

3-2017

Metadata (Book Review)

Gregory Wiedeman

University at Albany, State University of New York, gwiedeman@albany.edu

Follow this and additional works at: http://scholarsarchive.library.albany.edu/ulib_fac_scholar

Recommended Citation

Wiedeman, Gregory, "Metadata (Book Review)" (2017). *University Libraries Faculty Scholarship*. 96.
http://scholarsarchive.library.albany.edu/ulib_fac_scholar/96

This Book Review is brought to you for free and open access by the University Libraries at Scholars Archive. It has been accepted for inclusion in University Libraries Faculty Scholarship by an authorized administrator of Scholars Archive. For more information, please contact scholarsarchive@albany.edu.

Metadata. By Marcia Lei Zeng and Jian Qin. 2nd ed. Chicago: Neal-Schuman, 2016. 555 pp. Glossary, index. Softcover. \$84.00.

The second edition of *Metadata* by Marcia Lei Zeng and Jian Qin is a substantial leap forward for graduate texts, and if information science faculty plan to rely on one primary text on the subject, this is a worthy choice. Unfortunately, this said, *Metadata* might not be for archivists. Although the writing can be dense at times, substance-wise it is comprehensive, thorough, and cutting edge. Its pragmatic focus often makes complex topics readily accessible for students, and the authors' integration of open-linked data principles throughout the text is supremely effective. However, a practicing archivist will notice a persistent divide between the assumptions made by the text and the way metadata is handled in archives. The central conflict is between archival description and the more specific "metadata" covered in this book that envisions flat storage defined by the "one-to-one principle." This gulf needs to be bridged, as both communities can learn from one another.

Metadata is a very effective core text for graduate coursework. Its main strengths are its use of practical and relatable examples and its focus on the process of and decision making in applying metadata standards, rather than on the standards themselves. Readers get important real-world knowledge of what it is like to establish a metadata program for a digital project. The authors also incorporate recent scholarship throughout the text. They provide a clear and understandable overview of the many advances that have changed how we understand and make use of metadata.

The focus on cutting-edge literature is another major strength of the book, and it solidifies *Metadata* as a scholarly text that makes a real contribution to the field by providing an overview of how all this research ties together. It seems that the authors also intend the book to be a resource for practicing professionals. The substance of the text should lend it to this, but the somewhat dense writing can undermine this goal. This is not a book to be read straight through, but should instead be studied and discussed in sections. The dense text can be somewhat understandable considering its detailed coverage and advanced subject matter. The book seems to meet the needs of current graduate information studies programs that offer coursework in metadata, cataloging, or the subject with the awkward moniker, "digital libraries." Yet, it would seem that the skills and practices that the text details well would be useful for a wide range of courses that do not focus solely on metadata. As effective as it can be, it might be difficult to casually incorporate a text as large and broad as *Metadata* into other subjects where it is greatly needed.

Metadata's coverage of linked data is particularly effective. The book takes time to address how the traditional focus on *records* is being replaced by *statements* as the fundamental unit of metadata and then consistently reminds readers of the relevance of these ideas later in the text (pp. 78–79). Here the authors' commitment to practical examples lends itself well. Their coverage of Schema.org, in particular, is relevant, engaging, and useful (pp. 128–29).

The authors consistently focus on pragmatism over idealism, accepting that ". . . there

will always be a gap between the individual's practice and the general best practices" (pp. 214, 349). Chapter 4 focuses on the decision-making process for selecting and applying metadata structure standards, and it correctly assumes that, in most cases, information professionals are adapting monolithic standards to their own local realities (pp. 157–59). This gives readers a feel for what it is like to work with metadata in the real world. *Metadata* accepts that working with standards is ad hoc, but still encourages strict regulation, simplicity, and a firm reliance on functional requirements rather than a "sum of the parts" approach. The authors offer up the concept of application profiles for how users can standardize their local use while providing machine-readable documentation, while also being open to alternatives (p. 22). They also offer a basic overview of systems analysis and data modeling at a time when this work is becoming more and more relevant in LAM communities (pp. 160–173).

Metadata is very much a product of current graduate study in the field of library and information studies, and it continues some of the biases that have long been prevalent in the profession. It calls the establishment of Dublin Core a "historically-significant workshop," when it is still readily ignored outside of libraries (p. 16). XML is featured extensively—and almost exclusively—throughout the text, and the authors seem to think that all metadata (apart from perhaps MARC) is and should be expressed in XML (pp. 131–33). The similar JSON is only mentioned offhandedly as JSON-LD once or twice. The text also often underestimates the important distinction between metadata as serialized text and metadata stored in systems and governed by data models. Most important, the book continues the widely held practice of disproportionately focusing on structure standards over content standards. The authors discuss the importance of content standards often throughout the text, even once describing how a lack of effective best practices on the content side can jeopardize the most thorough application of metadata schemas (p. 246). Yet, most notably, they open the large final reference chapter on metadata standards by declaring content standards to be out of scope (p. 401). Still, I overstate these criticisms here, as the same issues seem to be even more prevalent in just about every information studies text. The authors also deserve credit for always using popular and relevant technology in their examples—a standard that similar works do not always match.

The most glaring issue with *Metadata* is also more of a criticism of the divide between libraries and archives in how each views metadata in theory and uses metadata in practice. The major conflict here is the assumption in the text that all metadata should follow the one-to-one principle. This idea that every object must have a corresponding record—or, preferably, a statement—runs contrary to archival practices that use hierarchical metadata to describe objects, in context, at whatever level is appropriate given resource constraints. It is certainly true that one collection always demands one record. Yet series, file, or even item descriptions are metadata records as well and need to be treated as such when it is resource appropriate.

Information professionals outside of archives may not see archival description as metadata. Typical archival practices certainly do not apply the same level of rigor at lower levels as a flat "digital library" does. This is because archives integrate resource

management into their descriptive practices, as not every archival object deserves the same level of care in a reality where staff time has limits. While its largest chapter focuses on metadata creation, *Metadata* contains no discussion of resource management.

While readers of *Metadata* can certainly learn from archival practices, archivists can also learn from flatlanders in other cultural heritage institutions. There is a widespread assumption in archives—with backlogs and ever-expanding volumes of records—that detailed lower-level description is almost always absurd. Additionally, one-to-one style rigor is also not always applied—even at the collection level. Most glaringly, there are few, if any, true archives access systems that effectively handle lower-level archival objects as records while maintaining the collection-level context that is needed. Just because a lower-level archival object needs to be presented together with its upper-level relationships, does not mean it cannot be discoverable.

Finally, while criticism of academic publishing is outside the expertise and scope of this review, it must be mentioned that the cost of the book is steep at \$84. This is notable because the book is published by a professional organization with an interest in open-access principles.

Gregory Wiedeman
University Archivist
University at Albany, SUNY