

Contents

1	The Model	1
	Giovanni Maria Sacco	
1.1	Exploratory Search	3
1.2	Dynamic Taxonomies Defined	4
1.2.1	Concepts	6
1.2.2	Relationships Among Concepts	7
1.2.3	Reduced Taxonomies and Exploration	8
1.3	Implications for Schema Design	11
1.4	Advantages	13
1.5	Application Areas	14
1.6	Faceted Search and Dynamic Taxonomies	15
1.7	Book Roadmap	16
2	Faceted Taxonomy-Based Sources	19
	Yannis Tzitzikas	
2.1	Introduction	19
2.2	Taxonomies and Partially-Ordered Sets	20
2.3	Faceted Taxonomies	21
2.4	Taxonomy-Based Sources	24
2.5	On Intervals	26
2.6	Modeling Interaction	28
2.6.1	Zoom Points	29
2.6.2	Zoom-in	30
2.6.3	Zoom-Side	31
2.6.4	Zoom-out	32
2.7	Restriction	33
3	Comparison with Other Techniques	35
	Giovanni Maria Sacco, Sébastien Ferré, and Yannis Tzitzikas	
3.1	Structured Access and Information Retrieval	35
3.1.1	Queries on Structured Data	38
3.1.2	Information Retrieval	41
3.2	Static Taxonomies	46
3.2.1	Static Monodimensional Taxonomies	48
3.2.2	Static Multidimensional Taxonomies with no Concept Composition Capabilities	48
3.2.3	Multidimensional Taxonomies with Concept Composition Capabilities	49

3.2.4	A Comparison	51
3.2.5	Taxonomy Pruning in Dynamic Taxonomies	52
3.2.6	Relaxing the Assumptions	54
3.2.7	Experimental Results	55
3.3	Decision Trees	57
3.4	Formal Concept Analysis	59
3.4.1	Data Model	59
3.4.2	Information Access	62
3.4.3	Conclusion	64
3.5	Semantic Web	64
3.5.1	Description Logics and OWL Ontologies	66
3.5.2	Semantic Web and Exploratory Search	70
3.5.3	Conclusions	72
4	User Interface Design	75
	Moritz Stefaner, Sébastien Ferré, Saverio Perugini, Jonathan Koren, and Yi Zhang	
4.1	Principles	75
4.2	Challenges	77
4.3	Navigation Modes	78
4.3.1	Zoom-in	81
4.3.2	Zoom-out	82
4.3.3	Shift	82
4.3.4	Pivot	83
4.3.5	Slice and Dice	84
4.3.6	Range Selection	85
4.3.7	Querying by Examples	85
4.4	Design Patterns	86
4.4.1	Selection Management	87
4.4.2	Revealing Hierarchy	88
4.4.3	Facet Management	90
4.4.4	Keyword Search	91
4.4.5	Filter Summary and History Navigation	93
4.4.6	Animated Transitions	93
4.4.7	Visualizing Proportions	94
4.5	Extensions and Related Approaches	95
4.5.1	FaThumb	96
4.5.2	Browsing Related Entities	97
4.5.3	Resource Analytics	98
4.5.4	Out-of-turn Interaction	99
4.6	Personalizing Faceted Search	102
4.6.1	Introduction	102
4.6.2	Related Work: Personalized Search and Filtering	103
4.6.3	Personalization Based on Collaborative Filtering	104
4.6.4	K-Nearest Neighbors Based on Item–Item Similarity or User–User Similarity	105

4.6.5	Singular Value Decomposition	106
4.6.6	Personalization Using Content Based Filtering	107
4.6.7	An Ontological Approach	109
4.6.8	Evaluation Regime	110
4.6.9	Conclusions	112
5	Extensions to the Model	113
	Giovanni Maria Sacco and Sébastien Ferré	
5.1	Data Mining	114
5.1.1	Association Rule Mining	114
5.1.2	Dynamic Taxonomy Foci and Association Rules	114
5.1.3	Integrating Association Rules with Dynamic Taxonomies	116
5.1.4	An Example	118
5.1.5	Father–Son Correlation in the Taxonomy	120
5.1.6	General Association Rules	120
5.1.7	Side-by-Side Comparison	122
5.2	Structured Objects	123
5.3	Virtual Concepts	126
5.4	Logics	127
5.4.1	From Taxonomies to Logics	127
5.4.2	From Logics to Dynamic Taxonomies	130
5.5	Web Ontologies	133
5.5.1	Re-defining Extensions and Dynamic Taxonomies	133
5.5.2	Additional Navigation Modes	136
5.6	Fuzzy Dynamic Taxonomies	138
5.7	Miscellanea	139
5.7.1	Predefined Foci for Personalization and Access Control	139
5.7.2	SAES Facets	140
5.7.3	Popularity, Recommendations, and Authoritativeness	142
5.7.4	Augmenting IR Recall	143
6	Engineering Taxonomy-Based Sources	145
	Yannis Tzitzikas	
6.1	Compound Terms Composition Algebra (CTCA)	145
6.1.1	Motivation	145
6.1.2	The Algebra in Brief	146
6.1.3	Deriving Navigational Trees from CTCA Expressions	154
6.1.4	Tackling the Taxonomy Evolution Problem	155
6.1.5	Expression Mining and Other Applications	159
6.2	Adaptation of Taxonomy-Based Sources Through User Feedback	161
6.3	Mapping Taxonomy-Based Sources	163
6.4	Distributed Taxonomy-Based Sources	169
6.4.1	Mappings and Mediators	169
6.4.2	Distributed Query Evaluation	172
6.5	Synopsis and Bibliographic References	174

7	Taxonomy Design	175
	Wisam Dakka, Panagiotis Ipeiritis, and Giovanni Maria Sacco	
7.1	General Guidelines for Taxonomy Design	175
7.1.1	Design ‘in the Small’	180
7.1.2	Design ‘in the Large’	186
7.2	Automatic Construction from Text Information Bases	190
7.2.1	Problem Overview	190
7.2.2	Supervised Facet Extraction for Collections of Text-Annotated Items	192
7.2.3	Unsupervised Facet Extraction for Collections of Text Documents	195
7.2.4	Evaluating Our Supervised Facet Extraction Technique	202
7.2.5	Evaluating Our Unsupervised Facet Extraction Technique	205
7.2.6	Further Discussion and Future Work	212
7.2.7	Conclusion	212
8	System Implementation	215
	Giovanni Maria Sacco, Yannis Tzitzikas, and Sébastien Ferré	
8.1	Architecture and Implementation Strategies	215
8.1.1	Logical Architecture	216
8.1.2	Computing the Focus	218
8.1.3	Computing the Reduced Taxonomy	220
8.1.4	Presentation Strategies	221
8.1.5	Physical Storage Structures for the Extension	222
8.1.6	Experimental Data	228
8.1.7	Further Performance Enhancements	234
8.2	Implementation over a Relational Database Management System	237
8.3	Case Studies: Existing Systems	241
8.3.1	Flexplorer and Mitos	245
8.3.2	FASTAXON	247
8.4	Formats and Protocols	250
8.5	Composition of Taxonomies with Logic Components	253
8.5.1	Logics	254
8.5.2	Logic Functors	256
8.5.3	Combining Attributes, Concrete Domains, and Taxonomies	256
8.5.4	Reconstructing the Description Logic $\mathcal{ALC}$	258
8.5.5	Conclusion	260
9	Applications and Experiences	263
	Giovanni Maria Sacco and Sébastien Ferré	
9.1	Introduction	263
9.2	E-commerce	265
9.2.1	The Thinning Game	266
9.2.2	The End Game	269
9.3	Multimedia Information Bases	272

9.3.1	Combining Conceptual Access with Low Level Multimedia Features	274
9.3.2	Monodimensional vs. Multidimensional Clustering for Low Level Features	275
9.3.3	Representing Low Level Multimedia Features	277
9.3.4	Examples of Exploration	279
9.4	Diagnostic Systems	282
9.4.1	Computer-Assisted Medical Diagnosis	282
9.4.2	Diagnosis Through Dynamic Taxonomies	283
9.4.3	Application of Dynamic Taxonomies to the Diagnosis of Rare Diseases	284
9.5	Digital Libraries and News Systems	288
9.6	E-government	290
9.7	File Systems	294
9.7.1	Implementation	294
9.7.2	Applications	295
9.7.3	Related Works	297
9.8	Geographical Information Systems	298
9.8.1	Data Model and Querying Language	298
9.8.2	Interface and Implementation of GEOLIS	300
9.8.3	Experiments	301
10	Conclusions	303
	Giovanni Maria Sacco and Yannis Tzitzikas	
	References	307
	Index	323
	Appendix A Color Images	329