Understanding Metadata

Danielle Janney

School of Information, College of Arts and Sciences, University of South Florida

LIS 6726: Metadata

Professor Xiying Mi

January 28, 2022

Metadata

The current information landscape has an overwhelming quantity of (digital and physical) data being added to it every day, ranging from crucial to frivolous importance. Regardless of desirability, data will exist and potentially be useful to someone, somewhere. For data to be most useful, it needs metadata to give it context. Metadata basically defined is "data about data," more thoroughly it is "structured, encoded data describing the characteristics of … information-bearing entities" (Zeng & Qin, 2016, p. 7). Metadata is created to aid in information discovery, usability, shareability, and preservation.

According to the National Information Standards Organization (NISO), there are three main types of metadata: descriptive, structural, and administrative. Descriptive metadata describes information entities using keywords, author, title, and other elements to facilitate discovery within catalogs, collections, curations, and indexes. The importance of descriptive metadata is reflected in the popularity of web-based search engines: search results are as effective as they are because of embedded metadata within a webpage that describes its content. Keywords entered in a search browser will return an index of results; for instance, searching "ant" on Google yields a link to the Apache Ant programming tool and then a Wikipedia article on the insect. This simple search highlights the dynamic power of metadata and the plethora of information indexed for internet consumers. Structural metadata shows how a multi-page document goes together and maintain a relationship between entities that make up a whole. An example would be a scanned book pages. Administrative metadata is for managing an information object; it indicates how an information object may be used (copyright), its location, and its acquisition/creation dates. Technical, preservation, and rights management metadata fall under the umbrella of administrative metadata.

Metadata is not without standards or control, there are numerous domains and institutions that have established structures and regulated semantic rules for their metadata

records to follow. The Dublin Core Metadata Element Set is a general 15-element structure anyone may apply to personal or institutional web documents, and does not enforce abiding by controlled vocabularies for element values. To contrast, the Cataloging Cultural Objects (CCO) builds upon existing metadata standards (such as the Visual Resources Association Core Categories), and regulates the values, content, organization, and format of metadata fields in order to produce consistent records. When working to produce a metadata record, it is important to understand the standards being upheld--something as trivial as the display of a date may need to be formatted in a specific manner: 11/06/1989 is 06 Nov 1989 is November 6, 1989, but only one may be used if the metadata schema is explicit. If my birthday was (for some bizarre reason) considered a work of art, then its date would be 1989-11-06 because the CCO follows the ISO 8061 international standard.

It is becoming more apparent there is a scale of metadata practices, some schemas or structures are flexible and others more rigid. But what is good metadata? What can humans do on their end to ensure maximum impact of effort? NISO lists six principles of metadata in *A Framework of Guidance for Building Good Digital Collections*:

1. Good metadata conforms to community standards in a way that is appropriate to the materials in the collection, users of the collection, and current and potential future uses of the collection.
2. Good metadata supports interoperability.
3. Good metadata uses authority control and content standards to describe objects and collocate related objects.
4. Good metadata includes a clear statement of the conditions and terms of use for the digital object.
5. Good metadata supports the long-term curation and preservation of objects in collections.

6. Good metadata records are objects themselves and therefore should have the qualities of good objects, including authority, authenticity, archivability, persistence, and unique identification.

The first three principles work together. The first explains that an organization or institution should analyze its collection and decide what metadata structures support it best; the second principle encourages optimizing object metadata records by leaving no detail out, regardless of its collection's descriptors; and the third principle orders controlled vocabularies to follow suite of the metadata structures used, while keeping in mind the target audience, languages spoken, and up- and outdated figures of speech. The latter three principles are more administrative and emphasize maintenance and copyright responsibilities, and integral transparency of data values in all aspects of metadata creation.

Imagining a library, archive or museum's collection online, one wonders where the metadata exists. Unlike a simple navigational website where metadata might be hidden from view in the header or body, LAM collections metadata are stored in databases or repositories. Currently in libraries, the widely-used Machine Readable Cataloging (MARC) records are being replaced by the Bibliographic Framework (BIBFRAME), an entity relationship model that will open the door to the future of Linked Data. This gradual change indicates a desire for faster, more user-friendly information discovery.

Metadata in a highly digitized world is transforming the world's access to information. Search engines at personal and institutional levels relay thousands of query results in less than a second. Interoperability of metadata schema within collections face the challenge of effective crosswalks while preparing for the next phase of Linked Data and entity relationship models. Metadata is not a new concept, but how it is interacted with is breaking ground to unite the information landscape.

References

Baca, M., Harpring, P., Lanzi, E., McRae, L., & Whiteside, A. (Eds.). (2006). *Cataloging cultural objects: A guide to describing cultural works and their images*. American Library Association. https://vraweb.org/wp-content/uploads/2020/04/CatalogingCulturalObjectsFullv2.pdf

National Information Standards Organization. (2007). *A framework of guidance for building good digital collections*. NISO Press. https://www.imls.gov/sites/default/files/publications/documents/framework3.pdf

National Information Standards Organization. (2004). *Understanding metadata*. NISO Press. https://www.lter.uaf.edu/metadata_files/UnderstandingMetadata.pdf

Zeng, M. L., & Qin, J. (2016). *Metadata* (2nd ed.). American Library Association.