

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Library Philosophy and Practice (e-journal)

Libraries at University of Nebraska-Lincoln

April 2010

Digital Images, Management and Metadata: The Long Tail, or the Order of Order?

Caroline Zentner Raasch
raaschca@aol.com

Follow this and additional works at: <https://digitalcommons.unl.edu/libphilprac>



Part of the [Library and Information Science Commons](#)

Raasch, Caroline Zentner, "Digital Images, Management and Metadata: The Long Tail, or the Order of Order?" (2010). *Library Philosophy and Practice (e-journal)*. 348.
<https://digitalcommons.unl.edu/libphilprac/348>

Digital Images, Management and Metadata: The Long Tail, or the Order of Order?

Caroline Zentner-Raasch
San Francisco, California USA

Introduction

Weinberger (2007) explains how knowledge can no longer be contained within neat, small boxes called catalogues. The range of resources and information on the Internet and the World Wide Web has changed forever the landscape of information access and retrieval. The questions explored in this article include why the digitization of images is a complex issue, and the relationship of metadata to the digitization of images.

Survey of the Literature

Weinberger (2007) explores three distinct forms of knowledge building, which he calls principles of organization: the First Order of Order, the Second Order of Order, and the Third Order of Order. An example of the First Order of Order is the physical order of books, journals, DVDs, and so on, in a library. An example of the Second Order of Order is a library catalog, where the First Order of Order is separated into categories that are searchable by a system such as author, title, subject, etc. The Third Order of Order is based on the digital order of information, where the ability to search for an item can be adjusted each time a search is performed. Weinberger says, "The basic fact that order often hides more than it reveals has sometimes itself been hidden within the art and science of organizing our world" (2007).

Bartholomae and Petrosky (2003) examine how images can expand the reading experience and knowledge of college students. By asking students to consider the role of images as part of the text, the authors show how changing the sequence of articles within a book can lead to new insights about reading images and text together. The context of an image changes when its relationship to the text changes.

Cornell University Library/Research Department (2009) defines metadata as a means of describing a variety of "attributes of information objects and gives them meaning, context, and organization." This tutorial discusses three areas of classification: descriptive, structural, and administrative, and says that "these categories do not always have well-defined boundaries and often exhibit a significant level of overlap" (Cornell University Library / Research Department, 2009). Issues such as static versus dynamic metadata elements, the internal versus the external functions of each element, and interoperability are examined during the tutorials in the web site.

Casey (2008) states that, "the future of bibliographic control will be collaborative, decentralized, international in scope, and Web-based." She examines the dynamic nature of the current metadata landscape and the rapid changes facing metadata and bibliographic control in the World Wide Web environment. Connaway, Radford, Dickey, Willams, & Confer (2008) explore the difference age makes in searching for materials and the thinking skills used for metadata searches. The authors say, "A major challenge facing today's libraries is to develop and update both traditional and digital collections and services to meet the needs of the multiple generations of users with differing approaches to information

seeking.” Libraries are challenged to provide access to digital content via the Internet, and user expectation for immediate delivery of diverse formatted materials continues to increase. Age-related methods for searching push beyond current understanding of the digital environment.

Smith (1999) discusses fundamental questions about images and digitization. These include methods for selection, providing reference service in the digital environment, providing access to the educational value of an image collection, and metadata for images. Smith says, “Digitizing these types of primary source materials offers teachers at all levels previously unheard-of opportunities to expose their students to the raw materials of history.” Users of digitized image collections “expect higher levels of functionality of digital objects than they do of library materials, in part because there is no online equivalent to a reference specialist available” (Smith, 1999). Improvement in image quality as well as quantity has changed searching expectation and retrieval issues.

OCLC/RLG Working Group on Preservation Metadata (2001) discusses how preservation issues are combined with metadata issues for digitized resources. The report states that, “metadata is commonly understood as an amplification of traditional bibliographic cataloguing practices in an electronic environment.” Due to the growth in digital collections, including images, film, and sound files, the need for longterm metadata access is important. The OCLC/RLG Working Group calls for 16 elements that are essential for future access. These elements are: date, transcriber, producer, capture device, capture details, change history, validation key, encryption, watermark, resolution, compression, source, color, color management, color bar/grayscale bar, and control targets. Questions about how metadata for “born digital” images, digital audio files, and web pages include: What type of comprehensiveness is required? What is the best practice for structure? How can metadata be broadly applicable over multiple migrations?

In 2002 the Working Group explored how the Open Archival Information System (OAIS) can provide a framework for standardized digital preservation metadata forum. Key issues are: “the distinction between structure and elements” and “the ‘leaves’ of the hierarchical tree—collectively defined what is referred to preservation metadata elements” (OCLC/RLG Working Group on Preservation Metadata, 2002). This report examines the framework for preservation metadata to be embedded in the OAIS information model. The objects can be either physical or digital, and are called data objects. Knowledge of representation information is necessary for understanding the “aboutness” of a string of bits. And a combined data object and representation information forms an “information object” that carries the information needed to understand each bit stream. The four classes of information object are: content information, preservation description information, packaging information and descriptive information (OCLC / RLG Working Group on Preservation Metadata, 2002). Of these four components, the Content Information and Preservation Description Information are central to the development of preservation metadata.

Dale (2004) discusses how metadata for still images has advanced in the past few years as the digitization of still images has grown. There are several reasons for this: the high cost of digitization, the need for libraries, museums, and cultural institutions to ensure long-term preservation for digitized materials, and the cost of ensuring the long-term availability of the collections. Dale says, “In the digital environment, the ability to manage and preserve information over time will be dependent upon metadata – both the kind of metadata and the level of detail collected” (Dale, 2004). She says that cooperative working groups have made significant advances in this area, which is making the cost of preservation metadata more economically feasible.

Dale and Waibel (2004) discuss issues of long-term access and the role of metadata in digital image preservation. Dale and Waibel state that, “technical metadata is only a subset of the complete suite of preservation metadata necessary to achieve the long-term viability of a digital asset. It has often been called the first line of defense against losing access.” They also point out that “technical metadata assures that the information content of a digital file can be resurrected” and “in its entirety, technical

metadata supports the management and preservation of digital images throughout the different stages of their life-cycles" (Dale and Waibel, 2004).

The National Information Standards Organization (2008) discusses metadata standards for digitized information resources, saying that "metadata is structured information associated with an object for the purposes of discovery, description, use, management, and preservation" (NISO, 2008). Six principles of good metadata are discussed. These include adherence to community standards, support for interoperability, authority control and content standards, statement of conditions and terms of use, supports long-term curation and preservation, and since "metadata records objects themselves," it should have the qualities of good objects, including authority, authenticity, archivability, persistence, and unique identification." (NISO, 2008).

Blossom (2005) discusses the long tail of content and libraries. The concept of the long tail was introduced by Anderson (2004). The "long tail" refers to statistical distribution and the "tails" at either end of a statistical curve. The long tail is the statistical representation of the acquisition and use of many different resources by only a few people. Blossom points out that since the publication of Anderson's article the topic of the long tail of information resources has been picked up by librarians, corporate executives, and information professionals. He says, "the huge portions of content thought to be of residual value to companies catering to mass audiences is turning out to be both powerful and profitable to a wide range of audiences." Librarians and information professionals should examine "both online and corporate models for tips as to how to manage the content that matters most to highly contextual audiences" (Blossom, 2005).

Storey (2005) examines how libraries have already begun the shift to a more interactive relationship with users. Storey says, "Libraries need to embrace the new digitization and networking capabilities inherent in The Long Tail, which create some intriguing possibilities." He believes that future of libraries will be found within the long tail network.

Conclusion

Digital images and metadata issues are complex and interrelated. The understanding of the long tail and the order of order together provide a sound foundation for examination of management of metadata in relation to digital images. The literature shows the importance of each step of the metadata process, why each step is critical to the success of a project, and how different the view can be depending upon where you are standing. Metadata is much more than data about data; it expresses the "aboutness" of the data and all the forms in which "aboutness" can be explained and accessed

Sources Cited

Anderson, C. (2004). The long tail. *Wired*. Available: <http://www.wired.com/wired/archive/12.10/tail.html>

Bartholomae, D., & Petrosky, A. (2003). *Ways of reading: words and images*. Boston : Bedford / St Martin's.

Blossom, J. (2005). Riding the long tail: Libraries confront the world of infinite content supply and demand. *Shore.com*. Available: <http://www.shore.com/commentary/newsanal/items/2005/20050627longtail.html>

Casey, D. 2008. The future of bibliographic control: A report from the Library of Congress. NASIG Pre-conference forum, June 4, 2008. Available: <http://www.niso.org/news/events/agenda/casey>

Connaway, L., Radford, M., Dickey, T., Williams, J., & Confer, P. (2008). Sense-Making and synchronicity: Information seeking and communication behaviors of millennials and baby boomers. *Libri*, 58, 2. Available: <http://www.oclc.org/research/publications/archive/2008/connaway-libri>

Cornell University Library/Research Department (2009). Moving theory into practice: Digital imaging tutorial. Available: <http://www.library.cornell.edu/preservation/tutorial/contents.html>

Dale, R. 2004. Introduction. *RLG DigiNews*, 8, 5. Available: <http://worldcat.org/arcviewer/1/OCC/2007/07/10/0000068892/viewer/file1.html>

Dale, R. & Waibel, G. 2004. Capturing technical metadata for digital still images. *RLG DigiNews*, 8, 5. Available: <http://worldcat.org/arcviewer/1/OCC/2007/07/10/0000068892/viewer/file1.html>

National Information Standards Organization. (2008). Framework of guidance for building a good digital collections, metadata. NISO. Available: <http://framework.niso.org/node/24>

OCLC/RLG Working Group on Preservation Metadata (2001). *Preservation metadata for digital objects: A review of the state of the art*. Retrieved from http://www.oclc.org/research/pmwg/presmeta_ep.pdf

OCLC/RLG Working Group on Preservation Metadata (2002). *Preservation, metadata and the OAIS information model: A metadata framework to support the preservation of digital objects*. Dublin, OH: OCLC. Available: http://www.oclc.org/research/pmwg/pm_framework.pdf

Smith, A. (1999). *Why digitize?* Washington, D.C.: Council on Library and Information Resources. Retrieved from <http://www.clir.org/pubs/reports/pub80-smith/pub80.html>

Storey, T. (2005). The long tail and libraries. OCLC Newsletter, April/May/June. Available: <http://www.oclc.org/news/publications/newsletters/oclc/2005/268/downloads/thelongtail>

Weinberger, D. (2007). *Everything is miscellaneous: The power of the new digital disorder*. New York: Times Books.