

**Metadata literacy: An analysis of metadata awareness in
college students and librarians**

Erik Mitchell

University of North Carolina – Chapel Hill

Table of contents

1. Introduction	3
2. The role of metadata	4
3. The role of literacy	5
4. The social impact of information technology.....	6
5. Information participants	6
6. What is metadata-literacy?	7
7. Research Purpose.....	8
8. Literature review and proposal outline.....	9

1. Introduction

While research questions the added value of metadata in search (Hawking & Zobel, 2007; Hunter, 2003), and the sustainability of traditional metadata practices (Calhoun, 2006), other studies point to the centrality of metadata in learning environments and digital libraries (Shreve & Zeng, 2004; Zeng & Smith, 2003). This review of literature and research proposal investigates the relationships between metadata and literacy and the use of metadata in participant driven information environments. This review of literature and research proposal views these topics with an eye focused on the roles of authorship and community participation in online information environments. The information age has seen a convergence between traditional literacy skills, digital social interaction, and digital organization and information management. The proposed research investigates what role metadata plays for students engaged in these tasks and how they relate to the students' concept of information literacy.

Related to these areas of research are the developments of new information environments within the past decade. Social communities such as Facebook, SecondLife, and Flickr use varying levels of metadata both at system and user assigned levels to create interesting and valuable environments. Other areas such as Web 2.0 technologies, open source development tools and more accessible development languages such as PHP: Hypertext Processor (PHP) and Ruby on Rails are blurring the lines between software developer and user. This trend is being mirrored in the information organization environments of Flickr, Del.icio.us and other related sites where users create content, organize it, and harvest it for new uses.

Reflective of these developments the research communities addressing emerging information environments, metadata creation/use, and information literacy are very active. Despite this activity, there has been little research which investigates the extent to which the awareness of metadata and document models impact the user experience in these information environments. This awareness is discussed in the following review of literature as metadata literacy. Further, while information literacy has been widely researched over the past 30 years, the research community is still searching for a unified model from which to investigate this concept. One of the key challenges in investigating literacy concepts in general is a research gap between disciplines which are interested in this concept. The education and information science fields for

example are both interested in discovering more about the role of literacy in information use and learning but tend to ask questions on opposite sides of the research arena.

This review of literature and research proposal investigates the relationship between these three areas of research by asking what constitutes metadata-literacy, what value it has, and to what extent students and librarians possess metadata-literacy skills.

2. The role of metadata

Metadata is defined in many ways. The definition of metadata used in this review of literature and proposal is “[Metadata is] structured data about an object that supports functions associated with the designated object” (Greenberg, 2003). This definition includes the elements of contextualization and intended use, both of which are key in understanding how information is used. In this review and proposal, information organization (IO) is defined as the process of ordering, surrogation, or description information and information objects. As Wright (2007, p. 23) discusses, organization and categorization have been precursors to information artifact generation from the earliest examples of human society. Wright’s observations are based on the work of Rosch (1978) and current thought in information organization theory (Weinberger, 2007) which view elements of information organization such as representation, surrogation, and categorization as key tasks in the process of world observation and cognition (Rosch, 1978, p. 28). This review of literature further takes the stance that representations of these organization structures are key elements in our documents and information systems.

Metadata is less central to the use and structure of traditional information artifacts such as books, paintings, or stories. While certain elements such as descriptive metadata (author, title, publication information), and categorical metadata (topics, dates, relationships) have played an important role in print resources, the role of metadata in these artifacts is changing in a digital environment. First, and perhaps most notably, the concept of primary use is diminishing. Information artifacts are increasingly generated to be re-used and transformed. Second, notions of authorship are changing. Informal modes of collaborative and community authorship are beginning to re-emerge as the technological platforms to enable them are becoming more widespread. Finally, the digital document enables new types of metadata to be recorded, often automatically and integrated into documents. This means that new ways of managing and using these documents are possible.

Metadata serves these changing roles by providing conceptual models and encoding systems to store this information. This means that new documents are being created in which a key element is the presence of metadata. The extent to which users are aware of and understand the role of metadata in these information artifacts and the role that they (the users) play in creating, using, and storing this metadata is central to the concept of metadata literacy.

3. The role of literacy

Literacy has emerged as an important concept for the discussion of learning in the information age. There are numerous works on the history of literacy and its role in shaping history (Stock, 1983; Wright, 2007). Wright (2007) for example discusses the rise of the concept of literacy in western society in Medieval Europe arguing that literacy had a far reaching impact in the middle ages a documents become representative of contracts, agreements, and social norms (p. 107). This fact, he asserts, meant that even those who could not read or write were affected by these documents. Wright claims that this growth was a form of stigmergy in which the presence of documents had far-reaching impacts that laid the groundwork of an information revolution far before the moveable type printing press was invented. The importance of literacy continues to be a central theme in current standards and much of the work conducted in this area points to this fact. Likewise many information theories point to the essential role that information and social context plays in modern society. Chatman's 'outsider' perspective (1996), and Dervin and Nilan's SenseMaking perspective (1986) both include statements about the personal and social impact of information use. Both of these models include questions about the role and perspective of the participant in determining how information is sought, harvested, created, and used. These tasks coincide with the common notions of literacy investigated in this proposed research.

It does not need to be argued that the role and uses of these skills must be re-considered and translated for new environments. Reports such as the Horizon Report (Educause, 2007, 2008) discuss how changes in information technology impacts information skills and reflects on the social implications of technology adoption. Wright (2007, p. 46) documents the relationships between societal structure and their information artifacts in discussing the extent to which information artifacts reflect community knowledge, morals, perspectives, and worldviews. He asserts that the creation of digital objects in networked electronic environments serve the same roles. Both the horizon report and Wright observe that a participant's level of information

literacy in the digital environment impacts their ability to engage in these emerging social constructs.

4. The social impact of information technology

As early as the computer began gaining ground, theorists including Turing, Licklider, and Bush began discussing what impact these devices would have on individual and social use of information. Vannevar Bush's Memex machine (1945) is widely regarded as a foundational theoretical construct for personal information management. Licklider and Taylor (1968) discussed two uses of computers in social contexts - social communication mechanism and mental modeling tool which are of growing importance in the current computing environment. From these foundation ideas a computing movement developed that in parts was focused on distributed social networks, collaborative development, and shared authorship. Nearly congruent with the development of the internet and personal computing on a large scale, both digital libraries and open source software movements developed with a focus on community participation and open social networks (Borgman, 1996; Raymond, 2001).

In recent years, the web has become an environment more open to user-created programs using applications like Ning (<http://ning.com>), Yahoo Pipes (<http://pipes.yahoo.com>), and Google Gadgets (<http://desktop.google.com>). Development environments such as Python (<http://www.python.org>) and Ruby (<http://www.ruby-lang.org/>) focus on abstracting technical details from program interaction to the point that writing web applications are now as much about the data and logic layers as about the technical approaches. Indeed, it is now possible to represent program logic in metadata structures such as the Resource Description Framework (w3c, 2004) (RDF) and have those structures turned into platform specific applications automatically. Pointing to the need to represent social information and socially shared information Wang et al. discuss the use of RDF and the Semantic Web (Wang, Carley, Zeng, & Mao, 2007, p. 82). While this level of automation is not the current standard, parallels in history can be seen in the development of the codex and the impact it had on the scroll, the impact of cheap paper on book cost, and the impact of the magnetic tape as a personal recording medium.

5. Information participants

Understanding the role of literacy and metadata in digital documents is related to an active view

of the information user. Use of technology and digital documents includes roles of seeking, retrieving, creating, and processing, all of which are tasks whose functions have new implications in digital contexts. In their 2004 study, the Pew Internet Trust found that 87% of American teenagers and 66% of American adults use the Internet (Lenhart, Madden, & Hitlin, 2005, p. 2). Another Pew Trust survey from 2005 showed continued growth, with 72% of adults online ("Demographics of Internet Users", 2005). The 2005 survey of Internet Activities also indicated that 90% of these users used Internet search engines in information seeking. The only more popular activity indicated was e-mail (91%) ("Internet Activities", 2005). The Statistical Abstract of the United States indicates similar usage numbers and patterns in their study from 2004 showing 62% of adults with Internet access at home (*Statistical Abstract of the United States*, 2006, p. 751). Despite the widespread use of technology however, many studies are also finding a gap between 'technology' familiarity and actual information literacy with regards to that technology (Rowlands et al., 2008; Yan, 2008). This review of literature and research proposal investigate participant literacy with a goal of investigating this gap.

6. What is metadata-literacy?

This proposed research casts a wide net across the current landscape of participants and systems being used in information-rich environments. It focuses specifically on metadata creation and use and asks what skills and resulting literacies are required to be proficient in this specific context. Literacy continues to be a widely researched topic in education and information science disciplines. While there exists many literacy standards (information literacy, visual literacy, social literacy, foundational literacy), there is no focus on literacy skills associated specifically with the notion of data and metadata and its role in digital documents. This is reflected in the fact that many of the literacies as defined in the literature focus more on a specific technology or use and not on the underlying documents or information artifacts. While information organization has been considered to be the realm of either organization experts such as catalogers and indexers or individuals engaged in personal information management, emerging social-centric systems are creating an environment in which users are collaboratively creating/using/harvesting organization and data structures. One of the key changing concepts in relation to these emerging tasks is role of the information consumer as author in an information process. Metadata-literacy adopts this view of the participant and takes as its base the assertion

that information organization and related document structures are central to information creation and use.

Information literacy has been represented by several models. Some models emphasize conceptual literacies (e.g. the ability to obtain, analyze, and create knowledge) (Horton, 2007). Other models emphasize skills such as searching, aggregating, and reading (Eisenberg, 2006). Still other models emphasize relationships between entities (Bruce, 1997) and broad literacy categories (Hughes & Shapiro, 1996). In the educational arena the discussion of literacy focuses more on educational outcomes. Bloom's Taxonomy (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956) and various literacy evaluation frameworks (NETS, 2007; SITES, 2006) investigate these skills and concepts from a student achievement perspective. This review of literature does not seek to consolidate these models but instead asks which components of each apply to metadata related skills, concepts, and uses.

7. Research Purpose

The purpose of the research informed by the following literature reviews is to investigate how participants view metadata and to what extent they are comfortable engaging with metadata. The connections between information theory, computer science, and learning theory centered around information literacy will be pursued in the literature review. One connection in particular, between metadata creation and use and the role that these tasks play in an individual's information environment, is of particular importance. This literature review will investigate awareness of participants of metadata in their information environments. It will focus specifically on electronic environments and ask participants to think generally about their role in document creation, use, and re-use.

The proposed study will approach the question of how participants perceive metadata and what impact they see it having on their information environment by using a mixed-methods approach. Further it will investigate how participants perceive authorship roles by asking them about their familiarity with and attitude towards metadata creation (such as tagging). In order to get a balanced view between "expert" and "novices" two participant groups will be investigated. By looking at the knowledge and perspectives of both students and instructors on the topic of metadata literacy, this research seeks to understand how metadata literacy is viewed from a more holistic perspective.

This proposed study will be conducted online using the Qualtrics platform. Given the subject matter, the context of the research, and the technological access and competence of the population this is an appropriate use of this method. It is anticipated that while participants may not have formulated clear concepts of the role of metadata in their digital information environments and will not have a generalizable understanding of metadata, they will both be able to think specifically about metadata tasks and grasp of the impact of metadata in their information experience. It is unknown whether or not there will be significant differences between the two primary populations (students and teachers).

8. Literature review and proposal outline

In order to inform this research, a review of literacy models will be used to identify tasks and perspectives with regards to information activities. First, a review of information literacy models and an investigation of the implications of using IL as a model for participant/document interaction will lead to a way of evaluating generalized literacies. Second, a review of metadata purposes and roles will introduce ways in which participants may use metadata. Third, a review of current participant perspectives will inform the current information practices of the participants. In the fourth literature review the concept of metadata-literacy will be investigated using the literacy evaluation framework. Finally, the fifth literature review will investigate methods by which the concept of metadata-literacy can be investigated. In investigating how these individuals perceive literacy, digital information, and metadata, this research will contribute to a grounded theory of metadata-literacy.

References

Bloom, B. S., Engelhart, M., Furst, E., Hill, W., & Krathwohl, D. (Eds.). (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals*. New York: David McKay.

Borgman, C. L. (1996). Social Aspects of Digital Libraries. Retrieved 2-Nov, 2006, from http://is.gseis.ucla.edu/research/dl/UCLA_DL_Report.html

Bruce, C. (1997). *The seven faces of information literacy*. Adelaide: Auslib Press.

Bush, V. (1945). As We May Think, *The Atlantic Monthly*.

Calhoun, K. (2006). The Changing Nature of the Catalog and its Integration with Other Discovery Tools; Final Report, Prepared for the Library of Congress.

Chatman, E. A. (1996). The impoverished life-world of outsiders. *Journal of the American Society for Information Science*, 47(3), 193-206.

Demographics of Internet Users. (2005). Pew Internet & American Life Project.

Dervin, B., & Nilan, M. (1986). Information needs and uses. *Annual Review of Information Science and Technology*, 21, 3-33.

Educause. (2007). *The Horizon Report: 2007 Edition*. Boulder, CO: Educause.

Educause. (2008). *The Horizon Report: 2008 Edition*. Boulder, CO: Educause.

Eisenberg, M. (2006). A Big6™ Skills Overview. Retrieved March 5, 2008, from <http://www.big6.com/showarticle.php?id=16>

Greenberg, J. (2003). Metadata and the World Wide Web. In M. A. Drake. (Ed.), *Encyclopedia of Library and Information Science*. New York: Marcel Dekker.

- Hawking, D., & Zobel, J. (2007). Does topic metadata help with Web search? *Journal of the American Society for Information Science and Technology*, 58(5), 613-628.
- Horton, F. W. J. (2007). *Understanding Information Literacy: A Primer*. Paris: United Nations Educational, Scientific and Cultural Organization.
- Hughes, S., & Shapiro, J. (1996). Information Literacy as a Liberal Art, *Educom Review* (Vol. 31).
- Hunter, J. L. (2003). A Survey of Metadata Research for Organizing the Web. *Library Trends*, 52(2), 318-344.
- Internet Activities. (2005, 5 Dec 2005). Retrieved 15 Feb, 2006,
from http://www.pewinternet.org/trends/Internet_Activities_12.05.05.htm
- Lenhart, A., Madden, M., & Hitlin, P. (2005). *Teens and Technology: Youth are Leading the Transition to a Fully Wired and Mobile Nation*: Pew Internet & American Life Project.
- Licklider, J. C. R., & Taylor, R. W. (1968). The Computer as a Communication Device. *Science and Technology*.
- NETS. (2007). National Educational Technology Standards for Students. Retrieved 21 January 2009,
from http://www.iste.org/Content/NavigationMenu/NETS/ForStudents/2007Standards/NETS_for_Students_2007_Standards.pdf
- Raymond, E. S. (2001). *The cathedral and the bazaar : musings on Linux and Open Source by an accidental revolutionary*. Cambridge, Mass.: O'Reilly.
- Rosch, E. (1978). *Cognition and categorization*. Hillsdale, N.J. New York: L. Erlbaum Associates ; distributed by Halsted Press.

- Rowlands, I., Nicholas, D., Williams, P., Huntington, P., Fieldhouse, M., Gunter, B., et al. (2008). The Google generation: the information behaviour of the researcher of the future. *Aslib Proceedings*, 60(4), 290-310.
- Shreve, G. M., & Zeng, M. (2004). Integrating Resource Metadata and Domain Markup in an NSDL Collection.
- SITES. (2006). *SITES 2006: The use of ICT in teaching and learning*.
- Statistical Abstract of the United States* (125th ed.). (2006). Washington DC: U.S. Census Bureau.
- Stock, B. (1983). *The Implications of Literacy*. Princeton, NJ: Princeton UP.
- w3c. (2004). Resource Description Framework (RDF): Concepts and Abstract Syntax. Retrieved March 12, 2008, from <http://www.w3.org/TR/rdf-concepts/#section-triples>
- Wang, F.-Y., Carley, K. M., Zeng, D., & Mao, W. (2007). Social Computing: From Social Informatics to Social Intelligence. *IEEE Intelligent Systems*, 22(2), 79-83.
- Weinberger, D. (2007). *Everything is Miscellaneous: The power of the new digital disorder*. New York: Times Books.
- Wright, A. (2007). *Glut : mastering information through the ages*. Washington, D.C.: Joseph Henry Press.
- Yan, Z. (2008). Undergraduate students' mental models of the Web as an information retrieval system. *Journal of the American Society for Information Science and Technology*, 59(13), 2087-2098.
- Zeng, M., & Smith, T. (2003). Building Semantic Tools for Concept-based Learning Spaces: Knowledge Bases of Strongly-structured Models for Scientific Concepts in Advanced

