

# Obsah

<b>O autorech</b> . . . . .	V
<b>Předmluva</b> . . . . .	VII
<b>1 Data, informace, znalosti?</b> . . . . .	1
1.1 Co jsou data, informace a znalosti . . . . .	1
1.1.1 Data . . . . .	2
1.1.2 Informace . . . . .	2
1.2 Znalosti . . . . .	4
1.3 Informační gramotnost . . . . .	4
1.3.1 Z historie . . . . .	4
1.3.2 Informační přehlčení? . . . . .	5
1.3.3 Formy informací . . . . .	6
1.3.4 Informace a komunikace . . . . .	7
1.3.5 Gramotnost a informační gramotnost . . . . .	7
1.3.6 Informační gramotnost a Internet . . . . .	9
1.4 Kolik je informací? . . . . .	10
Literatura ke kapitole 1 . . . . .	14
<b>I Práce s informacemi</b> . . . . .	17
<b>2 Informační proces</b> . . . . .	19
2.1 Základní pojmy . . . . .	19
2.2 Informační zdroje . . . . .	22
2.2.1 Typy a charakteristiky informačních zdrojů . . . . .	22
2.2.2 Knihovny . . . . .	24
2.2.3 Specializované firmy . . . . .	26
2.3 Vyhledávání dokumentů . . . . .	26
2.4 Popis dokumentů . . . . .	30
2.5 Charakteristika obsahu dokumentů . . . . .	32
2.6 Automatická charakteristika obsahu . . . . .	35

2.7	Boolovský model . . . . .	39
2.7.1	Princip boolovského modelu . . . . .	39
2.7.2	Úplnost a přesnost v boolovském modelu . . . . .	40
2.7.3	Zvyšování přesnosti a úplnosti . . . . .	44
2.8	Invertovaný soubor . . . . .	46
2.9	Další přístupy k vyhledávání textových dokumentů . . . . .	49
2.9.1	Rozšiřování boolovského modelu . . . . .	51
2.9.2	Topic . . . . .	54
	Literatura ke kapitole 2 . . . . .	56
<b>3</b>	<b>Zpracování specifických jevů přirozeného jazyka v automatizovaných informačních systémech . . . . .</b>	<b>58</b>
3.1	Automatizace indexování a vyhledávání textových informací . . . . .	58
3.2	Překonání jazykových bariér . . . . .	66
3.3	Standardizace textů . . . . .	73
3.4	Komunikace s počítačem v přirozeném jazyce . . . . .	77
	Literatura ke kapitole 3 . . . . .	88
<b>II</b>	<b>Práce se znalostmi . . . . .</b>	<b>91</b>
<b>4</b>	<b>Práce se znalostmi . . . . .</b>	<b>93</b>
4.1	Expertní systémy, znalostní systémy . . . . .	93
4.1.1	Zrození znalostního inženýrství . . . . .	93
	DENDRAL . . . . .	93
4.1.2	Charakteristické rysy . . . . .	94
4.1.3	Struktura . . . . .	96
4.1.4	Typy expertních úloh . . . . .	97
4.1.5	Diagnostické úlohy . . . . .	99
4.1.6	Generativní úlohy . . . . .	101
4.1.7	Některé slavné systémy . . . . .	102
	Mycin . . . . .	102
	TEIRESIAS . . . . .	104
	PROSPECTOR . . . . .	105
	R1/XCON . . . . .	105
4.1.8	Volba aplikační oblasti . . . . .	106
4.2	Znalosti a jejich reprezentace . . . . .	107
4.2.1	Pravidla . . . . .	109
4.3	Odvozování . . . . .	112
	OPS5 . . . . .	114
	Odvozování v diagnostických expertních systémech první generace . . . . .	115
4.3.1	Zpětné řetězení . . . . .	118
4.3.2	Přímé řetězení . . . . .	119
4.3.3	Agenda . . . . .	120
	AM . . . . .	120
4.3.4	Tabule . . . . .	121
	Hearsay . . . . .	121

4.4	Práce s neurčitostí . . . . .	121
4.4.1	Algebraické pojetí . . . . .	122
4.5	Získávání znalostí od experta . . . . .	124
4.5.1	Problém získávání znalostí . . . . .	124
4.5.2	Spolupráce s expertem a příprava na ni . . . . .	126
4.5.3	Verbální techniky elicitace znalostí . . . . .	128
4.5.4	Neverbální techniky elicitace znalostí . . . . .	130
4.6	Modelování znalostí . . . . .	133
4.6.1	Od transferu znalostí k modelování . . . . .	133
4.6.2	Modely řešení problémů . . . . .	137
4.6.3	Znalostní ontologie . . . . .	142
	Literatura ke kapitole 4 . . . . .	147
<b>5</b>	<b>Od dat ke znalostem . . . . .</b>	<b>151</b>
5.1	Základní rysy KDD . . . . .	151
5.2	Metody získávání znalostí . . . . .	155
5.2.1	Symbolické metody . . . . .	156
	GUHA . . . . .	160
	KEX . . . . .	161
5.2.2	Případové usuzování . . . . .	164
	CBR Express . . . . .	165
5.2.3	Neuronové sítě . . . . .	166
5.3	Automatické získávání znalostí z textů . . . . .	172
5.3.1	Základní problémy . . . . .	172
	Literatura ke kapitole 5 . . . . .	176
<b>III</b>	<b>Informace a Internet . . . . .</b>	<b>179</b>
<b>6</b>	<b>Internet!?. . . . .</b>	<b>181</b>
6.1	Historické ohlédnutí . . . . .	181
6.2	Co je tedy Internet? . . . . .	183
6.3	Informační služby Internetu . . . . .	184
6.3.1	Telnet . . . . .	184
6.3.2	Gopher . . . . .	185
6.4	Z historie vyhledávání informací v prostředí Internetu . . . . .	186
	Literatura ke kapitole 6 . . . . .	186
<b>7</b>	<b>World Wide Web – základ přístupu k informacím v Internetu . . . . .</b>	<b>188</b>
7.1	Základní charakteristika . . . . .	188
7.2	Z historie . . . . .	188
7.3	Klient/server . . . . .	189
7.4	Možnosti využití WWW . . . . .	191
7.5	Stavební kameny . . . . .	192
7.5.1	Hypertext . . . . .	192
7.5.2	Jazyk HTML . . . . .	194
	Základní princip . . . . .	194
	HTML a SGML . . . . .	195



	Další vývoj . . . . .	196
7.5.3	URL (Uniform Resource Locator) . . . . .	198
	URL pro FTP . . . . .	200
	URL pro Telnet . . . . .	201
	URL pro e-mail . . . . .	201
	URL pro Gopher . . . . .	202
	URL pro HTTP . . . . .	202
	URL pro Usenet News . . . . .	203
7.5.4	Protokol HTTP . . . . .	203
7.6	Klienti služby WWW . . . . .	204
7.6.1	Klienti WWW . . . . .	204
7.6.2	Klient WWW jako hypertextový prohlížeč . . . . .	205
7.6.3	Grafické prohlížeče . . . . .	206
7.7	Záložky . . . . .	207
7.7.1	Účel a princip . . . . .	207
7.7.2	Záložky nebo kopie? . . . . .	207
7.7.3	Organizování záložek . . . . .	209
7.7.4	Zastarávání záložek . . . . .	209
7.8	Tvorba dokumentů pro službu WWW . . . . .	211
	Literatura ke kapitole 7 . . . . .	212
<b>8</b>	<b>Informace a Internet . . . . .</b>	<b>214</b>
8.1	Úspěšnost vyhledávání aneb co jsou relevance, přesnost a úplnost . . . . .	215
8.2	Internet jako zdroj informací . . . . .	217
8.2.1	Internet jako seskupení počítačů . . . . .	218
8.2.2	Internet jako soubor programů . . . . .	220
8.2.3	Internet jako soubor zdrojů . . . . .	220
8.2.4	Internet jako knihovna . . . . .	222
8.3	Co je správné? . . . . .	224
8.4	Je Internet ideálním prostředím pro vyhledávání informací? . . . . .	225
	Literatura ke kapitole 8 . . . . .	227
<b>9</b>	<b>Vyhledávání informací v Internetu jako problém . . . . .</b>	<b>229</b>
9.1	Některá hledání jsou neúspěšná . . . . .	230
9.2	Vyhledávání v Internetu jako umění svého druhu . . . . .	232
9.3	Velikost Internetu . . . . .	232
9.3.1	Velikost prostoru WWW . . . . .	233
	Jak lze určit velikost prostoru WWW . . . . .	233
9.3.2	Charakteristika typického uživatele . . . . .	234
9.3.3	WWW dokumenty a vyhledávání . . . . .	235
9.4	Velikost WWW prostoru? . . . . .	236
9.4.1	Interpretace . . . . .	237
	Poznámky k veřejné indexovatelnosti v odhadech . . . . .	237
	Poznámky k časové platnosti odhadů . . . . .	240
	Poznámky k odhadování, přesnosti a nepřesnosti . . . . .	240
9.4.2	Má to smysl? . . . . .	241
9.5	Základní metody vyhledávání v Internetu a příčiny problémů . . . . .	242
9.5.3	Status Internetu a zklamání uživatele při vyhledávání . . . . .	242

9.5.4	Vyhledávací stroje a předmětové katalogy . . . . .	243
9.5.5	Jak se vyhledává a co zabírá čas při vyhledávání . . . . .	245
9.6	Vyhledávací služby versus tradiční on-line služby . . . . .	247
	Literatura ke kapitole 9 . . . . .	250
<b>10</b>	<b>Předmětové katalogy . . . . .</b>	<b>252</b>
10.1	Principy fungování . . . . .	252
10.1.1	Hledisko uživatele . . . . .	252
10.1.2	Hledisko údržby katalogu . . . . .	253
10.2	Výhody a nevýhody předmětových katalogů . . . . .	254
10.3	Vývoj a trendy předmětových katalogů . . . . .	256
10.3.1	Vyhledávání . . . . .	256
10.3.2	Portály . . . . .	256
10.4	Příklady předmětových katalogů . . . . .	257
	Literatura ke kapitole 10 . . . . .	258
<b>11</b>	<b>Vyhledávací stroje . . . . .</b>	<b>259</b>
11.1	Jde o hity . . . . .	259
11.2	Principy fungování . . . . .	260
11.2.1	Robot . . . . .	261
	Strategie procházení . . . . .	262
	Problémy robotů . . . . .	264
	Odmítnutí robota . . . . .	265
11.2.2	Index . . . . .	266
11.2.3	Vyhledávací stroj . . . . .	267
11.2.4	Problém aktuálnosti . . . . .	268
11.3	Jednoduché dotazy jsou rychlé a hloupé . . . . .	269
11.4	Volba klíčových slov je základ úspěchu . . . . .	270
11.4.1	Základem je přemýšlet . . . . .	270
11.4.2	Podstatná jsou podstatná jména a objekty . . . . .	271
11.4.3	Zkracování slov a užívání zástupných znaků . . . . .	272
11.4.4	Nalezení správné úrovně . . . . .	273
11.4.5	Synonyma . . . . .	275
11.4.6	Fráze . . . . .	275
11.5	Jak se vyhodnocují jednoduché dotazy? . . . . .	276
11.6	Jednoduché dotazy a default operátor . . . . .	276
11.7	Pokročilé dotazy . . . . .	277
11.7.1	Logické základy . . . . .	278
	Přehled boolovských prvků . . . . .	278
	Operátor AND . . . . .	279
	Operátor OR . . . . .	281
11.7.2	Pokročilejší operátory . . . . .	282
	Operátor NEAR . . . . .	282
	Operátory BEFORE a AFTER . . . . .	283
	Operátor AND NOT . . . . .	283
11.7.3	Jak formulovat pokročilejší dotaz . . . . .	284
	Používání závorek . . . . .	285
	Kombinace konceptů pro pokročilé dotazy . . . . .	287

	Interpunkce a velikost písmen . . . . .	287
	Zpřesňování dotazu . . . . .	288
11.8	Dotazovacích jazyků je mnoho . . . . .	288
11.9	Použití filtrů . . . . .	290
11.9.1	Filtr uzlu . . . . .	292
11.9.2	Filtr data . . . . .	293
11.9.3	Jiné druhy filtrů a voleb vyhledávání . . . . .	294
11.10	Jak řadí vyhledávací stroje dokumenty . . . . .	294
11.10.1	Proč jsou výsledky řazeny . . . . .	294
11.10.2	Nevýhody automatického řazení . . . . .	295
11.10.3	Obvyklá kritéria pro stanovení relevance . . . . .	296
11.10.4	Řazení a spamming . . . . .	298
11.10.5	Jiné faktory relevance . . . . .	299
11.11	Vyvarovat se zbytečných chyb . . . . .	301
11.11.1	Překlepy . . . . .	301
11.11.2	Zbytečné termíny . . . . .	302
11.11.3	Ignorovaná slova a speciální znaky . . . . .	302
11.11.4	Alternativní pravopis . . . . .	302
11.11.5	Příliš mnoho termínů, synonym . . . . .	303
11.11.6	Nevhodně sestavené nebo komplikované dotazy . . . . .	303
11.12	Vyhledávací strategie aneb jak hledat efektivně . . . . .	306
11.12.1	Obecné vyhledávání . . . . .	306
11.12.2	Specifické vyhledávání . . . . .	307
11.12.3	Kombinace obecného a specifického vyhledávání . . . . .	307
11.12.4	Inkrementální vyhledávání . . . . .	308
11.12.5	Řetězové vyhledávání . . . . .	310
11.12.6	Vyhledej-a-skoč . . . . .	310
11.12.7	Obecné přednosti funkce Find . . . . .	311
11.12.8	Kategoriální vyhledávání . . . . .	312
11.12.9	Vyhledej-a-seřaď . . . . .	312
11.12.10	Souhrn . . . . .	313
11.13	Poznání používaného vyhledávacího systému . . . . .	313
11.13.1	Drobné varování – dynamický vývoj vyhledávacích služeb . . . . .	313
11.13.2	Překrytí, pokrytí a pružnost . . . . .	315
11.13.3	Boolovské dotazy – ano nebo ne? . . . . .	317
11.13.4	Některé zajímavé rysy chování . . . . .	317
	Problémy s počty . . . . .	318
	Mýtus milionových výsledků . . . . .	319
	Problémy při zpracování . . . . .	320
11.14	Příklady vyhledávacích strojů . . . . .	320
	Literatura ke kapitole 11 . . . . .	321
<b>12</b>	<b>Metavyhledávací stroje . . . . .</b>	<b>323</b>
12.1	Výhody a nevýhody . . . . .	323
12.2	Principy fungování . . . . .	325
12.3	Metavyhledávání – řešení typu klient . . . . .	327
12.4	Metavyhledávání – řešení typu server . . . . .	328

12.4.1	Základní charakteristika . . . . .	328
12.4.2	Pohled z hlediska uživatele . . . . .	328
12.5	Příklady metavyhledávacích systémů . . . . .	330
	Literatura ke kapitole 12 . . . . .	330
<b>13</b>	<b>Metadata . . . . .</b>	<b>331</b>
13.1	Meta-úvod . . . . .	331
13.2	Co jsou metadata? . . . . .	333
13.3	Potřeba metadat . . . . .	335
13.4	Vyhledávání a metadata . . . . .	336
13.5	Problémy a překážky . . . . .	339
13.6	Standardizace metadat . . . . .	340
13.6.1	Schéma metadat . . . . .	340
13.6.2	Standardy . . . . .	341
13.6.3	Metadatové systémy . . . . .	342
13.6.4	Dublin Core . . . . .	343
	Vznik a vývoj . . . . .	343
	Základní charakteristika Dublin Core . . . . .	344
	Základní prvky Dublin Core . . . . .	345
13.6.5	Jiné projekty . . . . .	346
13.6.6	Kódování metadat . . . . .	347
	SGML . . . . .	348
	Text Encoding Initiative (TEI) . . . . .	348
	HTML . . . . .	349
	Warwick Framework . . . . .	352
	XML . . . . .	352
	RDF . . . . .	353
13.7	Praktické problémy s metadaty . . . . .	358
13.7.1	Výběr elementů, subelementů a schématu . . . . .	358
13.7.2	Vytvoření a uložení metadat . . . . .	359
13.7.3	Řízená terminologie . . . . .	360
13.7.4	Identifikace zdrojů . . . . .	361
	Požadované vlastnosti . . . . .	361
	Uniform Resource Locator (URL) . . . . .	363
	Uniform Resource Identifier (URI) . . . . .	363
	Uniform Resource Name (URN) . . . . .	364
	Persistent Uniform Resource Locator (PURL) . . . . .	365
	Digital Object Identifier (DOI) . . . . .	366
13.8	Budoucnost metadat . . . . .	367
	Literatura ke kapitole 13 . . . . .	368
<b>14</b>	<b>Hodnocení informací získávaných z Internetu . . . . .</b>	<b>370</b>
14.1	Proč hodnotit? . . . . .	370
14.2	Kvalita informací . . . . .	374
14.3	Služby zaměřené na recenzování a hodnocení . . . . .	375
14.4	Subject gateways, virtuální knihovny . . . . .	377
14.5	Kritéria hodnocení . . . . .	379
14.5.1	Účel zdroje, cílová skupina uživatelů . . . . .	381

14.5.2	Obsah . . . . .	381
	Odkazy na jiné zdroje . . . . .	382
	Zrcadla . . . . .	382
	Povaha obsahu . . . . .	383
	Kvalita psaného jazyka . . . . .	383
14.5.3	Autorita . . . . .	383
14.5.4	Presnost . . . . .	384
14.5.5	Aktuálnost, správa zdroje . . . . .	385
14.5.6	Dostupnost zdroje . . . . .	387
	Rychlost přístupu . . . . .	387
	Softwarová omezení . . . . .	387
	Jiná omezení přístupu . . . . .	388
	Náklady . . . . .	389
	Copyright . . . . .	389
	Spolehlivost přístupu, snadnost nalezení zdroje . . . . .	390
	Zabezpečení přístupu . . . . .	390
14.5.7	Prezentace a uspořádání informací . . . . .	390
	Pomoc při nalezení informace v rámci zdroje . . . . .	390
	Obrázky, rámy, Java . . . . .	392
	Design . . . . .	393
14.5.8	Snadnost použití . . . . .	395
	Služby podpory uživatelů . . . . .	395
14.5.9	Srovnání s jinými zdroji . . . . .	395
14.5.10	Celková kvalita . . . . .	396
14.5.11	Seznam kontrolních otázek . . . . .	397
	Literatura ke kapitole 14 . . . . .	400
<b>15</b>	<b>Etika v prostředí Internetu . . . . .</b>	<b>403</b>
15.1	Etika . . . . .	403
15.2	Informační etika, počítačová etika, etika v počítačových sítích . . . . .	404
15.3	Faktory, které podporují neetické chování uživatelů . . . . .	407
15.4	Některé etické problémy . . . . .	408
15.4.1	Soukromí . . . . .	408
15.4.2	Vlastnictví . . . . .	410
15.4.3	Svoboda . . . . .	413
15.4.4	Hacking, cracking . . . . .	413
15.4.5	Technologie agentů . . . . .	414
15.4.6	Informační nerovnost . . . . .	415
15.4.7	Informační obsah . . . . .	415
15.5	Všeobecné principy etiky v Internetu . . . . .	417
15.5.1	Etika a právo . . . . .	417
15.5.2	Profesní kodexy . . . . .	418
15.5.3	Soubory norem organizací, firem . . . . .	420
15.5.4	Zásady práce v prostředí Internetu . . . . .	420
	Obecná pravidla . . . . .	420
	Specifická pravidla, netiketa . . . . .	423
15.6	Budoucnost informační etiky? . . . . .	423



Literatura ke kapitole 15 . . . . .	424
<b>16 Inteligentní systémy na Internetu . . . . .</b>	<b>426</b>
16.1 Popis prostředí WWW . . . . .	426
16.2 Podpora „brouzdání“ . . . . .	427
16.3 Metavyhledávání . . . . .	429
16.4 Podpora obchodování . . . . .	430
Literatura ke kapitole 16 . . . . .	433
<b>17 Trendy vyhledávání v prostředí WWW . . . . .</b>	<b>435</b>
17.1 Evoluce – pestřejší budoucnost pro uživatele . . . . .	436
17.1.1 Integrace vyhledávacích funkcí do prohlížečů . . . . .	436
17.1.2 Portály . . . . .	436
17.1.3 Partnerství, slučování, akvizice . . . . .	437
17.1.4 Služby hybridního typu . . . . .	438
17.1.5 Zvláštní služby v rámci vyhledávacích strojů . . . . .	438
17.1.6 Trend k jednoduchosti . . . . .	440
17.2 Přibližování – vsudypřítomný Web . . . . .	441
17.3 Nedostatek informací – výzva pro vyhledávací stroje . . . . .	441
17.4 Síla člověka . . . . .	442
17.5 Personalizace vyhledávání . . . . .	444
17.6 Vyhledávání nezávislé na prohlížeči . . . . .	444
17.7 Dotazy podle dokumentů . . . . .	446
17.8 Techniky pojmového vyhledávání . . . . .	446
17.9 Expanzivnější indexování . . . . .	448
17.10 Specializace vyhledávacích strojů . . . . .	450
17.10.1 Proč používat specializované vyhledávací služby . . . . .	451
17.10.2 Regionální vyhledávací služby . . . . .	451
17.10.3 Věcně specializované vyhledávací služby . . . . .	452
17.10.4 Multimediální vyhledávací služby . . . . .	452
17.10.5 Kde hledat specializované vyhledávací služby . . . . .	452
17.11 Neviditelný Web . . . . .	453
17.11.1 Jak zviditelnit neviditelné . . . . .	455
17.11.2 Budoucnost neviditelného Webu . . . . .	456
17.12 Nová generace vyhledávacích služeb? . . . . .	456
17.12.1 Zpracování přirozeného jazyka . . . . .	457
17.12.2 Inteligentní agenti . . . . .	459
17.12.3 Filtrování . . . . .	459
17.12.4 Vizualizace . . . . .	460
Literatura ke kapitole 17 . . . . .	460
<b>IV Přílohy . . . . .</b>	<b>463</b>
<b>A Slovník pojmů . . . . .</b>	<b>465</b>
<b>B Slovník pojmů z oblasti vyhledávacích nástrojů . . . . .</b>	<b>497</b>