálnou knižnicou. Systém manažmentu digitálnej knižnice je všeobecný softvérový systém, ktorý poskytuje vhodnú infraštruktúru na vytvorenie a administráciu systému digitálnej knižnice a integráciu ďalších špecializovaných funkcií.

Na základe dôkladnej pojmovej analýzy manifest definuje vzťahy medzi pojmami model digitálnej knižnice, model systému digitálnej knižnice a model systému manažmentu digitálnej knižnice. Tieto vzťahy sú schematicky naznačené na obrázku 9.2.



- **Digitálna knižnica (DL)**
  - Organizácia, zhromažďovanie, riadenie a uchovávanie digitálneho obsahu, používateľská komunita, špecializovaná funkcionálnosť, merateľná kvalita, predpísané pravidlá
- **Systém digitálnej knižnice (DLS)**
  - softvérový systém, architektúra podľa funkcií, používateľská interakcia
- **Systém manažmentu digitálnej knižnice (DLMS)**
  - generický softvér, vytváranie a administrácia DLS
  - modularita, rozšíriteľnosť, dátový sklad, systémový generátor

Obr. 9.2 Schéma základných pojmov digitálnej knižnice

V pojmovom priestore digitálnej knižnice sa definuje šesť prvkov: obsah, používateľ, funkcionálnosť, kvalita, stratégia, architektúra.

Obsah je množinou informačných objektov organizovaných v súbore. Používateľ je všeobecný pojem pre aktérov oprávnených využívať a vytvárať nové informácie.

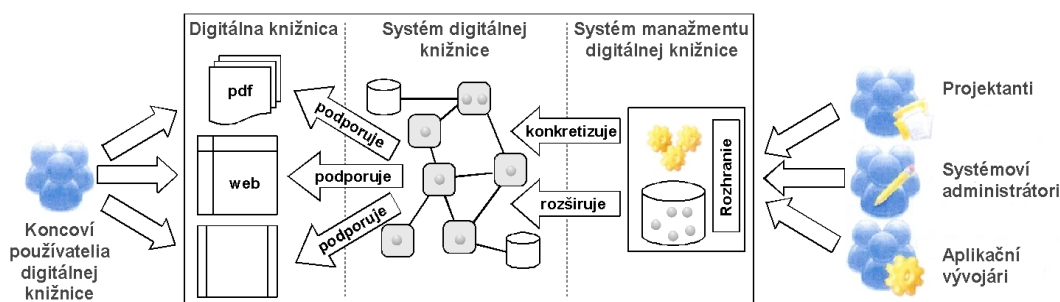
Funkcionálnosť predstavuje služby a súbor nadväzných funkcií digitálnej knižnice ako systému. Pritom možno definovať určité jadro funkcií ako napr. registrácia nového informačného objektu, vyhľadávanie, prezeranie, manažment informácií a ďalšie pomocné funkcie. Kvalita predstavuje také parametre, ktoré možno použiť na hodnotenie obsahu a správania digitálnej knižnice. Niektoré parametre možno charakterizovať ako objektívne, ktoré sa dajú merať, a iné ako subjektívne založené na experimentoch. Stratégia znamená súbor pravidiel, ktoré riadia interakciu medzi používateľmi a digitálnou knižnicou. Môže sa týkať bezpečnosti systému, systému riadenia digitálnych práv, riadenia záťaže systému a pod. Architektúra znamená organizovanie funkcionality a obsahu prostredníctvom hardvéru a softvéru digitálnej knižnice.

#### 9.4.2

### „Bohatá“ funkčnosť – rôzne prezentácie, základní aktéri

Zložky digitálnej knižnice spája princíp „bohatej“ funkčnosti. To znamená rozmanitosť informácií, množstvo interakcií a funkcií. Štruktúra informácií je bohatšia a jemnejšia s novšími nástrojmi ako napríklad taxonómie, ontológie, nástroje na vytváranie máp ako napr. pojmové mapy a tematické mapy. V niektorých systémoch sa uplatňuje paradigma adaptívneho fazetového prezerania. Umožňuje prinajmenšom tri možné pohľady na rovnaké údaje. Tieto pohľady môžu závisieť od hĺbky štruktúr informácií, pričom prepojenia (model vzťahov) sú viditeľné počas vyhľadávania.

Na úrovni aktérov v interakcii s digitálnou knižnicou sa určujú štyri hlavné roly: koncoví používatelia, projektanti, systémoví administrátori a aplikační vývojári. Koncovým používateľom môže byť ako tvorca obsahu, tak aj konzument obsahu, prípadne knihovník. Projektanti stanovujú funkčné a obsahové parametre systému. Môže pritom ísť o konfiguráciu funkcií (napríklad formát, dotazovací jazyk, používateľský profil, model dokumentu) alebo o konfiguráciu obsahu súvisiacu s externými zdrojmi. Pritom sa využívajú ontológie, klasifikačné systémy, súbory autorít, geografické slovníky. Administrátori digitálnej knižnice zabezpečujú výber softvérových zložiek a ich rozmiestnenie pri zabezpečení vhodnej úrovne kvality. Vývojári vyvíjajú jednotlivé zložky. Každá z týchto skupín používa odlišný aj spoločný slovník. Základom je súbor pojmov, ktoré rozpracúvajú ďalšie dokumenty o digitálnych knižniciach (referenčný model; The Digital Library Reference Model, 2006).



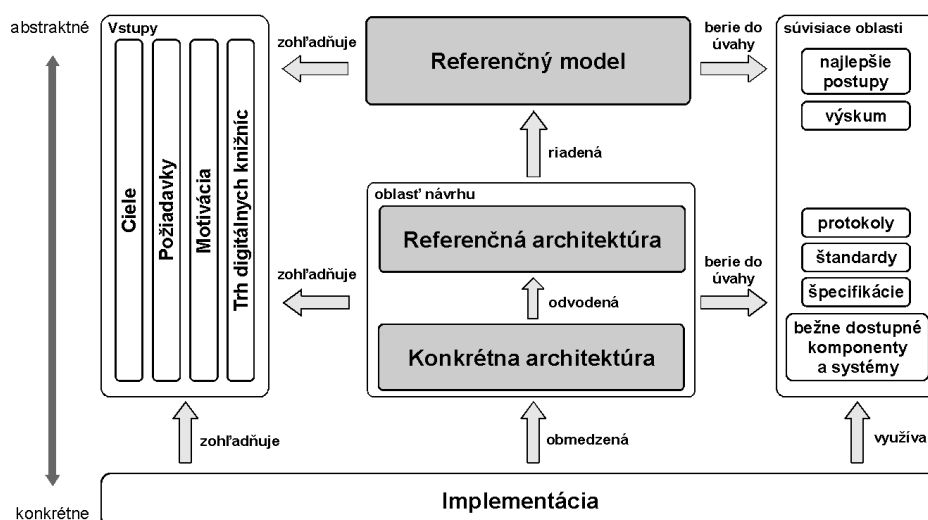
Obr. 9.3 Hlavné roly aktérov a tri vrstvy digitálnej knižnice podľa Manifestu DK

### 9.4.3 Referenčný rámec digitálnych knižníc

Dôležitou súčasťou manifestu je referenčný rámec digitálnych knižníc. Ide najmä o modelovanie základných pojmov a vzťahov medzi nimi, ktoré sú užitočné pri uvažovaní o digitálnej knižnici. Referenčný rámec sa skladá z troch zložiek: 1. referenčného modelu, 2. referenčnej architektúry a 3. konkrétnej architektúry.

1. Referenčný model je súbor vzájomne prepojených pojmov, ktoré opisujú časti ľudského poznania (napríklad vedecká oblasť). Referenčný model je definovaný ako minimálny súbor zjednocujúcich pojmov, axiómov a vzťahov v rámci určitej domény. Tieto pojmy sú nezávislé od špecifických štandardov a technológií. Sú rozpracované na vysokej úrovni abstrakcie, aby ich bolo možné pružne prispôsobiť konkrétnej implementácii. Cieľom referenčného modelu je konsolidovať rôznorodosť existujúcich prístupov do jedného celku.
2. Referenčná architektúra predstavuje vzor architektonického návrhu digitálnej knižnice na základe abstraktných pojmov a vzťahov z referenčného modelu. Je základom na implementáciu v systéme digitálnej knižnice smerom k špeciálnym aplikáciám.
3. Konkrétna architektúra zavádza dodatočné prvky na konkretizáciu referenčnej architektúry s konkrétnymi štandardmi a špecifikáciami.

Tieto zložky sú znázornené na obrázku 9.4.



Obr. 9.4 Zložky univerza digitálnej knižnice

Tieto tri zložky referenčného rámca umožňujú vývoj systému a tiež kompatibilitu a interoperabilitu konkrétnych digitálnych knižníc. Spolu s ďalšími súvisiacimi dokumentmi manifest zabezpečuje možnosť prepojenia pojmov a funkcií medzi rôznymi digitálnymi knižnicami a je otvorený pre vývoj nadväzných dokumentov.

Najčastejšie softvérové nástroje na vytváranie digitálnych knižníc a elektronických repozitárov sú D-Space, E-Prints, Fedora a i.

## 9.5 Ekologické aspekty digitálnych knižníc

Ekologické aspekty digitálnych knižníc vyplývajú z ich priameho včlenenia do prirodzenej informačnej činnosti človeka. Prioritou digitálnych knižníc nie je len prístupnosť, uchovávanie a prezentácia informácií, ale práve podpora komunikácie, interakcie a kolaborácie. V ekologickom aspekte je digitálna knižnica priestorom, ktorý sa vyvíja. Informačnými objektmi sú nielen texty alebo multimediálne formy informácií, ale aj senzorické údaje, mobilná komunikácia a samotní sociálni aktéri. Dôležitá je najmä vizuálna reprezentácia a spoločné delenie sa s informáciami medzi členmi komunity. Digitálna knižnica je tu preto „obývaný informačný priestor“ na tvorbu, manipuláciu a výmenu informácií v komunite.

Funkčné a obsahové parametre digitálnej knižnice (formát, dotazovací jazyk, používateľský profil, model relevancie) sú dynamické a možno ich prispôbovať konkrétnej situácii a potrebe. Podrobná štruktúracia informácií sa uplatňuje v nových nástrojoch organizácie poznania v elektronickom prostredí ako ontológie, taxonómie, pojmové a tematické mapy. Na to nadväzuje kontextové vyhľadávanie, adaptívne fazetové prezeranie a asociatívne algoritmy hodnotenia relevancie. Reprezentácie objektov sa približujú mentálnym reprezentáciám v ľudskej pamäti. Digitálne knižnice sú preto typickým produktom vznikajúcej „vedy 2.0“, ktorá integruje technologické inovácie so sociálnymi systémami (Schneiderman 2009). Veda 2.0 bude skúmať najmä vzťahy ľudí v sociálnych sieťach v informačných priestoroch. Za ekologické aspekty tu možno považovať skúmanie otázok dôvery k zdrojom a aktérom, empatie, zodpovednosti a súkromia. Ak sú interakcie zmysluplné, zodpovedné, podložené poznatkami a skúsenosťou, potom z nich vznikajú ekologické funkcie digitálnych knižníc, ktoré podporujú prirodzené ľudské informačné aktivity.

Informačná ekológia umožňuje vysvetliť fungovanie digitálnej knižnice metaforami nástroja, textu (obsahu komunikácie) a systému. Digitálne knižnice sú pre sociálneho aktéra nástrojom na spracovanie a využívanie informácií. Ako obsah komunikácie umožňujú konštrukciu významu v rôznych situáciách sprostredkovaných interakciami. Informačnou aktivitou je pritom ako formálna, tak aj neformálna komunikácia v komunite. Digitálne knižnice sú sociálno-technickým systémom, v ktorom sa ľudia a technológie navzájom prispôbujú. Ak sa doteraz vo väčšej miere ľudia prispôbovali technológiám, v ekologickom prístupe by sa mali technológie viac prispôbiť potrebám a správaniu človeka.

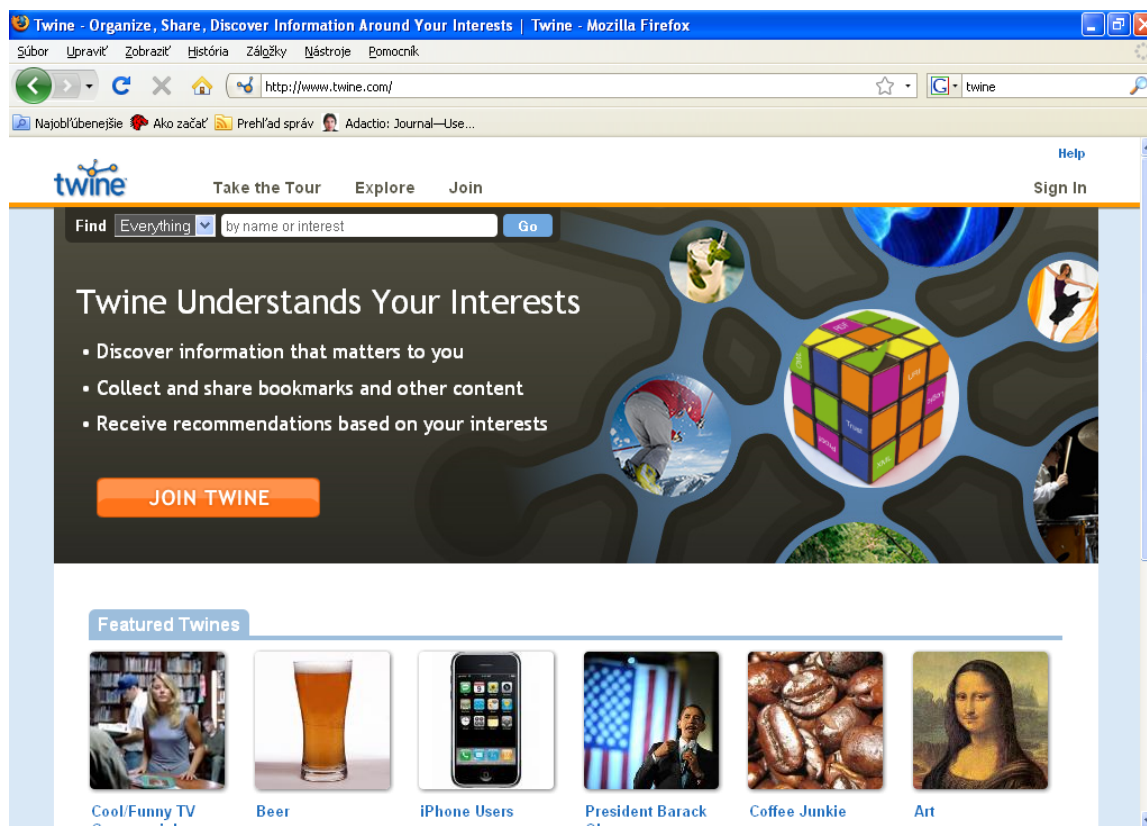
Základné funkcie vedeckej komunikácie sa transformujú aj do nového prostredia digitálnych knižníc (Sompel et al. 2004). Ide o registráciu výsledkov (zabezpečenie možnosti nadväznosti vedeckej práce), certifikáciu (priznanie platnosti a spoľahlivosti výsledkov), zabezpečovanie aktuálnosti poznania (informovanie o nových aktuálnych výsledkoch), archiváciu (uchovávanie poznatkov v čase), ocenenie (uznanie výsledkov a poznatkov aktérov vedeckej komunikácie). Tieto funkcie sa v digitálnom prostredí modelujú v repozitároch a digitálnych knižniciach.

Komplexný prístup k repozitárom pre výskum predstavuje európska koncepcia DRIVER (2009). Napríklad arXiv.org fungujúci od roku 1991, je reprezentantom vedeckej ekológie, v ktorej sa modelovo riešia toky informačných objektov od ich vzniku, cez sprístupňovanie a publikovanie až po archivovanie. Príkladmi špecializovaných vedeckých repozitárov sú aj CogPrints (kognitívne vedy) alebo Research Papers in Economics (RePec). V novších modeloch sa obsah na vstupe do digitálnej

knižnice rozloží a následne integruje v rôznych službách a funkciách ako vzdelávanie, sprostredkovanie, publikovanie, učenie alebo výskum (napr. projekt ORE – Object Re-Use and Exchange).

Informačná ekológia digitálnych knižníc môže podporiť premenu vedeckej komunikácie a kolaborácie. V ekologických digitálnych knižniciach dominuje pojmová infraštruktúra, napríklad pojmové mapy alebo ontológie a ich vizualizácia, ktoré umožňujú nielen efektívne vyhľadávanie, ale aj analyzovanie údajov a textov a ich premenu na poznatky. Digitálne knižnice sú tak súčasťou informačnej infraštruktúry a koncepcie „e-Science“ či „digitálnej vedy“ (Borgman 2007). Humanitní vedci tu môžu analyzovať využitie jazyka a originálne myšlienky. Historici vytvárajú viacdimenzionálne modely pamiatok, objektov, udalostí a súvislostí. Prírodní vedci integrujú údaje o molekulách, spektrách, chemických štruktúrach a vytvárajú modely. Hudobníci analyzujú témy, notové štruktúry a melódie. Sociálne vedy sa môžu orientovať na vyvodzovanie trendov v čase. Nové pojmové štruktúry možno vytvárať automaticky a porovnávať slovníky medzi rôznymi disciplínami.

Informačnú ekológiu digitálnych knižníc vidíme aj vo svetle trendov vytvárania nových verzií webu (web 2.0, web 3.0). Tento trend možno vysvetliť ako prechod od budovania zdrojov (stránok) k ich prepájaniu v sieťach a spolupráci (web 2.0). Web 3.0 súvisí so sémantickým webom, orientovaným na pojmové štruktúry a ich siete. Obsahuje aj nové nástroje organizovania, hodnotenia, filtrovania, personalizácie až po vyvodzovanie nových poznatkov zo štruktúr pojmov a tém (web 4.0). Vizionári predvídajú spoločné a zdieľané reprezentácie, ktoré by mali umožňovať kolektívnu skúsenosť a inteligentné využívanie informácií. Príkladom môže byť portál Twine ako online sociálna webová služba, ktorá integruje uchovávanie, tvorbu a objavovanie informácií (obrázok 9.5). Kombinuje funkcie digitálnych knižníc s funkciami wiki, fór, online databáz a aplikuje inteligentné programy na automatizované analýzy a dolovanie informácií a poznatkov.



Obr. 9.5 Sociálna webová služba Twine

## 9.6

### Zložky informačnej ekológie digitálnych knižníc

Na základe modelu informačného produktu (Steinerová 1998) určujeme v digitálnych knižniciach tieto zložky ich informačnej ekológie: poznatkovú, komunikatívnu, hodnotovú, významovú (reprezentatívnu) a technologickú zložku.

*Poznatkovú zložku* predstavujú informačné (digitálne) objekty a vzťahy medzi nimi. Informačné objekty sú organizované na rôznych stupňoch detailnosti štruktúry a prejavujú sa v pojmových štruktúrach, napríklad mapa pojmov a ich vzťahov alebo mapa domén. Ekologické digitálne knižnice poskytujú komplexnejšiu podporu kognitívnych aktivít používateľa ako filtrovanie, analyzovanie, interpretovanie, organizovanie a i. Mali by rešpektovať aj emócie pri selekcii a filtrovaní informácií.

*Komunikatívna zložka* vyplýva z možnosti vykonávať komunikatívne aktivity v prostredí digitálnej knižnice. V oblasti vedy a vzdelávania sa prezentuje a vytvára vedecký obsah. Komunikačné a informačné štýly študentov, učiteľov, vedcov a výskumníkov sa premietajú v preferenciách výberu formy a obsahu. Súčasťou komunikatívnej zložky sú aj informačné stratégie vyhľadávania informácií, napríklad kontextové a fazetové vyhľadávanie a navigácia používateľa. Komunikatívna zložka umožňuje také sociálne komunikačné aktivity v elektronickom prostredí ako dialóg, diskusie, „zdieľanie“ zdrojov, komentovanie, recenzovanie. V tejto zložke funguje sociálny softvér, ktorý podporuje rôzne stupne participatívnych aktivít – od blogov ku kolaboratívnemu písaniu. Funkcie sa zameriavajú na oznamovanie, triedenie informácií, učenie, vzdelávanie aj zábavu.

*Hodnotová zložka* súvisí so systémom hodnôt pri spracovaní informácií. Hodnotový systém založený na afektívnom informačnom správaní možno považovať za najdôležitejší ekologický aspekt. Integruje postupy a vzťahy v lokálnom prostredí. Digitálna knižnica funguje aj ako nástroj na vytváranie pridanej hodnoty, od opisných metadát cez analyzované a syntetizované informácie až po poznatky. Pridanou hodnotou sú aj nové možnosti simulácie objektov, udalostí a analýz údajov. Vo vedeckej komunikácii v elektronickom prostredí sa nanovo modeluje „hodnotová reťaz“ pri hodnotení kvality, recenzovaní a distribúcii výsledkov. Významovú zložku predstavujú najmä reprezentácie poznania, ktoré sú spojené s kontextom. Oddelenie reprezentácie objektu od jeho prezentácie umožňuje prispôbiť konkrétny objekt konkrétnej situácii. Napríklad štúdium originálu maľby možno upraviť podľa situácie – buď z hľadiska vnímania štruktúry alebo z hľadiska vnímania celku. Textové, číselné, auditívne, vizuálne a iné reprezentácie podporujú usudzovanie človeka, riešenie problémov, rozhodovanie.

*Technologická zložka* digitálnych knižníc v súčasnosti predstavuje „všadeprítomné“ elektronické prostredie. Vo fyzickej knižnici sa ľudia pohybujú v prostredí, triedia, vyberajú, manipulujú s dokumentmi. Podobné interakcie sa odohrávajú aj v digitálnej knižnici a ekologicky sa prispôbujú prostrediu. Technológie umožňujú objavovanie súvislostí a organizovanie vlastností digitálnych objektov. Oproti tradičnému informačnému prieskumu technológie podporujú vzdelávanie a výskum tak, že umožňujú mapovať témy, objavovať nové, vytvárať experimentálne analýzy, zoskupovať a triediť informácie, manipulovať s pojmami a modelmi. Inteligentné prostredie naviguje a podporuje intelektuálne aktivity.

## 9.7

### Trendy vývoja digitálnych knižníc a informačného prieskumu

Jedným z trendov je postupná premena digitálnych knižníc na tzv. „knowledge commons“. Možno ich vysvetliť ako spoločné poznatkové priestory informačných zdrojov. Ide o také prostredia, ktoré sa stanú centrom intelektuálnej činnosti, uľahčia komunikáciu a kolaboráciu medzi vedcami alebo inými skupinami ľudí. Mali by syntetizovať distribuované multimediálne dokumenty, sensorické

údaje a iné zdroje bez bariér prístupu k obsahu. Majú byť orientované na človeka s cieľom poskytovať používateľom personalizované zážitky a množstvo univerzálnych aj špecializovaných funkcií pri využívaní obsahu.

Digitálna knižnica ako pojem sa rozširuje k funkciám využívajúcim také vlastnosti elektronického prostredia, ktoré pri tradičných procesoch neexistovali. Preto je víziou digitálnej knižnice do budúcnosti koncepcia „obývaného informačného priestoru“. Technologicky podmienená kooperácia (computer-supported cooperative work) podporuje stretanie sa informácií a ľudí v elektronickom priestore.

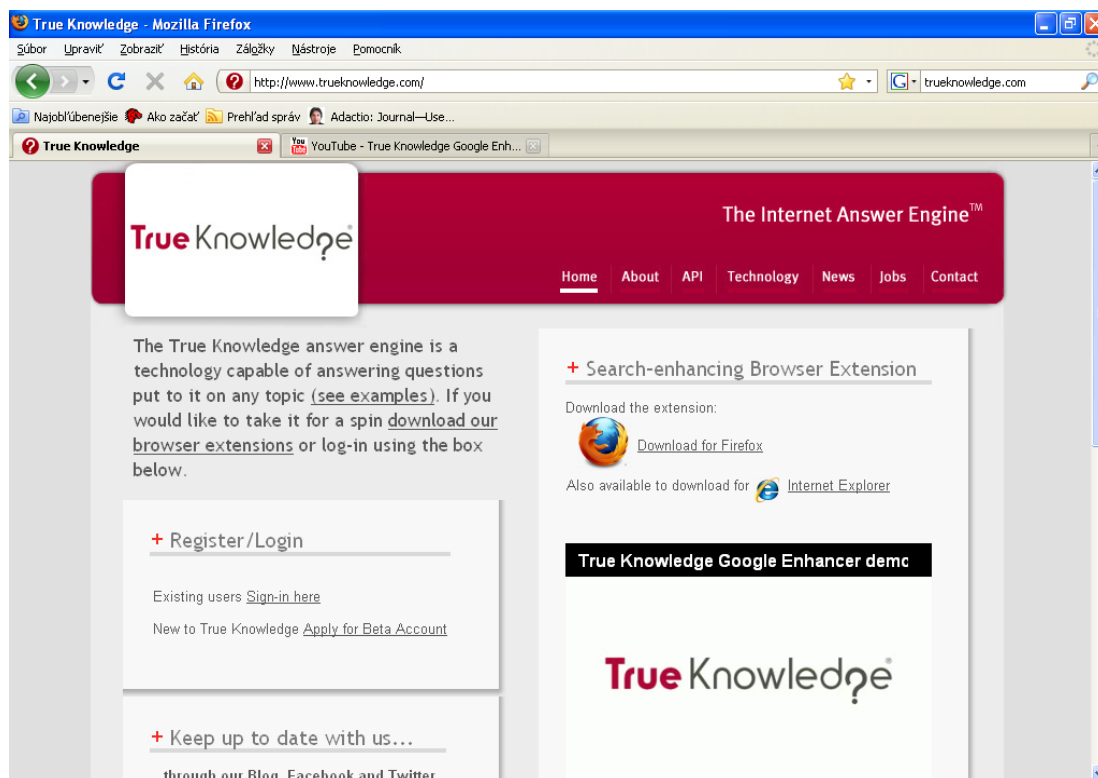
Digitálna knižnica v novej koncepcii umožňuje kolaboráciu a vykonávanie rôznych činností s informačnými objektmi, vzájomné informovanie sa o svojich aktivitách v digitálnom prostredí, poskytovanie rôznych pohľadov na informácie v kontexte ich použitia. Dôležitá je aj vizualizácia a vzájomné delenie sa s informáciami medzi členmi komunity používateľov. Víziou pre budúcnosť sú „prepojené komunity“ ako efektívne sociálne pracovné priestory umožňujúce tvorbu, výmenu a efektívnu vizualizáciu digitálnych informácií.

K trendom vývoja vyhľadávania informácií patrí:

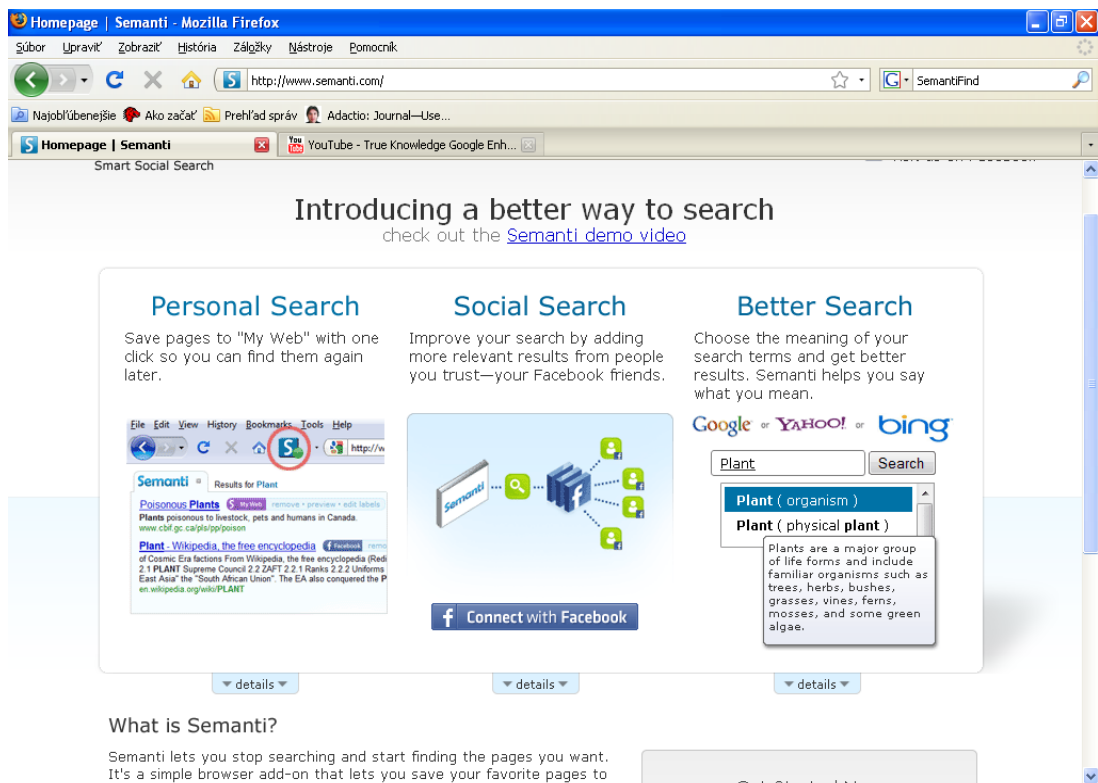
1. Personalizácia
2. Vizualizácia
3. Kontextualizácia
4. Participácia v komunitách.

Dôležitým trendom je aj tzv. sémantický prieskum. Nové prieskumové nástroje prostredníctvom kombinácie nových nástrojov organizácie poznania a práce s témami umožňujú objavovanie faktov a odkrývanie významov. Sémanticky orientované aplikácie využívajú novšie nástroje na kategorizáciu a klasifikáciu a inteligentné systémy sociálnych preojení. K takýmto nástrojom možno priradiť napríklad True Knowledge, Powerset alebo SemantiFind.

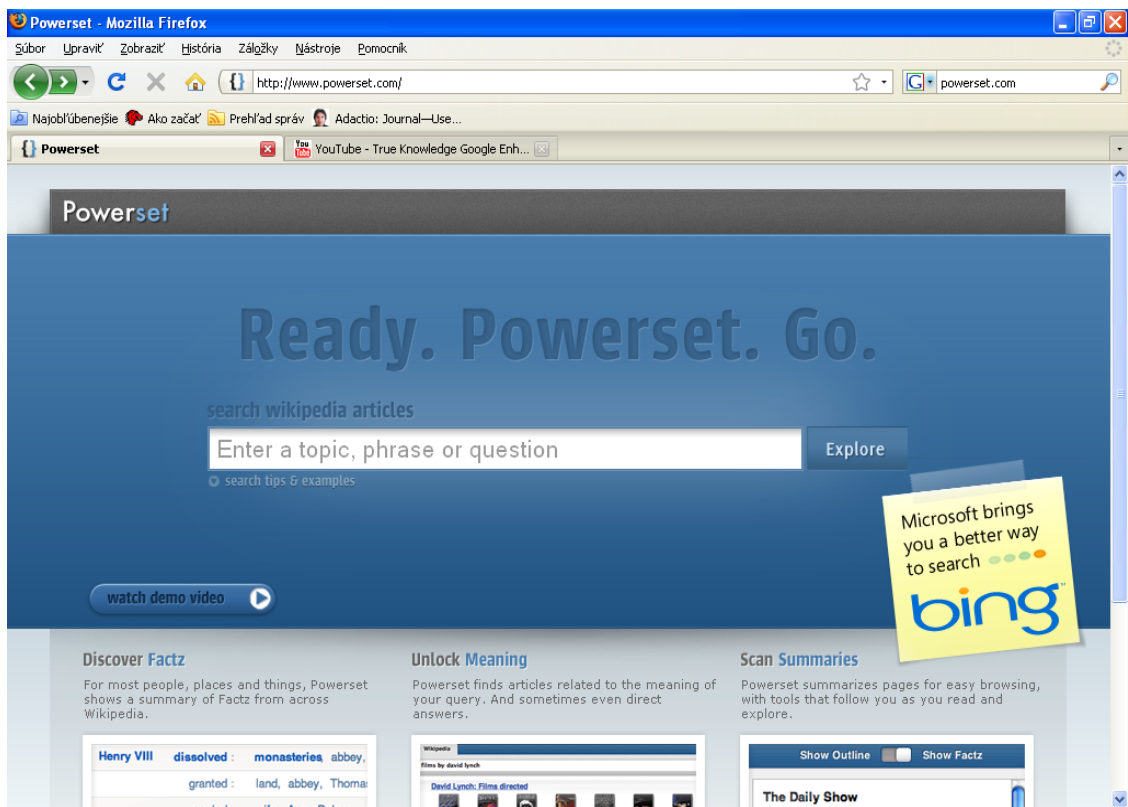
Na obrázkoch 9.6, 9.7 a 9.8 sú príklady týchto nástrojov.



Obr. 9.6 TrueKnowledge, internetový nástroj na odpovedanie

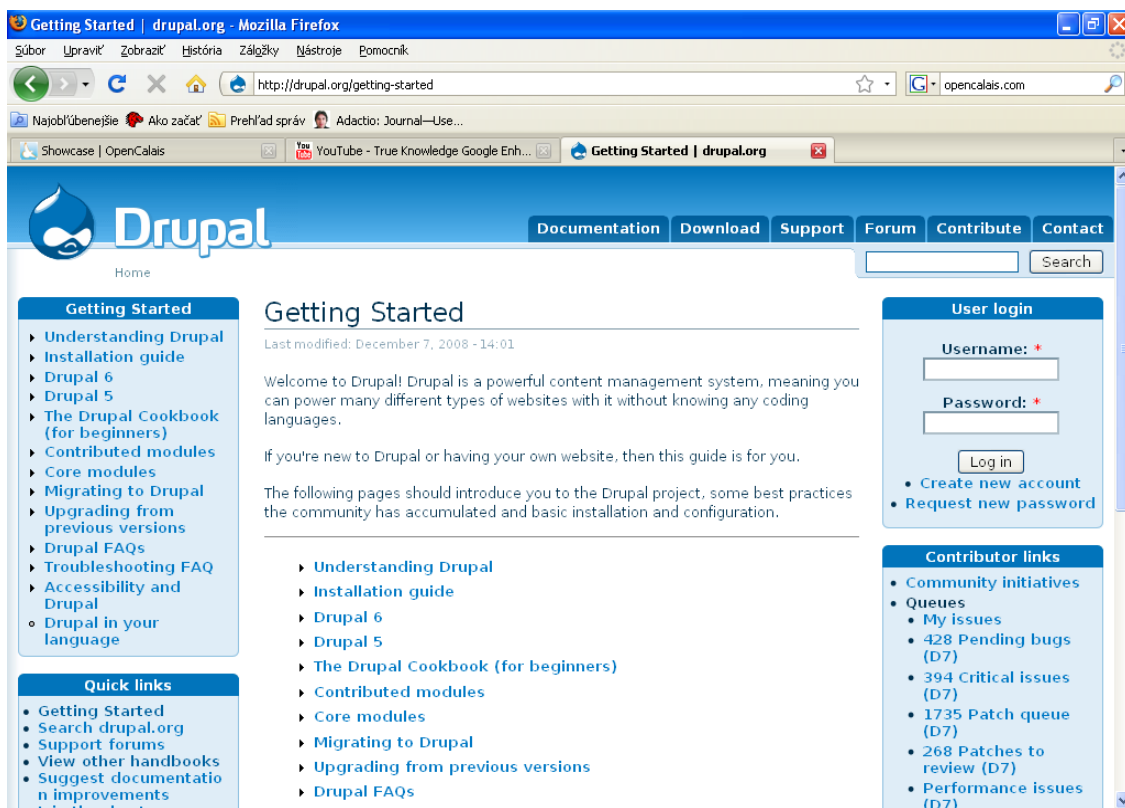


Obr. 9.7 Nástroj sémantického prieskumu SemantiFind



Obr. 9.8 Nástroj sémantického prieskumu Powerset

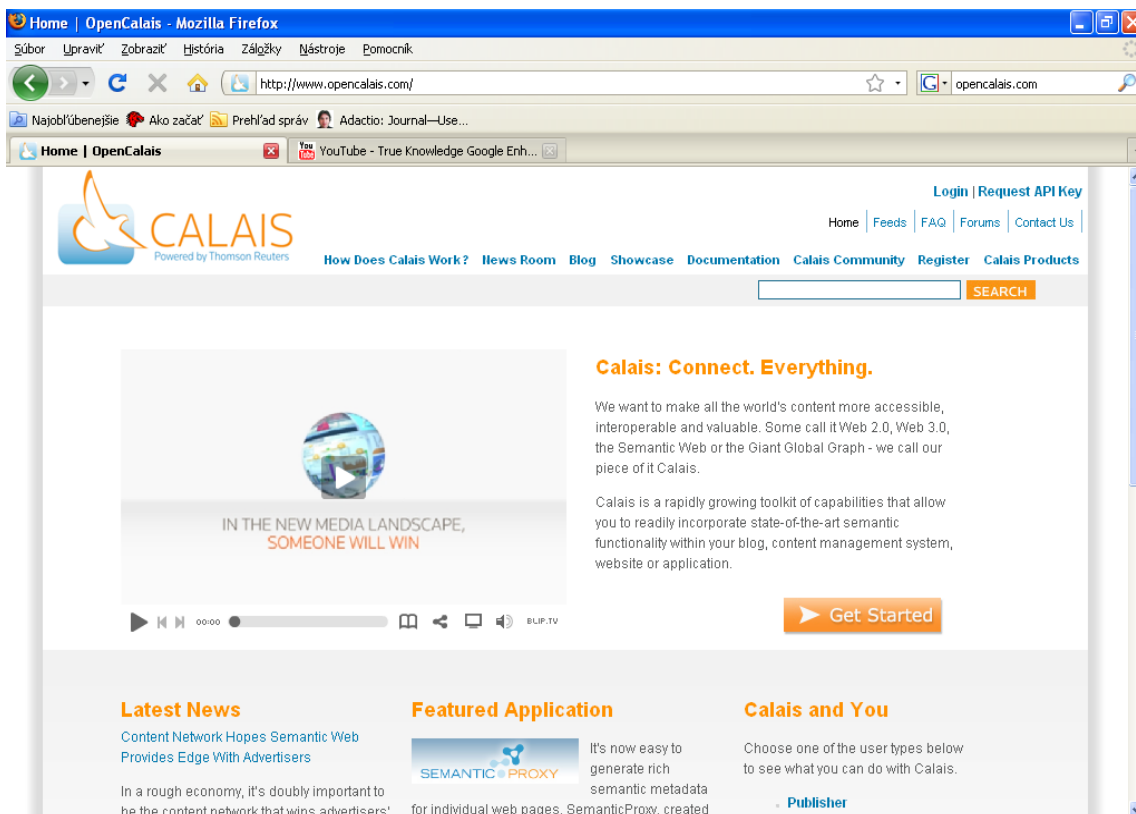
Trendom je aj integrácia prieskumových systémov a systémov na manažment obsahu a redakčných systémov. Príklad takéhoto systému je na obrázku 9.9 (systém Drupal).



Obr. 9.9 Systém CMS Drupal

Nástroje na spracovanie inteligentných objektov („smart objects“) predstavujú ďalší trend vývoja digitálnych knižníc a vyhľadávania. Sémantické algoritmy prepájajú pojmy a sociálne siete. Na obrázku 9.10 je príklad systému Calais využívajúci sociálne siete pri vyhľadávaní a riešení problémov.





Obr. 9.10 Systém Calais

Nové služby webu a digitálnych knižníc majú pomáhať nielen hľadať a využívať informácie, ale aj podporovať poznávanie a sociálne vzťahy. Informačná ekológia poskytuje novší rámec na vysvetlenie informačného prostredie človeka. Technológie sa musia viac prispôbovať potrebám človeka vrátane emócií, nálad či motivácie informačných aktivít.

V koncepte informačnej ekológie zdôrazňujeme „čistotu“ informačného prostredia sprostredkovanú inteligentnými nástrojmi kolektívnej pamäte a nástrojmi organizácie informácií.

Informačná veda musí sledovať trendy informačného správania v digitálnych knižniciach. V rozhraniach digitálnych knižníc sa budú tieto trendy odrzkaďovať vo funkciách ako okamžité použitie informácií, jednoduchosť, nelineárnosť a sociálnosť.

## 9.8 Zhrnutie

Teória a praktické projekty digitálnych knižníc sa intenzívne rozvíjajú v posledných 15 rokoch. Známe sú najmä americké iniciatívy na výskum digitálnych knižníc (DLI) a európske projekty DELOS). Digitálne knižnice sú zložitú sociotechnické systémy. Sú priestormi na ukladanie digitálnych objektov aj na komunikovanie a riešenie informačných potrieb komúnit. V súčasnosti sa v nich zdôrazňuje spolupráca, interakcia a aktivity človeka. Digitálne knižnice sú súčasťou vedeckej práce (e-science), ale aj informačných inštitúcií v elektronickom prostredí (knižnice, múzeá, galérie, školy, kancelárie). Problémy digitálnych knižníc súvisia s výberom digitálnych dokumentov a spôsobmi digitalizácie, identifikáciou častí digitálnych objektov, ich spracovaním prostredníctvom metadáto- vých štruktúr a dlhodobou archiváciou zdrojov.

Podľa Manifestu digitálnych knižníc sa digitálne knižnice delia na tri zložky: samotnú digitálnu knižnicu, systém digitálnej knižnice, systém manažmentu digitálnej knižnice.

Digitálne knižnice sú súčasťou informačnej ekológie, pretože prepájajú obsah, systém a sociálne komunity. V súlade s vedou 2.0 sa v nich prepájajú technologické inovácie so sociálnymi vzťahmi. Informačná ekológia digitálnych knižníc obsahuje poznatkovú, komunikatívnu, hodnotovú, významovú a technologickú zložku.

Medzi trendy vývoja digitálnych knižníc patrí personalizácia, kontextualizácia, sémantický prieskum, inteligentné nástroje interakcie a sociálna participácia v sieťach.

## Literatúra

- BEKIARI, Chryssoula – CONSTANTOPOULOS, Panos – DOERR, Martin. 2006. Information Patterns for Digital Cultural Repositories. In *ERCIM News*. July 2006, č. 66, s. 31 – 32.
- BORGMAN, Christine. 2007. *Scholarship in the Digital Age : Information, Infrastructure and the internet*. London : MIT Press, 2007. 336 s. ISBN 978-0-262-02619-2.
- CANDELA, L. et al. 2006. *The Digital Library Manifesto*. DELOS (Network of Excellence on Digital Libraries), 2006. 20 s. Thematic Priority: IST-2002-2.3.1.12. Technology-enhanced Learning and Access to Cultural Heritage.
- CHOO, Chun Wei. 2006. *The Knowing Organization : How Organizations Use Information to Construct Meaning, Create Knowledge, and Make Decisions*. Sec. ed. New York : Oxford University Press, 2006. 354 s. ISBN 0-19-517678-2.
- CONNAWAY, Lynn Silipigni. 2008. Make Room for the Millenials. In *NextSpace*. ISSN 1559-0011, 2008, no. 10, s. 18 – 19.
- DAVENPORT, Thomas H., PRUSAK, Laurence. 1997. *Information Ecology : Mastering the Information and Knowledge Environment*. New York : Oxford University Press, 1997. 255 s. ISBN 0-19-511168-0.
- DERVIN, B., REINHARD, C. D. 2007. How Emotional Dimensions of Situated Information Seeking Relate to User Evaluations of Help from Sources : An Exemplar Study Informed by sense-Making Methodology. In *Information and Emotion : the Emergent Affective Paradigm in Information Behavior Research and Theory*. Medford : Information Today, 2007, s. 51 – 84.
- FAST, Karl V., SEDIG, Kamran. 2005. The INVENT framework : Examining the role of information visualization in the reconceptualization of digital libraries. In *Journal of Digital Information* [online]. 2005, vol. 6, issue 3, article no. 362 [cit. 2009-02-08]. Dostupné na internete: <<http://journals.tdl.org/jodi/article/view/66/69>>.
- GIVEN, Lisa M. 2007. Emotional Entanglements on the University Campus : The Role of Affect in Undergraduates' Information Behaviors. In *Information and Emotion : the Emergent Affective Paradigm in Information Behavior Research and Theory*. Medford : Information Today, 2007, s. 161 – 175.
- GREŠKOVÁ, Mirka. 2008. *Kognitívne základy informačnej vedy : interakcia človeka s agentom na vyhl'adávanie informácií* [Dizertačná práca]. Univerzita Komenského v Bratislave; Filozofická fakulta; Katedra knižničnej a informačnej vedy. Bratislava : FiFUK, 2008. 139 s.
- INGWERSEN, Peter – JÄRVELIN, Kalervo. 2005. *The Turn : Integration of Information Seeking and Retrieval in Context*. Dordrecht : Springer, 2005. 448 s. ISBN 10 1-4020-3850-X.

- ISO 21127:2006, *Information and documentation – A reference ontology for the interchange of cultural heritage information*. Možnosť zakúpiť online na internete:  
<[http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=34424](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=34424)>.
- KUHLTHAU, Carol C. 1993. *Seeking Meaning : A Process Approach To Library And Information Services*. Norwood : Ablex, 1993. 199 s.
- LAMB, R., KLING, R. 2003. Reconceptualizing users as social actors in information system research. In *MIS Quarterly*. Special Issue. June 2003, vol. 27, no. 2, s. 197 – 235.
- LYNCH, Clifford. 2008. The Institutional Challenges of Cyberinfrastructure and e-Research. In *EDUCAUSE Review* [online]. Nov./Dec. 2009, vol. 43, no. 6 [cit. 2009-03-16]. Dostupné na internete:  
<<http://connect.educause.edu/Library/EDUCAUSE+Review/TheInstitutionalChallenge/47446>>.
- NAHL, D. 2007a. A discourse analysis technique for charting the flow of micro-information behaviour. In *J. Doc.* 2007, vol. 63, no. 3, s. 323 – 339.
- NAHL, D. 2007b. The Centrality of the Affective in Information Behavior. In NAHL, Diane – BILAL, Dania. *Information and Emotion: the Emergent Affective Paradigm in Information Behavior Research and Theory*. Medford : Information Today, 2007. ISBN 978-1-57387-310-9, s. 3 – 37.
- NAHL, Diane – BILAL, Dania. 2007. *Information and Emotion : the emergent affective paradigm in information behavior research and theory*. Medford : Information Today, 2007. 359 s. ISBN 978-1-57387-310-9.
- NARDI, B. A., O'DAY, V. L. 1999. *Information Ecologies : Using Technology with Heart*. Cambridge : MIT Press, 1999. 231 s. ISBN 0-262-14066-7.
- NICHOLAS, David et al. 2003 Digital information consumers, players, and purchasers : data and thoughts on information users and use in the new digital interactive environment. In *Aslib Proceedings* [online]. 2003, vol. 55, issue 1/2 [cit. 2009-02-08]. s. 23-31. Dostupné na internete pre predplatiteľov:  
<<http://www.emeraldinsight.com/Insight/viewContentItem.do?jsessionid=DF92261A337D9C3A9ED0D0F5D9306D08?contentType=Article&hdAction=lnkhtml&contentId=863975>>. ISSN: 0001-253X.
- PRABHA, Ch. et al. 2007. What is enough? : Satisficing information needs. In *Journal of Documentation* [online]. 2007, vol. 63, no. 1 [cit. 2009-02-08]. s. 74 – 89. Pre-print. Dostupné na internete:  
<<http://www.oclc.org/research/publications/archive/2007/prabha-satisficing.pdf>>. ISSN 0022-0418.
- RADFORD, Marie L., CONNAWAY, Lynn Silipigni. 2005. *Seeking Synchronicity : Evaluating Virtual Reference Services from user, Non-user, and Librarians perspectives. A research project funded by the National Leadership Grants for Libraries program of the Institute of Museum and Library Services (IMLS)* [online]. Rutgers, The State University of New Jersey and OCLC Online Computer Library Centre, Inc [cit. 2009-02-08]. Dostupné na internete: <<http://www.oclc.org/research/projects/synchronicity>>.
- ROWLANDS, I., FIELDHOUSE, M. 2008. *Information behaviour of the researcher of the future : a ciber briefing paper* [online]. University College of London : Centre for Information Behaviour Research, January 2008 [cit. 2009-03-17]. 35 s. Dostupné na internete:  
<<http://www.ucl.ac.uk/slais/research/ciber/downloads/ggexecutive.pdf>>.
- SCHMITZ, Dawn. 2008. *The Seamless Cyberinfrastructure : The Challenges of Studying Users of Mass Digitization and Institutional Repositories* [online]. Washington, D.C. : Council on Library and Information Resources, 2008 [cit. 2008-03-28]. Dostupné na internete:  
<<http://www.clir.org/pubs/archives/schmitz.pdf>>.
- SCHNEIDERMAN, Ben. 2008. Science 2.0. In *Science* [online]. March 2008, vol. 319 [cit. 2009-02-08]. S. 1349 – 1350. Dostupné na internete: <<http://www.sciencemag.org>>.

- SMITH, A. 2003. *New Model Scholarship : How Will it Survive?* [online]. Washington, D.C. : Council on Library and Information Resources, 2003 [cit. 2007-05-15]. Dostupné na internete: <<http://www.clir.org/pubs/reports/pub114/pub114.pdf>>. ISBN 1-887334-99-8.
- SOMPEL, H. van de et al. 2004. Rethinking Scholarly Communication : Building the System that Scholars Deserve. In *D-Lib Magazine* [online]. 2004, vol. 10, no. 9 [cit. 2009-02-08]. Dostupné na internete: <<http://www.dlib.org/dlib/september04/vandesompel/09vandesompel.html>>. ISSN 1082-9873.
- STEINEROVÁ, J. 1996. *Teória informačného prieskumu*. Bratislava : CVTI SR, 1996. 262 s. ISBN 80-65165-58-9.
- STEINEROVÁ, Jela. 1998a. *Tvorba informačných produktov*. Bratislava : CVTI SR, 1998. 130 s. ISBN 90-85165-73-2.
- STEINEROVÁ, Jela a kol. 2004. *Správa o empirickom prieskume používateľov knižníc ako súčasť grantovej úlohy VEGA 1/9236/02 Interakcia človeka s informačným prostredím v informačnej spoločnosti*. Bratislava : Filozofická fakulta UK, KKIV, 2004. 113 s.
- STEINEROVÁ, Jela. 2005a. *Informačné správanie : Pohľady informačnej vedy*. Bratislava : CVTI SR, 2005. 189 s. ISBN 80-85165-90-2.
- STEINEROVÁ, Jela. 2007a. Manifest o digitálnych knižniciach a využívanie informácií v informačnej spoločnosti. In *Ilib*. ISSN 1335-793X, 2007, roč. 11, č. 2, s. 15 – 19.
- STEINEROVÁ, Jela. 2007f. Relevance assessment for digital libraries. In *Mousaion*. ISSN 0027-2639, 2007, vol. 25, no. 2, s. 37 – 57.
- STEINEROVÁ, Jela – ŠUŠOL, Jaroslav – GREŠKOVÁ, Mirka. 2006. Information behaviour in relevance judgements. In *Využívanie informácií v informačnej spoločnosti*. Zborník z medzinárodnej konferencie. Bratislava, SR, 10.-11. 10. 2006. Bratislava : CVTI SR, 2006, s. 29 – 40.
- STEINEROVÁ, J., GREŠKOVÁ, M., ŠUŠOL, J. 2007. *Prieskum relevancie informácií : Výsledky rozhovorov s doktorandmi FiFUK*. Bratislava : CVTI SR, 2007. 150 s.
- STEINEROVÁ, J., ŠUŠOL, J. 2007. Users' information behaviour – a gender perspective. In *Information Research* [online]. 2007, vol. 11, no. 3, paper 251 [cit. 2010-01-10]. Dostupné na internete: <<http://InformationR.net/ir/11-3/paper251.html>>.
- STEINEROVÁ, Jela. 2008b. Seeking relevance in the academic information use. In *Information Research* [online]. 2008, vol. 13, no. 4, paper 380 [cit. 2010-01-10]. Dostupné na internete: <<http://informationr.net/ir/13-4/paper380>>.
- ŠUŠOL, Jaroslav. 2003. *Elektronická komunikácia vo vede*. Bratislava : CVTI SR, 2003. 156 s. ISBN 80-85165-88-0.
- TALJA, S., HARTEL, J. 2007. Revisiting the user-centered turn in information science research : an intellectual history perspective. In *Information Research* [online]. 2007, vol. 12, no. 4, paper colis04 [cit. 2010-01-10]. Dostupné na internete: <<http://InformationR.net/ir/12-4/colis/colis04.html>>.
- WILLIAMSON, K. 2005. Ecological Theory of Human Information Behavior. In Eds. FISHER, Karen E. – ERDELEZ, Sanda – McKECHNIE, Lynne E. F. *Theories of Information Behavior*. Medford : Information Today, 2005. ISBN 1-57387-230-X, s. 128 – 132.

## Zoznam webových sídel

Alexandria Digital Library  
<http://www.alexandria.ucsb.edu/>

arXiv.org  
<http://arxiv.org/>

Calais  
<http://www.opencalais.com/>

CIDOC Conceptual Reference Model  
<http://cidoc.ics.forth.gr/>

CIESIN (Americké centrum údajov pre interakcie človeka s prostredím)  
<http://sedac.ciesin.columbia.edu/wdc/#>

Collaborative Information Universe at Indiana University  
<http://iuni.slis.indiana.edu/>

DELOS (A Reference Model for Digital Library Management System)  
<http://www.delos.info/ReferenceModel/>

Digital Library Federation  
<http://www.diglib.org/>

DRIVER (Digital Repository Infrastructure Vision for European Research)  
<http://www.driver-repository.eu/>

Drupal  
<http://drupal.sk/>

D-Space  
<http://www.dspace.org/>

E-Prints  
<http://www.eprints.org/>

Europeana  
<http://www.europeana.eu/portal/>

Fedora  
<http://fedoraproject.org/>

Human Genome Project Information  
[http://www.ornl.gov/sci/techresources/Human\\_Genome/home.shtml](http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml)

Jednotná informační brána (JIB)  
<http://info.jib.cz/>

National e-Science Centre  
<http://www.nesc.ac.uk/>

National Science Digital Library (NSDL)  
<http://nsdl.org/>

National Virtual Observatory  
<http://www.us-vo.org/>

OCLC  
<http://www.oclc.org/reports/sharing/default.htm>

Open Archives Initiative (ORE – Object Re-Use and Exchange)  
<http://www.openarchives.org/ore>

Oxford e-Research Centre  
<http://www.oerc.ox.ac.uk/>

Perseus Digital Library  
<http://www.perseus.tufts.edu/hopper/>

Powerset  
<http://www.powerset.com>

SemantiFind  
<http://www.semanti.com/>

The European Library  
<http://www.theeuropeanlibrary.org/>

True Knowledge  
<http://trueknowledge.com>

Twine  
<http://www.twine.com/>

World Digital Library  
<http://www.wdl.org/en/>

## Zoznam ilustrácií

Obr. 1.1 Vzťahy pojmov informačné správanie, vyhľadávanie informácií a informačný prieskum.....	10
Obr. 1.2 Situácia základných zložiek pri vyhľadávaní informácií/informačnom prieskume.....	11
Obr. 1.3 Model informačného prieskumového procesu podľa Kuhlthau (Ingwersen, Järvelin 2005).....	14
Obr. 1.4 Štádiá vyhľadávania informácií (Ellis – Kuhlthau) (podľa Wilson, 1999).....	14
Obr. 1.5 Vývoj vzťahu medzi mechanizmami vyhľadávania a interakciou systému a človeka.....	16
Obr. 1.6 Interaktívne vyhľadávanie informácií (podľa Ingwersen, Järvelin 2005).....	18
Obr. 1.7 Saracevicov model rôznych vrstiev interakcií pri vyhľadávaní a využívaní informácií.....	20
Obr. 1.8 Epizodický model N. Belkina (podľa Belkin 1996).....	21
Obr. 1.9 Model interaktívneho informačného prieskumu ako multitaskingu.....	21
Obr. 2.1 Príklady informačných horizontov (študenti KKIV).....	26
Obr. 2.2 Informačný horizont (podľa Sonnenwald, D., Wildemuth, B., Harmon, G. L. 2001).....	26
Obr. 2.3 Horizont informačných zdrojov (podľa Savolainen 2008).....	27
Obr. 2.4 Rámec informačných postupov (podľa Savolainen 2008).....	29
Obr. 2.5 Grafické znázornenie prípadov formulácie dotazov v Booleovom modeli.....	31
Obr. 2.6 Vektorový priestorový model (podľa Grossman Frieder 2004).....	32
Obr. 2.7a, 2.7b Príklady invertovaného indexu (podľa Grossman, Frieder 2004).....	33
Obr. 2.8 Inferenčná sieť (podľa Grossman, Frieder 2004).....	34
Obr. 2.9 Posun vektoru dotazu ako reakcia na spätnú väzbu, aplikácia Rocchiovho algoritmu.....	35
Obr. 2.10 Zhlukovanie (podľa Manning et al. 2009).....	35
Obr. 2.11 Príklad tezauru (podľa Aitchison, Gilchrist, Bawden 1987 – 1997).....	37
Obr. 2.12 Príklad automatického vytvárania tezauru (Manning et al. 2009).....	38
Obr. 2.13 Model vyhľadávania informácií M. Batesovej – berrypicking/zbieranie lesných plodov.....	40
Obr. 2.14 Všeobecné vyhľadávanie, široká formulácia dotazu.....	41
Obr. 2.15 Špecifické vyhľadávanie, úzka formulácia dotazu.....	41
Obr. 3.1 SCOPUS – možnosti vyhľadávania.....	49
Obr. 3.2 SCOPUS – výsledky vyhľadávania.....	49
Obr. 3.3 SCOPUS – prehľad spracovávaných zdrojových dokumentov.....	50
Obr. 3.4 SCOPUS – výsledky vyhľadávania podľa autora.....	50
Obr. 3.5 SCOPUS – pokročilé vyhľadávanie.....	51
Obr. 3.6 WoK – vyhľadávanie.....	52
Obr. 3.7 WoK – vyhľadávacie pomôcky (abecedný prehľad zdrojových dokumentov).....	52
Obr. 3.8 WoK – výsledky vyhľadávania.....	53
Obr. 3.9 WoK – ladenie vyhľadávania.....	53
Obr. 3.10 Ladenie vyhľadávania.....	54
Obr. 3.11 SCOPUS – doladenie vyhľadávania.....	54
Obr. 3.12 Sledovanie citácií (podľa Bartošek 2000).....	55
Obr. 3.13 WoS – pokročilé vyhľadávanie.....	55
Obr. 3.14 WoS – pokročilé vyhľadávanie – zobrazenie výsledkov.....	56
Obr. 3.15 Vyhľadané citácie (krok 1).....	56
Obr. 3.16 Zobrazenie záznamu citovaného článku.....	57
Obr. 3.17 Vyhľadané citácie (krok 2) – zobrazenie vybraných záznamov.....	57
Obr. 3.18 Vyhľadávanie citácií – zobrazenie výsledkov.....	57
Obr. 3.19 Citačná analýza.....	58
Obr. 3.20 SCOPUS – Citation tracker.....	58
Obr. 3.21 WoK – nastavenie formátu záznamu.....	59
Obr. 3.22 WoK – nastavenie citačného alertu.....	60
Obr. 3.23 WoK – analýza výsledkov podľa krajiny.....	60
Obr. 3.24 WoK – história vyhľadávania.....	61
Obr. 3.25 SCOPUS – možnosti výstupu.....	61
Obr. 3.26 SCOPUS – rozšírené funkcie prístupné registrovaným používateľom.....	61
Obr. 3.27 WorldCat – rozšírené vyhľadávanie.....	63
Obr. 3.28 Súborný online katalóg UK – výsledky vyhľadávania.....	64
Obr. 4.1 Model distribuovaného dokumentového prieskumu (podľa Grossman, Frieder 2004).....	69

Obr. 4.2 Vyhľadávač Collarity.....	69
Obr. 4.3 Znáročenie algoritmu PageRank.....	70
Obr. 4.4 Rozdiely medzi web 1.0 a web 2.0 podľa O'Reilly ( <a href="http://www.oreillynet.com/">http://www.oreillynet.com/</a> ).....	71
Obr. 4.5 Mapa funkcií a služieb web 2.0 podľa O'Reilly ( <a href="http://www.oreillynet.com/">http://www.oreillynet.com/</a> ).....	72
Obr. 4.6 Web 2.0 ako prepojenie ľudí, zdrojov, služieb.....	73
Obr. 4.7 NSDL – Národná vedecká digitálna knižnica.....	74
Obr. 4.8 Pomocou iRazoo sa dajú hodnotiť a komentovať vyhľadané výsledky.....	78
Obr. 4.9 Kategória „technology“ v Mahalo.....	79
Obr. 4.10 Presúvanie výsledkov v Yoogle.....	80
Obr. 4.11 Príklad kustomizovaného vyhľadávača vytvoreného pomocou Google Custom Search.....	81
Obr. 4.12 Odporúčania hudby, podujatí a videí na Last.fm.....	82
Obr. 4.13 Viacúrovňová fazetová klasifikácia pomocou mSpace.....	84
Obr. 4.14 Momentka stránky z výsledkov na Ask.com.....	85
Obr. 4.15 Retrievr vyhľadáva fotografie a obrázky na základe nakresleného obrázku.....	86
Obr. 5.1 Pojmová mapa Relevancia (Steinerová, Grešková, Šušol 2007).....	92
Obr. 5.2 Vzťahy medzi relevantnými a vyhľadanými dokumentmi, odozva a presnosť (podľa Grossman, Frieder 2004).....	93
Obr. 5.3a Vzťahy medzi odozvou a presnosťou (optimálny, typický) (podľa Grossman, Frieder 2004).....	94
Obr. 5.3b Graf presnosť/odozva (podľa Manning et al. 2009).....	94
Obr. 5.4 Typy relevancie – pojmová mapa (Steinerová, Grešková, Šušol 2007).....	96
Obr. 5.5 Vývoj interpretácie relevancie v informačnej vede.....	97
Obr. 5.6 Relevancia v elektronickom prostredí – pojmová mapa (Steinerová, Šušol, Grešková 2007).....	98
Obr. 6.1 Kategorizácia emócií pri posudzovaní relevancie.....	111
Obr. 6.2 Kategorizácia emócií pri interakcii človeka so systémom.....	112
Obr. 7.1. Taxonómia delenia druhov organizmov.....	116
Obr. 7.2 Vplyv informačných štruktúr na vyhľadávanie informácií.....	120
Obr. 7.3 Polyreprezentácia poznania (podľa Ingwersen, Järvelin 2005).....	122
Obr. 7.4 Rozdiely medzi štruktúrovaným a neštruktúrovaným prieskumom.....	123
Obr. 7.5 Význam reprezentácií poznania pre vyhľadávanie.....	123
Obr. 8.1. Metadáta podľa Dublin core metadata initiative (DCMI), 2003.....	130
Obr. 8.2. Génová ontológia (Gene ontology consortium, 2003).....	131
Obr. 8.3 MyLifeBits – časový prehľad aktivít.....	132
Obr. 8.4 Vizualizácia folksonómie z Del.icio.us (Shaw, 2006).....	133
Obr. 8.5 Ontológia riešenia problémov v robotickom futbale (Cossentino et al., rok neznámy).....	134
Obr. 8.6 Porfýriov strom ako prvá sémantická sieť.....	136
Obr. 8.7 WordNet – záznam slova „technology“.....	136
Obr. 8.8 Bing Translator – strojový tlmočník na webe.....	137
Obr. 8.9 Výsledky z Google Translate po zadaní anglického dotazu „knowledge organization“ (organizácia poznania) preložené do francúzštiny.....	137
Obr. 8.10 Výsledky v Kartoo ako sémantická sieť.....	138
Obr. 8.11 Model odvodený z ontológie priebehu práce – „Workflow ontology“.....	139
Obr. 8.12 Vyhľadávanie autora „John Wilson“ pomocou GoPubMed, ktorý využíva ontológiu MeSH.....	140
Obr. 8.13 Zoznam výsledkov po zadaní „def:RDF“ do vyhľadávača Swoogle.....	141
Obr. 8.14 Zoznam podtried kľúčového slova „person“ vyhľadaný ONTOSEARCH2.....	142
Obr. 8.15 Zjednodušený model tvorby tematickej mapy z dokumentov a databáz (Garshol, 2002).....	143
Obr. 8.16 Príklad tematickej mapy 1 (Payley, 2006) – vzťahy medzi vedeckými paradigmami.....	144
Obr. 8.17 Príklad tematickej mapy 2 – zobrazenie súvislostí medzi témami v článkoch na BBC (2007).....	145
Obr. 8.18 Vizualizácia tematickej mapy talianskej opery pomocou funkcie Vizigator (2009).....	145
Obr. 8.19 Záznam zo slovníka Web and XML Glossary (Web, 2009), ktorý využíva tematické mapy.....	146
Obr. 8.20 Konceptuálna mapa zobrazuje vlastnosti pojmových máp (Novak, Cañas, 2008).....	147
Obr. 8.21 Mapa mysle – proces vytvárania mapy mysle (van Halen, 2009).....	148
Obr. 8.22 Mindomo – mapa sprehľadňujúca oblasť informačných technológií.....	148
Obr. 8.23 Mapa vytvorená pomocou Bubl.us– aplikácie na vytváranie pojmových máp a máp mysle.....	149
Obr. 8.24 Mappio – odporúčané mapy a jednoduché vyhľadávanie.....	150
Obr. 8.25 Vyhľadávanie verejne dostupných máp pomocou služby MindMeister.....	150
Obr. 8.26 CiteULike – zoznam kategórií priradených článkom používateľmi.....	152

Obr. 8.27 Fotografie ulíc Barcelony od používateľov Picasa na Google maps.....	152
Obr. 8.28 Pridávanie značiek k tváram a odstraňovanie nesprávne rozpoznávaných objektov.....	153
Obr. 8.29 Flickr – triedenie fotografií do setov a ďalšie možnosti organizácie.....	153
Obr. 8.30 Flickr – pridávanie poznámok na fotografiu a možnosť komentovania.....	154
Obr. 8.31 Blog registrovaný v adresári Technorati.....	155
Obr. 9.1 World Digital Library.....	161
Obr. 9.2 Schéma základných pojmov digitálnej knižnice.....	166
Obr. 9.3 Hlavné roly aktérov a tri vrstvy digitálnej knižnice podľa Manifestu DK.....	167
Obr. 9.4 Zložky univerza digitálnej knižnice.....	167
Obr. 9.5 Sociálna webová služba Twine.....	169
Obr. 9.6 TrueKnowledge, internetový nástroj na odpovedanie.....	171
Obr. 9.7 Nástroj sémantického prieskumu SemantiFind.....	172
Obr. 9.8 Nástroj sémantického prieskumu Powerset.....	172
Obr. 9.9 Systém CMS Drupal.....	173
Obr. 9.10 Systém Calais.....	174
Tab. 4.1 Prehľad základných typov vyhľadávacích nástrojov.....	76
Tab. 5.1 Rozdiely v podpore posudzovania relevancie medzi tradičným knižničným prostredím a elektronickým prostredím.....	100
Tab. 6.1 Rozdiely medzi pragmatickým a analytickým informačným štýlom.....	109
Tab. 7.1 Rozdiely medzi kategorizáciou a klasifikáciou (podľa Jacob 2004).....	117
Tab. 7.2 Príklad rámcovej reprezentácie poznania (rámec „Auto“.....)	122