

Contents

Preface	ix
Acknowledgments	xiii
1. ORGANIZING INFORMATION BY SUBJECT	1
Major Questions	1
Organization of Information	1
Organization of Information By Subject	7
Subject Indexing and Evaluation	14
Knowledge Organization Systems in the Online Environment	16
User Information Behavior (by David Elsweiler)	25
Acquiring Information	25
Organizing Information	29
Re-Finding Information	30
Summary	31
Chapter Review	33
Glossary	33
References	35
2. KNOWLEDGE ORGANIZATION SYSTEMS (KOSs)	
<i>Claudio Gnoli</i>	43
Major Questions	43
An Encompassing Notion: Knowledge Organization	43
Dimensions of Knowledge	44
A Long-Standing Enterprise	47
The Layers Forming Knowledge Organization	48
Types of KOSs	51
Titles	51
Keywords and Tags	52

Subject Headings Lists and Terminologies	53
Thesauri	55
Taxonomies	56
Classification Schemes	57
Ontologies	59
How to Choose an Appropriate KOS	60
Summary	61
Chapter Review	62
Glossary	62
References	63
3. TECHNOLOGICAL STANDARDS	67
Major Questions	67
Introduction	67
Standards for Design and Construction of KOSs	68
Underlying Web Standards: XML and RDF	69
Standards for Representation and Interchange of KOS	73
ISO 2709	73
ISO 2709 in XML	77
XML Schema of ISO 25964	79
SKOS	80
Representation Standards for Ontologies	85
Other KOS Representation Formats	87
Other Related Standards	88
Protocols, APIs, Querying Languages	90
Identification of KOSs and Their Elements	92
Terminology Registries	94
Summary	95
Chapter Review	97
Glossary	98
References	98
4. AUTOMATED TOOLS FOR SUBJECT INFORMATION ORGANIZATION: SELECTED TOPICS	107
Major Questions	107
Bibliometrics and Subject Representation (by Fredrik Åström)	107
Linking Documents: Mapping Research Fields through Document Clustering	108
Linking Concepts: Word Clusters for Identifying Topical Relations	112
Combining or Relating Co-Word/Text and Reference/Citation Analyses	114
Data and Tools	115
Bibliometrics and Information in Forms Other than Scientific Publications	116

Automated Subject Indexing and Classification	117
Introduction (by Ingo Frommholz)	117
Automated Text Categorization and Clustering (by Ingo Frommholz and Muhammad Kamran Abbasi)	117
Automated Text Categorization	118
Document and Category Representation	120
Tokenization, Stopword Elimination, and Stemming	120
Term Weighting	121
Document and Category Representation	122
Automated Assignment	123
Evaluation	125
Text Clustering	127
Document Representation	127
Clustering Functions	129
Partitional Clustering: <i>k</i> -Means	129
Hierarchical Clustering	130
Evaluation	131
Further Reading	131
Machine Learning on Text (by Dunja Mladenić and Marko Grobelnik)	132
Machine-Learning Techniques	132
Learning Approaches	134
Supervised Learning	134
Semi-Supervised Learning	135
Unsupervised Learning	136
Beyond Basics	137
Discussion	138
Summary	138
Chapter Review	140
Glossary	141
References	143
5. PERSPECTIVES FOR THE FUTURE	149
Major Questions	149
Two Perspectives on Information Organization: Computer Science and Library and Information Science (by Isto Huvila and Fredrik Åström)	149
The Future	155
Summary	158
Chapter Review	159
Glossary	159
References	160
Index	161