The INVENT framework: Examining the role of information visualization in the reconceptualization of digital libraries

Abstract. The objective of this paper is to show how information visualization can play an important and catalytic role in the reconceptualization of digital libraries as interactive knowledge environments. Information visualization has long been described as a beneficial and promising technology for digital libraries. Despite this, digital libraries infrequently rely on information visualization concepts and techniques. To date, digital libraries have been more concerned with fundamental issues, such as digitization, organization, preservation, and facilitating access through traditional search and browse interfaces. Currently, digital libraries are conceptualized as curated and searchable document repositories. But new research directions are attempting to reconceptualize them as interactive knowledge environments. This paper re-examines the role of information visualization in digital libraries in light of this new vision. In particular, we introduce a new conceptual framework for digital libraries called INVENT: INteractive Visual ENironmenTs. The INVENT framework emphasizes the importance of rich interaction with representations of information, especially visual representations, for supporting epistemic activities. There are six elements in the framework: digital objects, representations, activities, interactions, actors, and ecologies. This paper suggests that these elements should be conceptual cornerstones in the knowledge environment conceptualization of digital libraries.

1 Introduction

In their comprehensive paper reviewing the field of digital libraries, Fox and Urs wrote that “information visualization applied to digital libraries shows particular promise” (Fox & Urs, 2002, p. 543). This paper is interested in that promise: the role of information visualization in the development, evolution, and conceptualization of digital libraries.

One of the requirements for writing about digital libraries, judging from the literature, is to mention their inherent complexity and their enormous potential for developing and expanding human knowledge. Such observations are so pervasive they verge on cliché. Although we do not dispute them, they point to fundamental challenges for the evolution of digital libraries. Today, most digital libraries consist of little more than a document repository, a search engine, and an interface that supports keyword searching and hypertext browsing (Bieber et al., 2002; Feng et al., 2005; Marchionini, 1999; Paepcke, 1996). Digital libraries implemented in this manner hardly seem capable of fulfilling the ambitious goal of helping people manage, understand, and create knowledge.

For researchers, a primary concern is reducing the gap between digital libraries as they are, and digital libraries as they could or should be. For instance, Soergel (2002, para. 1) argues that “much of [digital library] practice is still at the stage of the ‘horseless carriage',” and Lynch (2003, para. 3) suggests that “we should recognize the limitations of a research program focusing on digital libraries as we understand them today.” Today, digital libraries are understood as curated, searchable, and networked document repositories (Larsen, 2004). This conceptualization has often been criticized for constraining the development of the field, which has led to many suggestions for overcoming these constraints (Larsen, 2004; Levy & Marshall, 1995; Marchionini, 1999; Soergel, 2002). The underlying theme of these suggestions is the need for a reconceptualization of digital libraries.

Digital libraries are entering a new phase in their development (Larsen, 2004; Soergel, 2002). This new phase aims to transform them from searchable document repositories to interactive knowledge environments. The goal is to develop digital libraries that do more than merely help people find relevant documents. They should also help people use those documents to create knowledge. This is, of course, a simplified description, but it captures the essence of the knowledge environment—digital libraries that actively guide, enable, and support users in their cognitive and knowledge work activities. Several frameworks have been proposed for achieving this vision, and they are examined in Section 4 (Beiber et al., 2002; Besser, 2002; Gonçalves et al., 2004; Larsen, 2004; Levy & Marshall, 1995; Marchionini, 1999; Soergel, 2002). For now, we draw attention to one of these documents, a report submitted to the National Science Foundation (NSF) by leading researchers about the future of digital libraries (Larsen, 2004). According to this report, the next phase of digital library research should strive to transform existing digital libraries into an “advanced digital knowledge environment that will enlighten and empower the next generation” (Larsen, 2004, p. ii). The report describes the next generation of digital libraries as ubiquitous, dynamic, interactive, collaborative environments that actively support learning, thinking, reasoning, problem solving, sense making, and knowledge creation. Although the NSF report differs from other perspectives on future digital libraries, it mainly differs with respect to technical details. At a conceptual level, there remains a common vision of digital libraries as interactive knowledge environments (for instance, see: Beiber et al., 2002; Börner & Chen, 2002a; Dillon, 2002; Marchionini, 1999; Soergel, 2002).

The knowledge environment model of digital libraries is radically different from the searchable repository model. In the repository model, digital libraries are designed to help users quickly and efficiently winnow a large document collection to a manageable set of relevant items. In the knowledge environment model, digital libraries are designed to support the full cycle of knowledge discovery, creation, and use. Clearly, this shift is ambitious and presents many challenges.

In this paper, we are specifically interested in the role that information visualization can play in the reconceptualization of digital libraries as knowledge environments. We consider information visualization to be more than a set of technologies and techniques for developing more usable interfaces to digital collections. We suggest that it can play a catalytic role in this reconceptualization process. To this end, this paper introduces the INVENT framework, which stands for INteractive Visual ENvironmenTs. The framework emphasizes how interaction with representations of information, especially visual representations, is essential for supporting cognitive and knowledge work activities. There are six conceptual elements in the framework: digital objects, representations, activities, interactions, actors, and ecologies. Section 5 describes the INVENT framework in detail.

This paper is primarily concerned with two topics: the reconceptualization of digital libraries as knowledge environments, and the role of information visualization in this reconceptualization. The INVENT framework, however, brings together research from a variety of fields, not just information visualization. These fields include interaction design, information architecture, human-computer interaction, knowledge management, and cognitive science, among others.

The remainder of the paper is divided into five parts. In the first part, we explore the transition of digital libraries from document repositories to interactive knowledge environments. In the second part, we review how information visualization has, to date, been applied to digital libraries. We also consider why visualization principles have had only a modest impact on digital library interfaces. In the third part, we compare seven conceptual frameworks for digital libraries. These frameworks have much in common with the ideas, goals, and motivations of information visualization. The fourth part, which forms the bulk of the paper, introduces the INVENT framework. This part explains each of the conceptual elements of the framework, primarily through the lens of information visualization. In the fifth and final part, we summarize the discussion and briefly outline future research suggested by the INVENT framework.

2 Digital libraries

Different communities and stakeholders emphasize different aspects of digital libraries (Bishop & Star, 1996; Borgman, 1999; Fox et. al., 1995; Fox & Urs, 2002; Levy, 2000). Computer scientists tend to see them in relation to databases, networks, retrieval engines, and other enabling technologies. Teachers are inclined to concentrate on their educational potential. Politicians commonly regard digital libraries as tools of overcoming digital divides and providing citizens with equal access to information resources. Librarians generally view them as extensions of the library-as-institution—a tool for “revitalizing their mission of accessing and disseminating information and knowledge” (Fox & Urs, 2002¸ p. 513).

Despite the diversity of opinions about the purpose and value of digital libraries, there is a shared understanding that like traditional libraries, digital libraries can and should play a central and multi-faceted role in developing and expanding human knowledge. Nevertheless, after more than a decade of research and development, there is recognition that the ability of digital libraries to fulfill this ambitious role has been, in large part, constrained by their current conceptualization (Bawden & Rowlands, 1999a; Bieber et al., 2002; Levy & Marshall, 1995; Larsen, 2004; Marchionini, 1999; Soergel, 2002).

Digital libraries were initially conceptualized as curated, searchable, networked repositories of digital resources (Larsen, 2004). This remains the prevailing conceptualization, but critics have argued that it has a limited vision and an inhibiting effect on innovation (Dillon, 2002; Levy & Marshall, 1995; Lynch, 2003; Marchionini, 1999; Soergel, 2002). To cite an early example of such criticism, Levy and Marshall (1995) examined the assumptions behind digital libraries and concluded that digital libraries were "limited by a largely unexamined and unintended allegiance to an idealized view of what libraries have been, rather than what they actually are or could be" (Levy & Marshall, 1995, p. 77-8). To cite a more recent example, Lynch has observed that digital libraries are still conceptualized in relation to traditional libraries, and subsequently, "much of the work on digital libraries has emphasized modernization—applying technology to do *what we have always done*, only more efficiently and effectively" (Lynch, 2003, para. 2; emphasis added). What is it that libraries have always done? Traditionally, libraries select, collect, organize, manage, store, preserve, and facilitate access to information (Borgman, 1997). Although these activities are important, they emphasize developing, maintaining, and improving a collection of digital resources instead of how people work with those resources to pursue various knowledge-related goals.

An emerging conceptualization of digital libraries, central to the theme of this paper, emphasizes what people do, or might do, with digital resources (Dillon 2002; Larsen, 2004; Soergel, 2002). In this conceptualization, digital libraries are more than searchable document repositories. They are environments that help people create knowledge from the digital resources in the repository. Some key ideas in this emerging conceptualization include: "information uses beyond mere information retrieval must be accommodated" (Budhu & Coleman, 2002, para. 17); "libraries connect people and information; digital libraries amplify and augment these connections" (Marchionini & Fox, 1999, p. 219); digital libraries should be "a workspace with rich content and powerful tools where people can work independently or collaborate with others to learn and to solve their information problems" (Marchionini, 1999, p. 40); the interfaces to a digital library "should be thought of as a complete workspace in which a variety of tools are available for flexibly interacting with richly-conveyed information" (Rao et al., 1995, p. 39); digital libraries "must go beyond [providing access to information] and support new ways of intellectual work" (Soergel, 2002, para. 4).

These examples are united by the idea that digital libraries should be more than tools for efficiently finding information—they should also be interactive environments that enable, support, and guide cognitive and knowledge work activities. In other words, digital libraries are shifting away from a classical document repository conceptualization and towards a knowledge environment conceptualization. A subtle but key aspect of this shift is that the knowledge environment is not a replacement for the document repository. Instead, the knowledge environment builds on the document repository, broadening the kinds of activities that digital libraries support and emphasizing the importance of interaction in carrying out those activities. For this reason, achieving the knowledge environment requires an *evolutionary* reconceptualization of digital libraries.

We suggest that there are two key transitions involved in the evolutionary reconceptualization of digital libraries: from information to knowledge, and from document repositories to knowledge ecosystems.

- *From information to knowledge*. The current conceptualization of digital libraries emphasizes the collection, organization, preservation, management, and retrieval of information resources, often on a large scale. Moreover, the main user activity in this view is finding resources relevant to a particular information need. In contrast, the emerging conceptualization of digital libraries emphasizes knowledge and, in particular, how interaction with information resources can actively support knowledge discovery, creation, and use. In this view, a primary goal of digital libraries is helping users convert information into knowledge.
- *From document repositories to knowledge ecosystems*. The current conceptualization of digital libraries also emphasizes the importance of document repositories as curated, searchable, and network accessible collections of digital resources. Frequently, these repositories are discrete and largely independent entities. In comparison, the reconceptualization of digital libraries envisions them as knowledge environments, or more broadly, as knowledge ecosystems. There are two important aspects to this view. One is a shift from individual, isolated collections to more interoperable, interconnected repositories. Another is the cooperative, interdependent, and evolving relationship between users and the digital library environment. In this sense, the reconceptualization of digital libraries requires an ecological perspective.

These two transitions are interrelated, and the reconceptualization of digital libraries needs to address them both. With regard to the first transition, the purpose of libraries, digital or otherwise, has long been framed in terms of education, learning, and knowledge, but manifested in terms of information, documents, and access (Bieber et al., 2002; Levy, 2000). With regard to the second transition, the potential of digital libraries lies less in their individuality and autonomy, and more in their interoperability and ability to aggregate information from multiple sources (Besser, 2002; Paepcke et al., 1998). Overall, both transitions require digital library technologies that actively and more directly support a wider spectrum of cognitive and knowledge work activities. Currently, organizing information and facilitating physical access (e.g., searching) takes precedence over supporting cognitive access (e.g., understanding), even though a cognitive perspective of access has long been recognized as important to achieving the goals of libraries and other information systems (Buckland, 1991; Ingwersen, 1992; Marchionini, 1995). Unfortunately, traditional libraries are constrained in how they can help users create knowledge from library resources. For example, it is not feasible for libraries to provide individual tutors to help people find, use, and understand library materials. Accordingly, designing to facilitate physical access is understandably pragmatic. Digitization, information retrieval, and networking technologies have certainly eased physical access, but they have been of limited value in helping people use retrieved information to engage in knowledge discovery, creation, and use. This is precisely the challenge for the next phase of digital library research (Larsen, 2004). Meeting this challenge requires a conceptual framework that understands and concentrates on the two transitions outlined above. We present such a framework in section 5 of this paper.

The long-term success of digital libraries depends on how these two transitions are realized. This paper suggests that information visualization can play a central role in these transitions. In fact, information visualization researchers articulated a similar vision for digital libraries in the mid-1990s.

3 Information visualization

The field of information visualization aims to improve human abilities for processing and understanding information by tapping into the powerful human visual processing system (Card et al., 1999; Spence, 2001; Ware, 2004). Information visualization is concerned with three interrelated topics: visual representations, interaction, and cognition. Visual representations encode information in graphical form, and make effective use of human visual perception. They can dramatically improve a person's ability to take in large volumes of information, find complex patterns, understand intricate structures, and create new knowledge (Tufte, 1990). But information visualization is concerned with more than representations and visual perception. It is also concerned with how people interact with representations. Interaction, in the context of information visualization, is the action a user performs on a representation and the subsequent response. Adding interaction to a visual representation can significantly enhance a user's ability to explore, query, navigate, transform, manipulate, and work with different elements and features of the representation (Dix & Ellis, 1998; Sedig & Liang, in press; Sedig & Sumner, in press). The role of interaction, therefore, is to support and enhance various cognitive activities such as reasoning, problem solving, decision making, and sense making. In short, appropriate interaction can lead to new insight, discovery, and understanding (Card et al., 1999; Spence, 2001). Overall, information visualization has significant application wherever there is complex information that people need to navigate, manage, search, understand, reason with, and generally interact with in order to acquire and create knowledge. Therefore, the motivation and rationale for applying information visualization to digital libraries is well-founded (Börner & Chen, 2002b).

During the early stages of digital library research and development, information visualization was the basis for many experimental digital library interfaces. The most prominent and well-known prototypes were those developed at Xerox PARC (Rao et al., 1995; Hearst, 1996; Hearst et al., 1996). Digital libraries, especially in the PARC research, were conceptualized as interactive visual environments for working with information and creating knowledge—i.e., "information workspaces" (Rao et al., 1995, p. 29). Rich interaction was considered vital, having "consequences that permeate the design of all components of the digital library" (Rao et al., 1995, p. 38). In this preliminary fusion of information visualization with digital libraries, both interaction and visual representations were viewed as critical to constructing digital libraries as knowledge environments. Conventional information retrieval played an important role, but the vision was expanded to encompass the full spectrum of knowledge work.

This expanded vision, however, has not yet materialized. Today, digital libraries that rely on information visualization techniques are the exception, not the rule. For example, techniques such as the fisheye lens are rarely used, even though their potential value for building rich interfaces to library systems has long been recognized (e.g., see Kinnucan, 1992). More importantly, most digital libraries are not environments in which users can conduct information work (Rao et al., 1995; Larsen, 2004; Soergel, 2002). A typical digital library provides search and browse facilities so that users can find relevant documents quickly and efficiently (Paepcke, 1996; Marchionini, 1999; Feng et al., 2005). Yet digital libraries could also help users engage in a wider range of knowledge-based activities such as analyzing, problem solving, sense making, decision making, interpreting, and modeling. This is not to dismiss the value of non-digital approaches to information work, such as printing a document from a digital library and then annotating it with a pencil (Sellen & Harper, 2002; Marshall, 2003). Rather, it is to highlight that there remain untapped possibilities for digital libraries to enhance how people interact with digital information artifacts, and that this interaction can radically improve how well digital libraries support knowledge discovery, creation, and use. Achieving these possibilities is the main motivation for applying information visualization techniques to digital libraries.

Why has information visualization not played a more significant role in the development of digital libraries to date? We suggest two major factors. The first factor is that the fields of information visualization and digital libraries needed to separately address basic and independent research questions. For information visualization, this basic research included developing methods for creating visual representations, exploring diverse interaction techniques, and understanding how interaction with visual representations supports cognition (Card et al., 1999; Spence, 2001). For digital libraries, this basic research has been largely concerned with document digitization, content organization, metadata creation, digital preservation, repository interoperability, and developing an infrastructure to support the continued evolution of digital libraries (Arms, 2000; Fox & Urs, 2002; Lesk, 1997). Moreover, when information visualization was applied to digital libraries, the primary motivation was testing and validating visualization techniques, not addressing specific digital library research questions. Similarly, when digital libraries sought to incorporate visualization techniques, the rationale was more about improving digital libraries than about addressing foundational research questions in information visualization. So although the rationale for applying information visualization techniques to digital libraries was (and still is) theoretically sound, the preoccupation with fundamental and distinct research questions made a successful combination during these early stages unlikely.

The second factor is the emergence of the Internet, and the Web in particular, as the preferred mechanism for accessing and distributing information. Because digital libraries strive to be universally accessible, the Web has become their primary distribution mechanism (and often the only one). However, the Web introduces severe technical limitations, the most notable being bandwidth constraints and reduced interaction capabilities. For this reason, information visualization tools are normally implemented as local clients, Java applets, or other non-Web technologies. Consequently, the limitations of the Web have slowed the deployment and constrained the effectiveness of information visualization techniques in digital libraries.

These two factors are now much less significant. Neither factor needs to further impede the use of information visualization techniques in digital libraries. For one reason, numerous prototypes and working systems demonstrate the advances made by both fields. In addition, there is a sizable body of empirical research literature to guide future development. Furthermore, the technical limitations of the Web are quickly lessening. Web browser technology has improved and network bandwidth has increased, so the Web can more easily support richer client-side interaction. In sum, the factors described above should no longer hinder the application of information visualization to digital libraries.

The reduced influence of these factors renews the potential for information visualization to be an essential ingredient in the reconceptualization of digital libraries. Furthermore, the ideas, principles, and motivations that informed early information visualization research are a fixture of new directions for digital library research (Börner & Chen, 2002a; Dillon, 2002; Larsen, 2004; Soergel, 2002). For example, the first item in Soergel's (2002, para. 6) research framework is that "a digital library should provide access to materials and objects *and* to the tools needed to process and present these materials in ways that serve the user's ultimate purpose" (emphasis in original). Soergel's main concern here is how the user's "ultimate purpose" cannot be achieved solely by supporting information retrieval activities. Yet, as we have shown, supporting a wide range of cognitive and knowledge activities has always been a prominent concern for information visualization. Interaction is another connection point between the benefits of information visualization and the goals of digital libraries. For example, studies of what people do with paper documents—gathering, reading, reviewing, organizing, annotating, clipping—highlight both the importance of interaction in knowledge work, and the limitations of modern digital libraries for supporting this work (Marshall, 2003; Sellen & Harper, 2002). Interaction, so central to information visualization, continues to be a weak spot in digital libraries, despite its oft-stated importance (e.g., Coleman & Oxnam, 2002; Levy & Marshall, 1995; Paepcke, 1996; Rao et al., 1995). If digital libraries are to become full-fledged knowledge environments, they must dramatically expand the range of activities they support and substantially improve the interaction techniques used to enable these activities. These topics are of long-standing interest to information visualization.

To date, information visualization has played a minor, ad-hoc role in the implementation and conceptualization of digital libraries. But as we have shown, information visualization is closely-aligned with the emerging view of digital libraries—supporting the full cycle of knowledge discovery, creation, and use through interaction with representations of digital information. This implies that information visualization should play an increasingly major, systematic role in the implementation and conceptualization of digital libraries. In this paper, we are specifically interested in how the ideas, principles, motivations, and techniques of information visualization can play a catalytic role in the evolutionary reconceptualization of digital libraries.

4 Models and frameworks

As shown in the preceding discussion, defining and conceptualizing digital libraries has been, and continues to be, a difficult endeavor (Bawden & Rowlands, 1999a; Borgman, 1999; Fox & Urs, 2002; Larsen, 2004). This has led researchers to propose a variety of conceptual frameworks as a means of identifying, clarifying, and unifying the ideas and principles that motivate digital library research and development (Bawden & Rowlands, 1999b; Beiber et al., 2002; Besser, 2002; Gonçalves et al., 2004; Levy & Marshall, 1995; Marchionini, 1999; Larsen, 2004; Soergel, 2002). Surveying these frameworks reveals the diversity of the field, as well as the challenge of conceptualizing it in an integrated manner.

In this section, we describe, analyze, and compare seven representative conceptual frameworks for digital libraries. We also examine the perspectives, goals, and key principles of these frameworks in light of the emerging view of digital libraries to underscore the importance of information visualization in the reconceptualization of digital libraries.

Underlying assumptions. Levy and Marshall (1995) presented a conceptual framework for digital libraries with three interrelated components: documents, technology, and work. They used this framework to explore assumptions about digital libraries and reveal important, yet largely unrecognized, implications for research and development. In particular, they questioned three common assumptions about digital libraries: they contain only fixed, permanent documents; they are used by people working alone; they involve only digital technologies. Levy and Marshall showed how these assumptions prevent digital libraries from being able to support the full scope of knowledge work. Furthermore, their analysis highlighted the restrictions placed on digital libraries conceptualized in terms of conventional information retrieval.

The Sharium. Marchionini (1999) argued that digital libraries should emphasize the role of group problem solving and knowledge sharing. This conceptualization of the digital library is called the “sharium.” The central premise of the sharium is that providing people with information, and the tools necessary for working with that information, does not guarantee that they will be able to solve their information problems. Because human beings are social creatures, social interaction plays a key role in understanding information and creating knowledge. In the sharium concept, digital libraries are interactive, virtual environments designed to support collaboration between all users of a digital library, including librarians. This conceptualization of digital libraries underscores the importance of interaction, specifically social interaction, when conducting knowledge-based activities.

Collaborative knowledge evolution. Beiber et al. (2002) proposed an approach to digital libraries called the Collaborative Knowledge Evolution Support System (CKESS). CKESS is described as “a digital repository in which a community of users collaborate to share and evolve their knowledge” (Beiber et al., 2002, p. 12). Although this framework has similarities with the sharium (Marchionini, 1999), it is distinguishable in two respects. First, where the sharium emphasizes the importance of digital libraries being collaborative workspaces for problem solving and knowledge creation, CKESS emphasizes the importance of digital libraries being collaborative workspaces judiciously attuned to the local work practices of a particular community. In this sense, the framework takes an ecological approach to digital libraries by addressing the complex and ever-evolving relationships between information resources, technology, people, and practice (Nardi & O’Day, 1999). Second, CKESS also emphasizes that the user community should guide the evolution and improvement of the digital repository. In modern digital libraries, the curatorial activities of librarians have chief responsibility for shaping the repository. In CKESS, the collaborative activities of the community have chief responsibility. In this model, the community generates tacit knowledge through interaction with the repository and through collaborative activities. This tacit knowledge is then used to collectively modify and improve the explicit knowledge captured in the repository. CKESS is designed to smooth the transition from explicit to tacit knowledge (and vice versa) through four knowledge conversion processes: socialization, externalization, combination, and internalization (Choo, 1988; Nonaka & Konno, 1998; Nonaka & Takeuchi, 1995). In sum, CKESS envisions digital libraries as constantly evolving repositories that enable collaborative knowledge work, facilitate knowledge conversion processes, and integrate with the local work practices of a particular community.

Interoperable repositories. Besser (2002) described a conceptual framework that emphasized the importance and advantages of interoperable document repositories over isolated collections. Interoperability refers to either synchronous searching of multiple repositories or asynchronous aggregation of information from multiple repositories into a single repository by intelligent agents (Fast & Campbell, 2001; Paepcke et al., 1998). From a user’s perspective, interoperability provides a useful technical abstraction layer because it allows the user to perform searches independent of the location, interface, and structure of individual repositories. The benefits of this abstraction have a significant precedent—Web search engines provide effectively the same service by aggregating information from multiple Web sites and making it searchable through a single, universal interface. The difference is that whereas Web search engines must address the complex nature of Web documents, digital libraries can rely on formal interoperability protocols such as the OAI protocol for metadata harvesting or Z39.50 (Fast & Campbell, 2001). But even if such protocols are widely adopted, interoperability alone is insufficient for achieving the full vision for digital libraries.

Broadening the vision. Seorgel (2002) presented a conceptual framework with three “overarching guiding principles” and eleven “specific themes.” The framework explicitly sought to redirect digital library research and development by broadening the conceptualization of digital libraries. Soergel's framework is an assembly of diverse ideas that either require substantial research or have been weakly implemented in existing digital libraries. The framework does not define a set of specific conceptual elements, does not explain how they are related, and does not apply them to particular digital library scenarios. Instead, the framework integrates and synthesizes ideas from across the digital library field to develop "a structured vision for what digital libraries can be" (Soergel, 2002, Conclusion section).

The 5S model. Gonçalves et al. (2004) developed the 5S model, which aspires to place digital libraries on a more substantial and rigorous theoretical foundation. The model has five "fundamental abstractions": streams, structures, spaces, scenarios, and societies. These abstractions form the basis of a formal, mathematical, descriptive language for digital libraries that is roughly akin to formal models for relational databases and information retrieval systems. The 5S model can also function at a conceptual level, and the authors present several case studies to demonstrate the model's analytical value. Even so, the primary motivation of the 5S model is to provide a means for describing digital libraries rigorously, unambiguously, and formally.

The ubiquitous knowledge environment. The NSF report (Larsen, 2004) on future research directions for digital libraries includes a conceptual model based on three elements: user, interaction, and information store (i.e., repository). Each of these elements is comprised of several diverse sub-elements, such as cognitive completion, querying, and interpretive capture. The linchpin of the model is interaction because it mediates "a continuous relationship between the itinerant user and the ubiquitous information store" (Larsen, 2004, p. 2). This notion of a "continuous relationship" underscores the importance of interaction for enabling sophisticated user activities. The central premise is that if digital libraries aim to support knowledge work, they cannot be mere electronic filing cabinets—they must be environments that are interactive, continuously available, and deeply integrated with work practices. The conceptual model in the NSF report is not intended to be a complete or rigorous theoretical foundation for digital libraries. Instead, it is designed to highlight key elements of the ubiquitous knowledge environment concept as a means of framing the next phase of digital library research.

These frameworks illustrate how the conceptualization of digital libraries has evolved. Early frameworks often served to highlight deficiencies, limitations, and assumptions of the initial document repository conceptualization. Recent frameworks have addressed these criticisms by becoming more comprehensive, usually by incorporating additional concepts and emphasizing latent ones.

We draw attention to three key trends in the evolution of these models and frameworks. First, there is a shift from a system oriented, technology based perspective to a human oriented, activity based perspective. Much of the research literature is, and continues to be, predominantly technological in nature (Dillon, 2002). It stands in contrast to the models and frameworks described above in which the technical infrastructure is merely a starting point for serving the intellectual needs of users. Meeting these intellectual needs requires a human oriented, activity based perspective.

Second, there is an expansion of the activities that users could, or should, be able to perform in digital libraries. Although these frameworks acknowledge the importance of information retrieval activities, they consistently point out that knowledge work involves far more than conventional querying, searching, and browsing. For example, studies of how people interact with paper documents have shown that reading, in a knowledge work context, is typically interweaved with writing, annotating, navigating, cross-referencing, note-taking, and other activities (Sellen & Harper, 2002). As the preceding frameworks illustrate, a persistent and growing theme in the literature is that digital libraries should be designed to support a broad range of cognitive and knowledge work activities, yet they continue to be designed around traditional information retrieval activities (Beiber et al., 2002; Feng et al., 2005; Larsen, 2004; Paepcke, 1996; Soergel, 2002;). Consequently, the reconceptualization of digital libraries as knowledge environments requires a framework that characterizes the activities digital libraries could, and should, support.

Third, there is an increasing recognition that interaction is critical to supporting user activities in digital libraries (Börner & Chen, 2002a; Coleman & Oxnam, 2002; Dillon, 2002; Larsen, 2004). Currently, interaction is usually limited to supporting traditional querying, searching, and browsing (Feng et al., 2002). But interaction can also be used to support cognitive processes, helping people work with information as they strive to create knowledge. In other words, interaction can be more than a means of navigating the world and achieving goals—it can also be a means for making sense of the world and discovering goals (Kirsh, 1997). Viewed in this light, interaction is essential to any knowledge-centered conceptualization of digital libraries. Therefore, digital library researchers should address the relationship between interaction and cognition, working towards a framework that characterizes interaction techniques that support cognitive and knowledge work activities.

These three trends are supportive of and consistent with the ideas, goals, and motivations of information visualization. Information visualization employs rich interaction with visual representations to amplify cognition and support a wide range of cognitive and knowledge work activities (Card et al., 1999). Similarly, as digital libraries have evolved, there has been a greater emphasis on creating digital libraries as environments that support the full breadth of what users do, or could do, with digital resources. There are clear parallels between the two fields. Information visualization, however, has typically been applied to the *implementation* of digital libraries—developing visual interfaces to digital collections (e.g., Börner & Chen, 2002a). This paper applies information visualization to the *conceptualization* of digital libraries.

We suggest that information visualization can play a central role in reconceptualizing digital libraries as they transition from information to knowledge and from document repositories to knowledge ecosystems. This requires a conceptual framework that is based on the knowledge environment perspective of digital libraries and that articulates the roles and benefits of information visualization. The following section describes such a framework.

5 The INVENT framework for digital libraries

In his early and influential analysis of how digital technologies might be applied to libraries, Licklider (1965, p. 2) observed that “the items of basic interest are not the print or paper, and not the words and sentences themselves—but the facts, concepts, principles, and ideas that lie behind the visible and tangible aspects of documents.” Four decades later, this description succinctly captures both the essential vision for digital libraries and the role of information visualization in realizing this vision.

In this section we describe the INVENT framework, which conceptualizes digital libraries as interactive knowledge environments. INVENT stands for INteractive Visual ENvironmenTs. It is a conceptual framework that emphasizes interaction with representations of information, especially visual representations, for supporting cognitive and knowledge work activities. INVENT is a general framework for reframing digital libraries as knowledge environments, but it can also serve as a lens for understanding the role of information visualization in such digital libraries. It is through this lens that the INVENT framework will be explained—to clarify how information visualization can play a catalytic role in the reconceptualization of digital libraries. More specifically, the framework illustrates how the central concepts of information visualization (interaction, visual representations, and cognition) and its goals (supporting knowledge discovery, creation, and use) are closely aligned with the long-term goals of digital libraries.

The INVENT framework consists of six conceptual elements: digital objects, representations, activities, interactions, actors, and ecologies. Digital objects capture and store information in a structured, digital form. Representations are the perceptible forms of digital objects, with visual representations being of particular interest. Activities refer to what people do in digital libraries as they work to transform information into knowledge. Interactions are the actions that users perform on representations, coupled with the response to those actions. Actors are any entity that interacts with a digital library. Lastly, ecologies refer to the cooperative, interdependent, and evolving relationships between actors and digital knowledge environments. The remainder of the paper explains each of these elements in greater detail.

5.1 Digital objects

Digital objects capture and store information in a structured, digital form. They are the information resources in a digital library. Since there is no universally accepted term for the items in a digital library (Arms, 2000), we use “digital object” to encompass many synonymous and related terms, including document, text, resource, and artifact. It is important to note that digital objects, in this framework, do not necessarily correspond to individual files, database records, or other system-level entities. They are simply the information resources that constitute a digital library, and should be conceived independently from any implementation details.

We have defined digital objects in terms of two primary characteristics. The first is that digital objects capture and store information. The second is that they are imbued with structure. The nature and role of this structure is particularly important and central to how digital objects are conceptualized in the INVENT framework.

By structure, we mean that a digital object's information content has a deliberate form and arrangement. Structure serves to identify, highlight, and clarify the essential features, properties, and relationships of both individual objects and a collection of objects. In other words, structure makes explicit what is otherwise implicit, tacit, or latent. It is a mechanism for asserting a digital object's key characteristics, while also describing how it relates to other objects.

We use the term structure in a general sense to include a wide range of established methods, techniques, and standards for describing, organizing, and systematizing a collection of information resources (Svenonius, 2000; Fensel et al., 2003). Examples of structure include ontologies, document surrogates, indexes, controlled vocabularies, metadata schemas, and document markup languages. Of course, these examples are diverse—an ontology is unlike a document surrogate. But they are also alike because they give form and arrangement to a digital object's information, which facilitates access to that information.

Clearly, structuring information is a major activity of libraries. The history of libraries can be viewed as an evolution of techniques for creating information structures. Multiple factors have influenced this evolution, the most important being the growth of library collections, the adoption of new technologies, and the information needs of library users. Historically, libraries have developed structures that improve access to the intellectual content of books (subject headings, for example). These structures, to a large degree, are designed to overcome the intrinsic difficulties of organizing books and other physical objects. Today, digital technologies are enabling new ways of creating, manipulating, and transforming the structure of information resources. But crucially, the ability of digital technologies to support user information needs is highly dependent on the structure of digital objects.

Information retrieval technology illustrates the connection between the information needs of users and the structure of digital object. In one sense, information retrieval systems are independent of structure because they allow users to override and subvert the fixed, pre-determined structure of a collection—keyword searching being the most obvious example of this strategy. But in another sense, information retrieval systems are extremely dependent on the structure of the digital objects being indexed. Indeed, the effectiveness of a retrieval system depends not on the algorithm themselves, but on how well the algorithms take advantage of available structure. The evolution of Web search engines illustrates this dependency. The first generation of Web search engines were descendants of classical information retrieval systems, such as Dialog and Lexis-Nexis. These classical systems were designed to index carefully controlled document sets that often featured controlled vocabularies, subject headings, special indexes, and other forms of structure. The Web, however, is a radically different search environment. It lacks traditional forms of structure and has highly unreliable metadata. These and other characteristics of the Web made it necessary to re-examine long-held assumptions about information retrieval in very large document systems (Lynch, 2001). In particular, Web search engines had to abandon their reliance on traditional forms of document structure and develop new techniques that used the Web's link structure (Kleinberg, 1999; Page & Brin, 1998).

The Web search engine example highlights two important lessons for the role of structure in digital libraries. The first lesson is that traditional methods of creating and manipulating structure are not necessarily well suited to the complexities of the emerging information landscape. For instance, the rationale underlying the Semantic Web initiative is that digital objects with rich, granular, and semantically sophisticated structure are essential for achieving the possibilities of a networked information infrastructure (Berners-Lee, 1999; Berners-Lee et al., 2001; Fensel et al., 2003; Lu et al., 2002). This includes the possibilities for digital libraries. In other words, structure is important, but new kinds of structure will be required as digital libraries transition from searchable repositories to interactive knowledge environments.

The second, and less obvious lesson, is that there is a vital link between the structure of a digital collection and the activities a knowledge environment can support. As discussed earlier in this paper, digital libraries were initially conceptualized in terms of conventional information retrieval, whereas they are now being conceptualized as environments that support cognitive and knowledge work activities. In this latter, activity-centric view, digital libraries are tools for selecting, identifying, locating, sorting, merging, filtering, organizing, understanding, and extracting relevant concepts, patterns, and relationships from digital objects. How well digital libraries can support these activities depends on how digital objects are structured.

To more clearly see the connection between structure and activities, let us consider the example of a horticultural digital library and a novice gardener named Walter who is learning about cacti. The digital library is, for simplicity, a homogenous collection where each digital object provides information about a single species of plant—the common name, the botanical name, photos and illustrations, light requirements, recommended humidity levels, and so forth. Walter wants to find cacti that make suitable houseplants. A retrieval-oriented structure, such as a controlled vocabulary, could adequately support this activity. Using the controlled vocabulary structure, Walter might start by finding the entry for houseplants, and then narrowing the results to cacti. But what if Walter needs to do more than learn which cacti make good houseplants? Walter could easily need to identify specific characteristics (for instance, recommended light and humidity levels), compare those characteristics, progressively filter out cacti that do not meet certain criteria, consider other types of houseplants that do meet his criteria, merge relevant information into a single document, and so forth. How can digital libraries effectively support these kinds of activities?

The cacti scenario suggests that if digital libraries want to support advanced user activities, then they need to maximize the structure of digital objects. Therefore, each object in the horticultural collection should have structures that allow the identification, selection, and extraction of specific information—the common name, for example, or optimal humidity levels. The design strategy is to encode as much information in the system as possible. Moreover, there is a link between structure and activities. If Walter wants to compare the light requirements for different cacti, then the objects need to explicitly identify that information and provide ways of manipulating that structure. Otherwise, Walter will need to use external aids to do the comparison. He might rely on his memory. More likely, he will record the information using some external tool, maybe a spreadsheet or perhaps on paper, thereby pushing the intellectual work outside the digital library.

The maximizing structure approach has several deficiencies. It assumes it is possible, and feasible, to design digital objects with enough structure to support all possible uses of the objects. In this approach, good design hinges on predictive ability. The designer must predict all possible forms, uses, and combinations of information in the system and then structure the objects accordingly. Moreover, what the system knows about a digital object becomes more important than what the user knows. These deficiencies are not an argument against structure. Rather, they point out that although there is a strong link between structure and what users can do with digital objects, structure alone is insufficient for creating interactive knowledge environments.

In the INVENT framework, structure serves as a starting point, not an end point. A digital object's structure allows it to be decomposed, analyzed, combined, and transformed to dynamically generate new structures. Critically, users control much of this process. Their actions create a progressive chain of ad-hoc structures that emphasize, reveal, hide, manipulate, and transform the various features, properties, and relationships of digital objects. The value of these ad-hoc structures is representing information in ways that are better suited to the individual, temporal, and contextual needs of users (Peterson, 1996). In this framework, these dynamically generated and context-sensitive structures are called *representations*. We describe representations further in the following section.

5.2 Representations

Representations, in this framework, are the perceptible forms of digital objects, and they play a crucial role in the activities of users. Where digital objects capture and store information in the repository, representations form the external, visible, and tangible manifestations of the features, properties, and relationships of that information. Furthermore, by dynamically decomposing, combining, and transforming digital objects into representations, digital libraries can support activities that would be difficult or impossible if users could only work with the original and unmodified objects. In this section, we provide several examples that demonstrate the value of separating digital objects from their possible representations, with particular emphasis on the value of visual representations.

In the INVENT framework, users do not interact with digital objects directly—they interact with representations of digital objects. To illustrate the value of representations, and especially the usefulness of distinguishing between digital objects and representations, consider a digital library of Italian Renaissance paintings. If the activity involves studying the finer details of an individual painting, then a high resolution digitization of the painting is desired. But if the activity involves understanding the distribution of colors in the painting, then a histogram will reveal patterns that are difficult to discern from the original painting, and impossible to discern accurately. In the first situation, the object is represented with a high degree of fidelity and a low degree of abstraction. In the second situation, the object is represented in a highly abstract manner, but a manner much better suited to the task. The essential concept illustrated by this example is not the complexity of a given transformation or the degree of abstraction, but the value and importance of transforming digital objects into alternate representations. Again, the structure of a digital object serves as a starting point for supporting user activities, not an end point—hence the need for a clear conceptual separation between digital objects and their possible representations.

Representations can take many forms: textual, algebraic, auditory, visual, and so on (Card et al., 1999; Norman, 1993; Peterson, 1996; Spence, 2001; Tufte, 1990). Digital libraries and the information systems from which they descend have always used representations to support user activities. Textual representations are the most common: indexes, abstracts, thesauri, document surrogates, search results, and so on. These representations are, and will remain, important to digital libraries. However, alternate forms bring unique benefits and, in the appropriate situation, they can be vastly superior (Card et al., 1999; Norman, 1993). Furthermore, if digital libraries seek to enhance human reasoning, problem solving, decision-making, sense making, and other knowledge-centered activities, then representations must play a core role in their conceptualization. This is because representations are essential for performing complex knowledge activities: "we do our thinking on the representations, sometimes on representations of representations. This is how we discover higher-order relationships, structures, and consistencies in the world or, if you will, in representations of the world. The ability to find these representations is at the heart of reasoning, and critical to serious literature, art, mathematics, and science" (Norman, 1993, pp. 51-52). If digital libraries are merely well organized document collections coupled with a search engine, then representations play a more limited role in their conceptualization. But if digital libraries are able to inject higher levels of abstraction into the process of interacting with information resources by dynamically generating representations in response to user actions, then they will be better able to support user's cognitive and knowledge work activities.

Visual representations are a particularly powerful means for supporting human cognition, and provide many benefits to mental processes (Card et al., 1999; Spence, 2001; Tufte, 1990; Ware, 2004). These benefits include increased memory and mental processing resources, improved understanding of complex patterns, greater information density, and simplified problem solving through perceptual inference. In general, visualization enables a shift from cognitive processes to perceptual ones. For instance, visual representations in digital libraries can shift the mental workload “from slow reading to faster perceptual processes” (Börner & Chen, 2002b, p. 2). The advantage of this shift is not reducing the need for cognition, but developing a better balance between cognition and visual perception. Still, this does not mean that visualizations obviate the need for text or other representational forms. Rather, the challenge is finding the most appropriate representation for a particular task, be it textual or visual or some other type of representation. A good representation reduces cognitive effort, often dramatically, though a poor representation can have the opposite effect (Norman, 1993).

We provide two examples to show how visual representations can effectively support sophisticated user activities in digital libraries. For the first example, let us consider a digital library of important European novels from the sixteenth and seventeenth centuries. If the activity is reading one of these novels, then the full text is obviously the most appropriate representation. But if the activity is anything other than reading the full text, an alternate representation may have pronounced advantages. For instance, many texts have multiple editions, and studying the differences between these editions is an important activity for literary scholars. Comparing multiple editions of a work to understand how it evolved would be an enormous undertaking for a scholar forced to use only the textual representations. The Interactive Timeline Viewer (ItLv) is a digital library visualization tool that simplifies comparing a base text with multiple variant texts (Monroy et al., 2002). The ItLv was used to explore six early editions of *Don Quixote*. By visually representing differences among various editions, scholars could interactively explore the text and discover important patterns in the novel’s evolution. The ItLv is an example of how interactive visualizations can dramatically simplify an activity that scholarly digital libraries could support. It illustrates how visual representations allow users to perform complex tasks efficiently, and how interaction with representations can lead to deeper understanding and insight.

For the second example, let us consider the difficulty of understanding the relationship between documents in an information space. Visual representations can help users develop mental maps of the topics in the database, reveal relationships among authors, and discover other important patterns. Morris et al. (2003) visualized emerging research fronts in scientific literatures to help users understand the ebb and flow of new developments within a particular domain, identify new research topics, and gain insight into how research findings move between different fields. Lin et al. (2003) visualized author co-citation patterns in academic journals to reveal groupings and intellectual connections between different authors. Smith and Fiore (2001) visualized electronic discussion forums to highlight conversational structure. In each of these three examples, visual representations are used to reveal important patterns, trends, and relationships that would be cognitively difficult or unreasonably time-consuming to discern by examining the original, unmodified digital objects.

A key challenge for digital libraries is developing techniques to generate meaningful and contextually appropriate representations. In the realm of information visualization, numerous techniques have been developed for generating visual representations applicable to digital libraries (Börner et al., 2002a; Börner et al., 2003; Card et al., 1999;

Chen, 1999; Hearst, 1996; Hearst et al., 1996; Rao et al., 1995). These representations can be used to enhance various user activities, such as navigating the information space, monitoring the space for new developments, understanding the significance of certain authors, tracking changes in the citation patterns of a document collection, and revealing emerging research fronts in a scientific domain. In other words, representations do not stand alone—they are closely linked to the activities that digital libraries aim to support.

5.3 Activities

Activities, in the INVENT framework, refer to the knowledge-centric tasks and processes that digital libraries could, or should, support. Some examples of these activities include searching, reading, note taking, organizing, locating, identifying, ranking, correlating, and analyzing. In other words, activities are what people do in digital libraries as they try to create knowledge from the information contained in digital objects.

Throughout this paper we have mentioned two general types of activities. The first type is information retrieval activities, namely searching and browsing, which are predominant in modern digital libraries (Feng et al., 2005). The second type is "cognitive and knowledge work activities," an expression we have used to emphasize the limitations of traditional information retrieval activities and to describe activities required by the knowledge environment concept. Of course, digital libraries should support both types. Embracing the implications of the knowledge environment model does not eliminate the need to support retrieval activities, but it does require digital libraries to dramatically improve their support for cognitive and knowledge work activities.

Although we have made a distinction between these two types, they are not mutually exclusive. Searching and browsing are an indispensable part of knowledge work. But so are reading, analyzing, note taking, and organizing (Sellen & Harper, 2002). Furthermore, these and many other activities are bound up with the goal of transforming information into knowledge. For this reason, dividing activities into two camps, one old and one new, may cause unnecessary confusion as we work towards a more complete understanding of the relationship between digital libraries and their users.

To rectify this, we introduce a new term—epistemic activities—to encompass all activities that support, guide, and enhance the knowledge-oriented work of digital library users. In this sense, epistemic activities include retrieval activities, but also any other activity related to achieving epistemic goals. The key characteristic of epistemic activities is how they facilitate the process of transforming information into knowledge. This process was identified in Section 2 as one of two key transitions in reshaping digital libraries as knowledge environments. Accordingly, designing for epistemic activities should be considered a major challenge for digital libraries.

Reframing digital libraries in terms of epistemic activities has important implications for their design, and for the design process. During the design process, digital library designers (i.e., librarians) are faced with a difficult problem: how to structure digital objects to meet the contextually and temporally sensitive information needs of individual users. The standard solution is to design an information architecture that supports most user activities, most of the time (Rosenfeld & Morville, 2002). As a result, digital libraries are designed for a generic user, not a specific user. For example, card-sorting is a common design technique for describing and categorizing the objects in a digital collection (Faiks & Hyland, 2000; Rosenfeld & Morville, 2002). In a card-sorting exercise, several users are each given a set of index cards. The cards are a representative sample of objects from the collection. Each user independently organizes the cards in whatever manner makes the most sense to him or her. The results are combined and analyzed to design an organization scheme, or structure, for the entire collection. The strategy here is to create a structure that reflects the commonalities among users, while ignoring individual differences. In this design approach, whether based on card-sorting or other techniques, how a particular user understands the world is washed away in favor of how the average user understands it.

If this is how digital libraries are designed, how can they support epistemic activities? On one hand, the design process deliberately and necessarily strips out idiosyncratic, contextual, and temporally sensitive information from the collection's structure. On the other hand, supporting epistemic activities requires that digital libraries be designed around user's individual, contextual, and temporal information needs.

To answer this question, let us reconsider, by way of example, the limitations of digital libraries designed around conventional searching and browsing. Alice wants to learn about the cinematic techniques used in the movie Citizen Kane, so she goes to a digital library about movies. All the resources in this library have been organized into a browsable hierarchy with twelve top-level categories. Alice identifies at least three categories that might contain the information she needs: Film Titles, Directing, and Cinematography. How should Alice proceed? If this is a typical digital library, Alice must choose between two basic activities: searching and browsing. If she wants to move closer to her goal by browsing the categories, she needs to align her knowledge of cinema to how information about cinema is stored, structured, and represented by the collection. Alice believes cinematic techniques probably belong in the cinematography hierarchy. If cinematic techniques are categorized this way, then Alice moves closer to her goal. But if cinematic techniques are categorized in the directing hierarchy, then she must retrace her steps and try an alternate path. Alice has one model of the world, the digital library has another model. To achieve her goals, she needs to map her model to the library's model. In other words, how the library understands cinema is more important than how Alice understands it, even if she is an expert in the subject. Her knowledge, which is individual and contextual, is overridden by the knowledge encoded in the system, which is fixed and based on a hypothetical typical user. Instead of the environment adapting to her knowledge, Alice must adapt to the environment.

In many cases, Alice will have trouble aligning her knowledge of cinema with the structure of the digital library. If this happens, her main option is abandoning the formal structure of the collection in favor of keyword searching. This is a common strategy, particularly now that searching the Web has become such a common activity (Neilsen, 2001). For example, in a study comparing how university students perceived the differences between Web search engines and library catalogues, Fast and Campbell (2004) found a strong preference for Web searching (see also, Campbell & Fast, 2005). Although participants admired the catalogue and considered it, in many respects, superior to the Web, they felt that searching the Web was easier, faster, and more likely to succeed. One of the most important factors behind this preference was that students often found it extremely difficult to make sense of the catalogue's structure. The information was organized, but not in a way that they could easily understand. As one participant said, the catalogue "was like a little maze." Students overwhelmingly preferred the messiness of the Web, largely because the search engine helped them move closer to their epistemic goals without having to first align their knowledge with the structure of the environment.

Nevertheless, no matter how well a digital library supports searching and browsing activities, it will still be limited in its ability to help users accomplish all of their epistemic goals. Many of these goals require more than finding relevant information. They frequently involve activities that convert found information into useful, actionable knowledge. Let us return to our example of Alice, who wanted to learn about the cinematic techniques used in Citizen Kane. At a high level, her epistemic goal is clearly defined, but at more specific levels it is ill defined. Which cinematic techniques is she interested in: lighting, framing, music, or camera angles? How many techniques? Will she require written descriptions, still images, or sample videos? Does she want to analyze the techniques within Citizen Kane specifically, or in relation to other films by Orson Welles? How might the digital library help Alice make sense of the information she finds, make decisions about different courses of action, create mental models of the information space, interpret the nuances of specific cinematic techniques, relate her existing knowledge of cinematic framing devices to the techniques used in Citizen Kane, and so forth?

These questions capture the essential characteristics of epistemic activities, and the need for digital libraries to support them. If digital libraries are restricted to conventional models of searching and browsing, they will have a limited capacity for supporting user's epistemic needs. Alice wanted to do more than find documents describing cinematic techniques. She also wanted to select, identify, locate, sort, merge, filter, organize, understand, and extract important concepts, patterns, and relationships from the collection. Allowing users to perform such activities is central to the knowledge environment view of digital libraries.

Designing digital libraries to guide users through the complexities of knowledge work by helping them perform epistemic activities is, clearly, an enormous challenge. But this challenge must be met if digital libraries are to be reshaped as knowledge environments. Furthermore, information visualization provides an invaluable set of technologies, principles, and techniques for supporting epistemic activities. Visualization tools support such activities by generating contextually meaningful visual representations coupled with a rich array of interactive features (Card et al., 1999; Spence, 2001). So far, our discussion has emphasized the value of visual representations to digital libraries. But information visualization is equally concerned with the cognitive benefits of interaction with these representations. In visualization systems, interaction is more than a means of navigating complex knowledge environments. It is also, and more importantly, a means of modifying, organizing, and adapting the environment to individual, contextual, and temporal epistemic needs. We turn our attention to interaction in the following section.

5.4 Interactions

Interactions are the actions performed on a representation, and the subsequent responses or reactions of the representation to those actions (Sedig & Liang, in press; Sedig & Sumner, in press). In this sense, interactions can be understood as a mediating layer between representations and activities. Users interact with representations as they go about their activities in the digital library. Note that we have not described interactions in relation to digital objects. In this framework, interactions are performed on representations that have been generated from digital objects, not the digital objects themselves.

It is important to distinguish between activities and interactions. Activities refer more to what people do with the resources in a digital library, whereas interactions refer more to how people carry out those activities. To clarify the difference, consider the physical world example of preparing a meal. The activity is cooking. But accomplishing that activity involves many different actions: opening the fridge, finding the ingredients, grasping the tomato, arranging the counter space, adjusting the oven temperature, and so forth. Activities are broader and more general; they encompass interactions. Interactions are narrower and refer to how people physically act on the environment and how the environment responds. In addition, activities may be accomplished by different interactions depending on various factors, one such factor being the local environment. For instance, browsing in a physical library involves walking down the aisles and scanning the shelves, whereas browsing in a digital library typically involves moving the mouse, clicking links, and scrolling the screen. In these cases, how the browsing activity is enacted depends in large part on the environment in which the activity occurs.

The actions in the cooking example may seem prosaic, but studies of how people accomplish everyday tasks have shown that modifying the environment is a common and effective strategy for simplifying an activity (Hollan et al., 2000; Kirsh, 1996; Maglio et al., 1999; Sellen & Harper, 2002). Kirsh (1995) describes the case of chef who needs to create a uniform and aesthetically pleasing arrangement of vegetables on a platter. An experienced chef will begin by placing the vegetables in rows, not piles. Arranging the pieces in rows is a better strategy because it is easier for people to estimate length than area or volume. The chef can use the rows to visually estimate how many pieces need to be arranged. More importantly, they allow her to better gauge the relative number of remaining pieces. By arranging the vegetables in rows, she is encoding information in the environment instead of keeping all the information in her memory. The important point is how the chef interacts with the environment to create a visuospatial arrangement that simplifies the activity. She no longer needs to solve the problem by relying on her mental abilities alone. Instead, she reduces the complexity of the problem by modifying the environment to make it more ‘cognitively congenial’ (Kirsh, 1996).

The vegetable platter scenario is not the only example of how interaction with the environment can simplify an activity. Scrabble players shuffle letters to find new word combinations (Maglio et al., 1999) and gin rummy players rearrange their cards to help them decide which card to play next (Kirsh, 1996). To cite an example more directly related to digital libraries, studies of how people use paper documents have found that even an activity as common as reading involves continuous arrangement of the work environment (Sellen & Harper, 2002). When people read, they constantly rearrange the elements in their visual field to better compare multiple documents, cross-reference important information, and navigate large articles. Similarly, observations of what people do in physical libraries have shown that “order-making” is especially common— i.e., gathering, rearranging, sorting, clipping, annotating, and manipulating documents (Marshall, 1998; Marshall, 2003). These studies highlight the importance of interaction in supporting cognitive processes, particularly how people continuously arrange, structure, organize, and modify their environment as means of improving their ability to perform cognitively complex activities[1].

Information visualization, as noted in Section 3, is concerned with the connection between visual perception, cognition, and interaction. The examples cited above demonstrate the importance and complexities of this connection. The chef, for instance, by arranging the vegetables in a particular way, used a combination of interaction and visual perception to reduce the cognitive complexity of the problem. Information visualization aims to extend the utility of interactive strategies through the benefits of abstract visual representations. In our discussion of representations, we showed how visual representations can help users accomplish activities that might otherwise be unreasonably time consuming, cognitively difficult, or have non-obvious solutions. Although static visual representations have many advantages (Tufte, 1990) they can be further enhanced through even simple interaction (Dix & Ellis, 1998; Sedig & Sumner, in press). Moreover, interaction can serve as an epistemic extension of static visualizations (Sedig & Liang, in press; Sedig & Sumner, in press). Without interaction, understanding a representation is largely dependant on perceptual processes. But with interaction, users can iteratively explore, organize, and discover important features, patterns, and relationships of the representation.

---

[1] Analyses of how people interact with their environments have found that human activities involve several distinct types of actions, including preparatory actions, maintenance actions, and *epistemic actions* (Kirsh, 1997). Epistemic actions should not be confused with epistemic activities, though they are related. In this paper, we have defined epistemic activities to mean activities related to how users convert information into knowledge. Epistemic actions, on the other hand, are an established concept in cognitive science. They are defined as “physical actions that make mental computation easier, faster, or more reliable” (Kirsh & Maglio, 1994, p. 513), and are an important part of many interactive experiences (e.g., Neth & Payne, 2002; Schwan, 2004). Epistemic actions are relevant to digital libraries and information visualization. But the literature on epistemic actions is primarily interested in contributing to theories of cognition, so applying the concept to the development of digital libraries is difficult. In this paper, we are specifically interested in developing a conceptual framework to guide the development of digital libraries. We introduced the term epistemic activities as a conceptual tool for thinking about the activities that digital libraries could support. In doing this, we have consciously borrowed from the concept of epistemic actions, but we view them as distinct concepts.

The following scenario illustrates some advantages of interaction with visual representations in digital libraries. Bob is a researcher entering a new and unfamiliar field of study. One of his most important activities will be developing a cognitive map of the major contributors to the field, important research questions, relevant journals, emerging research directions, and so on. In short, Bob needs to understand the structure of the information space. Visual representations of the features and relationships between documents in the digital library will help him make sense of that structure in ways that textual interfaces cannot, and numerous techniques exist for generating such representations (Börner et al., 2003; Chen, 1999; Lin et al., 2003; Morris et al., 2003; Smith & Fiore, 2001). However, the visual representation is only an initial step. The ability to interact with the representation is at least as important. Let us be more specific in how interaction can help Bob make sense of this unfamiliar research literature. Suppose that he knows the name of a prominent scientist in his new field of study. By using an author co-citation mapping tool (e.g., Lin et al., 2003) he can obtain a visual map showing all the scientists engaged in similar work and the relationships between their research. What next? Bob might probe the map to obtain further details about a particular author's publications and research interests. Another possibility is animating the map to see how the authoring relationships have evolved over time, or to trace the emergence of new research topics. He might also group two authors and then generate co-citation maps for each to visually compare similarities and differences in their research patterns. Probing, animating, and grouping are interactions that Bob can use to explore and make sense of the relationships in the digital library.

This scenario points to a fundamental difference between the role of interaction in a searchable repository model of digital libraries and a knowledge environment model. In a digital library based on the repository model, the overriding objective is helping users find what they are looking for as quickly, efficiently, and painlessly as possible. The design philosophy here is to creating a system that will allow the user to "get in and get out," which usually implies reducing the amount of interaction required to locate information. This shows the influence of classical information retrieval. In an ideal retrieval system, a query returns all of the documents relevant to the user's information need, and only the relevant documents. In theory, a single query should suffice. But this approach overestimates the ability of retrieval algorithms to extrapolate from user queries to their goals, while underestimating the importance of interaction in helping users to formulate their goals. In the words of Kirsh (1997, p. 86), "often we explore the world in order to discover our goals. We use the possibilities and resources of our environment to help shape our thoughts and goals [and] to see what is possible." In the scenario we just described, Bob wanted to understand the complex web of concepts, people, documents, and research findings in a particular field of scientific study. In a general sense his goal was clearly defined. But in a more specific sense it was not. Even if Bob is intimately familiar with the services of the digital library, the specific sequence of actions that will help him accomplish his goal is not clear. Certainly, no single query will suffice, nor will a static representation. The problem is too complex. The approach suggested by the INVENT framework is to dramatically expand the ways that Bob can interact with representations generated from the repository, so that he can use the digital library environment as a resource to guide and support his work.

We have argued that interaction is a central, yet under-appreciated concept in digital libraries. We are not alone in this view. For example, in their introduction to a special issue of the *Journal of Digital Information* concerning interactivity in digital libraries, Coleman & Oxnam (2002) underscore the importance of interaction by using the phrase "interactional digital libraries." Although a useful phrase for the purposes of that special issue, we hope this phrase does not become widely used because it suggests that some digital libraries may be interactive while others may not. It is not a question of whether digital libraries are, or should be, interactive environments. Instead, it is a question of how interaction can be used to achieve the goals of digital libraries and support user's epistemic needs.

An important and long-term research direction for digital libraries is categorizing and characterizing different interaction techniques to develop a prescriptive taxonomy of interactions with visual representations. It is one thing to say that digital libraries must be interactive or that zooming is a powerful interaction technique. It is quite another to have a taxonomy that offers guidance to interaction designers. Such a taxonomy would help designers know which interaction techniques are most appropriate (and why they are most appropriate) when designing a digital library for a specific community of users, who are engaged in certain activities and working with a particular collection of digital objects.

The four elements of the INVENT framework that we have described thus far—digital objects, representations, activities, and interactions—are directly related to information visualization. We now turn to the remaining elements: actors and ecologies. These elements extend the framework to incorporate the people who use digital libraries, their work practices, and the evolving and interdependent relationship between people and information resources. In so doing, the framework begins to place greater emphasis on concepts from other fields, notably knowledge management.

5.5 Actors

Actors perform actions in, on, and for the digital library. They are the entities who work with the resources of a digital collection. Though important, actors are unlike the four conceptual elements discussed thus far. The first two elements (digital objects and representations) described the role of information-bearing artifacts in digital libraries. The next two elements (activities and interactions) described the ways that people work with these artifacts. These four elements are connected by actors—the users that perform activities in digital libraries by interacting with representations of digital objects.

In this paper we have adopted a user-centered design philosophy. We have not, however, distinguished between different kinds of users. So far, we have employed the term 'users' in a general sense to mean end-users, or those whom the digital library is designed to serve. But end-users are not the only ones who interact with digital libraries. In this framework we define actors as *any* entity that interacts with a digital library. In addition, we distinguish between three general types of actors: users, librarians, and electronic agents.

Users are people who do work *in* a digital library. They are the people who approach a digital library with information needs, and interact with the digital library to fulfill those needs. Users work *in* the digital library in the sense that the digital library is an environment, and they work with the tools and resources that exist in this environment. Users might also be referred to as end-users, patrons, learners, clients, or readers—anyone for whom the digital library is designed, constructed, maintained, and enhanced.

Librarians are people who do work *on* a digital library. We use the term “librarian” in a broad sense to include anyone involved in designing, constructing, maintaining, and enhancing a digital library. Thus the term encompasses a multitude of labels: designers, programmers, managers, administrators, cataloguers, information architects, reference librarians, and so on. Libraries, digital or otherwise, are intimately connected to the activities of librarians. These activities include selecting, collecting, organizing, preserving, conserving, and facilitating access to the resources of the digital library (Borgman, 1997).

Electronic agents do work *for* the digital library. They are software tools that perform actions on digital objects. Electronic agents are typically conceptualized as relatively autonomous agents, such as web crawlers and metadata harvesters (Besser, 2002). In this framework, electronic agents are conceptualized more broadly to include everything from generating a full-text index for a collection of digital objects, to collecting metadata records from remote repositories and resizing digital images. In general, agents carry out automated, repetitive, and computationally intensive tasks that could not, or need not, be performed by users or librarians.

The essential distinction between users, librarians, and electronic agents is not the kinds of activities and interactions they perform, but the context and goals of those activities and interactions. For example, whereas all three types of actors may query a digital collection, the context in which a query is made and the goals that lie behind that query may be significantly different for each type of actor. Let us suppose that we have a digital library on Canadian literature and each of the three types of actors performs a simple keyword query for the author Margaret Atwood. The action (querying) and the details of the action (the phrase “Margaret Atwood”) will be the same for a user, a librarian, and an electronic agent, but the context and the goals of the action may be significantly different. A user might perform the query because he or she wants a poem by a Canadian author, and Margaret Atwood is the first name that comes to mind. A librarian might do the query because she is maintaining the library collection and needs to know what might enhance the Margaret Atwood resources. An electronic agent might do the query as part of automated process for generating reports on how frequently the Margaret Atwood resources are accessed.

At one level, the user, the librarian, and the electronic agent have performed the same action: querying the collection for the phrase Margaret Atwood. But at another level, they have performed those actions in a fundamentally different context. The user does the query while working *in* a digital library, the librarian does the query while working *on* a digital library, and the electronic agent does the query while working *for* a digital library. The actions are the same, but the context and goals that motivate those actions are different. Thus, how a digital library responds to a particular action may need to be different for different actors. Even if the digital objects relevant to a specific action are the same for each actor, the representation(s) generated in response to that action may vary depending on the type of actor. Furthermore, a specific action can occur in the context of various activities. This example underscores the importance of distinguishing between digital objects and their possible representations, and between activities and interactions.

A central premise of the INVENT framework is that if digital libraries are conceptualized as knowledge environments, then they must be able to support the individual and contextual needs of those who interact with the digital library. Just as digital libraries contain many types of digital objects, they should also be designed for many types of actors. We have differentiated actors in only the broadest terms because making fine-grained distinctions between different types of actors is beyond the scope of this paper, but we stress the need for digital libraries to make such distinctions—especially between different types of users. In addition, it is important to understand the relationship between actors and the digital library environment. This leads to the concept of ecologies, the next and final element in the framework.

5.6 Ecologies

Ecologies refer to the cooperative, interdependent, and evolving relationships between actors and their environment. The fundamental difference between environments and ecologies is that an environment is wholly external to the individual (i.e., actor), whereas ecologies encompass the connection between individuals and their environment. Adopting an ecological view of digital libraries requires more than enumerating the digital objects, representations, activities, interactions, and actors that comprise a digital knowledge environment. It also requires an understanding of the relationships between these elements, their interactions, and their co-evolution.

So far, our explanation of the INVENT framework has concentrated on how actors, especially users, engage in epistemic activities by interacting with representations of digital objects. Furthermore, we have placed particular emphasis on how the cornerstones of information visualization—visual representations, interaction, and cognition—are crucial to the next-generation of digital libraries. We have cited various examples of how information visualization techniques are deeply applicable to the next phase of digital library research. In addition, we have shown how interaction with visual representations supports the first key transition in the reconceptualization of digital libraries: from information to knowledge. But the second transition—from document repositories to knowledge ecosystems—has not received as much attention. The ecologies element, in large part, addresses this second transition.

In an ecological view, digital libraries are seen not as isolated technological artifacts, but as artifacts embedded within the local practices of a particular community. Consider the differences between a digital library for chemical engineers and an educational digital library for children. Although these are both digital libraries, they serve different user communities with different needs and abilities. In addition, they will be developed and maintained by different organizations, include different kinds of digital objects, represent those objects in different ways, and support different epistemic activities. Yet despite these differences they are still digital libraries. These differences stem from how digital libraries work best when they are an integral element of a particular knowledge ecosystem. In this respect, a digital library is quite different from other artifacts commonly associated with knowledge work. A pen, for example, is an effective and important tool for working with information. But where a pen is a generic object that can be mass-produced, a digital library must be deliberately tuned to the local ecology. Of course, even the pen can be considered part of an ecology, but it cannot adapt, evolve, and change. The pen will remain the same regardless of the ecology to which it belongs. In contrast, a digital library can be designed around the practices of a community. Moreover, it can adapt, evolve, and change alongside the community.

Another aspect of ecologies relates to group collaboration through digital libraries. Of the seven frameworks for digital libraries described in section 4, two of them emphasized the importance of collaboration and social interaction: the Sharium (Marchionini, 1999) and CKESS (Beiber et al., 2002). The Sharium conceptualizes digital libraries as social environments that facilitate group knowledge creation and problem solving. Interaction is not only between users and information, but also between groups of users, and between users and librarians. The digital library plays a mediating role in this interaction. CKESS takes a similar approach with an important difference. It specifically emphasizes how digital libraries can use social interaction not only as a means for evolving the community's knowledge, but also for evolving the knowledge as captured in the repository. In both of these frameworks, digital libraries are not static environments. They are dynamic environments designed to evolve as the user population grows and expands its knowledge. There is an inseparable connection between the state of the digital library and the knowledge of the community. They are bound to each other; each piece is inextricably linked to the other's evolution. In this sense, digital libraries that support social interaction further the ecological view of digital libraries.

We may also develop a wider view of digital libraries as ecologies by considering the connections between digital libraries and other information systems. The examples so far have emphasized the ecology formed by a single repository and a specific community of users. But digital libraries do not stand in isolation. They exist in conjunction with other digital libraries and as part of the larger information infrastructure of the Internet. At a technical level, interoperability protocols establish connections between isolated collections (Besser, 2002). More broadly, digital libraries are but one part of the information landscape that is emerging through the Internet. An important part of reconceptualizing digital libraries is establishing their niche in this global ecology.

Ecology completes our picture of digital libraries. The first two elements of the INVENT framework (digital objects and representations) tend to frame digital libraries in terms of artifacts. Here, knowledge is primarily conceptualized as a thing that can be externalized and captured in digital form. The next two elements of the framework (activities and interactions) define digital libraries mostly in terms of process. Knowledge, in this view, is dynamic and evolving. It cannot be completely, or even adequately, captured in a repository. Furthermore, if digital libraries intend to truly support knowledge creation, they must actively enable, guide, and support the process of transforming information to knowledge. The last two elements of the framework—actors and ecologies—emphasize how digital libraries are complex systems designed to support actor interactions with representations of digital objects. When all of these conceptual elements are taken together, the INVENT framework presents a model of a digital library and the knowledge it contains as neither an object, nor a process, nor a complex system, but all of the above (regarding various conceptualizations of knowledge, refer to Buckland, 1991; Choo, 1998; Allee, 1997).

6 Summary and future research

Throughout this paper, we have been concerned with two overlapping themes: the reconceptualization of digital libraries as interactive knowledge environments, and the role of information visualization in bringing about the transformative changes demanded by this reconfigured understanding of digital libraries. To this end, we have developed a framework that identifies and clarifies key elements in this vision. Importantly, the framework is descriptive, high-level, and conceptual. The intention is to stimulate thinking about the next generation of digital libraries, especially with regard to the role of information visualization. The framework is not, however, prescriptive and therefore it is not a tool for guiding implementation. Developing a more prescriptive framework, or taxonomy, should be an important, though long-term, objective of digital library research.

We draw attention to two of the research directions suggested by the INVENT framework. The first is developing a taxonomy of epistemic activities. The second, closely related to the first, is developing a taxonomy of interactions with visual representations. When researchers have applied information visualization techniques to digital libraries, they have been primarily concerned with generating visual representations from large collections of digital objects. They have been less concerned with how people interact with those representations, and how these interactions enable epistemic activities. As a result, the classical information retrieval paradigm continues to hold sway. Overcoming the limitations of this paradigm is a necessity for re-imagining digital libraries as knowledge environments. This clearly points to the need for taxonomies that characterize both epistemic activities and interactions in digital libraries.

Ultimately, digital libraries need to address fundamental questions about how computers can amplify human cognitive abilities. To capture, store, and organize information about the world is an insufficient goal. Digital libraries must also help people understand and make sense of the world as represented in digital collections. A founding principle of information visualization is how it can help people “use vision to think” (Card et al., 1999), and in this respect the field is closely aligned with the knowledge environment conceptualization of digital libraries. When Vannevar Bush described the Memex he provided much inspiration for the digital library field (Bush, 1945). It is worth remembering that he titled his paper “As we may think,” not “As we may search.”

References

Allee, V. (1997). *The knowledge evolution: Expanding organization intelligence* (Boston: Butterworth-Heinemann).

Arms, W. A. (2000). *Digital libraries* (Cambridge, Mass: MIT Press).

Bawden, D. and Rowlands, I. (1999a). “Digital libraries: Assumptions and concepts”. *Libri*, Vol. 49, No. 4, 181-191.

Bawden, D. and Rowlands, I. (1999b). “Digital libraries: A conceptual framework”. *Libri*, Vol. 49, No. 4, 192-202.

Berners-Lee, T. (1999). *Weaving the Web* (New York, New York: Harper Collins).

Berners-Lee, T., Hendler, J., and Lassila, O. (2001). “The Semantic Web”. *Scientific American*, Vol. 284, No. 5, 34-43.

Besser, H. (2002). “The next stage: Moving from isolated digital collections to interoperable digital libraries”. *First Monday*, Vol. 7, No. 6. http://www.firstmonday.org/issues/issue7_6/besser/

Bieber, M., Engelbart, D., Furuta, R., Hiltz, S. R., Noll, J., Preece, J., Stohr, E. A., Turoff M., and Van de Walle, B. (2002) “Toward virtual community knowledge evolution”. *Journal of Management Information Systems*, Vol. 18, No.3, 11-35.

Bishop, A.P. and Star, S.L. (1996). “Social informatics of digital library use and infrastructure”. In *Annual Review of Information Science and Technology*, edited by M. Williams (Medford, NJ: Information Today), Vol. 31, 301–401.

Borgman, C. (1999). “What are digital libraries? Competing visions”. I*nformation Processing & Management*, Vol. 35, No. 3, 227-243.

Borgman, C. (1997). “Now that we have digital collections, why do we need libraries?”. In *Proceeding of the 60th ASIS Annual Meeting*, edited by C. Schwartz and M. Rorvig (Medford, NJ: Information Today), Vol. 34, 27-33.

Börner, K., and Chen, C. (Eds.). (2002a). *Visual Interfaces to Digital Libraries* (Berlin: Springer-Verlag).

Börner, K., and Chen, C. (2002b). “Visual interfaces to digital libraries: Motivation, utilization and socio-technical challenges”. In *Visual interfaces to digital libraries*, edited by K. Börner and C. Chen (Berlin: Springer-Verlag), 1-9.

Börner, K., Chen, C., and Boyack, K. W. (2003). “Visualizing knowledge domains” In, *Annual Review of Information Science and Technology*, edited by B. Cronin (Medford, NJ: Information Today), Vol. 37, 179-255.

Brin, S. and Page, L. (1998), “The anatomy of a large-scale hypertextual Web search engine”. *Computer Networks and ISDN Systems*, Vol. 30, No.1-7, 107-117.

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

Budhu, M. and Coleman, A. (2002). “The design and evaluation of interactivities in a digital library”. D-Lib Magazine, Vol. 8, No. 11. http://www.dlib.org/dlib/november02/coleman/11coleman.html

Bush, V. (1945). “As We May Think”. *Atlantic Monthly*, Vol. 176, No. 1, 101- 108.

Campbell, D. G. and Fast, K. V. (2005) “Panizzi, Lubetzky and Google: How the Modern Web Environment is Reinventing the Theory of Cataloguing”. *Canadian Journal of Information and Library Science,* Vol. 28, No. 3, 25-38.

Card, S., Mackinlay, J.D., and Shneiderman, B. (1999) *Readings in information visualization: Using vision to think* (San Francisco: Morgan Kaufman).

Chen, C. (1999). “Visualising semantic spaces and author co-citation networks in digital libraries.” *Information Processing & Management*, Vol. 35, No. 3, 401-420.

Choo, W. C. (1998). *The knowing organization: How organization use information to construct meaning, create knowledge, and make decisions* (New York: Oxford University Press).

Coleman, A. and Oxnam, M. (2002) “Interactional digital libraries: Introduction to a special issue on interactivity in digital libraries”. *Journal of Digital Information*, Vol. 2, No. 4. http://jodi.tamu.edu/Articles/v02/i04/editorial/

Dillon, A. (2002) “Technologies of information: HCI and the digital library”. In *Human-Computer Interaction in the new millennium*, edited by J. M. Carroll (New York: ACM Press), 457-474.

Dix, A. and Ellis, G. (1998). “Starting simple–adding value to static visualization through simple interaction”. In *Proceedings of the 4th International Working Conference on Advanced Visual Interfaces* (AVI' 98, L'Aquilla, Italy, 1998), 124-134.

Faiks, A., & Hyland, N. (2000). “Gaining user insight: A case study illustrating the card sort technique”. *College & Research Libraries*, Vol. 61, No. 4, 349-357.

Fast, K. V and Campbell, D. G. (2004). “‘I still prefer Google’: University student perceptions of searching OPACs and the Web”. In *Proceedings of the 67th Annual Meeting of the American Society for Information Science and Technology,* (Providence, Rhode Island, November 13-18, 2004), Vol. 41, 138-146.

Fast, K. V. and Campbell, D. G. (2001). “The ontological perspectives of the Semantic Web and the Metadata Harvesting Protocol: Applications of metadata from improving

Web search". *The Canadian Journal of Information and Library Science*, Vol. 26, No. 4, 5-19.

Feng, L., Jeusfeld, M. A., and Hoppenbrouwers, J. (2005). "Beyond information searching and browsing: Acquiring knowledge from digital libraries". *Information Processing and Management*, Vol. 41, No. 1, 97-120.

Fensel, D. Hendler, J., Lieberman, H., and Wahlster, W. (2003) *Spinning the semantic web: bringing the World Wide Web to its full potential* (Cambridge: MIT Press).

Fox, E. A., Akscyn, R. M., Furuta, R. K. F., and Leggett, J. J. (1995) "Digital libraries". *Communications of the ACM*, Vol. 38, No. 4, 23-28.

Fox, E. A., and Urs, S. R. (2002) "Digital libraries". In *Annual Review of Information Science and Technology*, edited by B. Cronin (Medford, NJ: Information Today), Vol. 36, 503-589.

Gonçalves, M. A, Fox, E. A., and Watson, L. T., and Kipp, N. A. (2004). "Streams, structures, spaces, scenarios, societies (5S): A formal model for digital libraries". *ACM Transactions on Information Systems*, Vol. 22, No. 2, 270-312.

Hearst, M. A. (1996). "Research in support of digital libraries at Xerox PARC: Part I: The changing social roles of documents". D-Lib Magazine, May 1996. http://www.dlib.org/dlib/may96/05hearst.html

Hearst, M. A, Kopec, G., and Brotsky, D. (1996). "Research in support of digital libraries at Xerox PARC: Part II: Paper and digital documents". D-Lib Magazine, June 1996. http://www.dlib.org/dlib/june96/hearst/06hearst.html

Hollan, J., Hutchins, E., and Kirsh, D. (2000) "Distributed Cognition: Toward a New Foundation for Human-Computer Interaction Research". *ACM Transactions on Computer-Human Interaction*, Vol. 7, No. 2, 174-196.

Ingwersen, P. (1992). *Information retrieval interaction* (London: Taylor Graham).

Kinnucan, M. T. (1992). "Fisheye views as an aid to subject access in online catalogues". *The Canadian Journal of Information Science*, Vol. 17, No. 2, 25-40.

Kirsh, D. (1997). "Interactivity and Multimedia Interfaces". *Instructional Science*, Vol. 25, No. 2, 79-96.

Kirsh, D. (1996). "Adapting the Environment instead of Oneself". *Adaptive Behavior*. Vol. 4, No. 3-4, 415-452.

Kirsh, D. (1995). "The intelligent use of space". *Artificial Intelligence*, Vol. 73, 31-68

Kirsh, D., and Maglio, P. (1994). "On distinguishing epistemic from pragmatic action". *Cognitive Science,* Vol. 18, 513-549.

Kleinberg, J. (1999). "Authoritative sources in a hyperlinked environment". *Journal of the ACM*, Vol. 46, No. 5, 604-632.

Larsen, R. (2004). “Knowledge Lost in Information: Report of the NSF Workshop on Research Directions for Digital Libraries.” NSF Post Digital Library Futures Workshop, (Chatham, Mass., June 15-17, 2003). http://www.sis.pitt.edu/~dlwkshop/report.pdf

Lesk, M. (1997). *Practical digital libraries: Books, bytes, and bucks* (San Francisco, California: Morgan Kaufmann).

Levy, D. M. (2000). “Digital libraries and the problem of purpose.” *D-Lib Magazine*, Vol. 6, No. 1. http://www.dlib.org/dlib/january00/01levy.html

Levy, D. M., and Marshall, C. C. (1995). “Going digital: A look at assumptions underlying digital libraries.”. *Communications of the ACM*, Vol. 38, No. 4, 77-84.

Licklider, J. C. R. (1965). *Libraries of the future* (Cambridge, Mass.: MIT Press).

Lin, X. White, H. D. and Buzydlowski, J. (2003). “Real-time author co-citation mapping for online searching”. *Information Processing and Management*, Vol. 39, No. 5, 689-706.

Lu, S., Dong, M, and Fotouhi, F. (2002) “The Semantic Web: opportunities and challenges for next-generation Web applications.” *Information Research*, Vol. 7, No. 4. http://InformationR.net/ir/7-4/paper134.html

Lynch, C. (2003). “Reflections towards the development of a ‘post-DL’ research agenda”. NSF Post Digital Library Futures Workshop, Chatham, Mass., June 15-17, 2003. http://www.sis.pitt.edu/~dlwkshop/paper_lynch.html

Lynch, C. (2001). “When Documents Deceive: Trust and Provenance as New Factors for Information Retrieval in a Tangled Web”. *Journal of the American Society for Information Science & Technology*, Vol. 52, No. 1, 12-17.

Maglio, P., Matlock, T., Raphaely, D., Chernicky, B., and Kirsh D. (1999) “Interactive skill in Scrabble.” In *Proceedings of Twenty-first Annual Conference of the Cognitive Science Society* (Mahwah, NJ: Lawrence Erlbaum).

Marchionini, G. (1999). “Augmenting library services: Toward the sharium”. In *Proceedings of the International Symposium on Digital Libraries 1999* (Tuskuba, Japan, September 28-29, 1999), pp. 40-47. http://www.ils.unc.edu/~march/sharium/ISDL.pdf

Marchionini, G. (1995). *Information seeking in electronic environments* (Cambridge: Cambridge University Press).

Marchionini, G., and Fox, E. A. (1999). “Progress toward digital libraries: Augmentation through integration”. *Information Processing & Management*, Vol. 35, No. 3, 219-225.

Marshall, C.C. (2003). “Reading and interactivity in the digital library: Creating an experience that transcends paper”. In *Proceedings of the CLIR/Kanazawa Institute of Technology Roundtable* (Kanazawa, Japan, July 3-4, 2003).

Marshall, C.C. (1998). “The future of annotation in a digital (paper) world”. In *Successes and failures of digital libraries*,edited by S. Harum & M. Twidale (Urbana-Champaign: University of Illinois), 97-117.

Monroy, C., Kochumman, R., Furuta, R., and Urbina, E. (2002). “Interactive timeline viewer (ItLv): A tool to visualize variants among documents”. In *Visual interfaces to digital libraries*, edited by K. Börner and C. Chen (Berlin: Springer-Verlag), 39-49.

Morris, S. A., Yen, G., Wu, Z., and Asnake, B. (2003). “Time line visualization of research fronts”. *Journal of the American Society for Information Science & Technology*, Vol. 54, No. 5, 413-422.

Nardi, B. A., and O’Day, V. L. (1999). *Information ecologies: Using technology with heart* (Cambridge, Mass.: MIT Press).

Neth, H. & Payne, S. J. (2002). “Thinking by doing: Epistemic Actions in the Tower of Hanoi”. In *Proceedings of the Twenty-fourth Annual Conference of the Cognitive Science Society*, (Fairfax, Virginia, August 8-10, 2002), 691-696.

Nielsen, J. (2001, May 13). “Search: Simple and visible”. *Jakob Nielsen’s Alertbox*, May 2001. http://www.useit.com/alertbox/20010513.html

Norman, D. A. (1993). *Things that make us smart: Defending human attributes in the age of the machine* (Reading, Mass.: Addison-Wesley).

Nonaka, I. and Konno, N. (1998). “The concept of ‘ba’: Building a foundation for knowledge creation”. *California Management Review*, Vol. 40, No. 3, 40-54.

Nonaka, I. and Takeuchi, H. (1995). *The knowledge-creating company: How Japaneses companies create the dynamics of innovation* (New York, NY: Oxford University Press).

Paepcke, A. (1996). “Digital libraries: Searching is not enough”. *DLIB Magazine*, May 1996. http://www.dlib.org/dlib/may96/stanford/05paepcke.html

Paepcke, A., Chang, C. K., Garcia-Molina, H. and Winograd, T. (1998). “Interoperability for digital libraries worldwide”. *Communications of the ACM*, Vol. 41, No. 4, 33-43.

Peterson, D. (Ed.). (1996). *Forms of representation* (Exeter, UK: Intellect Books).

Rao, R., Pedersen, J. O., Hearst, M. A., Mackinlay, J. D., Card, S. K., Masinter, L., Halvorsen, P.K., and Robertson, G. C. (1995). “Rich interaction in the digital library”. *Communications of the ACM*, Vol. 38, No. 4, 29-39.

Rosenfeld, L. and Morville, P. (2002). *Information architecture for the World Wide Web, 2nd edition* (Cambridge, Mass.; O’Reilly).

Schwan, S. (2004). “The cognitive benefits of interactive videos: learning to tie nautical knots.” Learning and Instruction, Vol. 14, No. 3, 293-305.

Sedig, K. and Liang, H. (in press). “Interactivity of visual mathematical representations: Factors affecting learning and cognitive processes”. *Journal of Interactive Learning Research*.

Sedig, K, and Sumner, M. (In press). “Characterizing interaction with visual mathematical representations.” *International Journal of Computers for Mathematical Learning*.

Sellen, A. J. and Harper, R. H. R. (2002). *The myth of the paperless office* (Cambridge, Mass: MIT Press).

Smith, M. A. and Fiore, A. T. (2001). “Visualization components for persistent conversations”. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (Seattle, Washington, 2001), 136-143.

Soergel, D. (2002, December). “A Framework for digital library research: Broadening the vision”. *D-Lib Magazine*, Vol. 8, No. 12. http://www.dlib.org/dlib/december02/soergel/12soergel.html

Spence, R. (2001). *Information visualization* (New York: Addison-Wesley).

Svenonius, E. (2000). *The intellectual foundation of information organization* (Cambridge, Mass.; MIT Press).

Tufte, E. R. (1990). *Envisioning information* (Cheshire, Conn.: Graphics Press).

Ware, C. (2004). *Information visualization: Perception for design, 2nd edition* (San Francisco, CA: Morgan Kaufman).