Digital Library and Educational Initiatives at the University of Michigan

Gary M. Olson *1 and Daniel E. Atkins *2
School of Information *3
University of Michigan
Ann Arbor, Michigan USA

The new emerging digital technologies are transforming the handling of information, and as a result are changing the nature of human activities and organizations. There are a variety of visions of what the consequences will be (a recent sampling would include Lanham, 1993; Mitchell, 1995; Negroponte, 1995; Tapscott, 1996; or the contents of any issue of *Wired* magazine), everyone is in agreement that the changes will be profound and wide reaching.

At the University of Michigan we are involved in two broad responses to these changes. First, we have undertaken a number of research projects that explore new technologies and assess their effects on human activity. Second, we are creating new educational programs so we can prepare students for interesting and useful careers in these arenas. These two sets of activities are intertwined both via the people involved (faculty and students) and the ideas embodied in our approach to both. In this paper we describe these two areas of work, and illustrate our approach in the process.

Digital Libraries and Collaboratories

The widespread dissemination of information in digital form has profound implications for human activities and institutions. It is estimated that over 90% of all new data and information is now originally created in digital form: print-on-paper originates in word processors; professional audio is fully digital, and widespread digital video is on the way. Digital information, physically represented by electrons, photons, or magnetic fields, has vastly different properties than information stored on stone, papyrus, paper, or film. Digital information can be moved at nearly the speed of light; it can be stored at atomic scales of density; it can be copied perfectly over and over; and it has the property of "digital coherence."

Digital coherence refers to the fact that traditionally separate media such as text, image, audio, video, citations, and algorithmic information can now all be included in the same compound document. As we can already see in the crude beginnings of the World Wide Web, we are beginning to invent new genres – new socially agreed upon ways to create, disseminate, preserve, and use data-information-knowledge. These new types of documents may contain dynamic objects such as audio, video, or real-time data display and therefore have no full representation on paper. We are already well into the world of hybrid information - the coexistence of both print-on-paper and digital information resources to meet human needs.

These changes in how information is created, represented, disseminated, and retrieved is leading to new ways of organizing human activities. We have been exploring the components of a three-part vision of

new ways of supporting knowledge work. Figure 1 shows this vision schematically. The three components are (1) the support of people-to-people interactions, (2) access to stored knowledge bases in the form of digital libraries, and (3) the ability to interact with the physical world at a distance. In the following sections we describe two broad categories of research that involve these components. The first is work on digital libraries, and the second is work on collaboratories that bring all of the components together.

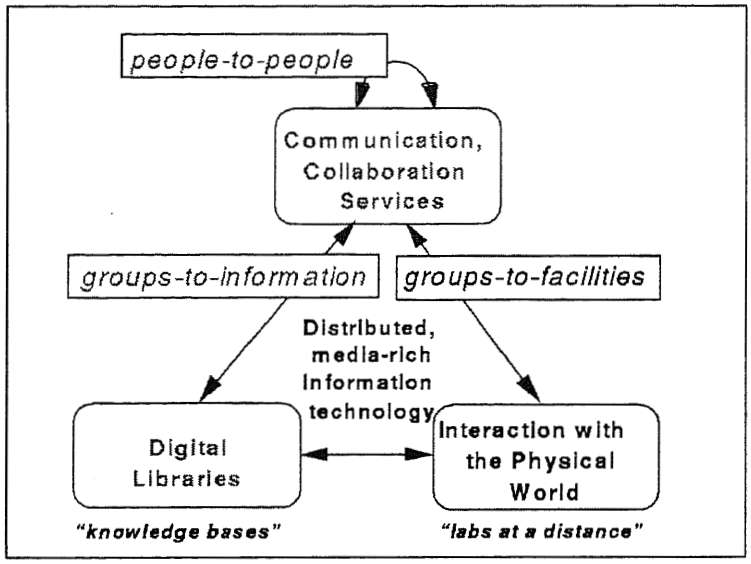


Figure 1. A vision of the new ways of supporting knowledge work.

Digital Libraries

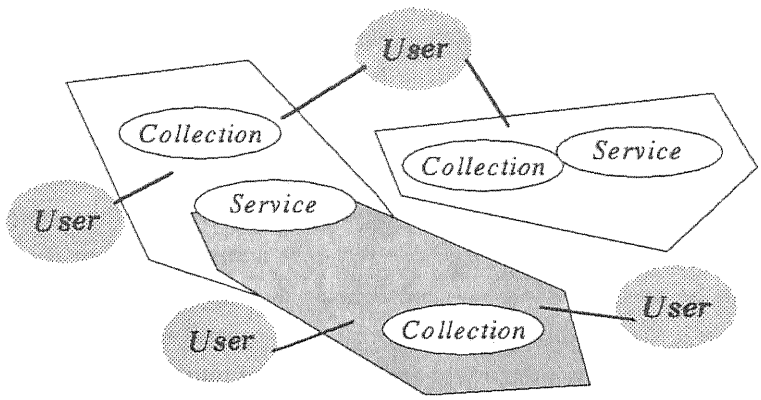
A multi-disciplinary team of faculty and students with roots in information and library studies, computer science, psychology, economics, and the university libraries was successful in winning a $4 million award through the Digital Library Initiative funded by the National Science Foundation (NSF), the Advanced Projects Research Agency (ARPA), and the National Aeronautics and Space Administration (NASA). An additional $4.5 million is being provided through industrial contributions and cost sharing. This project is creating an architecture and working prototype intended to broker access among millions of digital collections, services, and people in ways which are easy, relevant, and timely for humans to use. The testbed is focusing on collections in the areas of earth and space science and supporting users in public libraries, universities, and high school science education.

Digital library technology is changing at a rapid pace, making it inadvisable to fix specific designs for user interfaces, search engines, or collection structures. Instead, UMDL defines an architecture for the system that performs generic management tasks, such as allocating resources to user-requested tasks and brokering connections among collaborating system modules. UMDL specifies a language and protocol for communicating about informational or processing capabilities and interests, so that users and collections can be brought together appropriately. At that point, what they do with the connection is not determined at the level of our architecture.

Distributing tasks to large numbers of specialized, fine-grained agents promotes modularity, flexibility, and incrementality, as new services can come and go without disturbing the overall system. Limiting the

complexity of any individual agent simplifies control, promotes reusability, and provides a framework for talking the interoperability problems. Each module or agent supports a highly specialized library task, as well as a highly generalized communication interface. The advantage of this combination is that the task competence made possible by specialization can be brought to bear on a wide variety of situations, through flexible interaction with other agents.

For example, in formulating a query, it may be useful to generate synonyms for specified terms, in order to produce query variants that are likely to find relevant documents. In an alternate search scheme, synonyms might be useful in assessing how well some text matches an already formulated query. By encapsulating a general synonym service within a specialized thesaurus agent, we can provide the component functionality without committing to how it is to be employed system-wide.



UMDL is addressing many difficult issues of scale, decentralization, interoperability, and resource allocation. The approach has been to define very general mechanisms, and then test them with specific instances of software agents and protocols that use these mechanisms to provide library services. Atkins et al (in press) describe the project in greater detail.

UMDL is only the largest of a series of digital library projects at Michigan. As described in Atkins (1995), other projects are digitizing complete collections of major scholarly journals in humanities and social science (JSTOR, Journal Storage Project),creating a prototype of a database of art and architectural images (Art Image Browser), exploring the services and educational requirements for community information networks, building a widely-used Internet Public Library, and creating partnerships with colleagues in the arts and humanities, and local museums and schools to create pilot projects to increase and enhance accessibility to cultural heritage materials. Other digital library services are also available through a collaboration between the University Libraries, the Information Technology Division, and SI. Pointers to all of these projects are available through the School of Information home page (www.si.umich.edu).

We are also looking at how digital library technologies such as those being developed within UMDL

can be used in other domains. For example, fully electronic commerce is emerging as businesses begin to buy and sell their goods and services on digital networks. The variety of these goods and services offered will be enormous, since much existing commerce will move to the electronic medium, while, and perhaps more importantly, entirely new types of goods and services will emerge. We fully expect that the organization of business will radically change in response to both these goods and services and also in the way they are vended.

One specific area we have thought about is team formation. Ad hoc teams are an important building block of modern organizations (e.g., Waterman, 1990; Katzenbach & Smith, 1993). Instead of large, vertically integrated corporations, the trend is toward agile organizations that can easily change capabilities, products, and strategies as markets demand. The core constituents of agile organizations are dynamically organized teams. Team members come from both within and outside the agile enterprise. Within a team, each member provides some capability needed for a particular product or task. When coordinated activity is no longer in the interest of the parties, the teams are dismantled.

Most current business practice assumes that teams are formed based on experience. Typically, experience is built through relationships spanning years. Through these relationships, it is possible to fully determine the capability of organizations, and their credibility. Electronic commerce, however, flies in the face of current practice. Potential collaborators—corporations on the network—are known only through *electronic storefronts*. In electronic commerce, it is neither likely nor desirable that traditional team-forming methods be used. Rather, we need entirely new mechanisms, and the technologies being developed in UMDL seem ideally suited for this process. Of course, extensive work needs to be done to define the vocabulary for a system that would assist in the forming of teams. But this example shows that digital library technology has applications to a much wider arena.

The Collaboratory

A collaboratory is the "...combination of technology, tools and infrastructure that allow scientists to work with remote facilities and each other as if they were co-located." (Lederberg & Uncapher, 1989, p. 6) A National Research Council (1993) report defines a collaboratory as a "...center without walls, in which the nation's researchers can perform their research without regard to geographical location – interacting with colleagues, accessing instrumentation, sharing data and computational resources [and] accessing information in digital libraries." (National Research Council, 1993, p.7) A simplified form of these definitions describes a collaboratory as the use of computing and communication technology to achieve the enhanced access to colleagues and instruments provided by a shared physical location, but in a domain where potential collaborations are not constrained by temporal or geographic barriers.

The collaboratory idea appeared as scientists recognized the potential represented by expanding national and international computer networks. The first explicit discussion of collaboratories was at an NSF-sponsored workshop in 1989 convened by Joshua Lederberg and Keith Uncapher. This workshop gave the collaboratory concept visibility within the NSF and other relevant national scientific communities. The report of the workshop outlined a number of specific research priorities: enabling infrastructure to support collaboratories, construction of collaboratory testbeds in various scientific disciplines, and studies of the process of collaboration and the use of these testbeds by scientists (Lederberg & Uncapher, 1989). One outcome of the 1989 workshop was a series of further workshops in 1993 sponsored by the Computer Science and Telecommunications Board of the National Research Council to explore the

feasibility and utility of collaboratories for three disciplines: molecular biology, physical oceanography, and space physics. These fields were chosen for their heterogeneity in size, style of research, technical sophistication, and traditional sources of support. An important outcome of this activity was the NRC's (1993) report *National Collaboratories: Applying Information Technology for Scientific Research.* The report called for substantial support to develop, refine, and evaluate the collaboratory concept in realistic settings. Although not of the scope called for in the NRC report, at present there are numerous collaboratory-like facilities, mostly started after 1992. Many of these efforts have been funded in full or partially by grants from the High Performance Computing and Communication (HPCC) initiative; a $3 billion effort to produce the National Research and Education Network (NREN). The most mature projects are in atmospheric and space science, biology, medicine, and physics.

We have embarked on an ambitious series of collaboratory projects at Michigan. The oldest is the Upper Atmospheric Research Collaboratory (UARC). UARC supports a distributed community of space physicists by providing them with real-time control of instruments in Greenland, the ability to communicate with their colleagues over shared real-time data, and access to archived data. In UARC a half dozen instruments transmit data over the Internet to specially designed display programs. Scientists at ten sites around the world can view these displays by running the data viewing programs on local machines. A simple text-based "chat" window allows the scientists to share reactions with each other about the phenomena they are observing and to send instructions to the site crew in Greenland. Simple collaboration support such as data annotations, window sharing, and telepointing is also provided to assist the real-time interactions. In addition to the scientific research, communication and collaboration patterns of the scientists are being studied. A recent overview is in Olson et. al (in press).

While UARC is the oldest, it is by no means the only collaboratory project. A similar project is developing a testbed which will support collaboration in using medical images for diagnosis. The collaboration is between the primary health care physicians and health care specialists. Other projects recently initiated are exploring the application of "virtual co-location" technology to support automotive design teams on a global basis and the feasibility of a global collaboratory to support research in human virology. Once again, more details about all of these projects are available on the World Wide Web at www.si.umich.edu.

CREW – Our own collaboratory

We have created a collaboratory of our own, called the Collaboratory for Research on Electronic Work, or CREW, to support our wide range of projects exploring the new digital cultures. Sixteen University of Michigan faculty and a number of graduate students, postdoctoral fellows, visiting scholars, and support staff comprise CREW, and there is an extensive array of projects supported by both the federal government and industry. CREW has several international projects as well with partners in Europe and Asia.

The New School

The emergence of digital culture also creates new educational needs. We are in the process of creating a new school, called the School of Information, to support the educational needs of new generations of students. This school is the result of a merging of several streams of activity.

The first stream is the former School of Information and Library Studies. Formal education in librarianship at the University of Michigan began about 1926 with the creation of the Department of Library Science under the leadership of the Director of Libraries, William Warner Bishop. Dr. Bishop was motivated by enlightened self-interest to professionalize librarianship both here and elsewhere. Natural growth and demand for graduates led to the next step in the late 1960's. After a blue ribbon outside review in 1969, the Department became the School of Library Science with Russell Bidlack as the first dean.

About two decades later, in 1986, as part of a national trend to examine "library education," the faculty acknowledged a broader field of endeavor for the School by voting to change the name to the School of Information and Library Studies (SILS). Other similar schools also adopted broader names, some dropping explicit reference to "library" during this time.

In 1992, President James Duderstadt and Provost Gilbert Whitaker, of the University of Michigan, decided that SILS should take an even bigger step. They offered the School the opportunity to boldly move to educate information professionals for leadership in the new age of knowledge. They perceived both great challenges and opportunities as we move into an era in which quality of life is increasingly dependent upon knowledge, and digital technology is radically changing how this knowledge is created, disseminated, used, and preserved. Although the need for timely access to knowledge and information has never been greater, the institutions, systems and processes by which humans will meet these needs are rapidly changing. We were given, therefore, the dual challenge to not only re-invent ILS education, but also to help re-invent libraries.

As these recent changes were occurring at SILS, another group of faculty, many of whom were already participating in the various digital library and collaboratory projects described above, began discussing the need for a broad-based educational program that combined training in technology and human systems. Over and over we have seen that the key issues with digital technologies are not so much the technologies themselves but how they fit into human organizations and practices. Repeatedly organizations have made costly mistakes in adopting expensive technologies without proper up-front examination of the people and organizations that would use them (Landauer, 1995). Yet there is a vast body of knowledge about human cognitive, social, and organizational processes, and a large experientially based collection of design methods for carrying out user-centered design (Norman & Draper, 1986; Olson & Moran, 1996). This convinced us of the need to train people with the proper blend of technical skills, behavioral science knowledge, and design methods to make a difference.

This blend of skills is of course much the same set needed to train effective new age digital librarians. So in the spring of 1995 about a dozen faculty from various places at the University of Michigan (e.g., psychology, political science, economics, computer science, business administration) began in earnest talking with the existing SILS faculty about creating a new school to serve this wider range of needs. We have been making steady progress toward this goal, working out the details of curriculum, specific faculty appointments, and resources to create this new venture. This winter we anticipate receiving formal approval from the University to rename SILS the School of Information. The new School will focus initially on graduate studies at both the Masters and Doctoral levels. We anticipate having programs in librarianship, archives, human-computer interaction, information systems management, future systems architecture, and several other areas. We will also put together joint programs with other schools and colleges at Michigan to create programs in a wide range of specific domains. We expect it will take

several years to get all of these program up and running, but at least several of these will be in place by the fall of 1996 as the new era of education at the School of Information begins.

We are not only changing and enlarging the content of the school, but also its methods. We expect to be extensively involved in experimentation with new ways of learning. Much of our curriculum planning has focused on what all of these areas have in common, and we expect to have all students in the school share some common experiences. We also expect to have a project focus in much of our professional education. We also intend to have an active academic outreach program, in which we offer alumni and others who are not currently full-time students a range of educational opportunities. In some of our outreach we expect to use new distance learning technologies.

Summary

We at the University of Michigan are committed to an ambitious agenda of research, education, and outreach at the intersection of the human and technical sciences. Our projects on digital libraries and collaboratories are indicative of the kind of culture we are building: scholars from a number of disciplines coming together with partners outside of the University to explore how emerging digital technologies can be used for human good. This requires new ways of thinking, new kinds of alliances, and new varieties of human resources. Our new School of Information will be at the heart of these developments.

References

Atkins, D.E. (1995) The future of libraries and library schools. Paper presented at the Kanazawa Institute of Technology International Round Table for Library and Information Science, November 1995.

Atkins, Wellman, Birmingham, Soloway, et. al (in press) Building the University of Michigan Digital Library: Interacting Software Agents in Support of Inquiry-Based Education. *IEEE Computer*

Katzenbach, J.R., & Smith, D.K. (1993) *The wisdom of teams.* Cambridge, MA: Harvard Business School Press.

Landauer, T.K. (1995) *The trouble with computers.* Cambridge, MA: MIT Press.

Lanham, R.A. (1993) *The electronic word: Democracy, technology, and the arts.* Chicago: University of Chicago Press.

Lederberg, J., & Uncapher, K. (1989). *Towards a national collaboratory: Report of an invitational workshop at the Rockefeller University.* Washington, D.C.: National Science Foundation, Directorate for Computer and Information Science.

Mitchell, W.J. (1995) *City of bits: Space, place, and the infobahn.* Cambridge, MA: MIT Press.

National Research Council (1993). *National collaboratories: Applying information technology for scientific research.* Washington, D.C.: National Academy Press.

Negroponte, N. (1995) *Being digital.* New York: Knopf.

Norman, D.A., & Draper, S.W. (1986) *User centered system design.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Olson, G.M., Atkins, D.E., Clauer, R ., Weymouth, T., Finholt, T.A., Prakash, A., Rasmussen, C., & Jahanian, F. (in press). The Upper Atmospheric Research Collaboratory (UARC): Technology to support distributed team science. In G.M. Olson, J. Smith & T. Malone (Eds.), *Coordination theory and collaboration technology.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Olson, J.S., & Moran, T.P. (1996) Mapping the method muddle: Guidance in using methods for user interface design. In M. Rudisill, C. Lewis, P.B. Polson & T.D. McKay (Eds.), *Human-computer interface design*: San Francisco: Morgan Kaufmann.

Tapscott, D. (1996) *The digital economy: Promise and peril in the age of networked intelligence.* New York: McGraw-Hill.

Waterman, R.H., Jr. (1990) *Adhocracy.* New York: W.W. Norton.

*1 gmo@umich.edu

*2 atkins@umich.edu

*3 www.si.umich.edu