Book Review

Edward Swanson

Metadata. By Marcia Lei Zeng and Jian Qin. New York: Neal-Schuman, 2008. 365p. $65.00 softbound (ISBN 978-1-55570-635-7/1-55570-635-5).

Professors Marcia Lei Zeng and Jian Qin each have several years of experience as professional educators and researchers in the fields of knowledge organization and metadata (365). One fruit of their collective labors is the recently published *Metadata*, a comprehensive assessment of the theory and practice of metadata design and implementation. The authors have endeavored to take a "unique, enlightening, and holistic approach" (xvii) to the subject of metadata by fashioning a dual-purpose manual: a textbook for students and an "instructional guide" (xv) for metadata specialists.

Metadata is divided into four parts. The first is dedicated to basic principles and definitions and includes chapter 2, "Current Standards," a survey of the semantics and structures of various metadata schemes and content standards, many common in library and related institutions (e.g., Machine-Readable Cataloging (MARC), Encoded Archival Description (EAD), and Content Standard for Digital Geospatial Metadata (CSDGM)), others less popular (e.g., vCard and Sharable Content Object Reference Model (SCORM)). The Text Encoding Initiative (TEI) scheme, notably the metadata-centric header, is mentioned little in this chapter, with occasional references elsewhere in the volume. Notwithstanding the authors' statement that no partiality is intended with a scheme's exclusion (xvi), considering the popularity of TEI with digitized text and the historical role its header played as a model for EAD's own header, I still find this omission surprising.

Part 2 moves away from definitions to concentrate on constructs. Chapter 3, "Schemas—Structure and Semantics," offers an overview of element sets, application profiles, crosswalks, and best practices documentation. The following chapter, "Schemas—Syntax," discusses the encoding of metadata for computer manipulation while positing the advantages of Extensible Markup Language (XML). Chapter 5, "Metadata Records," brings together principles from the previous chapters to lay the conceptual groundwork for—and describe the encoding of—standards-based metadata records.

Metadata registries, repositories, and metadata sharing are the focus of "Metadata Services," the first chapter of part 3. Zeng and Qin continue their discussion of interoperability in chapter 8, "Achieving Interoperability," concentrating here on issues regarding metadata content and schemas that are distributed within or among metadata repositories. Between these two chapters lies "Metadata Quality Measurement and Improvement" (chapter 7), the content of which is self-evident from the title. I make special note of the sections therein devoted to (1) a description of the varying methods of analysis as a means to achieving quality metadata, and (2) the practical and, in my reading, implied ethical ramifications of poor metadata.

Part 4 consists of a single chapter devoted to assorted topics on metadata research, such as investigations on semantics and conceptual metadata modeling.

A pair of appendixes caps off the work. The first is an annotated list of various metadata schemes; the second is a short compendium of sources for controlled vocabularies, content standards, and best practices guides. These are followed by a well-rounded glossary, a rich bibliography (with many citations pointing to online resources), and a fine index.

Furthermore, Zeng and Qin offer a variety of aids throughout the work to assist the reader in learning specific concepts and practices. An array of helpful illustrations emphasizes or further delineates their points, typically taking the form of screenshots of applications, diagrams illustrating various principles, and tables of data. Each chapter ends with a selection of exercises for the student, all of which are also found on a companion website (www.metadataetc.org/book-website). The site likewise presents links to online information resources on metadata (duplicating the bibliographical references found at the end of each chapter in the book), quizzes, and further exercises. (These supplementary online materials are also available from the publisher on CD-ROM. I did not receive a copy of the disc for review.) The website in turn points to a wiki (www.metadataetc.org/wiki/index.php5?title=Main_Page) that supplies information on updates to the companion website, offers an online version of the book's glossary, and devotes a section to listing errata found in the textbook.

Metadata covers much ground, as the brief outline above demonstrates. Considering the scope of the work, I posit the question, Did Zeng and Qin succeed in their mission to create a suitable textbook for the metadata student? The answer is yes, with some qualifications.

First, a hazard of employing narrative form in textbooks is the sometimes brief, early presentation of a concept not solidly defined until later

in the text. This pitfall appears several times in *Metadata*. For example, the authors report on application profiles for education communities (49–50) well before they give a thorough account of the meaning and uses of application profiles (112–19). Often the reader can resolve any confusion with a consultation of the glossary. Many of the important terms the writers bring to bear in the text are defined here; definitions of lesser substance they confine only to the text body itself (e.g., the data encoding terms *wrapper* and *compound elements* and the concept of ontological modeling). When encountering such terms, the student must employ the index, thus establishing an inconvenient route to discovering the meaning of various unfamiliar expressions.

Next, Zeng and Qin acknowledge that their book does not provide a primer for encoding data for computer use (134). *Metadata* therefore requires some prerequisite knowledge of XML and Extensible Hypertext Markup Language (XHTML) to understand the encoding examples. This is one reason I caution against the neophyte, especially one lacking a technical background, using this text for self-study. For them, the writing style may be dense and at times opaque; for example, one section I find daunting—and the newcomer most likely also will—is the short but thickly technical explanation of the MPEG-7 standard, with which I was not previously familiar (77–79). Students here would benefit from the assistance of a mentor or practiced work colleague or consultation of other resources to understand the material put forward in such a concentrated manner.

These minor criticisms aside, *Metadata* and its companion websites offer an excellent foundation for advanced class instruction; the guiding hand of a knowledgeable classroom instructor is required to traverse some of the more difficult subject matter. Moreover, the writers explicitly state in the preface that "the text is not a step-by-step manual for creating metadata records" (xv–xvi). Most metadata course syllabi with which I am familiar call for extensive engagement with specific metadata content standards and records. Thus an instructor using this textbook in such a circumstance may find it challenging to provide in a single semester coursework on these standards in tandem with a full account of the readings offered in *Metadata*. An option to consider is a separate course for the practical application of metadata, which is alluded to by the authors (16).

The other audience to whom Zeng and Qin direct their work, metadata practitioners with appropriate background experience, will find this book a very good source for reference, review, and for expanding their knowledge on facets that lie outside their respective fields of expertise.—*Mark K. Ehlert (ehler043@umn.edu), MINITEX Library Information Network, Minneapolis.*