Contents

List of figures *xi*
List of abbreviations *xiii*
Preface *xvii*
About the author *xxi*

1 Introduction to subject headings and thesauri **1**
Introduction 1
Standards for controlled vocabularies 4
Precoordination and postcoordination 8
General do's and don'ts in selecting index terms 10
Subject headings 15
Thesauri 18
Creating and maintaining a controlled vocabulary 25
How to find subject headings and thesauri 26
Thesaurus software 26
Multilingual thesauri 29
Interoperability between vocabularies 30
What makes a good indexing system? 32
Notes 35
References 35

2 Automatic indexing versus manual indexing **39**
Introduction 39
Arguments against manual indexing 40

Bibliografische Informationen
http://d-nb.info/1027386075

digitalisiert durch DEUTSCHE NATIONAL BIBLIOTHEK

Is indexing by the author or editor a valuable alternative? 48
Arguments in favour of manual indexing 51
Some misconceptions about automatic indexing 57
Conclusion 59
Notes 60
References 60

3 Techniques applied in automatic indexing of text material 65
Introduction 65
Lexical analysis 66
The use of stop word lists 67
Stemming 68
Extracting meaningful word combinations 70
Index term weighting 72
Linking words and word combinations to a controlled vocabulary 73
Automatic classification 75
What can be expected of automatic text indexing? 75
References 77

4 Automatic indexing of images 79
Introduction 79
Images on the internet 80
Context-based indexing 84
Content-based indexing 89
Automatic image annotation 96
Mixed techniques 97
The purpose of it all 98

Notes 99
References 99

5 **The black art of indexing moving images** **101**
Manual indexing of moving images 101
Why index moving images automatically? 103
Indexing based on speech or text recognition 104
Keyframe indexing 108
The future of video indexing 110
Notes 112
References 112

6 **Automatic indexing of music** **113**
Introduction 113
Some examples of music retrieval 115
Indexing methods behind the retrieval 118
References 120

7 **Taxonomies and ontologies** **121**
The librarian's strained relation to taxonomies and ontologies 121
What are taxonomies and what are they used for? 125
Ontologies 131
The importance of taxonomies and ontologies 138
Notes 141
References 141

8 **Metadata formats and indexing** **143**
Introduction 143
What are metadata? 144
Metadata and the library world 145
Some important metadata standards 146

Bridges between standards 160
The benefits of metadata standards 162
What about indexing? 164
Notes 165
References 166

9 Tagging 167
What is tagging? 167
Why tagging? 168
Advantages and disadvantages of tagging 171
Towards a taxonomy of tagging 172
Tagging in the book and library world 175
User tags and author keywords 179
How tags are displayed 180
Conclusions 181
Notes 182
References 182

10 Topic Maps 185
Introduction 185
The TAO model of Topic Maps 187
The technical side of Topic Maps 188
Examples of Topic Maps 190
Are Topic Maps the future of indexing? 191
Notes 192
References 192

11 Indexing the web 195
Is it possible to index the web? 195
Manual web indexes 197
Bookmark sites 203

Evaluation of manual web indexing 205
Web indexing by search engines 206
How search engines work 207
Google's PageRank 211
What about indexing the 'deep web'? 215
Notes 218
References 218

12 The Semantic Web 221
Introduction 221
The criticism against the actual web 223
Planning Web 3.0 225
A timetable for the Semantic Web 234
The Semantic Web and traditional library instruments 237
Notes 242
References 243

Index 245