

# The role of information organization and metadata in digital documents

1	Introduction.....	3
2	The Information revolution.....	4
2.1	Overview .....	4
2.2	Defining the document.....	6
3	Information Organization tasks.....	12
3.1	Defining information organization.....	12
3.2	Resource Description .....	13
3.3	Resource classification and categorization .....	13
3.4	Participant perspective in categorization.....	16
4	Roles of metadata in information environments.....	19
4.1	Metadata definitions.....	19
4.2	Research in metadata creation and use.....	20
5	Metadata and the Web 2.0 technologies .....	23
5.1	Dublin Core metadata elements .....	27
5.2	RSS (Real Simple Syndication) metadata model.....	28
5.3	Resource Description Framework (RDF) .....	29

5.4	Creating metadata in community spaces .....	30
6	Concluding questions.....	34

# 1 Introduction

As discussed in the literature review on information literacy (IL), IL models include technical skills, information concepts, and usage contexts. Further, the literature review discussed the connection between information literacy models and constructivist and social constructivist approaches in modeling and teaching in each of these three aspects. The resulting framework for discussing forms of literacy includes the three concepts of pedagogy, information/learning theory, and role of the information environment as primary questions for evaluating the skills, concepts, and context of a specific element of IL. The IL review further found that there was insufficient attention given to the role of metadata and document structures in the digital environments for which many of the models were designed.

This literature review investigates elements of information organization (IO) and metadata that are commonly used in the digital information environment and seeks to answer the questions posed in the three-tiered IL framework outlined at the conclusion of the first literature review. In order to investigate the role of metadata in this framework, this review examines the relationship between the extended mind theory (Clark, 2001), social constructivism, and information organization. It will investigate this role by examining information environments and document models which support a fluid notion of the roles of contributor and consumer. Finally, this literature review discusses document models that are common in web environments including encoding models and content models. By identifying the nature of a metadata-rich information environment, the skills required to interact with that environment, and the theoretical foundation of the interaction, this review takes initial steps in identifying metadata literacy.

Given that IL curricula generally center on conceptions of the document, and particularly that social constructionism takes a post-modernist view of the document by deconstructing both its

structure and authorship, this review begins by looking at the impact that both historical and current conceptions of the document mean for IL and IO. The review continues by investigating information organization theories that are directly related to the IL skills reviewed in the first literature review and investigates connections between the current state of information organization theory, user perspectives on information, and the elements of organization that impact information literacy models.

## **2 The Information revolution**

### **2.1 Overview**

The latter part of the 20<sup>th</sup> century is commonly referred to as the “information age.” Key thinkers in information science such as Taylor and Ranganathan developed a theoretical grounding during the 1960s that coincided with technological advances. The IS profession still looks the ideas and contributions of Vannevar Bush, Licklider, and Turnig in forming information theory to guide current projects. Throughout the late 1960s into the 1980s a revolution in computing occurred along side significant changes in our information environment. For example, Licklider and Taylor’s (1968) foundational article during this time discussed how computers enabled shifts both in how people communicated and in how people would be able to represent their mental models, create documents, and collaborate.

Detailed histories of this period have been written by Markoff (2005) and Wright (2007) which discuss the roles of key individuals, technologies, and movements. One movement of note during this time was the creation of the open source software license (Raymond, 2001). Much like the development of user-driven information organization systems, the open source development movement quickly identified roles and mechanisms for participant interaction. The open source trend and the user-contribution trend have been developed in a digital environment and have led to significant impacts in the information marketplace.

The two notions of collaboration and documents have been a common theme in computer and information science literature. For example, reviews of the history of research in digital libraries shows an early interest in community based collaboration (Borgman, 1999; Marchionini & Maurer, 1995). The development of systems which embrace these concepts has continued and there are now systems allowing users to be more engaged in information organization practices by enabling collaborative cataloging and classification tasks. These systems, including Flickr, Del.icio.us, and Connotea, have enabled participants to engage in the representation and surrogation of resources without an overtly structured framework. As a result, the popularity of these sites has encouraged formal discussions on this topic including the work done by the Future of Bibliographic Control working group from the Library of Congress (Library of Congress Working Group on the Future of Bibliographic Control, 2008). Calhoun's (2006) report to that body presents fundamental questions about the future of current cataloging and library-sponsored information retrieval applications that it elicited strong feedback across the library community (Mann, 2006, 2008). The working group's final report recommended substantial changes to the process of bibliographic control, to library systems, and to library education policies. These ideas focus on the advantages of institutional organization systems such as the Library of Congress Subject Headings (LCSH) versus popular organization systems, questions the relative value of digitization of historical content versus maintenance of the status quo, and focusing on the document as discussion between information producers and consumers versus document as a static record (Library of Congress Working Group on the Future of Bibliographic Control, 2008).

This significant change in how documents are created and organized has implications not only for library and information science professionals but also for any information consumer in that there is a new set of skills, conceptual knowledge, and appreciation of context that coincides with the use of digital documents. As such, the information revolution as noted in the beginning of this section is not limited to simply new ways of accessing information or a change in the scope or pervasiveness of information in our everyday lives but also includes the development of

new information structures, conceptual foundations for those structures, and skills required to interact with them.

## 2.2 Defining the document

The evolution of our concept of the document has coincided with changes in categorization schemes, user based theories and technological advances. Buckland's "information as thing" concept (1991) is one of the most frequently cited perspectives which still serves as a basis for new theories. Buckland's (1997) article provides an overview of the origin of document theory in which he discusses perspectives from Briet, who asserted that an object became a document once it was analyzed, Otlet who viewed any object as a document, Duyvis who viewed the document as the material output of thought, and Ranganathan who viewed the document as only consisting of two dimensional textual based resource. Buckland (1997) compares these theories to the continuing evolution of documents in the digital realm pointing out that "documentalists increasingly emphasized whatever functioned as a document rather than traditional physical forms of documents" (p. 808). His claim that the "shift to digital technology would make this distinction more important" (p. 808) indicates a change in perspective from document as physical instance to document as instance of intellectual content which is now the widely held view. Buckland's definition includes both text-based physical objects including books and letters and less concrete examples including an antelope, or a model of a ship. While he does not explicitly include electronic formats in his 1991 work, he includes events as a different type of information. Lagoze (2000) extends this work on event-awareness in relation to resource description. Likewise, Greenberg expands the definition of the document by addressing the definition of objects "any entity, form, or mode for which contextual data can be recorded" (2003, p. 1876). Just as the documentalists movement of the early 20<sup>th</sup> century attempted to re-cast librarianship into 'knowledge management' (Buckland, 1997; Wright, 2007, p. 180), current trends in information science emphasize the relationship of information, document structure, and knowledge (Berners-Lee, 2006; Borland, 2007; Eriksson, 2007). Central to the

discussion are common ideas including an investigation of the nature of information and documents, an understanding of the information seeking process, and perspectives on the representation of information.

In his 1991 edition, Buckland also asks more generally “what is information?” His definition involves three states: process, knowledge, and thing (1991). A number of authors use this definition as a starting point. In discussing Buckland’s theory, Marchionini emphasizes the ability of information to be from any source including “bit, data, record, text, fragment, document, utterance, database, book, and library” (1995, p. 5). Buckland’s outline makes clear the differences between the instantiation of information and the information seeker in differentiating two key states of information first as thing and second as a process and establishes one of the key differences between categorization literature which tends to be thing or document focused versus user interaction literature which tends to be communication process focused. Generically stated, the cataloger seeks to extract information from a text and encode it in some external categorization system while the information seeker wishes to harvest that information for knowledge generation. While literature from both areas tends to not pick one view of information over the other, there is still a dramatic difference in the amount of attention paid to the document for organization and on information for extraction. Buckland (1991) gives attention to this theme in his book *Information and Information Systems*. “Since information and information handling is pervasive in human activities, an exploration of information systems that did not include the social, economic, and political context and the broad social role of information would be seriously incomplete” (1991, p. 9).

The concept of the document as encoded knowledge is well represented in the literature. Many of the models of IL reviewed in the information literacy literature review make specific mention of the relationship between technology and information most notably Bruce (1997) who positions information technology in relation to the interaction between the individual and information. Wright’s (2007) history of information draws direct connections between the evolution of information facets such as accuracy, timeliness, and ubiquitousness to the

development of the technology citing in particular the role that the codex and movable type printing press played in raising the role of information in the world. Holland (2006) goes further, arguing that language itself is a technological innovation “language, both verbal and written, provides tools for humans to alter and enhance their cognitive activities” (p. 95). These examples point to a symbiotic relationship in which the nature of information drives the technological encoding but also in which the resulting document structure guides the concept and representation process. Carr (2008) includes a wide number of examples of this stigmergic phenomenon. He documents the impact that the typewriter had on Nietzsche’s writing and the change in our conception of time that the pocketwatch had. These examples point to the idea that a document cannot be defined without considering the facets of technology, intellectual content and structure, and the role that technology plays in the representation of content in the encoded resource.

In relation to this changing concept of the document, social constructionism asserts that the view of information as an ‘information brick’ or as a noun-based state ignores the dynamics of user-centric information theory (Holland, 2006; Tuominen, Talja, & Savolainen, 2003, p. 563). This view asserts that neither the document nor the classification systems which represent it are fixed nor objective and that there is value in recognizing multiple perspectives in description and in recognizing the power in relationships between documents and users (Bates, 1998; Bowker & Star, 1999).

Related to the idea that digital documents include multiple authors and are embedded with organization structures, Green (2001) examines the relationships between information objects and knowledge organization and notes a growing sophistication between information type relationships. Green groups entity relationships into four types: bodies which include authors, publishers, and translators, bibliographic units which include both intellectual and physical entities, concepts/knowledge comprised of public or external knowledge and private or internal knowledge, and users with information needs (2001, p. 7). Green’s perspective provides an expanded method for viewing the relationships between document encoded information and user

based information interaction. An example of this the implementation of these relationships can be seen in Eriksson's (2007) work on semantic documents in which he combines ontologies with document markup.

Bringing the view of the participant into definitions of a document substantially changes the scope of the definition. For example, Green's (2001) work is an example of how information architecture seeks to bring the two concepts of document structure and participant information need together, creating a holistic approach to information organization, retrieval, and use. Rosenfeld and Morville's book on information architecture (IA) defines information as existing in the middle of knowledge "the stuff in people's heads" and data "facts and figures." They go on to discuss documents within the context of the "content" portion of their "Users, Content, Context" model (2002, p. 24). Additionally, Toms (2002) describes a model comprised of users, systems, and content in defining a document in terms of the system – content interaction "which is in essence the result of applying a set of computer processes to the microstructure and macrostructure of the text to provide access, storage, and update facilities" (p. 859). Finally, Oostendorp, Breure, and Dillon (2005) propose that information is also medium dependant, suggesting that in the digital medium misses some of the interface characteristics inherent in physical information artifacts.

We can say that one cannot judge a book by its cover, but this is false – we can and do make such judgments, often with good reason, based on size, wear and tear, annotation, price, date of publication, and so on. Indeed, in the book world, such attributes can prove very meaningful to a reader or user (2005, p. 313).

The researchers conclude that changes in publication, use, ownership, and the shift from printed text to integrated media will have a significant impact on how information is thought about. Wright (2007, p. 234) takes up a similar position, asserting that the concept of the document is grounded in a literacy tradition that focuses on the construct of authorship and authority and asks instead if web-based mediums are beginning to create documents based on oral

communication patterns rather than written patterns. Wright observes this change based on an observation that electronic texts such as instant messaging, blogs, and email are used in ‘conversational’ ways, allow a de-construction of authorship and authority, and utilize oral structures rather than written structures. In citing Amazon’s ability to represent literary, or fixed documents along with user created and managed documents along-side each other, Wright asserts that a new form of document is being created which allows participants to make use of both the written and oral traditions. This view of the document is in line with the definition of a resource offered by the W3C as “...anything that has identity. Familiar examples include an electronic document, an image, a service (e.g., "today's weather report for Los Angeles"), and a collection of other *resources*” (1999). Lagoze (2004) likewise observes that the simple nature of a digital document made accessible in a networked electronic environment fundamentally changes these roles and the methods by which metadata is created and kept for those documents. “My view, as presented in this paper, is that adaptation to the networked information context will require rather radical changes to the role of the catalog and the cataloging model” (2004, p. 3).

The implications for this shift both in how documents are created and how they are encoded is discussed by Ware and Warschauer (2005). They point to the ability to create “interactive written communication”, blurring traditional distinctions between author and reader. They also observe that electronic environments enable hypertext, this challenges both the linearity of printed texts and creates contextual classification structures. As Crook (2005, p. 511) observes, the use of contextual document structures is also a key element of creating an ‘active’ reader through decisions and choices in text interaction. The shift in document authorship has been paralleled with a shift in the content of digital documents to non-text media. As Wright points out, the dominance of text in the document may have had significant ramifications on other communication traditions including oral and symbolic traditions (2007, p. 39).

The differences noted above in viewing documents as complex objects (Ware & Warschauer, 2005), as evidence of knowledge (Borland, 2007), and as process of social discourse (Tuominen, et al., 2003) underscores changing conceptions of what a document is, how it is used, and what

authorship of the document means. This concept of the shifting document is important to this literature review for two reasons. First, there is a notion that metadata is becoming more central to the idea of a document and its context. Second the notion that authorship not only of the document but authorship of the categorization and contextualization of that document are changing. These changes in information creation, storage, and access have resulted in a re-evaluation of traditional processes and roles including the role of classification structures in encoding information, the role of catalogers and consumers of structured information, and the impact of the increase of scale and automation in creating and using these structures. Jacob (2004, p. 517) poses areas of research related to the interaction between a user and an information system in a categorization context: (a) Is communication between the information system and the individual influenced by the representation of resources? (b) Does the organizational structure of the information system cause the individual to adjust her internal cognitive structures? (c) Does the organization of resources contribute to the creation of a meaningful context for information? (d) Is the meaning of information influenced by the organizational structure of the information system? (e) What consequences follow from the different organizational structures?

As with the definition of information/documents, the shift to electronic and web-based interaction models is impacting theories on information literacy. As reviewed in this literature review, the evolution of new types of documents in digital environments is leading to theories (Ware & Warschauer, 2005) which accommodate fluid and iterative notions of authorship. The application of these theories can be seen in the use-focused metadata models and interaction methods reviewed in this paper. As information architecture literature asserts, the growth in use of contextual data, or metadata, in these documents is changing how participants interact with these documents (Morville, 2005). The implications of this change for participants are that a new set of literacies are required in order to understand how to interact with and create these documents.

## 3 Information Organization tasks

### 3.1 Defining information organization

Information organization (IO) is defined in this literature review as the process of ordering, surrogation, or description information and information objects. These three tasks have been identified by the author as broad tasks that both identify tangible elements of IO and have implications across areas of information research including retrieval, interaction, and personal information management. IO processes can involve the assignment of contextual metadata to documents, structuring of information objects via some document model, creation of new documents which serve specific roles, or creation of abstracted data structures including indexes, databases, and data-objects which serve the primary tasks of ordering, surrogation, and description. Information technology (IT) plays a key role in IO practices, influencing elements of creation and use of document derivatives. In the traditional library realm, organization tasks include annotation of documents directly creation of print or electronic surrogates in the form of catalog records, abstracts, and digital libraries, and the physical or virtual ordering or grouping of resources using the processes of categorization and classification. As noted above, the prevalent trend in IO research is to focus on the document or system surrounding the organization process. Research in this area tends to focus on methods of organization e.g. (Thibodeau & Barry, 2002; Weibel, 1995a; Yee, Swearingen, Li, & Hearst, 2003), implications of organization approaches e.g. (Araghi, 2005; Woodruff, Rosenholtz, Morrison, Faulring, & Pirolli, 2002), sustainability of organization approaches e.g. (Vuorikari, 2007; Weibel, 2005), and technologies of organization e.g. (Almasy, 2005; Bainbridge, Thompson, & Witten, 2003; Harter, 1996). These are all valuable areas of research in IO and metadata, but also tend to focus on a static-document or system-centric outcome. In contrast, this literature review is interested in the analysis of the use of metadata by information creators/consumers. In approaching the value of IO issues from the perspective of the value it lends to a user's experience and what skills, concepts, and competencies or literacies are required to engage in these practices, this review spends less time

examining the implications of complex IO structures and more time at the relationship of broad issues to use. The remainder of this section investigates the processes of description/surrogation, the act of categorization, and the interaction between documents and their creators/consumers as three primary examples of how the concept of metadata literacy can be examined. It examines each of these processes from the perspective of the general information participant.

### **3.2 Resource Description**

Resource description includes the processes of bibliographic description and element identification. This process is usually associated with cataloging and metadata practices and involve the recording of contextualized information from a resource. Research in this area has shown the importance of user-created organization schemes in file management (Barreau & Nardi, 1995; Barreau, 1995), and the role of metadata (Weibel, 1995b) while other research has demonstrated the inability of descriptive metadata to enhance web-search (Hawking & Zobel, 2007). Areas in which content description are considered to have value include accessibility (Chapman, 2007a), access for specific populations (Abbas, 2005), and personal information management (Jones, 2007). While the debate between the impact on retrievability of descriptive metadata is outside the scope of this work, it is worth pointing out that much recent research has focused on the impact of descriptive metadata on use and semantic web applications (Campbell, 2005; Dongwon, Peter Hoh, Fran, Young-Gab, & Doo-Kwon, 2005; Robertson & Greenberg, 2004).

### **3.3 Resource classification and categorization**

Classification and categorization philosophy emphasizes the importance of intended use and subjectivity of classification (Langridge, 1992), the benefit of collective classification (Mejias, 2004), and the difference between categorization and classification (Jacob, 2004, p. 538).

Classification is typically focused on grouping aspects of document aboutness onto a specific classification system. MacGregor and McCulloch (2006) describe aboutness as arising from the assignment “of subject headings according to their content”. In contrast, categorization focuses on identifying elements of a document but not in addressing primary topicality or locating the resource in the way that classification does. Within the confines of this literature review, the differences between classification and categorization are not as important as the implications of the use of these structures in digital documents. A clear debate in this area centers on who should be allowed to create classification and categorization systems (Guy & Tonkin, 2006; Marlow, Naaman, boyd, & Davis, 2006; Vuorikari, 2007). This debate includes evaluations of traditional classification structures such as hierarchies, trees, paradigms, and faceted systems (Kwasnik, 1999) and folksonomies (Shirkey, 2006). In investigating the application of these systems authors look at elements of intended use (Kwasnik, 1989), social and political elements (Bowker & Star, 1999, p. 324), and technical and metadata implications (Robertson & Greenberg, 2004). While a full review of the implications of folksonomy based categorization systems is outside the scope of this literature review, it is worth noting that user-created systems are a primary mechanism by which users are interacting with metadata in digital environments. Two primary issues are relevant to this work are how participants think of the process of categorization and classification and what approaches they employ to categorize and classify things.

In their description of Knowledge Organization Systems, Zeng and Chan (2004, p. 377) group classification structures into three types: term lists, classification and categorization schemes, and relational vocabularies. They differentiate forms of term lists by level of complexity: dictionaries (simple) to authority files (more complex relationships). They further separate traditional classification schemes, including Library of Congress (LCSH), Dewey Decimal (DDC), and Association for Computing Machinery (ACM), from newer models such as semantic networks and ontologies, while recognizing how these schemes rose from the combination of authority control and automatic processing. These tools can also be networked or linked together

to form larger structures called metathesauri or linked vocabularies (Robertson & Greenberg, 2004).

Kwasnik's (1999, p. 23) review of classification structures provides a sufficient framework to discuss common approaches to the creation of classification systems. In addition to laying out this framework, Kwasnik discusses elements of classification processes including serving as a way of expressing theories about the nature of a topic (p. 24), the inherent difficulty that comes along with refining complex concepts into a streamlined system (p. 46), and a growing need for flexibility and extensibility in the digital environment. Kwasnik's system underscores a shift in focus from the simpler world of bibliographic description for storage to resource generic description for multi-faceted retrieval and system interoperability. This change both in how classification occurs and who is responsible for creating and managing it is central to the argument that metadata literacy is an important element of information interaction.

Cognitive science literature points to the centrality of categorization processes as an element of cognition. Lakoff (1987) underscores the centrality of categorization in his work *Women, Fire, and Dangerous Things*.

Any time we either produce or understand an utterance of any reasonable length, we are employing dozens if not hundreds of categories: categories of speech sounds, of words, of phrases and clauses, as well as conceptual categories. Without the ability to categorize, we could not function at all, either in the physical world or in our social and intellectual lives (1987, p. 6).

Why users make use of these tools and what impact they have on learning and cognition is of particular importance to the concept of metadata-literacy. Some of the reasons for use include information seeking through advanced use of concept maps (Zeng & Chan, 2004, p. 377) description for finding and re-finding purposes (Bruce, Jones, & Dumais, 2004), and collaboration (Golder & Huberman, 2006). Research has pointed to issues of user motivation

(Guy & Tonkin, 2006) for tag selection purposes and some research has speculated on the cognitive load impact that creating and using folksonomies involves. While not focused on folksonomies, Marshall and Bly's (2004, p. 223) research on the sharing of encountered information found that users shared information to: (a) establish mutual awareness, (b) educate or raise consciousness, (c) develop rapport on common interests, and (e) demonstrate knowledge of recipients interests.

While the typical user may not think explicitly about the system and structure with which they categorize and classify information, research has taken time to enumerate and structure these approaches. For example the extent to which participants use metadata has been shown to be related to expertise. Hembrooke (2005, p. 868) found that level of participant knowledge in searching impacted the scope and efficiency of searching and that metadata feedback during search enabled better searching. Likewise, research has found that domain knowledge helps participants recognize relevant information during search (Ju, 2007) and that the use of scaffolding structures (e.g. metadata) is beneficial during searching (MaKinster, Beghetto, & Plucker, 2002). Despite a lack of formal research into why participants categorize and classify, and what goals they have when they engage in these tasks, the research reviewed has shown that these processes serve multiple roles including description, scaffolding, collaboration, and discovery and that participants may use a wide variety of approaches including structured, flat, hierarchical, and faceted systems.

### **3.4 Participant perspective in categorization**

While Kwasnik's (1999) work is useful in discussing how participant created classification structures might look, it does not address how participants are creating those structures. Kwasnik hints at the subjective and internal nature of classification systems while also stating that they serve a generalizing and knowledge documenting purpose. A key theorist who has added to the discussion of how these categorization and classification systems support the

documentation of knowledge is Rosch (1978). Rosch's prototype theory asserts that key elements of a thing or concept represent generalized prototypes which help users categorize and classify specific examples of those concepts. Rosch's theories have been used to explain how categorization and classification systems work (Weinberger, 2007; Wright, 2007). Wright's (2007, p. 22) discussion of prototype theory includes the work of Brown and Berlin in surveying taxonomies from multiple cultures. Weinberger (2007, p. 186) asserts that prototype theory supports the idea of cognitive efficiency in that it helps participants create links and associations between objects based on their categories.

Prototype theory asserts that there is a communally agreeable concept or prototype that can be used to represent categories in which specific items fall. This perspective is supported by the educational theory of Vygotsky (1977) who asserts that individuals learn from within a complex social context outside of which their knowledge is not grounded. The related theories of social constructivism build on this idea by asserting that the construction of these associations/categories is significantly governed by social influences and the process of dialog. Further, the extended mind theory (Clark & Chalmers, 1998) discusses the extent to which technology enhances our ability to create, maintain, and harvest knowledge and learning.

These perspectives are driving a shift from the use of categorization models for resource description to the use of them to facilitate user/information interaction. As an example of this, Rosenfeld and Morville (2002, p. 194) discuss two strikingly different perspectives for a thesaurus: one as a method for indexing and providing term control for documents, the other focused on providing a controlled vocabulary for user searching. Rosario, Hearst, and Fillmore's (2002) work on using metadata and faceted classification schemes to develop search interfaces is documented in a study in which they evaluated user-interaction levels with different interfaces to the same information. Their study sought to compare user success with interfaces based on differences between elements that supported either recognition or recall. They reported higher user-interaction levels (sense of control, task success, feature use and understanding) in systems that included faceted and tree-based navigation structures (2002, p. 6) and emphasized the

enhanced user success and satisfaction with metadata-based faceted browsing. They concluded with six recommendations for search systems: “Strive for consistency, offer informative feedback, offer simple error handling, permit easy reversal of actions, support user control, and reduce short-term memory load” (p. 8). One of the questions of the following section is to what extent users approach these systems equipped with a framework to use this metadata.

The key concept that emerges from this discussion of how categories are formed is the idea that libraries, documents, and information organization is increasingly influenced by social and technological structures discussed in the theories of social constructionism, and extended mind. For example, Pimentel (2007) discusses the process of information organization from the context of conversation theory (Scott, 2001) and observes that the realm of knowledge organization and representation is shifting from a static document model to a conversational model. Likewise Lanke’s (2007) work discusses the shift in libraries to document networks. Lanke’s work on libraries as places of “participatory conversations” emphasizes the role that Web 2.0 technologies (mashups, apis, blogs, beta-applications) play in turning information interaction into a two directional process. This research exists in contrast to the work completed on knowledge organization which focuses on traditional metadata issues of consistency, interoperability, and scalability (Zeng & Chan, 2004).

The implications of ‘conversation’ or social constructionist theories of information organization are in both the re-phrasing of the debate surrounding user-created organization structures and in a re-definition of the interaction between the document and the users. Pimentel observes that “conversation-based knowledge organization systems could transform an apparent ambiguity into a new kind of clarity” (2007, p. 5). As this field emerges and information systems adopt both the ‘conversation’ approach to documents and to user-created organization systems, the debate surrounding what literacies are required to work in these systems will also grow. The perspective of this literature review is that understanding how participants think about and use metadata is essential to being able to create effective information environments. In the next section of this review, specific roles of metadata in these environments will be examined.

## 4 Roles of metadata in information environments

### 4.1 Metadata definitions

Metadata definitions range from very general concepts of data about data and “the sum total of what one can say about any information object at any level of aggregation” (Gilliland, 2000) to more specific definitions such as “structured data about an object that supports functions associated with the designated object” (Greenberg, 2003). Metadata definitions sometimes focus on specific uses “structured information that feeds into automated processes” (Brand, Daly, & Meyers, 2003, p. 1) and sometimes on context “the sum total of what one can say about any information object at any level of aggregation” (Gilliland, 2000).

The conception of metadata can be as general as any contextual representation of an object or, in other words, anything about anything. Wright’s definition of information (for example) seems more appropriate for this perspective on metadata - “the juxtaposition of data to create meaning” (2007, p. 10). The establishment of this relationship, while on its surface, seems somewhat trivial, can be seen to have dramatic implications on data use and re-use in both print and electronic environments. Within the context of this literature review, the investigation of this relationship between content and context, information and language, symbols and their meaning, looks to the core relationship between ideas of literacy and their impact how individuals conceptualize and use information.

Metadata research includes management approaches (Chapman, 2007b; Halamka, 2008; LeBlanc & Kurth, 2008) innovative uses (Min-Yen & Yee Fan, 2008), metadata quality (Bruce & Hillman, 2004), metadata generation (Greenberg, Pattuelli, Parsia, & Robertson, 2001) and metadata interoperability and standardization (Greenberg, 2005; Zeng & Chan, 2006). Metadata is discussed as supporting the interaction between information organization and system design is discussed in information architecture (Morville & Rosenfeld, 2006; Rosenfeld, 2002). For example, significant design issues in internet applications include search and retrieval (Kwasnik,

Crowston, Nilan, & Roussinov, 2001), navigation (English, et al., 2002), accessibility (Harper & Bechhofer, 2007), and participant cognitive load (Furnas, 1997; Hert, et al., 2007). In this literature review, the research in metadata creation and use in participant created digital documents is of primary interest. Section 4.2 examines the research in this area in more detail.

## **4.2 Research on metadata creation and use**

Research on metadata creation and use by everyday participants has focused on a number of issues including quality of metadata, impact of metadata, and participant attitudes towards metadata creation. Research in the area of metadata creation and everyday use focuses on metadata quality, metadata functions and performance, and participant generated metadata. Research in this area is continuing to develop as evidenced by Zhang and Jastram's (2006) article which points to a lack of research on users' attitudes and competencies with regards to metadata. They assert that two major limitations of metadata studies with respect to users is their lack of investigation of metadata within the context of specific user groups and a lack of definition of the characteristics of these user groups (2006, p. 1101).

In studying the issue of participant created metadata, Zhang and Jastram (2006) found that metadata creators in the four studied fields (Library and Information Science, Business and Industry, Government/Organizations, and Information Technology) display different metadata creation behaviors (2006, p. 1120). They noted that metadata creators tended to emphasize subject style descriptors over other descriptive fields (e.g. date, publisher, etc) and that Business and Industry metadata creators tended to create more metadata than other fields (2006, p. 1120). Both Zhang and Jastram (2006) and Greenberg et al. (2001) found that authors both created reasonably good metadata and valued the metadata they created.

In investigating the extent of use of metadata by participants, Poore (1999) found that metadata is widely used in GIS circles and makes three key points regarding the impact of metadata on end user efficacy in system use. First, Poore draws on Latour's perspective of 'acting at a

distance,' or the ability to store context with data to make it more stable, contextualized, and enable aggregation (Poore, 1999). Second, Poore asserts that the evolution of metadata standards and encoding models from expert or system friendly systems to end-user friendly systems indicates the ability of metadata to be used by end users. Poore cites multiple uses of GIS metadata by end-users including discovery, re-use, and evaluation of information. Poore asserts that in this case, "Metadata were a process of communication between provider and user, a social process that extended the reach of spatial data" (1999). Finally, Poore observes that metadata is a forum for conversation and negotiation between creators and users. She says that while metadata enhances retrieval and use, it also requires great discipline and granularity on both the part of the cataloger and the searcher. She also observes that the use of hypertext and Z39.50 in the GIS community enabled metadata exchange to happen in a new, rapid, distributed way that had not been common several years before. She concludes, observing that metadata can serve a social connecting purpose in that "It could both recruit new converts by acting as a boundary object that could negotiate many different worlds and it could act at a distance by standardizing the delivery of spatial information remotely " (Poore, 1999).

A series of studies by Hert et al. (2007) investigated the role of metadata in enabling use of statistical data investigated the relationship of metadata to information retrieval, evaluation, and task completion. Specifically, Hert et al. found that metadata models grounded in user-centric mental models and oriented towards specific purposes and tasks are more effective than models focused on metadata models focused on data description alone (Hert, et al., 2007, p. 1280). Further, they found that issues of model complexity, granularity, and scope, particularly in a distributed environment with multiple information/metadata sources, contribute to effective metadata use. Finally, they suggest that users do not use metadata in isolation. While their study focused on statistical metadata, they found that users also used navigation metadata and task metadata as they interacted with the statistical system (Hert, et al., 2007, p. 1281).

In both Poore (1999) and Hert et al's articles (2007), the concept of scaffolding and how metadata supports cognitive work emerges. Hert et al. (2007) found that metadata scaffolds

learning and work through its ability to “enhance retrieval processes, improve information organization and navigation, and support management and preservation of digital objects” (2007, p. 1268). They use Jacob’s (Jacob, 2001, p. 89) definition of scaffolding which emphasizes a minimization of cognitive load through the provision of technology tools, knowledge, strategies or processes. The concept of scaffolding has strong connections both to the constructivist theories investigated in the information literacy review and in social constructionist theories investigated in the participant perspectives literature review. Scaffolding is seen from a number of perspectives as having significant impacts on both individual and community knowledge development. In the context of this literature review, scaffolding is created through the use of information organization and metadata skills and concepts.

This section of the literature review has discussed the impact of technology in changing how people interact with information. Just as the moveable type printing press fundamentally changed how people accessed information (Wright, 2007, p. 81) metadata serves a similar function in the current digital environment. Sites such as delicious, and flickr not only make extensive use of metadata, but also use metadata to define how participants interact with and create information. The review of research in this section shows that the issues related to metadata use differ from those of other metadata research areas. Investigating how metadata is used involves considering a holistic information environment that includes elements of individual/social contexts, information need/task context, and system/technical contexts. As was suggested in the review on information literacy, literacy models are helpful in grounding research on how metadata is used and the impact that it has on a participant’s information experience. The next section examines specific examples of how metadata is used in the creation of digital documents.

## 5 Metadata and the Web 2.0 technologies

As web-based information systems and classification methodologies have evolved, the gap between metadata creation and the use of it by the end user has begun to close. The polarized perspectives between highly structured metadata based interfaces versus loosely structured data are being re-examined as new systems use functionality that depend on a hybrid use of expert assigned metadata, automatic harvesting, and user-supplied content. The movement surrounding this development is commonly referred to as Web 2.0 or the read/write web (O'Reilly, 2005). This movement is comprised of three complimentary concepts. First, the technological foundation of Web 2.0 is grounded in web scripting languages, XML, readily exposed data-management applications, and service oriented architecture. Second, many Web 2.0 applications share fundamental philosophical assumptions about the value of user-driven information, open-source data, open source software, and the value of community. Third, web 2.0 applications are grounded in the concept of the 'data' web and as such create services that allow users to find, reuse, and remix metadata in their own sites. These concepts grounding web 2.0 inform the environment in which digital documents are created.

Lawrence Lessig (2004) refers to this movement in part as the remix culture which is concerned with the recombination of intellectual content to create new information objects. A key component of mashups and remix culture is the idea of data and metadata re-use and recombination. This is an emerging topic of research in metadata literature (Dushay & Hillman, 2003; Zeng & Chan, 2006; Zeng & Chan, 2004). While a complete review of web 2.0 applications is outside the scope of this literature review, some relevant examples of foundational web 2.0 technologies at the present time include:

- Facebook and the developer API - Facebook has over 67 million users and over 18,000 applications. Over 95% of Facebook users use at least one application (Facebook, 2008). Facebook applications are built on an open-data application programming interface (API)

which allows developers to harvest data from Facebook develop applications that can be used within the Facebook platform, and can interact with user profiles on-the-fly.

- Amazon Web Services - The Amazon web services platform comprises a suite of applications from data querying and harvesting to application and data hosting services (Amazon, 2008). Amazon web services have been designed as a scalable complete-service solution for developers. As a result a number of applications based on the Amazon data and service set have been developed.
- Flickr - Flickr's primary purpose is as a photo sharing site. Flickr currently has just over 2 Billion photos (compared to Facebook's 4.1 Billion photos) (Arrington, 2007). Flickr API services include a complete suite of data querying, harvesting, writing/deleting, and updating methods and have been implemented in several different scripting languages (Flickr, 2008).

Technology that serves as the foundation of these applications include Ajax (Garrett, 2005), Mashups (Wikipedia, 2006), web services (Berners-Lee, 2002) and linked data (Berners-Lee, 2006). These technologies are helping developers and users dissolve the barrier between information consumer and producer by creating technically simpler and more seamless methods of information interaction. XML-based encoding standards, are feeding these technologies and enabling the interoperability required to facilitate user contributed metadata models. Further, users themselves are demanding data formats that encourage interoperability. Interoperability is becoming an essential element of bibliographic management tools, non-proprietary music, instant messaging lists, and bookmark applications.

Technology has a history of being closely tied to information and learning systems. The review on information literacy included models which discussed the impact of technology on information discovery, use, and management. These models did not address are the complex nature of technology, document, and user interaction that is created in a Web 2.0 context. For

example, users of a bookmarking may be authors, consumers, and secondary contributors within a brief period of time. Two examples of this interaction are of particular interest to this literature review. First, the widespread use of RSS by web-based information providers is a visible intersection between structured metadata and popular use of that metadata. Second, collaborative or community driven content generation systems provide opportunities to investigate how participants think about metadata processes when engaged in this type of information work. Wang et al. (2007, p. 80) detail the structures supporting community software development. Their structure includes theoretical constructs such as computing theory, social psychology, and organization theory, technological infrastructure including web, databases, and applications such as communities, businesses, and interactive entertainment. In addition to these structures, the metadata model is a key element of community driven software.

The 2008 Horizon report in particular focused in on these types of technologies. The report pointed to four types of meta-services in its 2008 report: collaboration webs, data mashups, collective intelligence systems, and social operating systems (Educause, 2008). Previous Horizon reports have looked at social-networking, user-created content, and virtual-reality environments (Educause, 2007), social-computing, context-aware devices, and personal broadcasting (Educause, 2006). A 2005 Forrester Research report resulted in a highly cited chart which details specific social computing platforms including social networks, RSS, blogs, portals, comparison shopping sites, podcasts, wikis, and open source software (Vuorikari, 2007, p. 5). . Additionally, Liccardi et al (2007, p. 231) include a short enumerative list of social tools which bear mentioning: Tagging, Networking, Co-authoring, Sharing, Simulated life on the web and technologies.

As systems evolve, metadata services are becoming more embedded in a wide array of systems. In each of the examples above, specific information organization related tasks can be identified such as tagging, user-publishing, mashups/RSS, and collective filtering/context creation. By investigating the relationship between the three participant, the information system, and the

metadata and document model , this series of literature reviews seeks to investigate how concepts of literacy relate to participants' ability to effectively use these systems.

A metadata-rich document is defined as a document which is significantly composed of or contextualized by embedded metadata. Metadata tends to be more prominent in electronic environments. A printed book for example may contain some metadata on the title and verso pages, perhaps even the index, but the bulk of the work is structured narrative. Likewise, print journal articles and newspapers use a limited amount of metadata in their documents, often indicated by the special position of content in places that are commonly associated with the title, author, or abstract. In contrast, nearly any text-based electronic document is structured to some extent using metadata. HTML pages for example use a metadata standard which drives layout as well as determines structural context. A form of HTML called semantic or XHTML takes a somewhat more rigid approach to this standard, making a point to separate out the layout from the document and uses the XHTML tags to describe the structure of the document only. Other XML based documents often employ a semantic and syntactic based standard such as Dublin Core, EAD, or RSS. These documents derive a significant amount of meaning from the metadata model in which they are created and as such represent a new level of information for the document creator and user. This level of metadata in a document goes well beyond the tradition concept of descriptive metadata which forms many first or second generation electronic databases. Gilliland (2000) defines a taxonomy of these metadata types including Administrative, descriptive, preservation, technical and use metadata. Other types of metadata that have been defined include event-based (Lagoze, 2000), rights (Brand, et al., 2003), and geospatial (FGDC, 2008). The diverse approach to defining metadata types underscores the general agreement that metadata includes a number of different types and purposes. A third level of metadata is defined by Document Type Declarations (DTDs), Schemas (XSD), application profiles, and metadata registries which serve to control and coordinate the contents of metadata records. While metadata models focus on different levels of usage and scope, they also serve to generically define the structure and intended purpose of a metadata system. Detailed explication

of metadata-models are too specific for this literature review but it is useful to note two semantic metadata models and two syntactic models which have had significant impact on metadata-enabled information systems. The following sections examine specific examples of semantic and syntactic models including the Dublin Core model (Hillmann, 2005) in section 5.1 , the Real Simple Syndication (RSS) model in section 5.2 and the Resource Description Framework in section 5.3 .

## **5.1 Dublin Core metadata elements**

In general, the Dublin Core metadata elements represent resource description and representation in which metadata relationships are typed but essentially non-hierarchical. Advocates of the standard assert that it fills a low-barrier level of description which is essential for minimal description. The commonality of the core 15 elements has been argued to be readily accessible and easily grasped (Hillmann, 2005) and yet despite its popularity in libraries Dublin Core is not widely used outside of this arena.

The Dublin Core abstract model (Powell, Nilsson, Naeve, & Johnston, 2005) contextualizes the approach to resource description/surrogation using the Dublin Core metadata standard. The DCMI model is based on traditional resource description approaches, using named elements and attributes to describe facets of an information object. The DCMI model works primarily by description of and reference to external information objects but does also support relationship concepts and referencing of external entities. Dublin Core has been widely implemented in the library and information communities. For example, the University of Maryland Digital library used elements of Dublin Core (notably certain vocabularies) and combined it with more complex standards (such as MODS, LCSH) when necessary (Roper & Gueguen, 2006).

## 5.2 RSS (Real Simple Syndication) metadata model

The RSS metadata model is an RDF based framework encoded in XML. The purpose of the RSS standard is to enable easy harvesting of regularly updated content from information service providers. RSS has been viewed as an alternative to visiting web-pages for information (Pilgrim, 2002), a means for keeping up to date on rapidly changing information (Cohen, 2004), and a data standard for interoperable data processing and knowledge discovery (Thelwall, Prabowo, & Fairclough, 2006; Yahoo, 2008). RSS is a comparable standard to the ATOM syndication format which provides identical functionality via a similar open source standard (Nottingham, 2005). Both RSS and ATOM are document models which use an encoding scheme with elements similar to Dublin Core (such as title, creator, date) but employ a channel or broadcast model in which multiple items are included in a single document. This document is made available as a 'feed' which is made available for download via a web server.

RSS is increasingly used on the internet but research with students in information has shown that there is still little awareness of RSS and its purpose with only 30% of students being aware of RSS prior to the beginning of the course (Mitchell & Smith, 2008). Despite this lack of familiarity, support for RSS is growing. As RSS usage has grown, a number of services have grown to enable mass adoption of this technology. First, many web sites now make much if not all of their content available as RSS feeds. The New York Times for example makes much of its content available via RSS ("New York Times RSS feeds," 2009). Similarly search engines such as Google and Yahoo offer RSS as both a content delivery mechanism in search alerts or a data enhancement mechanism. Likewise, a number of sites offer RSS encoded data as part of their API data set. Both Facebook and Flickr for example output data in RSS feeds. Second, a number of readers and aggregators have been developed to display these feeds. Internet Explorer 7, Firefox version 3, and Safari automatically detect and display RSS feeds with advanced interactive features including, title/link/narrative differentiation, search features, and other embedded display features. Further, a number of aggregators exist online to help users view and

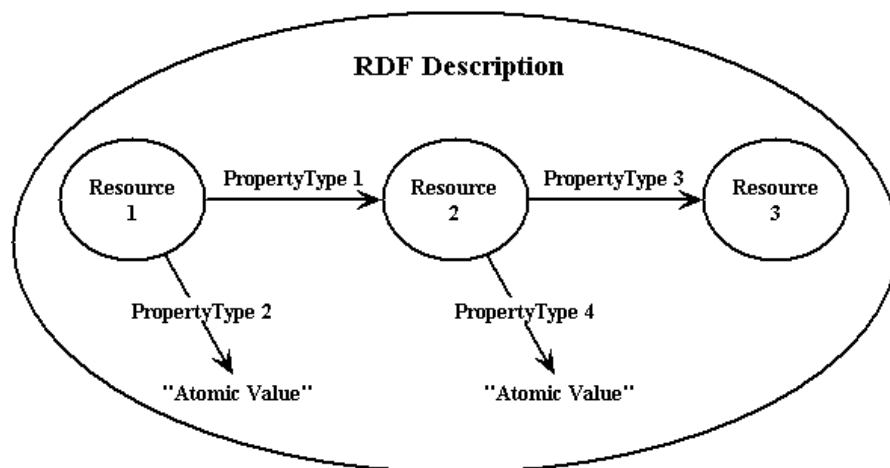
manage their RSS feeds. Finally, sites are being created which help users aggregate and harvest RSS in order to create new data feeds or services. Two examples in particular include Yahoo Pipes (Yahoo, 2008) and Intel's MashMaker (Intel, 2009). These services offer a graphical interface for importing, manipulating, and making available for use multiple data sets. While RSS remains a peripheral enabling technology, there is research into its various uses. A search of EbscoHost's Academic Search Premier for the subject "RSS Feeds" found over 200 articles about RSS from a number of different disciplines. Many of the articles served as basic introductions to the concept while others suggested ways in which RSS could serve traditional information needs including current awareness services (Neilson, 2008), directed advertising (Smith, 2006), and customization for data exchange between businesses (Brandl, 2006). RSS serves as a good example of a metadata-rich document which is publically accessible and usable. Little research has been done that investigates the extent to which users understand either the role of RSS feeds, the structure and syntax of them, or how they can be used or customized for particular use. It is also appealing for research because as a medium it is readily used and viewed in a number of ways.

### **5.3 Resource Description Framework (RDF)**

The RDF framework builds on the concepts of metadata and xml but establishes the additional component of typed relationships. While traditional XML encoded metadata standards relied on hierarchical representations of data, RDF uses the triple concept which includes the three elements of subject, predicate, and object to describe a specific relationship between an object and its contextualizing information (w3c, 2004). This type of relationship allows a structure more closely associated with faceted classification systems than traditional metadata systems. While RDF helps deconstruct the representation experience by removing the requirement of uniform granularity in a metadata record, it uses a very exact data encoding mechanism that removes ambiguity from the system. In brief, the RDF model adds more granular description,

data automation, and a faceted structure to XML based models. Figure 1 from Miller (1998) shows the relationship between a resource, a property, and a value.

**Figure 1 RDF Model**



The discussion of the two metadata standards of Dublin Core and RSS along with the RDF encoding model demonstrates the complex relationship between standards, encoding models, and relationship models. While these systems tend to be hidden from the user in information systems, they still impact how the system is used. As has already been discussed, RSS and ATOM serve as important examples of how structured metadata can be used to create new types of information applications. These applications include common metadata tasks such as read/harvest, re-use, recombine, and annotate. The remainder of this section examines how these tasks are approached.

## **5.4 Creating metadata in community spaces**

The study of metadata within the social/community space takes a different perspective from research which focuses on the use or individual creation of metadata-rich documents. Examining metadata creation in social/community spaces is important given the growing use of public and

social information spaces to create, store, and share contextual data. The development of social software in recent years has led to a number of varying definitions including: “Any type of computing application in which software serves as an intermediary or a focus for a social relation” (Schuler, 1994, p. 29), and “computational facilitation of social studies and human social dynamics as well as the design and use of ICT technologies that consider social context” (Wang, et al., 2007, p. 79). As evidenced by the broad definitions, social software and individual’s motivations for using it are too big to cover in depth here. The focus of this section is on the types of metadata creation services that have been implemented in a social software environment. Liccardi et al. (2007) for example, discuss a wide ranging area of research with regards to the use of social networks in learning. The areas of research discussed range from the process of forming and maintaining social networks, implications of social networks on learning experiences, and implications of the use of technology in these contexts. This literature review focuses on two points with regards to social networks; the role of information communication technology (ICT) in social networks and the role of social networks in enhancing learning objectives.

Liccardi et al. (2007) point to the relationship between (ICT) and social networks and their use in education including non-class time networking, increased ownership of class content, and a sense of larger community membership (2007, p. 231). Liccardi et al. also discuss the potential positive impacts of metadata tasks such as tagging, collective intelligence/recommender systems, and digital library/PIM systems on learning experiences. While they claim that ICT and social networks possess this potential, they also caution that these approaches also increase the impact of issues of student efficacy with technology, impact of limited interaction including asynchronous communication, and lack of verbal and visual cues, and limit) group interaction (2007, p. 230).

This literature review has already discussed the general nature of user-generated and managed metadata-rich information systems. Specific tasks that are enabled in these systems include item storage, tagging, rating, evaluating, managing, and preserving. Collaborative tagging tasks

enable both user and social centered organization including personal information management, community building, and collaborative description. Guy and Tonkin (2006) discuss the contrast between the activities of collaborative description and personal information management in tagging “Still, possibly the real problem with folksonomies is not their chaotic tags but that they are trying to serve two masters at once; the personal collection, and the collective collection” (2006, p. 12). As noted above, Liccardi et al. (2007) see collective description as building personal relevance, an essential element of collaborative tagging. Community building practices are common across several sites, but tend to be emphasized on user-centered sites such as MySpace. Although most systems must choose to explicitly enable these types of functions there are communities such as the community on Twitter.com which have decided to employ their own tag system based on simply using a word prefixed with a ‘#’ sign to indicate the tag. This tagging method allows the aggregation of tweets via third-party systems. An example of the use of this system for a widespread community categorization system was the encouragement of an National Public Radio program to have listeners tag their content in multiple systems with #dctrip09 and #ininaug09 to indicate elements of conversation related to the 2009 presidential inauguration (Carvin, 2009).

Personal information management tasks can be found both in the use of personal description tags in flickr and del.icio.us and in more scholarly site such as Connotea and CiteULike. These sites allow users to categorize electronic resources and share them with other users. Connotea lists its primary tasks as “Organize, Share, and Discover” (“Connotea,” 2006). Connotea in particular enables extended personal management by providing data export options that include bibliographic management software formats. In addition to allowing personal description of tags, Connotea provides contextual linking, indicating the number of cross references on sites and number of comments from other users. As with personal information management, collaborative tagging is a central function of these systems. Of particular interest is the University of Pennsylvania’s attempt to get user descriptions of existing library resources (Penn Tags). The PennTags site uses a popular tag representation method referred to as “Tag Clouds.” The display

at the top of the screen ranks popularly used tags by order and size. This method can be found on several sites and is a popular method of automatic extraction of data from collaborative tagging systems. Another example of the intersection of resource-centered representation and user-generated classification is the set of services offered by LibraryThing (Spalding, 2008). LibraryThing includes user contributed descriptive metadata, surrogate metadata, use metadata, and aggregates and distributes this data using standard book numbers, URIs, and automation techniques.

The second element of interest to this literature review is the role of social networks in enabling learning and documenting knowledge. These two concepts are key components of evaluating the outcomes of the use of ICT in constructivist-centric learning environments. This idea is discussed here as the process of community building. Community building practices can be seen in almost every site but are most seen on specific community sites such as MySpace (<http://www.myspace.com>) and Facebook (<http://www.facebook.com>). These community practices are supported by a wide variety of metadata types. In addition to the metadata types discussed by Gilliland (Gilliland, 2000) such as descriptive metadata, authorship/ownership metadata, and technical metadata there are other types including event metadata, use metadata, user metadata, and relationship metadata, all of which help construct online communities. Instances of the type of information constructed with these types of metadata include general community concepts such as friends, related interests, co-cited resources, and personal profiles. While these applications present a number of metadata creation and use opportunities, it is unclear whether or not their metadata models are explicit to either the end-users of the system or if they are common across multiple platforms.

As this section demonstrates, there are a number of applications which enable the creation of metadata by end-users. In some cases metadata creation serves individual needs while in others metadata creation serves both direct community interests and indirect interests. For this reason, discovering the impact of participant familiarity with metadata is important to creating effective information systems.

## 6 Concluding questions

The arrival of an ‘information age’ has coincided with significant changes in our conception of the document both from perspectives of format/content and from authorship/readership. Further, the perspectives of social constructionism and extended mind emphasize a symbiotic relationship between individuals/communities, technology tools, and information structures and serves as an informing theory behind the motivation and uses behind these metadata-tasks. The research in each of these areas of document content/structure, authorship, and classification represent large bodies of work in themselves and there can be no attempt to pursue the intricate details in this review. Suffice it to say that format changes are challenging traditional organization and management models and creating the need for new metadata schemes, increasing user interactions and the interest in user-focused systems is questioning the relevance of and authority of previous organization systems, and the nature and relationships of individuals and communities are under intense scrutiny.

The historical context provided in this review help both ground current perspective and remind us that there are common trends in research. The work by Vannevar Bush (1945) and Nelson (2002) have seen in part widespread adoption and in part a lack of understanding. Both Nelson (2002) and Berners-Lee (Laningham, 2006) have commented on the success and failures of their conceptions of the web. This reminder, that research and practice are often combined only through compromise is also seen in the different perspectives surrounding the document, the participant, and the participations in the reviewed literature. While organization literature clearly recognizes the iterative process of interaction at the base of current information seeking models, it decidedly focuses on using information to guide document description. In contrast, the use of categorization practices in information retrieval and use applications rarely discussed the possible impact on document classification and tended to think only about the user in an iterative fashion. Information Architecture literature took the most effort to demonstrate the cross

discipline dependence that exists between the use of categorization systems to describe and facilitate retrieval.

This literature review found that while there is significant research in the metadata field, much of it focuses on standards and systems that are far more complex than the common information consumer contends with in their everyday information use. Likewise, it found that little research focuses on the user experience of metadata and rather looks at the outcomes of metadata use in specific environments. The literature also indicates that the evolution of the document along with the changing nature of our information environments is making metadata more available to these creators/consumers and that a number of web sites are creating environments in which these users' roles are changing to include tasks which include metadata creation, enhancement, and use. The review of metadata-centric web services demonstrated how diverse the list of metadata-related concepts and tasks can become. Metadata tasks, skills, and concepts

The information literacy review concluded with an analytical framework that allowed the investigation of literacy concepts from the perspective of skills, concepts, and contexts along the axes of pedagogy, theoretical foundations, and the related role of the information environment. In order to position this research to use that model, this literature review investigated four broad metadata tasks that will be used in the evaluation of participant metadata-literacy. Each broad task has suggested skills and concepts pulled from the literature review which will help identify elements of literacy for these areas. Hert et al. (2007, p. 1268) suggests four perspectives that are valuable creating metadata models: content, user, organizational, and technical. Although this review does not focus on the technical aspects and the organizational aspects are in essence broken out into their own section, the two elements of user and content perspective will prove valuable in discussing metadata tasks involving item description, item classification, knowledge representation, and information use and re-use. Each of these four organization activities describes an intersection between a participant and a metadata-rich document and represents broad categories where information organization is closely tied to concepts of participation and literacy. This intersection is comprised of both a mental model of the information being encoded

and a technological foundation for that encoding. The metadata-literacy literature review looks at two questions in relation to these activities. First, how do participants think about the four above activities during document interaction and second what skill or conceptual knowledge can be generalized from observing their interaction with documents during exercises in these organization activities. The purpose of this review was to lay the groundwork for understanding the relevance of information organization and metadata research to the current information environment. Particular attention was paid to how and to what extent users interact with metadata (as opposed to traditionally displayed information) and what relationships those interactions have to organization theory. In the following literature reviews on participant perspective and metadata literacy these interactions and relationships will be investigated in more depth.

## References

- . (Penn Tags). Retrieved 24 Apr, 2006, from <http://tags.library.upenn.edu/>
- Abbas, J. (2005). Creating metadata for children's resources: issues, research, and current developments. *Library Trends*, 54(2), 303(315).
- Almasy, E. (2005). Tools for Creating Your Own Resource Portal: CWIS and the Scout Portal Toolkit. *Library Trends*, 53(4), 620-636.
- Amazon. (2008). Welcome to Amazon Web Services. *Amazon.com* Retrieved 12 March 2008, from <http://www.amazon.com/gp/browse.html?node=3435361>
- Araghi, G. F. (2005). Users Satisfaction Through Better Indexing. *Cataloging & Classification Quarterly*, 40(2). Retrieved from [http://www.haworthpress.com.libproxy.lib.unc.edu/store/E-Text/View\\_EText.asp?sid=8AAM21Q55HBR9JR14F4ST11SUC500439&a=3&s=J104&v=40&i=2&fn=J104v40n02%5F03](http://www.haworthpress.com.libproxy.lib.unc.edu/store/E-Text/View_EText.asp?sid=8AAM21Q55HBR9JR14F4ST11SUC500439&a=3&s=J104&v=40&i=2&fn=J104v40n02%5F03)
- Arrington, M. (2007, 13 November 2007). 2 Billion Photos on Flickr. *TechCrunch* Retrieved 12 March 2008, from <http://www.techcrunch.com/2007/11/13/2-billion-photos-on-flickr/>

- Bainbridge, D., Thompson, J., & Witten, I. H. (2003). *Assembling and enriching digital library collections*. Paper presented at the Proceedings of the 3rd ACM/IEEE-CS joint conference on Digital libraries.
- Barreau, D., & Nardi, B. A. (1995). Finding and Reminding: File Organization from the Desktop. *SIGCHI Bulletin*, 27(3), 39-42.
- Barreau, D. K. (1995). Context as a factor in personal information management systems. *Journal of the American Society for Information Science*, 46(5), 327-339.
- Bates, M. J. (1998). Indexing and access for digital libraries and the internet: Human, database, and domain factors. *Journal of the American Society for Information Science*, 49(13), 1185-1205.
- Berners-Lee, T. (2002, 24 July 2002). Web Services. W3C Retrieved 12 March 2008, from <http://www.w3.org/DesignIssues/WebServices.html>
- Berners-Lee, T. (2006, 2006-07-27). Linked Data. 2008-09-09, from <http://www.w3.org/DesignIssues/LinkedData.html>
- Borgman, C. L. (1999). What are digital libraries? Competing visions. *Information Processing & Management*, 35(3), 227-243. Retrieved from [http://dx.doi.org/10.1016/S0306-4573\(98\)00059-4](http://dx.doi.org/10.1016/S0306-4573(98)00059-4). doi:[http://dx.doi.org/10.1016/S0306-4573\(98\)00059-4](http://dx.doi.org/10.1016/S0306-4573(98)00059-4)

Borland, J. (2007). A Smarter Web. *Technology Review*, 110(2), 64-71.

Bowker, G. C., & Star, S. L. (1999, 1999). Sorting things out classification and its consequences.

from <http://ezproxy.wfu.edu:3000/login?url=http://www.netlibrary.com/urlapi.asp?action=summary&v=1&bookid=13186>

Brand, A., Daly, F., & Meyers, B. (2003). *Metadata Demystified*. Hannover, PA: Sheridan Press.

Brandl, D. (2006). RSS finds pertinent manufacturing data. *Control Engineering*, 53(1), 16-16.

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

Bruce, H., Jones, W., & Dumais, S. (2004). Keeping and re-finding information on the web:

What do people do and what do they need? *Proceedings of the American Society for Information Science and Technology*, 41(1), 129-137.

Bruce, T., & Hillman, D. (2004). The continuum of metadata quality: Defining, expressing, exploiting. In D. Hillman (Ed.), *Metadata in Practice* (pp. 248-256). Chicago ALA Editions.

Buckland, M. (1991). *Information and information systems*. New York: Greenwood Press.

Buckland, M. K. (1997). What is a document? *Journal of the American Society for Information Science*, 48(9), 804-809.

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: Draft 2B.

Campbell, D. G. F., Karl V. (2005). Academic Libraries and the Semantic Web: What the Future  
May Hold for Research-Supporting Library Catalogues. *The Journal of Academic  
Librarianship*, 30(5), 382-390.

Carr, N. (2008). Is Google Making Us Stupid? (cover story). *Atlantic Monthly* (10727825),  
302(1), 56-63.

Carvin, A. (2009). Moving Forward With Inauguration Report. Retrieved 25 Jan 2009,  
from [http://www.npr.org/blogs/inside/2009/01/moving\\_forward\\_with\\_inaugurati.html](http://www.npr.org/blogs/inside/2009/01/moving_forward_with_inaugurati.html)

Chapman, A. (2007a). Resource discovery: catalogs, cataloging, and the user. *Library Trends*,  
55(4), 917(915).

Chapman, J. W. (2007b). The Roles of the Metadata Librarian in a Research Library. *Library  
Resources & Technical Services*, 51(4), 279-285.

Clark, A. (2001). Reasons, robots and the extended mind. *Mind & Language*, 16(2), 121-145.

Clark, A., & Chalmers, D. (1998). The extended mind (Active externalism). *Analysis*, 58(1), 7-19.

Cohen, S. M. (2004). RSS. *Computers in Libraries*, 24(2), 28-28.

Connotea. (2006). Retrieved 23 Apr, 2006, from <http://www.connotea.org/>

Crook, C. (2005). Addressing research at the intersection of academic literacies and new technology. *International Journal of Educational Research*, 43(7-8), 509-518.

Dongwon, J., Peter Hoh, I., Fran, J., Young-Gab, K., & Doo-Kwon, B. (2005). A message conversion system, XML-based metadata semantics description language and metadata repository. *Journal of Information Science*, 31(5), 394-406.

Dushay, N., & Hillman, D. (2003). *Analyzing metadata for use and re-use*. Paper presented at the DC-2003 Dublin Core Conference: Supporting Communities of Discourse and Practice . Metadata Research and Applications.

Educause. (2006). *The Horizon Report: 2006 Edition*. Boulder, CO: Educause.

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

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

- English, J., Hearst, M., Rashmi, S., Swearingen, K., & Yee, K.-P. (2002). Flexible Search and Navigation using Faceted Metadata. Retrieved from [http://www.rashmisinha.com/articles/flamenco\\_S\\_02.pdf](http://www.rashmisinha.com/articles/flamenco_S_02.pdf)
- Eriksson, H. (2007). The semantic-document approach to combining documents and ontologies. *International Journal of Human-Computer Studies*, 65(7), 624-639.
- Facebook. (2008, 2008). Facebook statistics. *Facebook* Retrieved 13 March, 2008, from <http://www.facebook.com/press/info.php?statistics>
- FGDC. (2008, 25 Apr 2005). FGDC Geospatial Metadata. Retrieved 25 Jan 2009, from <http://www.fgdc.gov/metadata>
- Flickr. (2008, 2008). Flickr Services. Retrieved 12 March 2008, from <http://www.flickr.com/services/api/>
- Furnas, G. W. (1997). *Effective View Navigation*. Paper presented at the Proceedings of ACM CHI 97, Conference on Human Factors in Computing Systems. Retrieved from <http://www.si.umich.edu/~furnas/Papers/CHI97-EVN.2.pdf>
- Garrett, J. J. (2005, 12 Mar 2006). Ajax: A New Approach to Web Applications. Retrieved 12 Mar, 2006, from <http://www.adaptivepath.com/publications/essays/archives/000385.php>

- Gilliland, A. J. (2000). Setting the Stage. *Introduction to Metadata: Pathways to Digital Information* Retrieved February 28, 2008, from [http://www.getty.edu/research/conducting\\_research/standards/intrometadata/setting.html](http://www.getty.edu/research/conducting_research/standards/intrometadata/setting.html)
- Golder, S. A., & Huberman, B. (2006). The Structure of Collaborative Tagging Systems. *Journal of Information Science*, 32(2), 198-208
- Green, R. (2001). Relationships in the Organization of Knowledge: An Overview. In C. A. Bean & R. Green (Eds.), *Relationships in teh Organizatiton of Knowledge*. Boston: Kluwer Academic.
- Greenberg, J. (2003). Metadata and the World Wide Web. In M. A. Drake. (Ed.), *Encyclopedia of Library and Information Science*. New York: Marcel Dekker.
- Greenberg, J. (2005). Understanding Metadata and Metadata Schemes. *Cataloging & Classification Quarterly*, 40(3/4), 17-36.
- Greenberg, J., Pattuelli, M. C., Parsia, B., & Robertson, W. D. (2001). Author-generated Dublin Core Metadata for Web Resources: A Baseline Study in an Organization. *Journal of Digital Information*, 2(2).

Guy, M., & Tonkin, E. (2006). Folksonomies: Tidying up Tags? *D-Lib Magazine*. Retrieved from <http://www.dlib.org/dlib/january06/guy/01guy.html>

Halamka, J. D. (2008). Navigating Info Overload. *Computerworld*, pp. 26-26. Retrieved from <http://ezproxy.wfu.edu:3000/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=28788103&site=ehost-live>

Harper, S., & Bechhofer, S. (2007). SADIE: Structural semantics for accessibility and device independence. *ACM Trans. Comput.-Hum. Interact.*, 14(2), 10.

Harter, S. P. (1996). *What is a Digital Library? Definitions, Content, and Issues*. Paper presented at the KOLISS DL '96. Retrieved from <http://php.indiana.edu/%7Eharter/korea-paper.htm>

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.

Hembrooke, H. A., Granka, L. A., Gay, G. K., & Liddy, E. D. (2005). The effects of expertise and feedback on search term selection and subsequent learning. *Journal of the American Society for Information Science and Technology*, 56(8), 861-871.

Hert, C. A., Denn, S. O., Gillman, D. W., Oh, J. S., Pattuelli, M. C., & Hernández, N. (2007). Investigating and modeling metadata use to support information architecture

development in the statistical knowledge network. *Journal of the American Society for Information Science & Technology*, 58(9), 1267-1284.

Hillmann, D. (2005). Using Dublin Core. Retrieved 25 Jan 2009,  
from <http://dublincore.org/documents/usageguide/>

Holland, G. A. (2006). Associating social constructionism and extended cognition in information studies. *Journal of Documentation*, 62(1), 91-100.

Intel. (2009). Mashmaker. Retrieved 25 Jan 2009, from <http://mashmaker.intel.com/web/>

Jacob, E. K. (2001). The everyday world of work: two approaches to the investigation of classification in context. *Journal of Documentation*, 57, 76-100.

Jacob, E. K. (2004). Classification and Categorization: A Difference that Makes a Difference. *Library Trends*, 52(3), 515-540.

Jones, W. (2007). Personal Information Management. *Annual Review of Information Science and Technology*, 41,

Ju, B. (2007). Does domain knowledge matter: Mapping users' expertise to their information interactions. *Journal of the American Society for Information Science and Technology*, 58(13), 2007-2020.

- Kwasnik, B. H. (1989). *How a Personal Document's Intended Use or Purpose Affects Its Classification in an Office*. Paper presented at the 12th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, New York.
- Kwasnik, B. H. (1999). The Role of Classification in Knowledge Representation and Discovery. *Library Trends*, 48(1), 22.
- Kwasnik, B. H., Crowston, K., Nilan, M., & Roussinov, D. (2001). Identifying Document Genre to Improve Web Search Effectiveness. *Bulletin of the American Association of Information Science and Technology*, 27(2).
- Lagoze, C. (2000). *Business Unusual: How 'Event-Awareness' May Breathe Life into the Catalog*. Paper presented at the Bicentennial Conference on Bibliographic Control in the New Millennium, Library of Congress, Washington, DC.
- Lagoze, C. (2004). Bridging the past and future: Scholarly Communication in the 21st Century. Retrieved from <http://www.cs.cornell.edu/lagoze/papers/Tsukuba%202004.pdf>
- Lakoff, G. (1987). *Women, fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: U of Chicago P.
- Langridge. (1992). *Classification: Its Kinds, Elements, Systems, and Applications*. London: Bowker.

Laningham, S. (2006). *developerWorks Interviews: Tim Berners-Lee, IBM: IBM.*

Lankes, R. D., Silverstein, J., & Nicholson, S. (2007). *Participatory Networks: The Library as Conversation: American Library Association's Office for Information Technology Policy.*

LeBlanc, J., & Kurth, M. (2008). An Operational Model for Library Metadata Maintenance. *Library Resources & Technical Services, 52(1), 54-59.*

Lessig, L. (2004). *Free culture : how big media uses technology and the law to lock down culture and control creativity.* New York: Penguin Press.

Library of Congress Working Group on the Future of Bibliographic Control. (2008). *On the record: Report of the Library of Congress Working Group on the Future of Bibliographic Control.* Washington D.C.: Library of Congress.

Liccardi, I., Ounnas, A., Pau, R., Massey, E., Kinnunen, P., Lewthwaite, S., et al. (2007). The role of social networks in students' learning experiences. *SIGCSE Bull., 39(4), 224-237.*

Licklider, J. C. R., & Taylor, R. W. (1968). The Computer as a Communication Device. *Science and Technology.*

- Macgregor, G., & McCulloch, E. (2006). Collaborative Tagging as a Knowledge Organisation and Resource Discovery Tool. *Library Review* 55(5). Retrieved from <http://eprints.rclis.org/archive/00005703/>
- MaKinster, J. G., Beghetto, R. A., & Plucker, J. A. (2002). Why Can't I Find Newton's Third Law? Case Studies of Students' Use of the Web as a Science Resource. *Journal of Science Education and Technology*, 11(2), 155-172.
- Mann, T. (2006). *What is going on at the Library of Congress?* Paper presented at the AFSCME 2910. Retrieved from <http://www.guild2910.org/AFSCMEWhatIsGoingOn.pdf>
- Mann, T. (2008). *"On the record" but off the track.* Paper presented at the AFSCME 2910. Retrieved from <http://www.guild2910.org/WorkingGrpResponse2008.pdf>
- Marchionini, G. (1995). *Information Seeking in Electronic Environments*. Cambridge: Cambridge UP.
- Marchionini, G., & Maurer, H. (1995). The roles of digital libraries in teaching and learning. *Commun. ACM*, 38(4), 67-75.
- Markoff, J. (2005). *What the dormouse said : how the sixties counterculture shaped the personal computer industry*. New York: Viking Penguin.

Marlow, C., Naaman, M., boyd, d., & Davis, M. (2006). *HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, ToRead*. Paper presented at the Proceedings of Hypertext 2006, New York.

Marshall, C. C., & Bly, S. (2004). *Sharing encountered information: digital libraries get a social life*. Paper presented at the Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries.

Mejias, U. A. (2004). Ideant: A del.icio.us study - Bookmark, Classify and Share: A mini-ethnography of social practices in a distributed classification community. Retrieved from [http://ideant.typepad.com/ideant/2004/12/a\\_delicious\\_stu.html](http://ideant.typepad.com/ideant/2004/12/a_delicious_stu.html)

Miller, E. (1998). An Introduction to the Resource Descriptive Framework. *D-Lib Magazine*,

Min-Yen, K., & Yee Fan, T. (2008). Record Matching in Digital Library Metadata. *Communications of the ACM*, 51(2), 91-94.

Mitchell, E., & Smith, S. (2008). *Bringing Information Literacy into the Social Sphere*. Paper presented at the LITA 2008.

Morville, P. (2005). *Ambient Findability*. Cambridge: O'Reilly.

Morville, P., & Rosenfeld, L. (2006). *Information Architecture for the World Wide Web*.

- Neilson, C. (2008). Current Awareness on a Shoe String: RSS at the HQC. *Internet Reference Services Quarterly*, 13(1), 57-67.
- Nelson, T. (2002, 12, May 2002). I DON'T BUY IN. Retrieved 11 April, 2008, from <http://hyperland.com/buyin.txt>
- New York Times RSS feeds. (2009, Jan 25, 2009). Retrieved 25 Jan 2009, from <http://www.nytimes.com/services/xml/rss/index.html?scp=1-spot&sq=rss&st=cse>
- Nottingham, M. (2005, Dec 2005). The Atom Syndication Format. Retrieved 25 Jan 2009, from <http://tools.ietf.org/html/rfc4287>
- O'Reilly, T. (2005, 30 September, 2005). What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. *O'Reilly* Retrieved 12 March, 2008, from <http://www.oreilly.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>
- Oostendorp, H. v., Breure, L., & Dillon, A. (Eds.). (2005). *Creation, Use, and Deployment of Digital Information*. Mahwah, NJ: Lawrence Erlbaum
- Pilgrim, M. (2002, 18 December, 2002). What is RSS? *O'reilly XML.xom* Retrieved 12 March, 2008, from <http://www.xml.com/pub/a/2002/12/18/dive-into-xml.html>
- Pimentel, D. M. (2007). *Exploring classification as conversation*. Paper presented at the Proceedings of the North American Symposium on Knowledge Organization.

Poore, B. S. (1999). A metanarrative of metadata: The social construction of the FGDC=s Content Standard for Digital Geospatial Metadata.

Powell, A., Nilsson, M., Naeve, A., & Johnston, P. (2005). DCMI Abstract Model. Retrieved 28 July, 2006, from <http://dublincore.org/documents/abstract-model/index.shtml>

Raymond, E. S. (2001). *The cathedral and the bazaar : musings on Linux and Open Source by an accidental revolutionary*. Cambridge, Mass.: O'Reilly.

Robertson, W. D., & Greenberg, J. (2004, October 11-14, 2004). *Architecting a Cross-Disciplinary Thesaurus for the Semantic Web*. Paper presented at the Metadata across Languages and Cultures: Proceedings of the International Conference on Dublin Core and Metadata Applications.

Roper, J., & Gueguen, G. (2006). Not so different after all: Creating access to diverse objects in digital repositories.

Rosch, E. (1978). *Cognition and categorization*. Hillsdale, N.J. New York: L. Erlbaum Associates ; distributed by Halsted Press.

Rosenfeld, L. (2002). Information architecture: Looking ahead. *Journal of the American Society for Information Science and Technology*, 53(10), 874-876.

- Rosenfeld, L., & Morville, P. (2002). *Information Architecture for the World Wide Web*.  
Cambridge, MA: O'Reilly.
- Schuler, D. (1994). Social computing. *Commun. ACM*, 37(1), 28-29.
- Scott, B. (2001). Gordon Pask's Conversation Theory: A Domain Independent Constructivist  
Model of Human Knowing. *Foundations of Science*, 6(4), 343-360.
- Shirkey, C. (2006). Ontology is Overrated: Categories, Links, and Tags Retrieved 20 Mar, 2006,  
from [http://shirky.com/writings/ontology\\_overrated.html](http://shirky.com/writings/ontology_overrated.html)
- Smith, S. (2006). Advertising On-Demand. *EContent*, 29(8), 19-19.
- Spalding, T. (2008). *What is social cataloging and where is it going?* Paper presented at the  
LITA 2008.
- Thelwall, M., Prabowo, R., & Fairclough, R. (2006). Are raw RSS feeds suitable for broad issue  
scanning? A science concern case study. *Journal of the American Society for Information  
Science and Technology*, 57(12), 1644-1654.
- Thibodeau, S. G., & Barry, R. K. (2002, 2002). Development of the Encoded Archival  
Description DTD. Retrieved 24-OCT-2005, from <http://www.loc.gov/ead/eaddev.html>

- Toms, E. G. (2002). Information interaction: Providing a framework for information architecture. *Journal of the American Society for Information Science and Technology*, 53(10), 855-862.
- Tuominen, K., Talja, S., & Savolainen, R. (2003). Multiperspective digital libraries: The implications of constructionism for the development of digital libraries. *Journal of the American Society for Information Science and Technology*, 54(6), 561-569.
- Vuorikari, R. (2007). *Folkosnomies, Social Bookmarking, and Tagging: State-of-the-art*: European Schoolnet.
- Vygotsky, L. (1977). *Mind In Society: The Development of Higher Psychological Processes*. Cambridge: Harvard UP.
- w3c. (1999, 24 May, 1999). Web Characterization Terminology & Definitions Sheet. Retrieved 13 March, 2008, from <http://www.w3.org/1999/05/WCA-terms/>
- 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.

Ware, P. D., & Warschauer, M. (2005). Hybrid literacy texts and practices in technology-intensive environments. *International Journal of Educational Research*, 43(7-8), 432-445.

Weibel, S. (1995a). Metadata: The Foundations of Resource Description. *D-Lib Magazine*.

Weibel, S. (1995b). Metadata: The Foundations of Resource Description. *D-Lib Magazine*,

Weibel, S. (2005). Border Crossings: Reflections on a Decade of Metadata Consensus Building.

*D-Lib Magazine*, 11(7/8). Retrieved

from <http://www.dlib.org/dlib/july05/weibel/07weibel.html>

Weinberger, D. (2007). *Everything is Miscellaneous: The power of the new digital disorder*.

New York: Times Books.

Wikipedia. (2006). Mashup - Web Application Hybrid. *Wikipedia* Retrieved 16 Apr, 2006,

from [http://en.wikipedia.org/wiki/Mashup\\_\(web\\_application\\_hybrid\)](http://en.wikipedia.org/wiki/Mashup_(web_application_hybrid))

Woodruff, A., Rosenholtz, R., Morrison, J. B., Faulring, A., & Pirolli, P. (2002). A Comparison

of the Use of Text Summaries, Plain Thumbnails, and Enhanced Thumbnails for Web

Search Tasks. *Journal of the American Society for Information Science and Technology*,

53(2), 172-185.

Wright, A. (2007). *Glut : mastering information through the ages*. Washington, D.C.: Joseph Henry Press.

Yahoo. (2008). Yahoo Pipes, *Yahoo Pipes*: Yahoo.

Yee, K.-P., Swearingen, K., Li, K., & Hearst, M. (2003). *Faceted Metadata for Image Search and Browsing*. Paper presented at the Proceedings of ACM CHI 2003. Retrieved from <http://bailando.sims.berkeley.edu/papers/flamenco-chi03.pdf>

Zeng, M., & Chan, L. M. (2006). Metadata Interoperability and Standardization – A Study of Methodology Part II. *D-Lib Magazine*, 12(6).

Zeng, M. L., & Chan, L. M. (2004). Trends and issues in establishing interoperability among knowledge organization systems. *Journal of the American Society for Information Science and Technology*, 55(5), 377-395.

Zhang, J., & Jastram, I. (2006). A study of the metadata creation behavior of different user groups on the Internet. *Information Processing & Management*, 42(4), 1099-1122.

