

UCLA

Papers

Title

What Can Studies of e-Learning Teach Us about Collaboration in e-Research? Some Findings from Digital Library Studies

Permalink

<https://escholarship.org/uc/item/6cj1756r>

Journal

Center for Embedded Network Sensing, 15(4)

Author

Borgman, C L

Publication Date

2005-12-06

What can studies of e-Learning teach us about collaboration in e-Research? Some findings from digital library studies

Christine L. Borgman
Professor & Presidential Chair
Department of Information Studies
Graduate School of Education & Information Studies, UCLA
235 GSEIS Bldg, Box 951520
Los Angeles, CA 90095
[http://is.gseis.ucla.edu/cborgman/
borgman@gseis.ucla.edu](http://is.gseis.ucla.edu/cborgman/borgman@gseis.ucla.edu)
1-310-825-6164
1-310-206-4460 (fax)

Paper submitted to:
Computer Supported Cooperative Work: An International Journal
Special Issue on Collaboration in e-Research

ABSTRACT	2
INTRODUCTION	3
USERS AND USES OF SCIENTIFIC DATA.....	4
BACKGROUND: ADEPT AND CENS PROJECTS.....	6
Alexandria Digital Earth Prototype Project (ADEPT).....	6
Center for Embedded Networked Sensing (CENS).....	7
DIGITAL LIBRARIES AND COLLABORATIVE WORK.....	10
Differences in Use Based on Discipline or Specialty	11
ADEPT Findings.....	11
CENS Findings.....	12
Incentives to Use e-Learning and e-Research Technologies.....	13
ADEPT Findings.....	14
CENS Findings.....	16
Differences in Use of Information by Role	18
ADEPT Findings.....	18
CENS Findings.....	20
Selecting and Sharing Information.....	20
ADEPT Findings.....	21
CENS Findings.....	22
Functionality and Architecture Requirements	24
ADEPT Findings.....	24
CENS Findings.....	26
DISCUSSION AND CONCLUSIONS.....	27
ACKNOWLEDGEMENTS.....	31
REFERENCES	32

ABSTRACT

E-Research is intended to facilitate collaboration through distributed access to content, tools, and services. Findings from two large, long-term digital library research projects are used to illustrate ways in which access to such resources does and does not facilitate collaboration. Both the Alexandria Digital Earth Prototype Project (ADEPT) and Center for Embedded Networked Sensing (CENS) project on data management leverage scientific research data for use in teaching. Two types of collaboration are considered: direct collaboration, in which faculty work together, and indirect or serial collaboration, in which faculty use or contribute shared content such as teaching resources, ontologies, or research data. Implications for collaboration in e-Research are divided into five categories: (1) differences in use based on discipline or specialty, (2) incentives to use e-Learning and e-Research technologies, (3) differences in use of information by role, (4) selecting and sharing of information, and (5) functionality and architecture requirements. Reuse and repurposing of content from research to teaching are proving to be even more complex than anticipated. Better tools and services to manage content can improve capture, management, and preservation. Making content more shareable increases the likelihood that it will be shared. Significant barriers and disincentives to sharing exist, including scientific priority, intellectual property, lack of standards, and the effort to implement systems compared to perceived value. Personal digital libraries offer a middle ground between private control and public release of content. We are just beginning to understand *how* e-Research can facilitate collaboration. The next step is to understand *why*.

INTRODUCTION

Cyberinfrastructure, as it is known in the U.S., or e-Research, as it is known in the U.K. and Europe, promises to facilitate scholarly collaboration by providing access to shared data and document repositories, tools, and services. Considerable progress is being made on building the technical framework, on establishing standards for interoperability, and on the construction of digital libraries to store scholarly content. However, relatively little research has been done to determine how and whether these technologies will facilitate collaboration or enable access to new forms of knowledge. Evidence from prior social studies of science suggests that the adoption of such information technologies is a complex and not always successful process (Almes, Birnholtz et al., 2004; Carr, 1999; Foster & Gibbons, 2005; Hughes, 2004; Kline & Pinch, 1999; Kubicek & Dutton, 1997; MacKenzie & Wajcman, 1999; Woolgar, 2003; Wouters, 2004).

One of the main drivers of e-Research is the “data deluge” (Hey & Trefethen, 2003). The volume of scientific data being generated by highly instrumented research projects (linear accelerators, sensor networks, satellites, seismographs, etc.) is so great that it can only be captured and managed using information technology. The amount of data produced far exceeds the capabilities of manual techniques for data management, and thus the need for control of these data is another essential driver of e-Research (Lord & Macdonald, 2003). Once these data are captured and curated, they can be shared over distributed networks. If these same data can be made available for other applications such as e-Learning, many opportunities arise for economic and political leverage of the investments in e-Research.

This article draws upon findings of two long-term digital library projects in the U.S. on the intersection between e-Research and e-Learning. In the Alexandria Digital Earth Prototype Project (ADEPT) (1999-2005), we studied the use of research-based digital libraries of primary scientific data for teaching at the undergraduate level. Within the Center for Embedded Networked Sensing (CENS), our project (2002-2008) has dual goals of developing an infrastructure for the management of research data by scientific teams and making these data useful for teaching at the middle school and high school levels. Our research in ADEPT and CENS confirms the need for studies of collaborative work in e-Research and e-Learning and for iterative design and evaluation of the technology. This article reflects upon how our findings can inform the design of e-Research infrastructure.

USERS AND USES OF SCIENTIFIC DATA

Sharing data is a core element of scientific collaboration. It is a complex social process involving trust, incentives, disincentives, risks, and intellectual property (Arzberger, Schroeder et al., 2004; Bishop, Van House & Battenfield, 2003; Bowker, 2005; David, 2003; David, 2004; David & Spence, 2003). Data sharing between scientists is a complex and little-studied area (Borgman, in press; Bowker, 2005; Hilgartner & Brandt-Rauf, 1994). Data sharing between scientists, teachers, and students has received even less attention; our research appears to be among the first.

Scientists who collaborate with each other tend to have similar disciplinary knowledge and analytical skills. Such similarities cannot be assumed when the same scientific data are shared with teachers and students (Enyedy, 2003). To serve these two communities with one set of resources, two potential conflicts must be addressed. One is that scientists and students collect and analyze data for different purposes. Scientists' primary goal is the production of knowledge, while students' primary goal is to learn the concepts and tools of science. For students, "doing science" is a means to learn new content and skills. In the ideal case, students also will generate data that contributes to knowledge in their classroom and to scientific knowledge.

The second conflict is that scientists, teachers, and students bring far different skill sets and epistemologies of science to the use of scientific data. As part of their graduate study and research training, scientists have learned practices to select, collect, organize, analyze, store, and disseminate data. Scientific practices reflect a tacit understanding about the nature of science, researchable questions, knowledge claims, and evidence necessary to support claims. By comparison, teachers and students at the middle and high school levels generally lack deep subject knowledge, research methods expertise, and knowledge of data management practices. Students in introductory university courses have only slightly more knowledge of scientific practices than do high school students. For scientific data from e-Research projects to be useful for learning, teachers and students need considerable assistance to bridge the gaps in knowledge and skills between them and the scientists who produced the data.

BACKGROUND: ADEPT AND CENS PROJECTS

Some background on the goals of the ADEPT and CENS projects will aid in explaining the implications of our findings.

Alexandria Digital Earth Prototype Project (ADEPT)

The central goal of the ADEPT project (<http://is.gseis.ucla.edu/ADEPT/>) was to make geo-spatial information resources intended for research purposes 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 (Hill & Janee, 2004). ADEPT is a set of services associated with the ADL intended to enable faculty (i.e., members of academic staff who teach and conduct research) to construct lectures and assignments using content from the ADL and other sources, enable teaching assistants in lab sessions to use the information resources assembled by the supervising faculty member, and enable students 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, teaching assistants, students), and two purposes (research and teaching).

The research and development of ADEPT was conducted from 1999 to 2004; the education and evaluation component of the project continued to 2005. We studied the design, deployment, and adoption of prototype ADEPT learning environments at two

university campuses. Research methods included classroom observations, interviews with faculty, students, teaching assistants, and developers; analysis of teaching materials (lectures, assignments, exams); and analysis of available metadata standards (Borgman, 2004b; Borgman, 2004a; Borgman, 2004c; Borgman, 2004d; Borgman, 2004e; Borgman, 2005a; Borgman, Gilliland-Swetland et al., 2000; Borgman, Leazer, Gilliland-Swetland & Gazan, 2001; Borgman, Leazer et al., 2004; Borgman, Smart et al., 2005; Borgman, Smart et al., 2004; Champeny, Borgman et al., 2004; D'Avolio, Borgman et al., 2005; Gazan, Leazer et al., 2003; Leazer, Gilliland-Swetland, Borgman & Mayer, 2000; Mayer, Smith, Borgman & Smart, 2002). Research on system design and architecture also is documented elsewhere (Ancona, Freeston, Smith & Fabrikant, 2002; Hill & Freeston, 2003; Hill, Janee, Dolin, Frew & Larsgaard, 1999; Hill, Carver et al., 2000; Hill & Janee, 2004; Janee & Frew, 2002; Janee, Frew & Hill, 2004; Smith, Ancona et al., 2003; Smith & Zheng, 2002). This article reflects on the implications of our findings for collaborative work in e-Research. Further work on the ADEPT software and on broader system deployment continues under a joint U.S. (National Science Foundation) and U.K. (JISC) funded project (*DialogPlus: Digital Libraries in Support of Innovative Approaches to Learning and Teaching in Geography*, 2003).

Center for Embedded Networked Sensing (CENS)

As a large, multidisciplinary research collaboration among multiple universities that involves sharing heterogeneous data collections, the Center for Embedded Networked Sensing (<http://www.cens.ucla.edu>) fits the U.K. definition of e-Science (*e-Science Core*

Programme, 2004). CENS is a National Science Foundation Science and Technology Center based at UCLA that includes dozens of cooperating scientists, technologists, educators, and teachers (middle school and high school). The Center was launched in August, 2002, with funding to 2007, and renewable to 2012. CENS investigators manage many additional grant projects through the Center.

CENS is developing embedded networked sensing systems and applying this technology to scientific applications. These are large-scale, distributed, systems composed of smart sensors and actuators embedded in the physical world. They monitor and collect information on such diverse subjects as plankton colonies, endangered species, soil and air contaminants, medical patients, and buildings, bridges, and other human-made structures. A central goal of embedded networked sensing systems is the ability to reveal previously unobservable phenomena. The researchers in CENS are investigating fundamental properties of these systems, developing new enabling technologies, and exploring novel scientific and educational applications. Computer scientists, engineers, and scientists (e.g., biology, geology, seismology, environmental sciences, marine sciences) are collaborating to design and deploy these systems. As the Center has evolved, scholars in related fields have joined our projects, including participating faculty from statistics, law, architecture, design, and film.

The CENS education and data management teams have undertaken a broader scope of work than did the ADEPT education and evaluation team. Our scope in CENS includes both design of an infrastructure for managing the scientific data and the design of tools

and methods for making these data useful for educational applications. The content in CENS consists of real-time data generated by scientific research projects, rather than resources already collected into a digital library. The effort on this project was accelerated by an additional grant from the National Science Foundation specifically for data management and educational research (Sandoval & Borgman, 2004-2008).

Research questions addressed in this article focus on collaborative use of scientific data. Our findings to date are reported in papers and talks (Borgman, 2004b; Borgman, 2004a; Borgman, 2004c; Borgman, 2004d; Borgman, 2004e; Borgman, 2005a; Borgman, 2005b; Borgman & Enyedy, 2005; Shankar, 2003).

CENS' sensor networks currently are deployed to study habitat biology, water quality, seismology, contaminant transport, marine microorganisms, and several other topics. Habitat biology and water quality research are current foci of our data management and education research; we also have interviewed scientists in other CENS areas. Habitat data generated by sensors at an ecological reserve in the mountains east of Los Angeles (*James San Jacinto Mountains Reserve*, 2004) can be monitored in real time or analyzed as datasets over selected time periods. For prototyping purposes, we deployed a similar set of habitat sensors near one of the participating schools. Scientists, teachers, and students (grades 7 through 12) have access to these data in real time and access to archives of previously generated data.

DIGITAL LIBRARIES AND COLLABORATIVE WORK

Behavioral studies of digital libraries have focused largely on individual users, following the traditions of information-seeking research; only a few studies have examined collaborative aspects of these technologies (Agre, 2003; Bishop et al., 2003; Borgman, 2000a; Borgman, 2000b; Borgman, 2003a; Case, 2002; Ellis, 1989; Kuhlthau, 1991; Lynch, 2003; Marchionini, Plaisant & Komlodi, 2003; Star, Bowker & Neumann, 2003; Van House, 2003; Van House, Bishop & Battenfield, 2003). Our research questions in ADEPT focused initially on individual users, but collaborative issues quickly arose. Collaboration was a central concern of our CENS research on data management from the start.

Lessons from research on ADEPT and CENS that provide insights into collaborative aspects of e-Research can be divided into five categories, which are used to organize the remainder of this article: (1) differences in use based on discipline or specialty, (2) incentives to use e-Learning and e-Research technologies, (3) differences in use of information by role, (4) selecting and sharing of information, and (5) functionality and architecture requirements. Research supporting these findings is discussed in context rather than presented as a separate literature review.

Differences in Use Based on Discipline or Specialty

It will come as no surprise to those schooled in collaborative work or social studies of science that the use of data and information varies by discipline or specialty (Case, 2002; Meadows, 1998).

ADEPT Findings

Due to the content of the introductory courses on geography that we studied, most of the research subjects in ADEPT were physical geographers; the rest were human or cultural geographers. In the one study that compared these two specialties, we found differences in degree of collaboration and in data sources. Physical geographers were more likely to participate in large, collaborative projects and their research was more data-driven.

Several of them used data sets produced by agencies such as NCAR (*National Center for Atmospheric Research*). Human geographers in our sample tended to work on their own and to write sole-authored scholarly books; their research was more concept-driven than data-driven (Borgman et al., 2005). Fry (2003) also found that social/cultural geographers (comparable to the human geographers in our studies) tended to work individually rather than as part of collaborative teams.

Information resources used by geographers in both specialties were similar, however. All of the faculty we studied rely heavily on maps and spatial images, as would be expected for a field in which physical location and spatial orientation are organizing metaphors.

When asked to define “primary data,” faculty from all specialties responded (in varying ways) that primary data were anything raw and unprocessed, such as sensor data and field notes, while secondary data were analyzed or processed or interpreted in some way (Borgman et al., 2005).

CENS Findings

Collaboration is one of the factors that influences sharing of data. Research collaborations frequently are based on sharing expensive instrumentation or resources (Finholt, 2002; Olson & Olson, 2000; Sonnenwald, 2006). Agreements about sharing data are central to establishing collaborations (David & Spence, 2003). The degree of automation in data collection and analysis also is associated with the likelihood of sharing data (Pritchard, Carver & Anand, 2004). Thus research specialties that are more collaborative and make more use of instrumentation are more likely to use e-Research technologies. Our initial findings in CENS support this hypothesis.

We spent the first year of the CENS project assessing the state of data management practices and available methods in each of the participating research groups (Shankar, 2003). We found a wide disparity across CENS in practices, tools, resources, and data archives. Habitat ecology and seismology were selected as initial topic areas due to their different histories of data management. Seismic data have been collected via automatic instruments since the since the early 1970s, and a common metadata standard has been in use since 1988 (*Standard for the Exchange of Earthquake Data (SEED)*, 2004),

maintained by a global organization (*Federation of Digital Broad-Band Seismograph Networks (FDSN)*, 2004). The FDSN assigns network codes to provide uniqueness to seismological data streams. Many of these data are contributed to a community repository (*Incorporated Research Institutions for Seismology (IRIS)*, 2004), which has established policies and practices for the use of those data. Thus the seismology community has a long history of distributed, collaborative, and standardized management of its data.

Habitat ecologists, in contrast, tend to work alone or in small groups. Spreadsheets are their preferred tool for data analysis, with data models specific to each project. These spreadsheets are used to produce graphs, charts, and tables for their research publications. Data may not be maintained beyond the end of a study. As this community makes more use of the CENS sensor network technology, they are beginning to standardize their data collection and analysis to a greater degree (Borgman & Enyedy, 2005; Borgman, Enyedy, Wallis & Sandoval, 2006). A community resource for data management does exist, known as the Knowledge Network for Biocomplexity. Their first software products and tools became available in 2001 (*Ecological Metadata Language (EML)*, 2004; *Knowledge Network for Biocomplexity (KNB)*, 2004). The CENS scientists we are studying are not making much use of KNB as yet.

Incentives to Use e-Learning and e-Research Technologies

Both the ADEPT and CENS 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 (Enyedy & Goldberg, 2004; Sandoval & Reiser, 2003). While a

laudable goal, scientists and teachers are very busy people. Few are interested in dabbling with technology for its own sake. Rather, they will adopt a technology only if it offers sufficient advantages to justify the investment in learning, in changing associated practices, and in costs of the technology itself (Rogers, 1995). Research is a collaborative activity in most scientific fields. Teaching is usually a solo activity. These differences appear to influence incentives to use e-Learning and e-Research technologies (Borgman, 2004d).

ADEPT Findings

While the overall ADEPT project was based on the premise that faculty would make more use of primary data in teaching undergraduate geography courses if they had better tools to mine digital libraries and to extract these data in forms useful for instruction, our team treated that premise as a research question. We assessed incentives and criteria for adoption of technology throughout the ADEPT research (D'Avolio et al., 2005).

Geography is a technology-intensive field, and we found participating ADEPT faculty to be sophisticated users of technology in their research. Although most expressed interest in experimenting with new instructional methods, few of them employed computer-based technology in their teaching. The most common reasons given were that too much advance planning was required for computer-based instruction and that too much assistance would be required to install and support the equipment (Borgman et al., 2000; Borgman et al., 2001; Leazer et al., 2000).

Having anticipated these barriers, we built substantial resources into the ADEPT grant to lower the effort required to use the technology. Our hypothesis was that if we could provide sufficient assistance in course development to trade for the instructors' time in participating in the research, we could persuade faculty to teach with the ADEPT prototypes. Our investment of staff resources afforded the opportunity to conduct formative evaluation and to contribute iterative assessments to the design and implementation teams (Borgman et al., 2001; Gazan et al., 2003).

In the 2002-2003 academic year, we conducted an extensive assessment of a full prototype deployment of the ADEPT system in two sections of an introductory course in physical geography (Champeny et al., 2004). The course was taught in the fall and spring quarters by the same instructor; the system was refined during the interim winter term. The ADEPT "digital learning environment," as implemented at that time, had three components that could be used in various combinations: the "object collection" of primary source modules (instructors could contribute their own or use those in the collection), the "lecture composer" to organize lecture outlines with these objects embedded, and the "knowledge base" of concepts and relationships (these were a form of ontology (Hovy, 2003)).

The instructor of this course used all three components, devoting most of his efforts to developing a knowledge base. He requested that approximately 1,000 concepts be created for the 10-week course. At an estimated one-hour of labor per concept, the ontology for

this one introductory course would have required the equivalent of 25 weeks of full-time labor (Champeny et al., 2004). As the course went on, the effort became unsustainable for the instructor and his graduate students and he did simplify his requirements. This is not the first, nor likely the last, project to become mired in ontology construction (Goble & Wroe, 2004; Ribes & Bowker, 2004).

When recruiting other faculty members to use ADEPT in their classrooms, we demonstrated the various options without offering judgments as to their usefulness. Two instructors implemented prototypes in 2003-04 at a different university campus, and were presumed to be unfamiliar with the experience of the instructor discussed above. These two faculty members used the tools to construct lectures and embed primary data resources; neither instructor chose to construct an ontology nor to use concepts already developed. The ADEPT ontology was intended for “serial collaboration:” each instructor could use the knowledge base concepts contributed by prior instructors and could add his or her own concepts to the database. The massive investment in knowledge base construction was justified by the potential for re-use. This aspect of ADEPT did not succeed, at least in the short time remaining in the project for other instructors to implement the system.

CENS Findings

Although we are at a much earlier stage in the CENS project, it is becoming clear that scientists’ interest in data management is associated with instrumented data collection.

The first two to three years of CENS were devoted to engineering work in developing sensor technology and networks. While scientists in application domains such as seismology, biology, habitat ecology, and environment have worked closely with the engineering teams since the beginning of the Center, only recently has the technology matured enough to produce a steady stream of scientific data. The scope of the impending data deluge is becoming apparent, lending urgency to improving capture and management mechanisms. The methods and research questions in ecology were changed substantially by the introduction of remote sensing technology via satellites in the 1980s and 1990s (Kwa, 2005). Our research in data management may provide insights to how these research areas will evolve with the introduction of embedded sensing networks.

Middle and high school teachers are users, rather than producers, of scientific data and thus have different incentives to use e-Research technologies than do scientists. A goal of CENS is to teach with the same primary data used by participating scientists rather than to use “canned” datasets with pre-defined questions and answers, as is the case with most science learning projects. Few teacher education programs cover inquiry learning or the use of primary data sources in their core curriculum. Accordingly, these approaches rarely are employed in teaching at the middle and high school levels. Given the overhead of training teachers in these methods, we sought participation in CENS by innovative teachers who already are familiar with the approach.

Better pedagogy alone is insufficient incentive for teachers to participate. Course content must be based on the California Science Standards. Teachers are required to teach to

these standards and students are tested on them; no incentives exist to teach topics that are not included in these standards. Because the environment is a central topic in these standards, we had little difficulty selecting appropriate modules in habitat biology and water quality, however.

Differences in Use of Information by Role

An unexpected finding in the ADEPT project was differences in the use of information by role of researcher or teacher. We are just beginning to explore these implications in CENS.

ADEPT Findings

In initial stages of the project we treated geographers as one sample population. As differences emerged in their use of information when researching and when teaching, we realized that we had begun the project with the naïve assumption that sharing primary scientific data between research and teaching would be largely a matter of providing good tools, because the same people are both the researchers and the teachers. However, it was in the transfer of data between research and teaching that we encountered the greatest disconnect in the ADEPT project.

The researcher-teacher role differences were manifested in three ways. One was in their use of information technology. As noted above, even the most technologically

sophisticated faculty members made little use of technology in their teaching. Their offices often contained two or more computer workstations, multiple hard drives, and other equipment such as scanners. Some of the faculty used supercomputers in their research. Yet most of them left these technologies behind in their offices, heading to class with chalk and overhead transparencies (Borgman, 2004d; Borgman et al., 2000; Borgman et al., 2001; Leazer et al., 2000). The second difference is in the use of data. They taught introductory courses from textbooks rather than from primary sources. Any research data used to illustrate their lectures usually came from their own research. These data usually were presented in synthesized forms (maps, images, tables) rather than as raw data for students to mine in course assignments (Borgman et al., 2001; Borgman et al., 2005; Gazan et al., 2003; Gilliland-Swetland & Leazer, 2001; Leazer et al., 2000).

The third comparison, which we examined most closely, was between information-seeking for research and teaching (Borgman, Leazer et al., 2004; Borgman et al., 2005; Borgman, Smart et al., 2004). Geographers sought information in support of their research in expected ways: they track the new literature in their fields, browse familiar sections of the library, bookmark favorite web sites, follow citation links, attend professional conferences, and receive sources and references from their scholarly peers. Searching for information in support of their teaching was more serendipitous. They may find useful items for teaching in the process of searching for research topics, and may also find research ideas or resources while gathering information for teaching.

CENS Findings

In our CENS studies of information and technology use, we have studied scientists only as researchers. In subsequent research, we may also look at how these scientists use their research data in teaching undergraduate and graduate courses. We are not yet studying the data and information uses of middle and high school teachers, but this also is a topic under consideration for future research. Teachers have far fewer skills than scientists in technology and in information management. They also have far less access to information technology. These differences are important considerations in making scientific data useful in e-Learning for middle and high school students.

Selecting and Sharing Information

Access to information and control of information are fundamental aspects of science. While “open science” drives the incentives to publish and to verify findings quickly through peer review (David, 2003), many disincentives also exist to sharing information, especially research data. Disincentives fall into four categories: (1) rewards for publication rather than for data management; (2) the amount of effort required to document data for use by others; (3) concerns for priority, including rights to control results or sources until publication of research; and (4) intellectual property, both control and ownership of one’s own information and access to information controlled or owned by others (Borgman, in press). Here I focus briefly on issues of priority and intellectual property that arose in the ADEPT and CENS projects, and how these factors influence collaboration in e-Research.

ADEPT Findings

We studied how faculty members select and use data and information for their research and for their teaching. We asked few specific questions about sharing information, but these issues arose throughout the project. Faculty members frequently copy materials from textbooks, journal articles, newspapers, and other copyrighted sources for use as illustrative examples in their teaching. When displayed in a classroom to a few dozen or a few hundred students, instructors commonly assume that use falls under educational provisions of U.S. copyright law. (Educational use provisions of copyright laws vary widely by country; these studies focused on use within the U.S.). When the same resources are stored on a website in digital form, even if access is restricted to enrolled students, narrower fair use guidelines associated with digital resources usually apply (*The Digital Millennium Copyright Act of 1998*). If instructors wish to make those same resources publicly available via a digital library, their use definitely is subject to copyright permission. Faculty members interviewed for ADEPT often felt that the effort required to manage the use of copyrighted resources in an e-Learning environment constrained their choice of teaching resources, compared to what they normally used with more traditional teaching methods (e.g., chalkboards, handouts, overhead displays, PowerPoint slides, maps, and objects such as rock samples). These constraints were a disincentive to using the ADEPT technology in teaching and to contributing resources to the collection for faculty.

The education and evaluation team strongly encouraged the builders of the ADEPT object collection to include metadata tags for copyright ownership on each object as it

was created, but the builders were inconsistent in doing so. If ownership of the object collections were clearly identified, it would be possible to distribute the ADEPT software with the public objects and restrict access to copyrighted materials. The lack of clear distinction hampered the dissemination of the ADEPT collections.

Geography faculty members' choices of data to use in teaching were influenced by intellectual property considerations, by availability, by familiarity, and by cost (Borgman et al., 2005). They draw upon their own data for a variety of reasons: these resources are familiar, available, usually do not require copyright permissions from others, and use does not incur additional monetary costs. Physical geographers often write grants to purchase datasets for their research. They may be able to use these same data in their teaching (or at least their own analyses of those data), but are unlikely to get instructional funds to purchase data for teaching alone.

CENS Findings

Issues of data sharing and ownership are explicit research questions in CENS. We are studying how scientists, teachers, and students determine their data requirements, their criteria for selecting and preserving data, their use of scientific data and how that use evolves over time, and incentives and disincentives to contribute data to repositories. Our research methods include attending workgroup meetings of scientific teams and analyzing their work products (datasets, websites, publications), interviewing individual

faculty and research groups, visiting research sites, and identifying appropriate data repositories, metadata standards, and structures.

Findings to date reveal that data such as meteorological measurements are considered non-proprietary and are made public immediately. The James Reserve, for example, posts data on local conditions every 15 minutes (<http://www.jamesreserve.edu/weather.html>).

They also have webcams on multiple observation points (e.g., bird feeders, nestboxes, moss growing) that Internet visitors can control remotely

(<http://www.jamesreserve.edu/WebcamFiles/RoboCam.html>). Access to other data from research projects at the site varies. Individual investigators are using small local arrays to study braken ferns, for example. Multiple investigators are obtaining ecological data from the James Reserve and elsewhere via three-dimensional robotic sensor data collection from the NIMS technology (*NIMS: Networked Infomechanical Systems*, 2005; Batalin, Rahimi et al., 2004; Sutton, 2003). Data from individual projects are less likely to be made public until a sufficient portion of the data are analyzed and published (Borgman & Enyedy, 2005; Borgman et al., 2006).

Seismology, with its long history of common standards, community repositories, and collaboration has addressed access policies more explicitly. Data contributed to IRIS (*Incorporated Research Institutions for Seismology (IRIS)*, 2004) are embargoed from public use for two years from the date of experiment before being released, although requests can be made to individual investigators for data to be shared sooner (<http://www.iris.edu/services/faq.htm#iris>). The CENS seismology research team

contributes its data to IRIS in the standard SEED format (*Standard for the Exchange of Earthquake Data (SEED)*, 2004). However, their embedded networked sensors are generating data at a higher rate (500 samples/second) than IRIS currently accepts, raising questions about how data management practices will keep pace with advances in data collection methods.

Functionality and Architecture Requirements

The differences in use based on research specialty and on role, varying incentives to use technologies, and varying practices in selecting and sharing information all have implications for the functionality and architecture requirements of e-Research. In both ADEPT and CENS, the implications cluster into database and metadata issues. I discuss briefly the requirements identified in each of these projects, then return to implications for collaboration in the discussion and conclusions.

ADEPT Findings

ADEPT would not be successful either as a single shared collection, due to varying roles and incentives, nor as independent collections for each faculty member, as the potential for re-use would be undermined. A middle ground was identified early in the design process (Janee & Frew, 2002) and its efficacy confirmed by findings of the education and evaluation team (Borgman, Leazer et al., 2004; Borgman et al., 2005; Champeny et al., 2004). The ADEPT architecture addresses faculty's reluctance to release their data by

enabling each instructor to gather his or her own resources into a “personal digital library.” Faculty members can choose to share, or not to share, their personal digital libraries with others and to make items visible or not visible in the shared collection. This approach resolves some problems of intellectual property rights, and enables faculty members to use their research data for teaching without necessarily contributing it to the common pool (Borgman, 2003b).

Metadata is the second architectural issue arising from differences in the use of geographic resources for teaching and research. The Alexandria Digital Library, on which ADEPT is based, includes sophisticated metadata and a gazetteer. These capabilities were extended greatly in the ADEPT project by the teams at University of California, Santa Barbara (Janee et al., 2004), providing extensive access by location (place name, latitude and longitude) and some access by concept. However, these access mechanisms proved insufficient for teaching purposes. Faculty, in their roles as teachers, asked for data, images, or simulations that would demonstrate concepts such as erosion or adiabatic processes, regardless of the physical location on earth. Further complicating matters, an instructor might use an image or dataset to illustrate multiple points in one or more lectures (Borgman et al., 2005). Anticipating all the ways that a document or object might be described is one of the classic problems of knowledge organization, and will not be solved by any single system (Svenonius, 2000).

CENS Findings

Similar divisions between public and private data are emerging in our CENS research. Some data always are public, some data are private initially and made public later, and some data remain private. The personal digital library model should be useful in CENS, but “personal” libraries often will be controlled by collaborative teams rather than by individuals. Teams could manage their own data and set criteria for which data, at what time, will be released to the community repository. This approach also may allow access to scientific data for educational applications sooner, especially if the data are provided at a higher degree of granularity than is required for research publications.

The mismatch of metadata structures between research and teaching applications is even more striking in CENS. Metadata models for the habitat biology community (e.g., *Ecological Metadata Language (EML)*, 2004) describe the data (e.g., time, date, sensor location), while educational metadata models (e.g., *ADEPT/DLESE/NASA metadata (ADN)*, 2001; *IEEE Learning Object Metadata (LOM)*, 2004) describe the educational activity (e.g., grade, level, resources required for the activity, time to perform the activity, educational standards, etc.). We found no overlap in data elements between these scientific and educational metadata formats. The formats cannot be reconciled because they serve different purposes: metadata in scientific applications describe the data, while metadata in educational applications describe the pedagogy. A major thrust of our current work is to explore ways to bridge this gap via layered models, filtering tools, or other methods.

DISCUSSION AND CONCLUSIONS

E-Research is intended to facilitate collaboration through distributed access to content, tools, and services. *How* e-Research will facilitate collaboration is an open question.

Our research did not begin with a focus on collaborative work in either the Alexandria Digital Earth Prototype Project or in the Center for Embedded Networked Sensing project on data management. As collaborative issues became ever more prominent, we shifted our emphasis to explore them. I have drawn upon the findings of these two large, long-term research projects to suggest ways in which distributed access to content, tools, and services does and does not facilitate collaboration.

Two types of collaboration arose in these projects. One is the more traditional model, in which faculty work together on research projects or with teams from other disciplines (in this case, information studies, psychology, and education) to build new tools and services. The second is indirect or serial collaboration, in which faculty contribute or use contributed content. In the latter case, faculty members collaborate indirectly by building upon prior work of others whom they may or may not know. The shared content may be teaching resources or ontologies (ADEPT) or research data (CENS).

In both the ADEPT and CENS projects, we found that degree of collaboration varies by research specialty. Research in some specialties involves multiple investigators and in others research is accomplished solo. This variance is not surprising or atypical. What is notable is the relationship between degree of research collaboration and sharing of

content. Research collaboration usually requires sharing instruments and sharing data; solo researchers have fewer occasions or incentives to share. However, as solo research begins to rely more on instrumentation such as sensor networks, the incentives to use standardized methods of data capture and data management increase. Standards depend upon consensus within scientific communities, and thus result from collaborative activities. Once captured in standard formats, these data are more amenable to processing by off-the-shelf tools and services, and also are more amenable to being shared. Thus instrumented data collection can lead to more collaboration in sharing content and in the use of common tools and services.

The relationship between better tools and more collaboration is by no means direct, however. Making data, teaching resources, or other forms of content easier to share does not mean that faculty *will* share. Researchers are concerned about the priority of their claims and will protect data and sources until publication. Data that are ancillary to research findings, such as meteorological records, are more readily and immediately shared. Faculty members are concerned about managing their data as intellectual property, whether for research or teaching purposes. Making better tools available to manage data increases the likelihood that data are managed in consistent ways, documented, and preserved. Even if data are closely held by investigators, better tools get those data into the pipeline, and the investigators may be willing to release them at some later time. Without such tools, data may never be captured or preserved for future sharing.

Our initial hypothesis in ADEPT was that better tools would promote the use of primary data in teaching, especially if the overhead in constructing resources was lowered substantially through staff support. We found that lowering the overhead was necessary to get faculty to participate in the research project by deploying prototypes in their teaching. The tools were attractive, but for reasons other than anticipated. Geographers were more interested in these tools for the purposes of managing their own data than for using the data in the digital library provided. When they did use primary data in teaching introductory courses, it was their own data. The digital library at the core of our study received little use for teaching purposes.

Of the three sets of tools and services offered for use in teaching geography courses, those for organizing lecture content and illustrating lectures with maps, images, and other objects were the most attractive to the most faculty members. The knowledge base, or ontology, to which one faculty member had devoted most of his efforts, received the least interest in reuse by other instructors. The potential lack of ontology reuse was a concern of the education and evaluation team from the early design phases. Ontologies are most effective when built collaboratively because they represent a community's collective understanding of concepts and relationships. No matter how sophisticated the ontology built by an individual, the community may not use it if they lack a sense of ownership.

Differences in the use of content, tools, and services by the same faculty when in the roles of researcher and teacher may have profound implications for the design of e-Research. In both the ADEPT and CENS projects, a key goal is to leverage the

investment in research data by making it useful for teaching purposes. Reuse and repurposing of content are proving to be even more complex than anticipated. The low level of technology use in teaching introductory undergraduate geography courses is not due to lack of technology expertise. Quite the contrary; these faculty make extensive use of technology in their research but choose not to use it in their teaching. We relieved disincentives of logistics and development time by providing staff support. Concerns remained for the unreliability of the technology, for the constraints imposed by intellectual property permissions, and the difficulty of identifying useful content for teaching.

The mismatch of metadata for research and teaching is an issue in both projects. Researchers search for content by concepts and, in geography, by location. Teachers search for examples that will illustrate concepts and processes. They often are less concerned with place or with specifics of the concept than with presentation criteria such as the clarity of the image and how well it will display in a classroom. Research materials rarely are described in these ways. Conversely, educational objects are described with pedagogical concepts and lack adequate description of the scientific content. Metadata standards for research and teaching are based on different epistemologies. They cannot be reconciled through the usual technique of “crosswalks” (Godby, Young & Childress, 2004) because minimal intersections exist.

Some promising directions for facilitating collaboration through e-Research have emerged from the ADEPT and CENS research projects. Personal digital libraries offer a

middle ground between private control and public release of content. Better tools and services to manage content will improve capture, management, and preservation. Making content more shareable increases the likelihood that it will be shared. We have identified important differences between research and teaching uses of content in distributed environments, which is a first step toward building tools and services to bridge the gap. We also are identifying interactions between degrees of collaboration in research, uses of instrumentation for data collection and analysis, and sharing of data. Many of the factors influencing collaboration are subtle, nuanced, and powerful. Much more research is needed to identify individual factors and the relationships amongst them. We are just beginning to understand *how* e-Research can facilitate collaboration. The next step is to understand *why*.

ACKNOWLEDGEMENTS

ADEPT and CENS are both large collaborative projects. The education and evaluation research in ADEPT described here has been conducted since 1999 in collaboration with Anne Gilliland-Swetland 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 was funded by National Science Foundation grant no. IIS-9817432, Terence R. Smith, University of California, Santa Barbara, Principal Investigator. We have been fortunate to have the participation of many talented students, all of whom are listed as co-authors on the ADEPT papers cited herein.

The data management and education research in CENS is conducted in collaboration with William A. Sandoval and Noel Enyedy, both of the Department of Education, and Jonathan Furner of the Department of Information Studies; all three are faculty of the Graduate School of Education & Information Studies, UCLA. Students and post-doctoral fellows who have participated on the data management team to date are Stasa Milojevic, Jillian Wallis, Kalpana Shankar, and Eun Park. Jillian Wallis is now employed on the project as Data Archivist. CENS is funded by National Science Foundation Cooperative Agreement #CCR-0120778, Deborah L. Estrin, UCLA, Principal Investigator. Additional research described here also is funded by NSF award number ESI-0352572, William A. Sandoval, Principal Investigator. We have enjoyed the extensive contributions to our work made by members of the education team, including (during the period covered in this paper) Sara Terheggen, Karen Kim, Joe Wise, Kathy Griffis, Andy Wu, Kelli Millwood, and Jeff Bockert. The James San Jacinto Mountains Reserve is part of the University of California and its director, Dr. Michael Hamilton, is a Co-Principal Investigator of CENS. The braken fern project is led by Prof. Phil Rundel and Dr. Eric Graham of UCLA. The CENS seismology team is led by Prof. Paul Davis, Department of Earth and Space Sciences, UCLA, who is also a co-PI on CENS.

REFERENCES

- ADEPT/DLESE/NASA metadata (ADN)*. (2001). <http://www.dlese.org/Metadata/adn-item/index.htm> Visited: 15 December 2004
- Agre, P. E. (2003). Information and institutional change: The case of digital libraries. *Digital library use: Social practice in design and evaluation*. A. P. Bishop, N. Van House & B. P. Battenfield. Cambridge, MA, MIT Press: 219-240.

- Almes, G., Birnholtz, J. P., Hey, T., Cummings, J., Foster, I. & Spencer, B. (2004). CSCW and Cyberinfrastructure: Opportunities and Challenges (panel session). *Proceedings of the Conference on Computer Supported Cooperative Work*, Chicago, IL, Chicago, IL. Association for Computing Machinery.
<http://www.acm.org/cscw2004> Visited:
- Ancona, D., Freeston, M., Smith, T. R. & Fabrikant, S. (2002). Visual explorations for the Alexandria Digital Earth Prototype. *Joint Conference on Digital Libraries*, Portland, Oregon, Portland, Oregon. Springer-Verlag.
<http://vw.indiana.edu/visual02/Ancona.pdf> Visited:
- Arzberger, P., Schroeder, P., Beaulieu, A., Bowker, G., Casey, K., Laaksonen, L., Moorman, D., Uhler, P. & Wouters, P. (2004). An International Framework to Promote Access to Data. *Science*, **303**(5665): 1777-1778.
- Batalin, M. A., Rahimi, M., Yu, Y., Liu, D., Kansal, A., Sukhatme, G. S., Kaiser, W. J., Hansen, M., Pottie, G. J., Srivastava, M. & Estrin, D. (2004). Call and Response: Experiments in Sampling the Environment. *Call and Response: Experiments in Sampling the Environment*, Los Angeles, Los Angeles.
http://www.ee.ucla.edu/%7Ekaiser/nims/Documents/Publications/SenSys_2004_Call_and_Response.pdf Visited: 6 December 2005
- Bishop, A. P., Van House, N. & Bittenfield, B. P., Eds. (2003). *Digital library use: Social practice in design and evaluation*. Cambridge, MA, MIT Press.
- Borgman, C. L. (2000a). Digital libraries and the continuum of scholarly communication. *Journal of Documentation*, **56**: 412-430.
- Borgman, C. L. (2000b). *From Gutenberg to the Global Information Infrastructure: Access to Information in the Networked World*. Cambridge, MA, The MIT Press.
- Borgman, C. L. (2003a). Designing digital libraries for usability. *Digital library use: Social practice in design and evaluation*. A. P. Bishop, N. Van House & B. P. Bittenfield. Cambridge, MA, MIT Press: 85-118.
- Borgman, C. L. (2003b). Personal digital libraries: Creating individual spaces for innovation. *NSF Workshop on Post-Digital Libraries Initiative Directions*, Chatham, MA, Chatham, MA.
http://www.sis.pitt.edu/~dlwkschop/paper_borgman.html. Visited: 19 January 2005
- Borgman, C. L. (2004a). *E-Science, Digital Libraries, and Knowledge Communities*. Public Lecture, Dept. of Computer and Information Sciences, University of Strathclyde.
- Borgman, C. L. (2004b). *E-Science, Digital Libraries, and Knowledge Communities*. Public Lecture, Oxford Internet Institute, University of Oxford.
- Borgman, C. L. (2004c). *E-Science, E-learning, and Digital Libraries*. Seminar, Oxford Internet Institute, University of Oxford.
- Borgman, C. L. (2004d). The Interaction of Community and Individual Practices in the Design of a Digital Library. *International Symposium on Digital Libraries and Knowledge Communities in Networked Information Society*, University of Tsukuba, Tsukuba, Ibaraki, Japan., University of Tsukuba, Tsukuba, Ibaraki, Japan. University of Tsukuba.

- Borgman, C. L. (2004e). *Science, Cyberinfrastructure, and Knowledge Communities: Leveraging Scientific Data for Educational Applications*. Science and Cyberinfrastructure Session, Society for Social Studies of Science (4S).
- Borgman, C. L. (2005a). Disciplinary Differences in E-Research: An information Perspective. Keynote Talk. *First International Conference on e-Social Science*, Manchester, UK, Manchester, UK. National Center for e-Social Science. http://www.ncess.ac.uk/events/conference/programme/presentations/ncess2005_borgman.pdf
- Visited: 10 August 2005
- Borgman, C. L. (2005b). The Social Embeddness of Embedded Networked Sensing. *MIT-OII Workshop on New Approaches to Research on the Social Implications of Emerging Technologies*, University of Oxford, University of Oxford. Oxford Internet Institute. <http://www.oii.ox.ac.uk/collaboration/?rq=specialevents/20050415> Visited: 11 April 2005
- Borgman, C. L. (in press). *Scholarship in the Digital Age: Information, Infrastructure and the Internet*. Cambridge, MA, The MIT Press.
- Borgman, C. L. & Enyedy, N. (2005). Designing from Within: A Case Study of Embedded Networked Sensing. *Society for Social Studies of Science (4S)*, Pasadena, CA, Pasadena, CA.
- Borgman, C. L., Enyedy, N., Wallis, J. & Sandoval, W. A. (2006). *Scientific practices in data collection, use, and sharing: A case study of embedded networked sensing (work in progress)*. UCLA.
- Borgman, C. L., Gilliland-Swetland, A. J., Leazer, G. H., Mayer, R., Gwynn, D., Gazan, R. & Mautone, P. (2000). Evaluating digital libraries for teaching and learning in undergraduate education: a case study of the Alexandria Digital Earth Prototype (ADEPT). *Library Trends*, **49**(2): 228-250.
- Borgman, C. L., Leazer, G. H., Gilliland-Swetland, A. J. & Gazan, R. (2001). Iterative design and evaluation of a geographic digital library for university students: A case study of the Alexandria Digital Earth Prototype (ADEPT). *Research and Advanced Technology for Digital Libraries: 5th European Conference, ECDL 2001, Darmstadt, Germany, September 2001: Proceedings*, New York, New York. Springer, Lecture Notes in Computer Science.
- Borgman, C. L., Leazer, G. H., Gilliland-Swetland, A. J., Millwood, K. A., Champeny, L., Finley, J. R. & Smart, L. J. (2004). How Geography Professors Select Materials for Classroom Lectures: Implications for the Design of Digital Libraries. *Joint Conference on Digital Libraries*, Tucson, AZ, Tucson, AZ. Association for Computing Machinery.
- Borgman, C. L., Smart, L. J., Millwood, K. A., Finley, J. R., Champeny, L., Gilliland, A. J. & Leazer, G. H. (2005). Comparing faculty information seeking in teaching and research: Implications for the design of digital libraries. *Journal of the American Society for Information Science & Technology*, **56**(6): 636–657.
- Borgman, C. L., Smart, L. J., Millwood, K. A., Finley, J. R., Champeny, L., Gilliland-Swetland, A. J. & Leazer, G. H. (2004). Getting ready for class: The information seeking behavior of geography professors in support of teaching. *Annual Meeting*

- of the American Educational Research Association*, San Diego, CA, San Diego, CA. American Educational Research Association.
- Bowker, G. C. (2005). *Memory Practices in the Sciences*. Cambridge, MA, MIT Press.
- Carr, H. V. (1999). *Technology Adoption and Diffusion*. United States Air Force War College. <http://www.au.af.mil/au/awc/awcgate/innovation/adoptiondiffusion.htm> Visited:
- Case, D. O. (2002). *Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior*. San Diego, Academic Press.
- Champeny, L., Borgman, C. L., Leazer, G. H., Gilliland-Swetland, A. G., Millwood, K. A., D'Avolio, L., Finley, J. R. & Smart, L. J. (2004). Developing a digital learning environment: An evaluation of design and implementation processes. *Joint Conference on Digital Libraries*, Tucson, AZ, Tucson, AZ.
- D'Avolio, L., Borgman, C. L., Champeny, L., Leazer, G. H., Gilliland-Swetland, A. J. & Millwood, K. A. (2005). From Prototype to Deployable System: Framing the Adoption of Digital Library Services. *Annual Meeting of the American Society for Information Science & Technology*, Charlotte, NC, Charlotte, NC.
- David, P. A. (2003). The Economic Logic of 'Open Science' and the Balance between Private Property Rights and the Public Domain in Scientific Data and Information: A Primer. *The Role of the Public Domain in Scientific Data and Information*. Washington, D.C., National Academy Press. <http://siepr.stanford.edu/papers/pdf/02-30.html> Visited:
- David, P. A. (2004). *Towards a cyberinfrastructure for enhanced scientific collaboration: Providing its 'soft' foundations may be the hardest part*. Oxford Internet Institute Research Reports:University of Oxford. 23 pages. <http://www.oii.ox.ac.uk> Visited: 15 January 2005
- David, P. A. & Spence, M. (2003). *Towards Institutional Infrastructures for E-Science: The Scope of the Challenge*. Oxford Internet Institute Research Reports:University of Oxford. 92 pages. <http://www.oii.ox.ac.uk> Visited: 15 September 2003
- DialogPlus: Digital Libraries in Support of Innovative Approaches to Learning and Teaching in Geography*. (2003). National Science Foundation (U.S.) and Joint Information Systems Committee (U.K.). <http://www.dialogplus.org/> Visited: 19 November 2004
- The Digital Millennium Copyright Act of 1998* (1998). www.copyright.gov/legislation/dmca.pdf Visited: Oct. 28, 1998 Visited
- e-Science Core Programme*. (2004). Engineering and Physical Sciences Research Council (EPSRC). <http://www.epsrc.ac.uk/ResearchFunding/Programmes/e-Science/default.htm> Visited: 18 November 2004
- Ecological Metadata Language (EML)*. (2004). <http://knb.ecoinformatics.org/software/eml/> Visited: 25 November 2004
- Ellis, D. (1989). A behavioural approach to information retrieval design. *Journal of Documentation*, **45**: 171-212.
- Enyedy, N. (2003). Knowledge construction and collective practice: At the intersection of learning, talk, and social configurations in a computer-mediated mathematics classroom. *The Journal of the Learning Sciences*, **12**(3): 361-408.

- Enyedy, N. & Goldberg, J. (2004). Inquiry in interaction: Developing classroom communities for understanding through social interaction. *Journal for Research in Science Teaching.*, **41**: 862-893.
- Federation of Digital Broad-Band Seismograph Networks (FDSN). (2004). <http://www.fdsn.org/> Visited: 25 November 2004
- Finholt, T. A. (2002). Collaboratories. *Annual Review of Information Science and Technology*. B. Cronin. Medford, NJ, Information Today. **36**: 73-107.
- Foster, N. F. & Gibbons, S. (2005). Understanding Faculty to Improve Content Recruitment for Institutional Repositories. *D-Lib Magazine*, **11**(1). <http://www.dlib.org/dlib/january05/foster/01foster.html> Visited:
- Fry, J. (2003). *The Cultural Shaping of Scholarly Communication with Academic Specialisms* University of Brighton. Brighton, UK. Pages.
- Gazan, R., Leazer, G. H., Borgman, C. L., Gilliland-Swetland, A. J., Smart, L. J., Ancona, D. & Nilsson, R. M. (2003). Use Scenarios in the Development of the Alexandria Digital Earth Prototype (ADEPT). *Annual Meeting of the American Society for Information Science & Technology*, Long Beach, CA, Long Beach, CA. Information Today.
- Gilliland-Swetland, A. J. & Leazer, G. H. (2001). Iscapes: digital learning environments for the promotion of scientific thinking by undergraduates in geography. *First ACM/IEEE-CS Joint Conference on Digital Libraries*, Roanoke, Virginia, Roanoke, Virginia.
- Goble, C. & Wroe, C. (2004). The Montagues and the Capulets. *Standards and Ontologies for Functional Genomics*.
- Godby, C. J., Young, J. A. & Childress, E. (2004). A Repository of Metadata Crosswalks. *D-Lib Magazine*, **10**(12). <http://www.dlib.org/dlib/december04/godby/12godby.html> Visited:
- Hey, T. & Trefethen, A. (2003). The Data Deluge: An e-Science Perspective. *Grid Computing – Making the Global Infrastructure a Reality*, Wiley.
- Hilgartner, S. & Brandt-Rauf, S. I. (1994). Data Access, Ownership and Control: Toward Empirical Studies of Access Practices. *Knowledge*, **15**: 355-372.
- Hill, L. & Freeston, M. (2003). Introduction to georeferencing in digital libraries, Tutorial. *Joint Conference on Digital Libraries*, Houston, TX, Houston, TX. <http://www.rice.edu/jcdl03/tutorials.html#geo1> Visited:
- Hill, L., Janee, G., Dolin, R., Frew, J. & Larsgaard, M. (1999). Collection metadata solutions for digital library applications. *Journal of the American Society for Information Science*, **50**(13): 1169-1181.
- Hill, L. L., Carver, L., Larsgaard, M., Dolin, R., Smith, T. R., Frew, J. & Rae, M. A. (2000). Alexandria Digital Library: User evaluation studies and system design. *Journal of the American Society for Information Science*, **51**(3): 246-259.
- Hill, L. L. & Janee, G. (2004). The Alexandria Digital Library Project: metadata development and use. *Metadata in Practice: a Work in Progress*. D. Hillman & E. Westbrooks. Chicago, American Library Association.
- Hovy, E. (2003). Using an ontology to simplify data access. *Communications of the ACM*, **46**(1): 47-49.
- Hughes, T. P. (2004). *Human-Built World: How to Think about Technology and Culture*. Chicago, University of Chicago Press.

- IEEE Learning Object Metadata (LOM)*. (2004). <http://ltsc.ieee.org/wg12/> Visited: 15 December 2004
- Incorporated Research Institutions for Seismology (IRIS)*. (2004). <http://www.iris.edu> Visited: 25 November 2004
- James San Jacinto Mountains Reserve*. (2004). University of California. <http://www.jamesreserve.edu/> Visited: 30 November 2004
- Janee, G. & Frew, J. (2002). The ADEPT digital library architecture. *Second ACM/IEEE-CS Joint Conference on Digital Libraries*. Portland, OR, ACM Press: 342-350.
- Janee, G., Frew, J. & Hill, L. L. (2004). Issues in georeferenced digital libraries. *D-Lib Magazine*, **10**(5). <http://www.dlib.org> Visited: 19 January 2005
- Kline, R. & Pinch, T. (1999). The social construction of technology. *The Social Shaping of Technology*. D. MacKenzie & J. Wajcman. Maidenhead, Open University Press.
- Knowledge Network for Biocomplexity (KNB)*. (2004). University of California, Santa Barbara. <http://knb.ecoinformatics.org/index.jsp> Visited:
- Kubicek, H. & Dutton, W. H. (1997). The social shaping of information superhighways: An introduction. *The Social Shaping of Information Superhighways: European and American Roads to the Information Society*. H. Kubicek, W. H. Dutton & R. Williams. New York, Campus Verlag, St. Martin's Press: 9-44.
- Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, **42**(5): 361-371.
- Kwa, C. (2005). Local ecologies and global science: Discourses and strategies of the International Geosphere-Biosphere Programme. *Social Studies of Science*, **35**(6): 923-950.
- Leazer, G. H., Gilliland-Swetland, A. J., Borgman, C. L. & Mayer, R. E. (2000). Classroom evaluation of the Alexandria Digital Earth Prototype (ADEPT). *ASIS 2000 Proceedings of the 63rd ASIS Annual Meeting, Chicago, IL, November 12-16, 2000*. D. Kraft. Medford, NJ, Information Today. **37**: 334-340.
- Lord, P. & Macdonald, A. (2003). *e-Science Curation Report. Data curation for e-science in the UK: an audit to establish requirements for future curation and provision*. JISC Committee for the Support of Research. 85 pages. http://www.jisc.ac.uk/uploaded_documents/e-scienceReportFinal.pdf Visited: 1 November 2004
- Lynch, C. A. (2003). Colliding with the real world: heresies and unexplored questions about audience, economics, and control of digital libraries. *Digital library use: Social practice in design and evaluation*. A. P. Bishop, N. Van House & B. P. Buttenfield. Cambridge, MA, MIT Press: 191-216.
- MacKenzie, D. & Wajcman, J., Eds. (1999). *The Social Shaping of Technology*. Maidenhead, Open University Press.
- Marchionini, G., Plaisant, C. & Komlodi, A. (2003). The people in digital libraries: Multifaceted approaches to assessing needs and impact. *Digital library use: Social practice in design and evaluation*. A. P. Bishop, N. Van House & B. P. Buttenfield. Cambridge, MA, MIT Press: 119-160.

- Mayer, R. E., Smith, T. R., Borgman, C. L. & Smart, L. J. (2002). Digital libraries as instructional aids for knowledge construction. *Educational Technology*, **42**(6): 38-42.
- Meadows, A. J. (1998). *Communicating Research*. San Diego, Academic Press.
- National Center for Atmospheric Research. (2004). <http://www.ncar.ucar.edu/ncar/> Visited: 2 December 2005
- NIMS: Networked Infomechanical Systems. (2005). <http://www.cens.ucla.edu/portal/nims/> Visited: 3 March 2005
- Olson, G. M. & Olson, J. S. (2000). Distance matters. *Human-Computer Interaction*, **15**(2-3): 139-178.
- Pritchard, S. M., Carver, L. & Anand, S. (2004). *Collaboration for knowledge management and campus informatics*. University of California, Santa Barbara. 38. http://www.library.ucsb.edu/informatics/informatics/documents/UCSB_Campus_Informatics_Project_Report.pdf Visited: 14 November 2005
- Ribes, D. & Bowker, G. C. (2004). Ontologies and the machinery of difference: Towards a sociology of knowledge representation. *Society for the Social Studies of Science*, Paris, Paris.
- Rogers, E. M. (1995). *Diffusion of Innovations*. New York, The Free Press.
- Sandoval, W. A. & Borgman, C. L. (2004-2008). *CENS Education cyberInfrastructure (CENSEI), NSF Award ESI-0352572*. National Science Foundation, Directorate for Education and Human Resources. <http://www.cens.ucla.edu> Visited: 20 May 2005
- Sandoval, W. A. & Reiser, B. J. (2003). Explanation-driven inquiry: integrating conceptual and epistemic supports for science inquiry. *Science Education*, **87**: 1-29.
- Shankar, K. (2003). *Scientific data archiving: the state of the art in information, data, and metadata management*. <http://cens.ucla.edu/Education/index.html> Visited: 19 January 2005
- Smith, T. R., Ancona, D., Buchel, O., Freeston, M., Heller, W., Nottrott, R., Tierney, T. & Ushakov, A. (2003). The ADEPT concept-based digital learning environment. *European Conference on Digital Libraries*, Trondheim, Norway, Trondheim, Norway. Springer.
- Smith, T. R. & Zheng, M. (2002). ADEPT Knowledge Organization Team. Structured models of scientific concepts as a basis for organizing, accessing and using learning materials. *Joint Conference on Digital Libraries*, Portland, Oregon, Portland, Oregon. ACM Press.
- Sonnenwald, D. H. (2006). Scientific collaboration: Challenges and solutions. *Annual Review of Information Science and Technology*. B. Cronin. Medford, NJ, Information Today. **40**: in press.
- Standard for the Exchange of Earthquake Data (SEED). (2004). Federation of Digital Seismographic Networks (FDSN). http://www.iris.edu/manuals/SEED_chpt1.htm Visited: 25 November 2004
- Star, S. L., Bowker, G. C. & Neumann, L. J. (2003). Transparency beyond the individual level of scale: Convergence between information artifacts and communities of practice. *Digital library use: Social practice in design and evaluation*. A. P. Bishop, N. Van House & B. P. Battenfield. Cambridge, MA, MIT Press: 241-270.

- Sutton, C. (2003). *UCLA Develops Mobile Sensing System for Enriched Monitoring of the Environment*. UCLA. <http://www.engineer.ucla.edu/stories/2003/nims.htm> Visited:
- Svenonius, E. (2000). *The Intellectual Foundation of Information Organization*. Cambridge, MA, The MIT Press.
- Van House, N. A. (2003). Digital libraries and collaborative knowledge construction. *Digital library use: Social practice in design and evaluation*. A. P. Bishop, N. Van House & B. P. Buttenfield. Cambridge, MA, MIT Press: 271-296.
- Van House, N. A., Bishop, A. P. & Buttenfield, B. P. (2003). Introduction: Digital libraries as sociotechnical systems. *Digital library use: Social practice in design and evaluation*. A. P. Bishop, N. Van House & B. P. Buttenfield. Cambridge, MA, MIT Press: 1-21.
- Woolgar, S. (2003). *Social Shaping Perspectives on e-Science and e-Social Science: the case for research support: A consultative study for the Economic and Social Research Council (ESRC)*. Economic and Social Research Council (ESRC). http://www.sbs.ox.ac.uk/html/faculty_area_sci_tech.asp; www.sbs.ox.ac.uk/downloads/e-SocialScience.pdf Visited:
- Wouters, P. (2004). What is the matter with e-Science? – thinking aloud about informatisation in knowledge creation. *Society for Social Studies of Science*, Paris, Paris. <http://www.nerdi.knaw.nl> Visited: 29 October 2004