

Massimo Melucci • Ricardo Baeza-Yates
Editors

Advanced Topics in Information Retrieval



Springer

Contents

1	Digital Libraries	1
	Maristella Agosti	
1.1	Introduction	1
1.2	Digital Libraries in the Beginning	2
1.3	Digital Library Systems	7
1.3.1	User Interface to Digital Library Systems	9
1.3.2	Significative Examples of Digital Library Systems	10
1.4	The Digital Library Manifesto and the DELOS Digital Library Reference Model	15
1.5	Digital Library Systems Become User-Centered Systems: Adding Advanced Annotation Functions to Digital Library Systems	19
1.6	Interoperability between Digital Library Systems	22
1.7	Evaluation of Digital Libraries	24
1.8	Conclusions	26
2	Scalability Challenges in Web Search Engines	27
	Berkant Barla Cambazoglu and Ricardo Baeza-Yates	
2.1	Introduction	27
2.2	Components	28
2.3	Objectives	29
2.4	Parameters	31
2.5	Scalability Issues	32
2.5.1	Single-Node System	32
2.5.2	Multi-Node Cluster	38
2.5.3	Multi-Cluster Site	42
2.5.4	Multi-Site Engine	44
2.6	Open Problems	47
2.6.1	Crawling	47
2.6.2	Indexing	48
2.6.3	Query Processing	48
2.7	Conclusions	49

3	Spam, Opinions, and Other Relationships: Towards a Comprehensive View of the Web Knowledge Discovery	51
	Bettina Berendt	
3.1	Introduction	51
3.2	Basics and Terminology	53
3.2.1	From Information Retrieval to Knowledge Discovery	53
3.2.2	Knowledge Discovery Phases	55
3.2.3	Modeling Tasks and Model/Pattern Structures	57
3.2.4	Learning Cycles and Knowledge Discovery	58
3.2.5	Web Mining: Content, Structure and Usage	60
3.3	A Short Overview of Web Content Mining on Text	61
3.3.1	Application and Data Understanding	61
3.3.2	Data Preparation	63
3.3.3	Modeling	65
3.3.4	Evaluation and Deployment	76
3.4	New Challenges: Web Mining and Privacy	77
3.5	Conclusions	79
3.5.1	Context	80
3.5.2	Learning Cycles	80
3.5.3	Definitional Power and Viewpoints	81
3.5.4	Tools and Access	81
4	The User in Interactive Information Retrieval Evaluation	83
	Peter Ingwersen	
4.1	Introduction	83
4.2	Research Frameworks, Models and Other Central Concepts	84
4.2.1	Research Design and IIR Research Setting Types	86
4.2.2	Central IIR Components	88
4.2.3	The Integrated Cognitive Research Framework for IR	89
4.3	IR Interaction—Research Designs with Test Persons	90
4.3.1	Search Job Design, Simulated Task Situations, Test Persons and Evaluation Measures	91
4.3.2	IR Interaction Ultra-Light Studies	95
4.3.3	IR Interaction Light	99
4.3.4	Naturalistic Field Investigations of IR Interaction—Exemplified	102
4.4	Conclusions	104
4.5	Classified Bibliography	105
5	Aggregated Search	109
	Mounia Lalmas	
5.1	Introduction	109
5.2	Background and Motivation	110
5.3	Vertical Representation	113
5.4	Vertical Selection	114
5.5	Result Presentation	115

5.6	Evaluation	120
5.7	Conclusions	123
6	Quantum Mechanics and Information Retrieval	125
	Massimo Melucci and Keith van Rijsbergen	
6.1	Introduction	125
6.2	Another View of Probability	126
6.3	Three Notions of Quantum Probability	133
6.3.1	Superposition	133
6.3.2	Interference	136
6.3.3	Entanglement	138
6.4	Quantum Probability and Information Retrieval	141
6.4.1	Abstract Vector Spaces	141
6.4.2	Using Quantum Mechanics in Information Retrieval	151
6.5	Conclusions	154
7	Multimedia Resource Discovery	157
	Stefan R�ger	
7.1	Introduction	158
7.2	Basic Multimedia Search Technologies	162
7.2.1	Piggy-Back Text Retrieval	162
7.2.2	Automated Annotation	163
7.2.3	Content-Based Retrieval	170
7.2.4	Fingerprinting	171
7.3	Challenges of Automated Visual Indexing	172
7.4	Added Services	173
7.4.1	Video Summaries	173
7.4.2	New Paradigms in Information Visualisation	175
7.4.3	Visual Search and Relevance Feedback	178
7.5	Browsing: Lateral and Geo-Temporal	181
7.6	Conclusions	185
8	Information Retrieval in Context	187
	Ian Ruthven	
8.1	Introduction	187
8.1.1	Contextual Information Seeking	188
8.1.2	Elements of Context	189
8.1.3	Practical Uses of Context	191
8.2	Views of Context	192
8.2.1	Socio-Technical Systems View	192
8.2.2	Contextually Sensitive Systems View	194
8.2.3	Engagement View	196
8.2.4	Summary	199
8.3	Examples of Contextual Information Retrieval	200
8.3.1	Personal Context	200
8.3.2	Social Context	202
8.3.3	Interaction Context	203

- 8.3.4 Task Context 205
 - 8.3.5 Summary 205
 - 8.4 Contextual Evaluation 205
 - 8.5 Conclusions 207
- 9 Digital Advertising: An Information Scientist’s Perspective 209**
James G. Shanahan and Goutham Kurra
 - 9.1 Introduction 209
 - 9.2 Online Advertising Background 210
 - 9.2.1 Purchase Funnel 210
 - 9.2.2 Types of Advertising 211
 - 9.2.3 Payment Models 212
 - 9.2.4 Market Numbers and Trends 213
 - 9.3 Sponsored Search-Based Advertising 213
 - 9.3.1 Sponsored Search Overview 214
 - 9.3.2 Organic Search Engine 214
 - 9.3.3 Sponsored Search Engine 219
 - 9.3.4 Research Challenges and Opportunities 223
 - 9.4 Contextual and Display Advertising 230
 - 9.4.1 Research Challenges and Opportunities in Contextual Advertising 230
 - 9.4.2 Display Advertising 234
 - 9.4.3 Research Challenges and Opportunities in Display Advertising 234
 - 9.5 New Trends and Issues 236
 - 9.6 Conclusions 237
- References 239**
- Index 271**