LIBRARY AND INFORMATION SCIENCE TEXT SERIES

THE ORGANIZATION OF INFORMATION

SECOND EDITION

Arlene G. Taylor

CONTENTS

List of Figures xvii

Preface xxi

Acknowledgments xxv

1 Organization of Recorded Information 1
The Need to Organize 1
The Nature of Information 2
The Nature of the Organization of Recorded Information 3
Organization of Information in Different Environments 7
Libraries 7
Archives 9
Museums and Art Galleries 11
The Internet 13
Digital Libraries 15
Information Architecture 18
Data Administration 19
Knowledge Management 21
Conclusion 23
Notes 23
Suggested Readings 25
General 25
Organization of Information in Libraries 25
Organization of Information in Archives/Manuscripts . . . 26
Organization of Information in Museums/Art Galleries 26
Organization of Information in the Internet 26
Organization of Information in Digital Libraries . . . 27
Organization of Information in Information Architecture 27
Organization of Information in Data Administration 28

Organization of Information in Knowledge Management 28

2 Retrieval Tools 29
The Need for Retrieval Tools 29
The Basic Retrieval Tools, Their Formats, and Their Functions 30
Bibliographies 30
Pathfinders 33
Catalogs 33
Purposes of Catalogs 34
Forms of Catalogs 36
Arrangements Within Catalogs 38
Indexes 40
Finding Aids 43
Registers 44
Search Engines and Directories 44
Conclusion 45
Notes 46
Suggested Readings 47

3 Development of the Organization of Recorded Information in Western Civilization 49
Inventories, Bibliographies, Catalogs, and Codification 49
Antiquity 49
Middle Ages 51
European Renaissance 52
From Inventories to Finding Lists to Collocating Devices 53
Period of Codification 55
Twentieth Century 57
Description 57
Subject Access 59
Verbal Subject Access 59
Classification 60
Special Materials 61
Archives 61
Museums and Art Galleries 62
Subject Access to Special Materials 62

Mechanization of Bibliography 62
The Documentation Movement 63
Library Automation 65
Conclusion 65
Notes 66
Suggested Readings 67

4 Encoding Standards 69
Encoding of Characters 70
Encoding of Records (Syntax) 70
Currently Used Examples of Standards for Encoding Records 71
MARC (MAchine-Readable Cataloging) 72
MARC 21 76
UNIMARC 78
SGML (Standard Generalized Markup Language) 78
XML (eXtensible Markup Language) 79
DTDs (Document Type Definitions) and XML Schemas 80
Frameworks 90
Warwick Framework 90
RDF (Resource Description Framework) 93
METS (Metadata Encoding & Transmission Standard) . . . 95
Semantic Web 96
Conclusion 97
Notes 97
Suggested Readings 100

5 Systems and System Design 103
Systems 103
Databases 104
Bibliographic Networks 106
Integrated Library Systems (ILSs) 108
Development of Online Public Access Catalogs (OPACs) 109
System Design 111
Organization of Information and System Design 112
Searching Methods 113
Retrieval Models 114

Standardization and Systems . . . 115
Display . . . 116
Basic Search Queries . . . 119
Initial Articles . . . 121
Truncation, Boolean Operators, and Proximity . . . 121
Punctuation . . . 122
Meta-Searching and Z39.50 . . . 123
User-Centered System Design . . . 125
Universal Design . . . 126
Multiple Languages/Scripts . . . 126
Other Aids for Users . . . 127
Authority Control Integration . . . 129
Conclusion . . . 131
Notes . . . 132
Suggested Readings . . . 135

6 Metadata . . . 139
The Basics of Metadata . . . 141
Metadata Schemas . . . 142
Metadata Characteristics . . . 143
Metadata and Cataloging . . . 144
Objectives of an Information System . . . 146
Types of Metadata . . . 147
Administrative Metadata . . . 147
Preservation Metadata . . . 148
Rights and Access Metadata . . . 149
Meta-Metadata . . . 150
Structural Metadata . . . 150
Implementations of Structural Metadata . . . 151
Descriptive Metadata . . . 152
Management Tools . . . 152
Application Profiles . . . 152
Metadata Registries . . . 153
Crosswalks . . . 154
Harvesting Tools and Templates . . . 155
Conclusion . . . 155
Notes . . . 156
Suggested Readings . . . 158

7 Metadata: Description . . . 159
Units to Be Described . . . 160

Finite vs. Continuing Resources 161
FRBR's Entities 162
Creation of Surrogate Records 163
Bibliographic and General Metadata Schemas 165
ISBD (International Standard Bibliographic Description) . . 165
Anglo-American Cataloguing Rules, Second Edition, 2002 Revision (AACR2R) 168
The Dublin Core (DC) 171
MODS (Metadata Object Description Schema) . . . 176
Domain-Specific Metadata Schemas 177
ISAD(G) (General International Standard Archival Description) 177
Archival *APPM (Archives, Personal Papers, and Manuscripts)* Records 180
EAD (Encoded Archival Description) 180
TEI (Text Encoding Initiative) Headers 181
GILS (Government [or Global] Information Locator Service) Records 182
FGDC (Federal Geographic Data Committee) *Content Standard for Digital Geospatial Metadata* . . . 184
VRA (Visual Resources Association) Core Categories for Visual Resources 185
CIMI XML Schema for SPECTRUM 188
ONIX International 189
Other Surrogate Record Types 189
Index Records 189
On-the-Fly Records 190
Environmental Influences in Descriptive Metadata Creation . . 190
Conclusion 192
Notes 192
Suggested Readings 196

8 Metadata: Access and Authority Control 201
Need for Attention to Access 201
Need for Attention to Bibliographic Relationships 203
Choice of Access Points 206
Primary Access Point 207
Main Entry Controversy 208
Justification for Main Entry 209
AACR2 Principles for Choosing Main Entry 214

Additional Access Points 220
Authority Control 220
Headings for Access Points 224
Principles for Choice of Personal Name 224
Principles for Form of Personal Name 226
Principles for Entry Word and Remaining Structure of Personal Name 227
Principles for Choice of Corporate Name 228
Principles for Entry Word and Form of Heading for Corporate Names 229
Principles for Choice of Uniform Title 230
Principles for Arrangement of Uniform Titles 230
International Authority Control 231
Conclusion 233
Notes 235
Suggested Readings 238

9 Subject Analysis 241
Purpose of Subject Analysis 242
Challenges in Conceptual Analysis 242
Cultural Differences 244
Differences in Methods Used 244
Consistency 245
Nontextual Information 246
Subject Analysis Process 247
Exhaustivity 250
Identification of Concepts 252
Topics Used as Subject Concepts 252
Names Used as Subject Concepts 252
Chronological Elements as Subject Concepts 253
The Concept of Form as Subject 254
Translating Concepts into Index Terms 255
Translating Concepts into Classification Notations 256
Conclusion 256
Notes 256
Suggested Readings 258

10 Systems for Vocabulary Control 261
Controlled Vocabulary Challenges 262
Specific vs. General Terms 262

Synonymous Concepts . . . 262
Word Form for One-Word Terms . . . 263
Sequence and Form for Multiword Terms and Phrases . . . 263
Homographs and Homophones . . . 264
Qualification of Terms . . . 265
Abbreviations and Acronyms . . . 265
Popular vs. Technical Terms . . . 265
Subdivision of Terms . . . 266
Precoordination vs. Postcoordination . . . 266
General Principles for Creating Controlled Vocabularies . . . 267
Specificity . . . 267
Literary Warrant . . . 268
Direct Entry . . . 268
General Principles for Applying Controlled Vocabulary Terms . . . 268
Specific Entry . . . 268
Number of Terms Assigned . . . 269
Concept Not in Controlled Vocabulary . . . 269
Index Terms for Names . . . 270
Controlled Vocabularies . . . 270
Mechanics of Controlled Vocabularies . . . 271
Subject Heading Lists . . . 274
Library of Congress Subject Headings (LCSH) . . . 274
Sears List of Subject Headings (Sears) . . . 275
Medical Subject Headings (MeSH) . . . 275
Thesauri . . . 277
Art & Architecture Thesaurus (AAT) . . . 278
Thesaurus of ERIC Descriptors . . . 278
INSPEC Thesaurus . . . 281
Ontologies . . . 282
Natural Language Processing (NLP) . . . 284
Keywords . . . 286
WordNet® . . . 287
Conclusion . . . 288
Notes . . . 290
Suggested Readings . . . 292
LCSH . . . 293
Sears . . . 294
MeSH . . . 294
AAT . . . 294

ERIC 294
Ontologies 295
NLP 295

11 Systems for Categorization 297
Theory of Categorization 297
Classical Theory of Categories 298
Prototype Theory 300
Bibliographic Classification 301
Hierarchical, Enumerative, and Faceted Classifications 302
Major Bibliographic Classification Schemes 305
Classification Concepts 307
Broad vs. Close Classification 307
Classification of Knowledge vs. Classification of a Particular Collection 308
Integrity of Numbers vs. Keeping Pace With Knowledge 309
Fixed vs. Relative Location 311
Closed vs. Open Stacks 311
Location Device vs. Collocation Device 312
Classification of Serials vs. Alphabetic Order of Serials . . 313
Classification of Monographic Series (Classified Separately vs. Classified as a Set) 314
Taxonomies 315
Classification on the Internet 316
Artificial Neural Networks (ANNs) 320
Conclusion 322
Notes 322
Suggested Readings 324
Classification in General 324
Specific Classification Schemes 326
Dewey Decimal Classification 326
Universal Decimal Classification 326
LC Classification 327
Colon Classification 327
Other Classification Schemes 327
Taxonomies 328
Classification and the Internet 328
Artificial Neural Networks (ANNs) 329

12 Arrangement and Display 331
Arrangement of Physical Information Packages 331
Libraries 331
Archives 333
Arrangement of Intangible Information Packages 334
Arrangement of Surrogates (Metadata) 335
Filing History 336
General Rules for Arrangement 337
Filing/Display Dilemmas 338
Conclusion 341
Notes 341
Suggested Readings 342
Arrangement of Information Packages 342
Libraries 342
Archives 342
Internet 342
Arrangement of Bibliographic/Surrogate Records 343

Conclusion 345
Note 346

Appendix: Subject-Analysis Application 347

Glossary 353

Selected Bibliography 385

Index 407