Structures for Organizing Knowledge: Exploring Taxonomies, Ontologies, and Other Schemas

June Abbas
To Charlie, Ben, and Clifford
My inspirations for organizing life
Contents

List of Figures and Tables xi
Preface xv
Acknowledgments xxi

PART I. Traditional Structures for Organizing Knowledge

Chapter 1. Introduction to Structures for Organizing Knowledge 3
   Introduction 3
   Framework for Exploring Organizing Behavior 5
      Scope 6
   Definitions 6
   Structure 6
   Information and Knowledge 8
   Conclusion 14
   Thought Exercises 14
   References 15
   Websites Noted in Text 16

Chapter 2. Historical Perspectives and Development of Structures for Organizing Knowledge 17
   Introduction and Scope of the Chapter 17
   Informal Structures 18
   Formal Structures 18
   Informal and Formal Structures 19
   History of Structures for Organizing Knowledge 20
   Early Taxonomists and Their Contributions 21
      Early Bibliographic Traces 21
      Later Developments in Library and Information Science 22
         Cataloging Codes 23
         Controlled Vocabularies 24
         Classification Schemes 26
   Contributions from Philosophy 29
   Contributions from Natural History 30
   Contributions from Cognitive Science 34
PART II. Personal Structures for Organizing Knowledge

Chapter 5. Structures for Organizing Knowledge in Personal and Professional Contexts

Introduction and Scope of the Chapter

Knowledge Structures and Knowledge Spaces within Personal Contexts

Defining Personal Knowledge Spaces

Research Focused on Personal Knowledge Organization

Organization of Personal Knowledge Spaces in the Home Environment

Knowledge Structures and Knowledge Spaces within Professional Contexts

Defining Professional Knowledge Spaces

Research Focused on Professional Knowledge Organization

Research on Paper-Based Knowledge Organization

Research on Digital Knowledge Organization
PART III. Socially-Constructed Structures for Organizing Knowledge

Chapter 6. Social Knowledge-Organizing Behaviors and Socially-Constructed Structures for Organizing Knowledge: Research and Discussion

Introduction and Scope of the Chapter

Introduction to Social Networking, Bookmarking, and Cataloging Sites

Definitions

History of Social Networking, Bookmarking, and Cataloging Sites

Examples of Social Bookmarking and Social Cataloging Sites

Social Bookmarking Sites: Delicious and Flickr

Social Cataloging Sites: LibraryThing

Research Related to Social Knowledge-Organizing Behaviors and Socially-Constructed Knowledge Spaces

Research Related to Exploration of Tags and Social Structures for Organizing Knowledge

Learning More about Site Use and Tagging Activities

Nature of Tags, Tagging, and Folksonomies

Comparing Tags and Folksonomies to Controlled Vocabularies

Conclusion

Thought Exercises

References

Websites Noted in Text

Chapter 7. Extending Our Thinking: Creating a Structure for Organizing Knowledge from Various Threads

Introduction and Scope of the Chapter

Review of Traditional, Personal, and Socially-Constructed Threads

Traditional Structures for Organizing Knowledge

Personal Structures for Organizing Knowledge

Socially-Constructed Knowledge Structures

Practical Applications of the Threads

Integration of Tags and Folksonomies within Library OPACs

Integration of Tags and Folksonomies within Controlled Vocabularies

Refining Folksonomies

The Missing Part of This Conversation
List of Figures and Tables

FIGURES

Figure 1.1. Ornithology Collection at the National Museum of Natural History 4
Figure 1.2. Bentley’s Snowflakes Metadata Scheme 7
Figure 1.3. Representing Bentley’s Snowflakes for Multiple Users 9
Figure 1.4. From Data to Wisdom 12
Figure 1.5. Abbas’s Model of Age-Appropriate Indexing 13
Figure 2.1. Examples of Different Forms of Knowledge-Organizing Structures 20
Figure 2.2. Example of a Cladogram 34
Figure 2.3. Berlin’s Folk Classification Hierarchy 38
Figure 3.1. ISO Standards Development Process 49
Figure 4.1. Linnaean Hierarchy for Homo sapiens 71
Figure 4.2. Article 1 of the Botanical Code 74
Figure 4.3. Rules 5a and 5b from the Bacteria Code 76
Figure 4.4. Rule 6 and Part of Rule 7 from the Bacteria Code 77
Figure 4.5. Rules 15 and 16 from the Bacteria Code 78
Figure 4.6. Table of Contents of the Zoological Code 80
Figure 4.7. Article 3 from the Zoological Code 81
Figure 4.8. Article 5 from the Zoological Code 82
Figure 4.9. Subject Cloud Interface on the Botanicus Digital Library Homepage 87
Figure 4.10. Sample Page from Botanicus Showing Results of Search for “Systema Vegetabilium” 88
Figure 4.11. Excerpt from Darwin Core XML Schema 89
Figure 4.12. uBio NameBank Homepage 90
Figure 4.13. uBio ClassificationBank Homepage 91
Figure 4.14. LSID Homepage 92
Figure 4.15. GBIF Data Portal Search for Palliolisenitis quinqueungulis 94
Figure 4.16. Overview Screen for Palliolisenitis quinqueungulis in the GBIF Data Portal 95
Figure 4.17. ITIS Report on *Palliolisentis quinqueungulis* Machado, 1960

Figure 4.18. Encyclopedia of Life Homepage

Figure 4.19. Record in the Encyclopedia of Life

Figure 4.20. ZooBank Record for *Tectodamaeus longus* Registered 9-17-2009

Figure 4.21. Record Structure for a ZooBank Publication Record

Figure 4.22. Descriptive and Subject-Related Elements of a Bibliographic Record

Figure 4.23. Example of a Bibliographic Record in a Library OPAC

Figure 4.24. MARC Record for *Library of Congress Subject Headings* by Chan

Figure 4.25. Dublin Core General Elements

Figure 4.26. EAD Finding Aid from the Library of Congress

Figure 4.27. Example from LCSH for Term “Cataloging” Showing Syndetic Structure

Figure 4.28. Example of Scope Note and Parenthetical Qualifier in Entry for Term “Speech Disorders” from LCSH

Figure 4.29. Example of Subject Heading and Relationships for the Term “Computer Adventure Games” from LCSH

Figure 4.30. Juvenile Subject Headings Entry for the Term “Poison Frogs”

Figure 4.31. Results of Search in DDC for the Term “Frogs”

Figure 4.32. “Frogs—Resource Economics” Traced

Figure 4.33. “Dendrobatidae” Traced

Figure 4.34. OCLC Authority File Record for “Dendrobatidae”

Figure 4.35. Correlation to LCSH within the OCLC Classification Web

Figure 4.36. Classification of “Dendrobatidae” in LCC

Figure 4.37. Bloom’s Taxonomy

Figure 4.38. Bloom’s Taxonomy Revised to Illustrate Affective Domain

Figure 5.1. Representation of Personal Knowledge Space

Figure 5.2. Categories of Taxonomies, Thesauri, and Ontologies from the Taxonomy Warehouse

Figure 5.3. Commerce & Industry Category on Taxonomy Warehouse

Figure 5.4. Vocabulary Details of the WAND Industrial Equipment and Supplies Taxonomy

Figure 6.1. Delicious Popular Bookmarks

Figure 6.2. Delicious Tag Cloud: Popular Tags
Figure 6.3. Flickr Features for Describing an Image 183
Figure 6.4. Personal Tag Cloud in Flickr 184
Figure 6.5. Personal Tag “Big List” in Flickr 185
Figure 6.6. Three Threads Merged in the LibraryThing Record Structure 186
Figure 6.7. LibraryThing Record for the Book Everything Is Miscellaneous 187
Figure 6.8. LibraryThing Details View for the Book Everything Is Miscellaneous 188
Figure 6.9. Adoption/Development/Use Cycle 189
Figure 6.10. Gaps in Socially-Constructed Structures Research 199
Figure 7.1. Knowledge Threads: Additional Strands to Consider 207
Figure 7.2. FRBR Model: First Group Entities 211
Figure 7.3. Knowledge Threads: Tapestry of Possibilities 223
Figure 8.1. MARC 5XX Notes Fields 231
Figure 8.2. Factors in Developing New Structure 234

TABLES

Table 4.1. Examples of Forms of Subject Headings in LCSH 116
Table 4.2. Notes within the DDC Scheme 124
Table 4.3. Colon Classification Classes 131
Structures for organizing knowledge are an always present, but often unrecognized, part of our everyday personal and professional lives. They are the backbones behind the systems we use for work, for play, and for both physical and intellectual exploration. They not only help organize and retrieve “objects” in physical spaces like libraries, but they also give us the frameworks we use for making new connections, understanding disciplines, and making sense of professional and personal knowledge spaces.

Very young children begin organizing and grouping objects in their physical spaces. Pot lids and plastic containers strewn on the kitchen floor may look haphazard to the harried parent, but when one looks at the underlying structure, it is possible to see how children group everyday objects according to size, shape, color, texture, use, or even the sound they make. These groupings might be viewed as the beginnings of how children initially make sense of—and learn about—physical and mental spaces by developing and applying structures for organizing knowledge.

Library and information science professionals use structures for organizing knowledge when they catalog and classify objects in their collections, when they develop databases, or when they design customized taxonomies for application within a business context. Organizing structures are used when searching online to find articles in a database or travel information on the web, when browsing the shelves in a library or a favorite bookstore, or when downloading music and creating playlists. For example, a person planning a special event might create one playlist that holds music intended for use at the event, or a teen might create a playlist for music to listen to while studying for a test. At various points during every day, everyone organizes information and knowledge and interacts with structures for organizing knowledge.

Purpose and Scope

*Structures for Organizing Knowledge: Exploring Taxonomies, Ontologies, and Other Schemas* explores and explains how we organize knowledge by looking at three broad questions: (1) How do people organize objects in personal and professional contexts so that they make sense and are useful? (2) What roles do categories, classifications, taxonomies, and other structures play in the process of organizing? (3) What do information professionals need to know about human
organizing behaviors in order to design useful structures for organizing knowledge?

**Organization**

To answer these questions, the book is organized using three major threads:

- *Traditional Structures for Organizing Knowledge*—Part I looks at structures used in libraries, such as MARC records, subject headings, and classification schemes, as well as traditional structures that may not be as familiar, such as those from natural science. The historical contributions to the organization of knowledge from fields such as library and information science, philosophy, natural science, and cognitive science are examined. Exemplars of how the structures have remained the same and/or have been adapted for use in the digital environment are also included in this section.

- *Personal Structures for Organizing Knowledge* are the focus of Part II. These are systems developed by individuals in both home- and work-related contexts. Several research streams from library and information science (knowledge organization and human information behavior) and human–computer interaction (personal information management) are introduced, and research in each area of personal knowledge structures is explored.

- *Socially-Constructed Structures for Organizing Knowledge*, or those that are beginning to emerge as the result of individual and collaborative uses of social bookmarking and social cataloging Web 2.0 sites, are examined in Part III. Research focused on these new environments is becoming more prevalent and providing information professionals with a glimpse into how people organize their own collections.

In Part I, which includes Chapters 1–4, the historical and multidisciplinary underpinnings of the development of traditional organizing structures such as categories, classification schemes, taxonomies, ontologies, controlled vocabularies, and others encountered along the way are reviewed. This study is not limited to those structures developed by library and information science, but, rather, a broad interdisciplinary net is cast in order to examine early structures that have informed current development of structures for organizing knowledge. Chapter 1 sets the stage for the explorations by introducing the reader to the concepts and theories that are central to understanding information, knowledge, and the representation and organization of information and knowledge. In Chapter 2 the contributions of multiple disciplines to the understanding of how and why people organize information and knowledge are examined. Conceptual frameworks and working examples of organizing structures of multiple disciplines such as philosophy, cognitive science, biology, and library and information science increase our understanding about human organizing behavior and application.
Organizing structures cannot be fully examined without also understanding the standards that are followed during development of the structures. Chapter 3 examines national and international standards related to the development and application of structures for organizing knowledge within several disciplines, such as those standards for thesaurus construction, controlled vocabulary development or subject headings list development, or specialized metadata schemes. Best practices for developing and applying organizing structures are also discussed.

Chapter 4 examines the principle functions of knowledge structures within disciplinary contexts through examination of different taxonomies in use and the standards developed to ensure the proper application of the knowledge structure. Chapter 4 presents an overview of the knowledge structures used within a sample of academic disciplines, including biology, zoology, social science, and library and information science. Examples of how the knowledge structures have translated into applications available in the digital environment, e.g., online databases for taxonomic identification, discipline-specific digital libraries of specimens or documents, and online archives, are reviewed as exemplars.

Part II addresses the questions of how people organize objects in personal and professional contexts. Chapter 5 outlines the research from fields, such as library and information science and computer science, related to how people organize information and knowledge in personal and professional contexts. Organizing structures such as categories, taxonomies, and ad hoc “artfully designed” organizing systems provide a means to organize the objects in professional and personal lives but also enable discovery and access of information and objects needed to satisfy information needs, to make purchases, to conduct work, or, quite simply, to live life. The development and use of organizing structures in everyday personal and professional knowledge spaces are explored through the lenses of personal information management (or PIM) and domain analysis.

Part III, composed of Chapters 6–8, introduces readers to socially-constructed knowledge organization. Chapter 6 discusses online social sharing sites, such as LibraryThing, Delicious, and Flickr. The Web 2.0 world has brought the professionally-constructed and personal organizing structures together in novel ways. Individuals now collect and organize personal information spaces online in social sharing sites for pictures, home collections of books, web bookmarks, and other multimedia such as videos, music files, podcasts, etc. Sites like Flickr allow for personal collections of pictures to be shared with friends, family, and online global contacts. LibraryThing provides an information space to organize and describe personal collections of books while also integrating professional organizing structures like the Library of Congress Subject Headings through what is called “tagmashing,” or combining personal organizing structures with preestablished subject access schemes developed by library and information science professionals. This Web 2.0 context can be thought of as a Petri dish for researchers to learn more about how people orga-
nize personal information spaces, how they describe their collections and personal objects, and how they develop organizing structures individually and collaboratively. Chapter 6 also presents a framework for examining the research and perspectives focused on tagging, bookmarking, and social sharing behaviors, along with research related to social construction of structures for organizing knowledge.

Last, in Chapters 7 and 8, ways in which to connect the three threads of (1) traditional professionally-developed organizing structures, (2) personal structures for organizing knowledge spaces, and (3) socially-constructed organizing structures are investigated. In Chapter 7 the practical applications of shared folksonomies, both the impacts these potentially new structures will have on everyday professional and work-related lives and the implications for library and information science and related professions and disciplines, are explored. Practical applications, such as how to integrate socially-constructed technologies (in this instance, tags and their resulting folksonomies) into existing structures for organizing knowledge, such as library OPACs or existing controlled vocabularies (e.g., subject headings lists, thesauri), have the potential to change how knowledge is represented in libraries and other information organizations, but little is known at this point about the efficacy of doing so. Methods for refining or enhancing folksonomies, such as automatic and manual methods of “tag gardening” or reviewing and augmenting personal and/or community-based folksonomies, also are being explored by researchers in information science and computer science. Chapter 7 reviews the research into these issues and discusses the potential of integrating socially-constructed knowledge structures into traditional knowledge structures.

Chapter 8 weaves the three threads together through the use of two envisioning exercises designed to help the reader think through some of the issues and potentials of merging the three threads into new structures for organizing knowledge. First, readers are asked to construct their image of today’s users, taking into account what they learned throughout the book and incorporating users’ uses of technology into the picture. Second, readers will envision what a new structure for organizing knowledge might look like if the three threads of traditional (represented by the MARC record structure), PIM insights, and socially-constructed structures (represented by LibraryThing’s Common Knowledge fields and others for readers to add) are woven together.

Each chapter follows the same structure: a box of bulleted “Focus Points” of issues and thoughts for the reader to consider, the main topics (as described earlier) along with helpful illustrations or “snapshots” of the different structures for organizing knowledge being discussed or used as examples, and a set of thought exercises designed to help the reader review and apply the topics being covered.
Audience

Structures for Organizing Knowledge is written for students of library and information science as well as for the practicing library and information science professional. The book is not meant to be a “how-to” guide for developing, applying, or implementing structures for organizing knowledge; rather, it is designed to present a conceptual discourse and to inspire thinking about taxonomic behavior, or how and why people organize knowledge, in various contexts. It also serves as a textbook on the historical development of structures for organizing knowledge and the current interdisciplinary theories and research related to the creation and application of structures for organizing knowledge. Furthermore, the book encourages readers to explore the new phenomenon of socially-constructed structures for organizing knowledge, such as those emerging in Web 2.0 social bookmarking and social cataloging sites, and to postulate how these socially-constructed structures for organizing knowledge may affect not only the existing structures already in use but also those developed in the future. Because library and information science professionals not only use structures for organizing knowledge but also develop and implement them, they must be thoroughly aware of how these structures function and how the structures are useful in differing contexts. To understand how to effectively develop and use structures for organizing knowledge, it is important that library and information science students and professionals understand the conceptual framework and rich history of structures for organizing knowledge, as well as the past and current understanding of organizing behaviors.

A secondary audience for the work is that of researchers in library and information science and related fields. Researchers will find the book’s discussions of classical and newly developing folksonomies and metadata schemes and the reviews of the research literature useful to their work or to spark new ideas for research.

The three threads explored in the book present avenues for expanding our thinking on classification and classification schemes, taxonomy and ontology development, and structures for organizing knowledge. Learning more about how people naturally organize, categorize, name, and use organizing structures for making sense of their physical and mental spaces is important as the impacts of Web 2.0 and Library 2.0 developments and changes in user expectations are evaluated. This book will create new connections and ideas as historical perspectives are revisited, as more is understood about why and how people are taxonomists, and as the potentials of socially-constructed taxonomies are explored.