The Interaction of Community and Individual Practices in the Design of a Digital Library

Christine L. Borgman, Professor & Presidential Chair
Dept of Information Studies, Box 951520,
Graduate School of Education & Information Studies,
University of California, Los Angeles,
Los Angeles, CA 90095-1520

Abstract

For a single digital library to serve multiple communities, design must be based on an understanding of practices in each of the communities and of relationships between those practices. This paper draws upon data from two large, collaborative research projects, the Alexandria Digital Earth Prototype (ADEPT) and the Center for Embedded Networked Sensing (CENS). Our findings suggest that users' roles, and the degree to which those roles are based on individual or community interests, is an important determinant of how and whether they will use a digital library. In turn, the effect of users' roles varies by the type of content in the digital library (i.e., primary vs. secondary sources) and by the user's level of expertise in the knowledge domain. Despite these complicating factors, digital libraries have the potential to align the activities of scientists and students by providing context for the scientific process and by representing data in ways that are useful for multiple purposes.

Keywords: Scholarly communication, knowledge communities, teaching, learning, infrastructure, digital libraries, uses, users, science, geography, biology, sensor networks, undergraduate students, high school, K-12 schools

1 Introduction

Digital libraries have become an essential technology for providing access to information in a networked world. They also provide a means for people to share knowledge within a community and for communities to share knowledge with each other. From an economic perspective, the more communities and purposes that can be served by a single digital library, the greater the potential cost-benefit ratio of the investment. From a socio-technical perspective, however, many traps and challenges await the unwary digital library designer who attempts to serve multiple communities with a single system [1-5]. Digital libraries are complex systems that support many activities associated with the seeking, use, creation, and sharing of information [6]. These activities can be performed in many ways. If one digital library is intended to serve multiple user communities, design must be based on an understanding of practices in each of the user communities and of relationships between those practices.

We have several decades of research on how individuals use information systems such as digital libraries. Much less research exists on information-related practices of user communities [3]. In this short paper I explore the interaction between individual activities and community-driven information practices in the design of digital libraries that are intended for multiple user communities. My analysis draws upon data from two large, collaborative research projects funded by the (U.S.) National Science Foundation, the Alexandria Digital Earth Prototype (ADEPT) (http://www.alexandria.ucsb.edu http://is.gseis.ucla.edu/adept/) and the Center for Embedded Networked Sensing (CENS) (http://www.cens.ucla.edu), and upon my own earlier research in scholarly communication, information seeking, and human-computer interaction. The symposium on digital libraries and knowledge communities, of which this paper is a part, provides an opportunity to speculate on explanations for some of the results of these two research projects.

2 Background: Goals of the ADEPT and CENS projects

The ADEPT and CENS projects are generating as many questions as answers about how to design digital libraries and other information services for multiple communities and multiple purposes. Comparing the questions and results arising from each of the projects leads to some useful insights about the influence of individual and community practices on the design of digital libraries. First, some background on the goals and status of the projects is necessary.

Our overarching goal in the ADEPT project is to make geo-spatial information resources that are produced and described for research purposes also usable for teaching and learning at the undergraduate level. The Alexandria Digital Library (ADL), constructed as part of the (U.S.) Digital Libraries Initiative Phase I (1994-1998), provides access to geo-spatial resources in many media via sophisticated searching mechanisms [7]. ADEPT is a set of services associated with the ADL intended to support multiple purposes: enable faculty to construct lectures and assignments using content from the ADL and other sources, enable teaching assistants (TAs) in lab sessions to use the information resources assembled by the supervising faculty, and enable stu-

dents to explore the lecture resources and to perform interactive assignments that utilize data, simulations, and other information resources assembled by their instructor. Thus ADEPT has three user communities (faculty, TAs, students), and two purposes (research and teaching). We are presently in the last year of a five-year research project (1999-2004) and have produced a series of papers on the design and deployment of ADEPT at two university campuses [8-12].

We have similar goals in the CENS project, except that we are working with real-time data being generated by scientific research projects rather than with resources that are already collected into a digital library. The data will be collected and organized as they are generated by sensors in the field. Our three communities – scientists, teachers, and students (grades 7 through 12) – all will have access to these data in real time and access to archives of previously generated data. Our goals are to understand the information management practices and requirements of the scientists, teachers, and students so that we can design digital library services to support them.

Other researchers on the CENS education team are concerned with designing teaching modules and studying inquiry-based learning. Our CENS research is in much more formative stages than is the ADEPT project. We are presently in the second year of the five-year grant to the Center (2002-2007), and are pursuing additional funding to explore in more depth some of the issues discussed here.

Both the ADEPT and CENS projects involve scholars in their role as scientists. In ADEPT we are working with geographers; in CENS we are currently working with biologists and will expand into other disciplines associated with the Center. Both projects involve primary data sources; i.e., data collected for research purposes that have received little, if any, prior analysis. In ADEPT we are studying how faculty select and use primary sources in their own teaching, whether from their own research, from ADL, or from other sources. In CENS we are studying how researchers, teachers, and students will collect, select, organize, and use primary data that are being generated continuously from field instruments. Location, as determined by geo-spatial coordinates (latitude and longitude) are essential data elements in both projects.

In ADEPT, researchers and teachers are the same people (i.e., university faculty) in different roles. In CENS, the researchers and the teachers are different people, and come from different communities of practice, with considerably different degrees of knowledge about the scientific domain. Both projects aim to facilitate inquiry learning, which is a method of involving students in scientific practices so that they gain a deeper epistemological understanding of science [13, 14]. Using primary sources to conduct research in ways that scientists do is an essential part of inquiry learning. The ADEPT software is being deployed in undergraduate courses, and mostly in introductory courses that meet general education requirements. CENS' students in grades 7 through 12 are likely to have even less science background than do the ADEPT undergraduates. Teaching assistants (TAs) are a user group specific to ADEPT. These are graduate students (usually at the PhD level) who conduct weekly laboratory sessions in large undergraduate courses for small groups of students. The TA's role is to reinforce concepts from the lectures given by the faculty, answer student questions, and distribute and mark course assignments.

While both projects focus on the use of primary data, in ADEPT we also are studying faculty use of secondary sources (e.g., published literature) as a means of understanding and comparing their practices for teaching and research [10]. A related aspect of the ADEPT project is to determine how faculty employ primary and secondary resources in their teaching.

3 Community and individual practices

With a decade or so of research on digital libraries behind us, the challenges of designing a system to support the richness of information practices within any one user community are now apparent [15]. Digital libraries often are "boundary objects" [16, 17] between user communities, because they serve different purposes for each of those communities. In this sense, digital libraries can be a means to translate practices, data, and knowledge. It may only be in the process of negotiating the design and implementation of a complex system such as a digital library that these differences in practices become apparent. The flexibility that digital libraries offer for searching, using, creating, and managing information can be destabilizing to a community (e.g., a scientific research group). For example, a digital library may enable activities that were the exclusive role of one person or group (e.g., senior scientists) to be performed by others (e.g., graduate students, external partners), and thus contribute to changes in roles and norms [4, 18]. Most studies of the relationship between digital libraries and community practices have focused on research communities in the sciences [15, 19, 20]. The ADEPT and CENS projects introduce two new variables to the interaction between digital libraries and communities: (1) the paired communities have great disparity in domain knowledge and data management skills, from high expertise (faculty and research scientists) to almost complete novices (students in grades 7-12

and in introductory undergraduate courses), and (2) the same digital libraries are intended to serve different, and potentially competing, purposes – research and teaching. We are encountering many of the same community practice issues reported in other research, plus new issues raised by the differences in requirements for research and teaching applications.

The differences in expertise and purposes of the user communities we are studying in ADEPT and CENS highlight the tension between individual activities and community-driven practices in the use of digital libraries. People often function in different roles and interact with their environment accordingly. In some roles, use of a digital library may be driven by individual interests. In other roles, use of a digital library reflects practices common to a community such as researchers, practitioners, teachers, or learners. I propose that the role in which a person interacts with a digital library is a significant determinant of his or her user behavior. A person using a digital library as a researcher, for example, will frame a query in different terms than when that same person is using a digital library as a teacher. Roles may be distinguished by the degree to which the orientation is more individual or more community-driven. The effect of the user's role also varies by the type of content in the digital library (i.e., primary vs. secondary sources) and by the user's level of expertise in the knowledge domain.

In our research on ADEPT and CENS, we are finding that scholarly communication appears to have the strongest community orientation and teaching the strongest individual orientation. Similarly, the use of secondary sources follows more community-based practices than does the use of primary sources. These findings are discussed in a sequence from the strongest community orientation to the strongest individual orientation.

3.1 Primary vs. secondary sources

3.1.1 Scholarly communication and digital libraries

Digital libraries for secondary source content such as scholarly publications have a long history, evolving from several decades of research on text-based information retrieval systems. Digital libraries of scholarly content now play a key role in scholarly communication – the means by which scholars in any field use and disseminate information through formal and informal channels. It is inherently a communal endeavor. Scholars write for each other, making them different from authors who write for the trade press or mass media. Scholars publish their work to stake their claim to their ideas and results and to get credit for their accomplishments. They cite the publications of others for reasons such as describing prior research on which they build; giving credit for prior theories, methods, or discoveries; or refuting prior findings [21, 22].

Scholarly publication has a formalized structure, usually negotiated within each field over long periods of time. Publication manuals such as that of the American Psychological Association [23] codify the format of articles (e.g., sections of a document and what should be included in each, such as the introduction, literature review, methods, results, discussion, and conclusions) and the format of bibliographic references. Journals and conferences specify the scope, format, and content of articles in their calls for papers. Instructions to reviewers specify criteria for acceptance. These structures continue to evolve through community-based processes such as editorial boards and program committees. Each discipline has its own practices, which may vary greatly between the sciences, engineering, medicine, social sciences, and humanities. The practices are neither simple nor stable, but all build upon an incentive structure that has evolved since the earliest print journals, which is to get credit for one's work through a peer review process. As new genres of electronic publication emerge, they build upon existing community practices [6, 22, 24]. Digital libraries now also support the informal processes of scholarly communication, through "pre-print" repositories such as arXiv.org and the circulation of working papers via other forms of institutional repositories such as California Digital Library's e-scholarship [25] or Dspace at MIT and elsewhere [26].

Digital libraries of scholarly content are boundary objects between those with greater and lesser expertise in a field. Experts view digital libraries as repositories for their work. They contribute to digital libraries, but may search them only rarely. Novices view digital libraries as sources of content. It is they who search digital libraries most often. Rather than searching databases, senior scholars rely on informal processes of scholarly communication to receive information about new work. By the time a journal article is published and appears in print or in a digital library, those "in the know" already know about it – they have received it directly from colleagues, as drafts or pointers to a repository, through presentations or discussions at conferences and colloquia, or through conversations or email. Thus senior scholars rely heavily upon their "invisible college" for access to information [27]. For this community, the publication is the formal record of the research. For the student, the high school teacher, or others new to a topic area, the bibliographic infrastructure of scholarly communication provides the means to discover and retrieve the records of research. They search indexing and abstracting services, catalogs, digital libraries, physical libraries, known journals, and the websites of known researchers or research

groups. Thus it is the novices who rely upon metadata and on the formal structures of documents to locate and retrieve information resources [28].

Publications (whether conference papers, journal articles, or books) embody the activities involved in research. These documents reflect the negotiation within a research team about what story will be told, how it will be told, the choice and format of data to support the results, conclusions and interpretations to be drawn, and who will receive authorship credit. Readers of a scholarly document interpret it in the context of other work on which it builds or with which it may conflict. Interpretations also are influenced by external factors such as readers' opinions of the journal, authors, laboratory, or funding agency [22].

Our studies of geographers in the ADEPT project confirm earlier research on scholars' information-seeking behavior in support of research. They report typical behaviors such as browsing library shelves, browsing personal collections, following citation references in articles, asking colleagues, visiting new book and new journal issues shelves in the library, and attending conferences [21]. All of our subjects use online sources, and many continue to be heavy users of campus libraries and print sources [10]. We have not yet asked similar questions of CENS researchers, but have no reason to believe that their research-related information seeking is atypical scholarly behavior.

3.1.2 Digital libraries of primary scientific data

Digital libraries of primary scientific data have a far shorter history than do digital libraries for secondary, published resources. Practices associated with primary scientific data also are less well developed. Yet the pressures and incentives to develop such digital libraries are considerable. Research groups have become larger and more geographically distributed, requiring shared repositories as a means to collect and organize their data. Shared repositories enable researchers to compare and combine data from multiple projects and to conduct longitudinal analyses that might not otherwise be possible or affordable. Funding agencies have encouraged, and sometimes required, researchers to make their primary data available to others. Some journals require that data sets be deposited along with the journal article that reports on those data. Funding agencies also may encourage or require research projects to make their data available for educational applications, from K-12 through university, which is among the drivers of the ADEPT and CENS projects.

For generations of scholars, the journal article or the book was the end product of research. The scholarly communication practices of each field provide a framework within which to document each project, survey, or experiment. Recently, however, the trend is toward scientific data and databases being the end product of scientific research [19]. Establishing community practices to provide access to primary scientific data sources is turning out to be extremely difficult, for a variety of socio-technical reasons [3-5, 15, 18-20, 29-31]. Scholars share data and collaborate on a day-to-day basis, but do so in social and organizational contexts where trust, practices, assumptions, and reciprocity are in place. Where journal articles provide a context for scientific data, digital libraries decontextualize those same data [4]. Scholars often are reluctant to post their data openly where it risks being misinterpreted or used out of context. They also may lack incentives to document and clean the data sufficiently for it to be used by others [19]. Building trusted institutions around digital libraries of primary data that encourage reciprocity appears to be a key requirement for success [4].

The communities of practice associated with secondary sources are much larger than those associated with primary sources of data. For example, those who work on various aspects of water quality may read and publish in the same journals, but take very different stances on the use of primary data depending on whether the data are intended for use in environmental studies or industrial applications [4]. Determining the optimal size of community that can share a digital library of primary source data is a research question worth pursuing. A small research team with a common set of instruments and a common set of goals has incentives to share data. Large, competing teams may have fewer incentives to share, even if they have the same instrumentation as other teams. A critical mass of users is required to develop software to support specific metadata formats for data management and data analysis. Highly specialized metadata formats to serve narrowly defined communities are expensive to maintain, but generic metadata formats may lack sufficient granularity for a research team's requirements.

The choice of metadata format is an important technical and economic matter. Metadata choices are also epistemic choices, however, for they determine how knowledge is conceptualized and interpreted. Designers of digital libraries have begun to consider the "epistemic cultures" of user communities as a means to explain some of these factors [4, 32, 33].

In the ADEPT project, we are studying the criteria by which users select resources from this repository and from other sources. ADEPT is an outgrowth of the Alexandria Digital Library, thus we started with an established repository. We wish to understand more about their knowledge processes, such as how these users judge the value, authenticity, or usefulness of an item for a particular purpose. These appear to be highly individualized, diverse

practices. We have determined that no single DL can provide the range of desired resources for our small sample of geography faculty. Complicating matters further, individual items are used and described differently on different occasions [10].

Among our first activities in CENS was to explore the data management practices of participating research teams, which span a wide range of science and technology disciplines [34]. At one extreme of data sophistication in CENS is the seismology community, which has a long history of highly instrumented data collection, and contributes its data to a shared repository (Incorporated Research Institutions for Seismology- IRIS) (http://www.iris.washington.edu/). This community appears to have constructed a trusted and shared repository within an institutional framework. By contrast, the habitat biology community is more typical of CENS research teams. The habitat research team captures and manages their data locally, using locally developed metadata models that build upon models from their constituent disciplines. We have chosen to focus our initial research efforts on the habitat monitoring project as it provides an opportunity to observe, and participate in, the development of data management models and practices for a new technology (embedded networked sensors) and because it provides educational opportunities suitable for our school community (grades 7-12 biology and physics).

3.2 Research [and, or, vs.] teaching[1]

Comparing the use of digital libraries in research vs. teaching reveals how individualistic information-related behavior can be. The ADEPT and CENS projects are among a very few to compare information-related activities for research and teaching. In both projects, the use of scientific data for research purposes is the reference point. The Alexandria Digital Library was constructed originally as a repository of research materials; ADEPT attempts to add services to make these resources useful for undergraduate teaching by those same researchers. In CENS, our goal is to make the same data streams from sensor networks concurrently available to scientists and to teachers and students in grades 7-12. In the latter case, the data must remain useful and usable to the scientists while also being useful for educational applications.

While a wide array of studies have involved scientists and scientific data in educational activities (more so in K-12 than undergraduate), the vast majority of those projects have avoided the dual data management problem by providing students with processed or "canned" selections of data for scripted activities. Scripted approaches contribute to our understanding of how inquiry learning can be accomplished, but do little to leverage the investment in scientific data production.

From the first stages of the ADEPT project we were struck by the individual nature of undergraduate teaching and the associated information-related behaviors. Observing three different instructors teaching the same introductory course in three consecutive academic terms, we found that they each used different textbooks, assignments, and examinations. Their choices of teaching methods and examples varied accordingly [8, 36]. Our current challenge lies in determining how to transfer tools and content from the instructors with whom we prototyped the system to a larger community.

As a starting point for the ADEPT research, we had naively assumed that sharing primary scientific data between research and teaching would be largely a matter of providing good tools, because the same people (university faculty) are both the researchers and the teachers. However, it is in the transfer of data between research and teaching that we have encountered the greatest disconnect in the ADEPT project. Most of the faculty that we studied (at two major research universities) are sophisticated users of technology in addition to being sophisticated researchers. Several of them participate in large projects gathering data from distributed sensor networks and satellites. Some build climate models on supercomputers. Some have Unix workstations on their desks. Even those at the more humanistic end of the research spectrum are sophisticated users of images and archives. Yet most of them leave most of their technology behind in their offices when they enter their undergraduate classrooms. Instead, they carry colored chalk, stacks of transparencies, and sometimes slides, maps, or samples such as rocks. They teach from textbooks rather than from primary sources. When they do bring primary research data into their teaching, it is usually from their own research, and it is usually in a synthesized form (maps, images, tables) rather than as raw data for students to mine in course assignments. Even if textbooks contain CD-ROMs of images, movies, and other supplemental examples, few of the faculty in our study look at them, much less employ them in class lectures or student assignments [8-12]. Thus digital libraries of primary sources appear to be boundary objects between the roles of faculty: these systems serve very different purposes to them in their roles as researchers and their roles as instructors.

One aspect of the tension between digital libraries for research and teaching is the ability to control one's own resources. Many of the faculty we interviewed for ADEPT expressed a desire to use their research data in teaching – especially slide collections – but were reluctant to release them for general

[1] With thanks to Pam Samuelson for this metaphor, in [35].

use by others [10]. We have addressed this problem in the architecture of ADEPT by enabling each instructor to construct his or her own "personal digital libraries." Individuals can gather their own resources into one place as a subset of the larger digital library. They can choose to share, or not to share, their personal digital libraries with other faculty. This approach also resolves some problems of intellectual property rights, as faculty commonly teach with maps and other examples for which they do not control copyrights [10, 38].

The degree of individual or community orientation is playing out differently in CENS, where researchers and teachers are two different communities, and where the educational application is at K-12 rather than university level. K-12 teachers have far less autonomy in their curricular choices than do university faculty. They must follow curricula prescribed by state and national standards. However, within those standards, they have a fair amount of freedom as to the design of classroom lessons. K-12 teachers commonly borrow good lessons and modules from other teachers, which is another indication that teaching is more community-based at the K-12 level than at the university level. As more university course syllabi are posted online, borrowing among faculty is likely to increase. This trend is worth watching, as it may be an indicator of a transition toward a less individualistic orientation in university instruction.

In CENS, the differences between research and teaching with respect to primary source data are apparent in the purposes for which those data are used and in the domain knowledge that each community brings to those data. Scientists' primary goal is the production of knowledge for their research community, while students' primary goal is to learn the concepts and tools of science. Scientists, as part of their graduate study and research training, have established discipline-specific practices to select, collect, organize, analyze, store, and disseminate data. These practices reflect a tacit understanding about the nature of science, appropriate questions for research, knowledge claims, and the types of evidence required to support knowledge claims. By comparison, K-12 teachers and students generally lack deep subject knowledge, research methods expertise, and knowledge of data management practices.

As noted earlier, research reports must be understood in a context, and putting primary source data in a digital library tends to decontextualize the data [4]. Providing context for scientific research questions is a central goal of inquiry learning approaches, which is our goal in ADEPT and in CENS. In education research, context can be framed in terms of the fidelity between "real" scientific practice and the tools and practices of classrooms. At one pole of educational researchers are those who advocate maintaining a high degree of fidelity between what students do and what scientists do by having the students get involved in live, ongoing scientific studies often "at the elbows" of real scientists doing their own research [39]. At the other pole are approaches that create self-contained learning communities within the classroom. They model the process of scientific investigation but do not work directly with scientists [40]. The design of the tools for accessing, analyzing and presenting scientific data will vary by the approach chosen. Inquiry learning addresses epistemology directly by assisting students to "think like scientists" rather than to learn "about science" from textbooks and small, often artificial, experiments [13, 14]. Thus the match between how science is done and how scientific data are represented in digital libraries may be as critical for science learners as for scientists.

Fidelity between the practice of science and the learning of science will require ways of describing data so that it is useful and usable for both purposes. Metadata formats for scientific data and for educational applications have emerged as a significant barrier to consistent representation in both the ADEPT and CENS projects. Our analysis of metadata models across CENS reveals that scientific metadata models (e.g., SensorML [41], Ecological ML [42]) describe the data, while educational metadata models (e.g., LOM [43], GEM [44], SCORM [45]) describe the scripted scientific educational activity. Almost no overlap in data elements exists between the metadata formats currently in use by the scientific and educational communities we are studying.

4 Discussion and conclusions

Characteristics of knowledge communities are an important consideration in the design of digital libraries. The difficulty lies in determining which characteristics to consider, and for what purpose. In comparing findings from two continuing studies on the design of digital libraries to support research and teaching, it appears that the user's role is an important determinant of how and whether they will use a digital library. Some roles, such as that of teaching, have a more individualistic orientation and others, such as research, have a more community orientation. The influence of these roles is stronger in digital libraries of primary data sources than in the use of secondary sources.

Digital libraries of secondary sources (i.e., published documents) can support large communities of scholars. Individuals who may vary widely in perspective and paradigm can agree on such matters as the editorial standards for a journal and the choice of bibliographic format. Digital libraries of primary sources (i.e., scientific data that have received mini-

mal analysis or interpretation) support much smaller communities of scholars. Sharing digital libraries of primary resources requires agreement on more micro-level details such as choice of metadata format, and thus requires more consensus on epistemological matters.

Others have noted that digital libraries often serve as boundary objects between scientific communities [3, 4, 19, 20]. In ADEPT, it appears that a digital library is in some senses a boundary object between the various roles an individual assumes. That is, an individual may search the digital library in very different ways depending upon whether the activity is part of research or part of teaching. It is in the attempt to design a digital library to serve both roles that these distinctions with regard to system use have become apparent. In other words, the differing uses of the digital library make the otherwise invisible boundary between these roles apparent. Our conclusions from the ADEPT project are that multiple forms of data representation and more extensive functionality are required if a single digital library is to serve both research and teaching.

When designing digital libraries for use by scientists, teachers, and K-12 students, differences in expert and novice behavior become apparent. Scientists bring a much deeper knowledge of the scientific process and domain to their use of a digital library of primary source data than do K-12 teachers and students. If the goal of a shared digital library is to align activities of scientists and students, as it is in inquiry learning projects such as CENS, then finding ways to represent "real scientific data" in ways interpretable by K-12 students and teachers is essential. Only in this way are we likely to increase the fidelity between science and science learning.

Improving fidelity through better data modeling will require substantially new approaches to metadata and to data management. At present, metadata models for scientific data describe the data, per se (e.g., time, date, sensor location, reading), while metadata models for learning science describe the learning experience (e.g., grade level, resources required for the activity, time to perform the activity, education standards met). We are finding minimal intersection in data elements between these two sets of metadata models in the communities under study. Maintaining the data in a useful form for the scientists is paramount. Thus adapting scientific metadata models to educational applications is more promising than is adapting educational models to scientific uses. The next phases of CENS research will address the identification and adaptation of appropriate metadata models for a specific application, which is habitat monitoring. In our current and final year of the ADEPT project, we are focusing on the transferability of digital library tools and resources between faculty members. The findings of these projects should contribute to further understanding of the influence of user roles, individual and community orientation, size of community, application (research vs. teaching), and domain knowledge in the design of digital libraries. Admittedly, these are ambitious goals, and we may continue to produce more questions than answers along the way.

Acknowledgements

The education and evaluation research in ADEPT described here is conducted in collaboration with Anne Gilliland and Gregory Leazer, both of the Department of Information Studies, Graduate School of Education & Information Studies, UCLA, and Richard E. Mayer, Department of Psychology, UC-Santa Barbara. ADEPT is funded by National Science Foundation grant no. IIS-9817432, Terence R. Smith, University of California, Santa Barbara, Principal Investigator. The educational research in CENS described here is conducted in collaboration with William A. Sandoval and Noel Enyedy, both of the Department of Education, Graduate School of Education & Information Studies, UCLA. CENS is funded by National Science Foundation Cooperative Agreement #CCR-0120778, Deborah L. Estrin, UCLA, Principal Investigator.

This paper also benefited from thoughtful discussions with Phil Agre of UCLA, Geoffrey Bowker and Susan Leigh Star of UC-San Diego, and Roy Pea of Stanford University. Prof. Bowker and Prof. Enyedy provided helpful comments on earlier drafts. Stasa Milojevic of CENS assisted in compiling bibliographic references.

References

1. Agre, P.E., *Information and institutional change: The case of digital libraries*, in *Digital library use: Social practice in design and evaluation*, A.P. Bishop, N. Van House, and B.P. Buttenfield, Editors. 2003, MIT Press: Cambridge, MA. p. 219-240.

2. Marchionini, G., C. Plaisant, and A. Komlodi, *The people in digital libraries: Multifaceted approaches to assessing needs and impact*, in *Digital library use: Social practice in design and evaluation*, A.P. Bishop, N. Van House, and B.P. Buttenfield, Editors. 2003, MIT Press: Cambridge, MA. p. 119-160.

3. Star, S.L., G.C. Bowker, and L.J. Neumann, *Transparency beyond the individual level of scale: Convergence between information artifacts and communities of practice*, in *Digital library use: Social practice in design and evaluation*, A.P. Bishop, N. Van House, and B.P. Buttenfield, Editors. 2003, MIT Press: Cambridge, MA. p. 241-270.

4. Van House, N.A., *Digital libraries and collaborative knowledge construction*, in *Digital library use: Social practice in design and evaluation*, A.P. Bishop, N. Van House, and B.P. Buttenfield, Editors. 2003, MIT Press: Cambridge, MA. p. 271-296.

5. Van House, N.A., A.P. Bishop, and B.P. Buttenfield, *Introduction: Digital libraries as socio-technical systems*, in *Digital library use: Social practice in design and evaluation*, A.P. Bishop, N. Van House, and B.P. Buttenfield, Editors. 2003, MIT Press: Cambridge, MA. p. 1-21.

6. Borgman, C.L., *From Gutenberg to the Global Information Infrastructure: Access to Information in the Networked World*. 2000, Cambridge, MA: The MIT Press.

7. Hill, L.L. and G. Janee, *The Alexandria Digital Library Project: metadata development and use*, in *Metadata in Practice: a Work in Progress*, D. Hillman and E. Westbrooks, Editors. 2004, American Library Association: Chicago.

8. Borgman, C.L., et al., *Evaluating digital libraries for teaching and learning in undergraduate education: a case study of the Alexandria Digital Earth Prototype (ADEPT).* Library Trends, 2000. **42**(2): p. 228-50.

9. Borgman, C.L., et al., *Iterative design and evaluation of a geographic digital library for university students: A case study of the Alexandria Digital Earth Prototype (ADEPT)*, in *Research and Advanced Technology for Digital Libraries: 5th European Conference, ECDL 2001, Darmstadt, Germany, September 2001 : Proceedings*, P. Constantopoulos and I.T. Solvberg, Editors. 2001, Springer: New York. p. 390-401.

10. Borgman, C.L., et al., *Comparing Faculty Information Seeking in Teaching and Research: Implications for the Design of Digital Libraries.* Journal of the American Society for Information Science & Technology, in press.

11. Gilliland-Swetland, A.J. and G.H. Leazer. *Iscapes: digital learning environments for the promotion of scientific thinking by undergraduates in geography*. in *First ACM/IEEE-CS Joint Conference on Digital Libraries*. 2001. Roanoke, Virginia.

12. Leazer, G.H., et al., *Classroom evaluation of the Alexandria Digital Earth Prototype (ADEPT)*, in *ASIS 2000 Proceedings of the 63rd ASIS Annual Meeting, Chicago, IL, November 12-16, 2000*. 2000. p. 334-340.

13. Sandoval, W.A. *Epistemic supports for collaborative science inquiry*. in *Poster presentation at CSCL'99, Computer Supported Collaborative Learning Conference*. 1999. Palo Alto, CA.

14. Sandoval, W.A. and B.J. Reiser, *Explanation-driven inquiry: integrating conceptual and epistemic supports for science inquiry.* Science Education, 2003. **87**: p. 1-29.

15. Bishop, A.P., N. Van House, and B.P. Buttenfield, eds. *Digital library use: Social practice in design and evaluation*. 2003, MIT Press: Cambridge, MA.

16. Star, S.L., *The structure of ill-structured solutions: Boundary objects and heterogenous distributed problem solving.*, in *Distributed Artificial Intelligence*, L. Gasser and M. Huhns, Editors. 1989, Pitman: London. p. 37-54.

17. Star, S.L. and J. Griesmer, *Institutional ecology, "translations," and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-1939.* Social Studies of Science, 1989. **19**(3): p. 387-420.

18. Bowker, G.C., *Mapping biodiversity*. International Journal of Geographical Information Science, 2000. **14**(8): p. 739-754.

19. Bowker, G.C., *Biodiversity datadiversity.* Social Studies of Science, 2000. **30**(5): p. 643-683.

20. Bowker, G.C. *Work and information practices in the sciences of biodiversity*. in *VLDB 2000, Proceedings of 26th international conference on very large data bases*. 2000. Cairo, Egypt: Kaufmann.

21. Meadows, A.J., *Communicating Research*. 1998, San Diego: Academic Press.

22. Borgman, C.L., *Digital Libraries and the Continuum of Scholarly Communication.* Journal of Documentation, 2000. **56**: p. 412-430.

23. *Publication Manual of the American Psychological Association*. 5th ed. 2001, Washington, D.C.: American Psychological Association.

24. Cronin, B. *Scholarly communication and epistemic cultures*. in *Scholarly tribes and tribulations: How tradition and technology are driving disciplinary change*. 2003. Washington, DC: Association of Research Libraries.

25. Huwe, T.K., *Social Sciences E-Prints Come of Age: The California Digital Library's Working Paper Repository.* ONLINE, 2002. **26**(September/October): p. 38-42.

26. Tansley, H., M. Bass, and M. Smith. *DSpace as an open archival information system: current status and future directions*. in *Research and Advanced Technology for Digital Libraries. 7th European Conference ECDL*. 2003. Trondhem, Norway: Germany: Springer.

27. Crane, D., *Invisible colleges: diffusion of knowledge in scientific communities*. 1972, Chicago: University of Chicago Press.

28. Agre, P.E., *Institutional circuitry: Thinking about the forms and uses of information.* Information Technology and Libraries, 1995. **14**: p. 225-230.

29. Bowker, G.C., *The Kindness of Strangers: Kinds and Politics in Classification Systems.* Library Trends, 1998. **47**(2): p. 255-92.

30. Bowker, G.C., et al., *A Room with a View [and] Cultural Analysis and What Designers Need To Know–A Case of Sometimes Too Much, Sometimes Too Little, and Always Too Late [and] Synthesizing Diverse Perspectives [and] Make Room for Ethnography in Design!* Journal of Computer Documentation, 1998. **22**(1): p. 11-30.

31. Bowker, G.C. and S.L. Star, *Social theoretical issues in the design of collaboratories: Customized software for community support versus large-scale infrastructure*, in *Coordination theory and collaboration technology*. 2001, Lawrence Erlbaum Associates, Inc., Publishers: Mahwah, NJ, US. p. 713-738.

32. Van House, N.A., *Science and technology studies and information studies*, in *Annual Review of Information Science and Technology*, B. Cronin,

Editor. 2004, Information Today: Medford, NJ. p. 3-86.

33. Knorr-Cetina, K., *Epistemic Cultures: How the Sciences Make Knowledge*. 1999, Cambridge, MA: Harvard University Press.

34. Shankar, K., *Scientific data archiving: the state of the art in information, data, and metadata management.* 2003. Retrieved September 26, 2003 from http://cens.ucla.edu/Education/index.html

35. Samuelson, P., *DRM [and, or, vs.] the Law*. Communications of the ACM, 2003. **46**(4): p. 41-45.

36. Borgman, C.L., Gilliland-Swetland, A.J., Leazer, G.H., Mayer, R., Gwynn, D., Gazan, R., & Mautone, P., *Evaluating digital libraries for teaching and learning in undergraduate education: a case study of the Alexandria Digital Earth Prototype (ADEPT).* Library Trends, 2000. **49**(2): p. 228-250.

37. Gazan, R., et al. *Use Scenarios in the Development of the Alexandria Digital Earth Prototype (ADEPT).* in *Annual Meeting of the American Society for Information Science & Technology*. 2003. Long Beach, CA: Information Today.

38. Borgman, C.L. *Personal digital libraries: Creating individual spaces for innovation*. in *NSF Workshop on Post-Digital Libraries Initiative Directions*. 2003. Chatham, MA.

39. Hay, K. and S. Barab, *Constructivism in practice.* The Journal of the Learning Sciences, 2001. **10**(3): p. 281-322.

40. Reiser, B.J., et al., *BGuILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms*, in *Cognition and instruction: Twenty-five years of progress*, S.M. Carver and D. Klahr, Editors. 2001, Erlbaum: Mahwah, NJ. p. 263-305.

41. *SensorML (Sensor Modeling Language).* http://vast.uah.edu/SensorML/

42. *EML (Ecological Metadata Language).* http://knb.ecoinformatics.org/software/eml/

43. *IEEE Learning Object Metadata (LOM).* http://ltsc.ieee.org/wg12/

44. *Gateway to Education Materials (GEM).* http://www.geminfo.org/Workbench/Metadata/index.html

45. *SCORM (Sharable Content Object Reference Model).* http://www.adlnet.org/index.cfm?fuseaction=scormabt