The role of digital libraries in the transformation of society: A perspective from the United States

Yi-Tzuu Chien*
School of Library and Information Science
University of Tsukuba
ytchien@slis.tsukuba.ac.jp

Abstract

Innovation in technology, testbeds, and tools development marks the first decade of digital library research. As an emerging field, this is a normal and expected path of growth. The vision of the libraries of the future, however, points to a role of helping to transform the society in virtually all aspects of life. This paper discusses the many challenges of the information society and argues for a re-orientation of the future research in order to facilitate that transformation. While the perspective presented here is mainly based on activities in the United States, future directions call for efforts across national, cultural borders.

Key words: digital libraries, societal challenges, transformation process, research agenda, local vs. global reach

As a field of study, Digital libraries research did not formally begin until about a decade ago, though many of the ingredients and related technologies were long in place then. In the early 1990s, the U.S. government launched a joint agency research program led by the National Science Foundation (NSF), along with the National Aeronautics and Space Administration (NASA) and the Defense Advanced Research Projects Agency (DARPA). Much of the academic research efforts in the U.S. under the auspices of that initiative and other efforts that followed around the world shared a "vision" for the library of the future.

The First Decade: from vision to technology innovation

This articulated vision for a "Digital Library" has the following characteristics:

- It is a distributed knowledge network, connecting information, technologies, and people.
- It is a set of enabling multi-media technologies transforming collections of materials from books, music, videos, museum artifacts and other content materials into digital information resources that are easily accessible.
- It is a set of new information services that enrich the society through improved capabilities in public education, health care, commerce, transportation, and the management of government among others.
- It is a library without walls, open 24 hours a day and is accessible where the network is.
- A digital library is simultaneously personal, public, and ubiquitous.

---

* The author is currently a Visiting Researcher in the School of Library and Information Science at the University of Tsukuba. He is a Vice President for Research with the World Technology Evaluation Center, Inc. in Baltimore, Maryland, USA; http://www.wtec.org.

Simple yet encompassing, this vision remains valid as a shared goal for the research community even a decade later today. In the span of the ten years, initially fueled by the joint-agency program in the U.S. and later by concerted efforts world wide, the research in digital libraries has taken on a deliberate path of growth, diversity, and transition. During the first phase of the research initiative, known as DLI-1 from 1994 - 1998 [1], the focus was on the core technologies. This strategy is to deal with the technical issues and infrastructures necessary for building digital libraries as a knowledge network, enabler of multi-media technologies, new and novel information services, and a ubiquitous platform for public information access. The program selected six major university-industry research consortia, exploring a variety of technical issues at the intersection of computing, communications and content management for next generation knowledge networks. These included technologies for searching digital video materials directly by content, by integrating speech, text, and image processing techniques; for managing multiplicity of information infrastructures and user interfaces; for enabling and facilitating collaborative work at a distance with novel software agents; for creating user-centered, scaleable spatial information systems; for handling heterogeneity and interoperation between dispersed knowledge networks; and for scaling digital repositories across multiple knowledge domains.

Transition to Human-centered Perspective

During the second phase (DLI-2, 1998 –2003) of the initiative [2], the goal and scope of the research were extended considerably when additional U.S. Federal agencies joined the sponsorship: The National Library of Medicine, the Library of Congress, and the National Endowment for Humanities. These additions helped strengthen the program's resources and broadened its scope of activity. Programmatically, DLI-2 began a shift of research from a pure technological focus in phase 1 to a more human-centered systems perspective. The goal is to have the research projects collectively address the entire life of digital libraries: information and knowledge creation in the form of new information objects and collections; new technologies for access, discovery, search and retrieval; enduring resources to allow long term use and continuous embellishment; and archival and preservation strategies and tools. This shift of emphasis reflects first of all the mission requirements of the new sponsoring agencies. It is also very much embraced by the research community as part of the maturing process of digital library research. Table 1 summarizes the many ways such a shift has been designed into the program: from research itself, development of testbeds, content and collections, infrastructures, to the overall societal context of applications.

Table 1. Transition from DLI-1 to DLI-2

Component	DLI-1 (1994 – 1998)	DLI-2 (1998 – 2003)
Research	• Foundations of digital content management in networked environments • Broad topics in computer and information sciences	• Refined agenda informed by DLI-1 and related efforts • Topics in step with convergence of computing, communications, and content technologies
Testbeds	• For Digital Library Research	• For Digital Library Research with added emphasis on interoperability, technology integration, and applications
Content/Collections	• Donated from industrial research partners and other private	• Donated from research partners and project sponsors

	organizations	• Increased emphasis on content development, metadata, type standards, intellectual property management
Infrastructure	• Limited operational capability • Build on first generation Internet	• Operational digital libraries with collections of value to domain and other "communities" users • Build on Next Generation Internet
Context	• Primarily for fundamental technology development and evaluation	• Understanding digital libraries in domain and social contexts • Digital Libraries as human-centered systems

A new component of DLI-2 is to initiate collaborative research projects with the international communities. This is a very important addition, as many of the digital libraries research and developments efforts outside of the U.S., notably in Japan and the U.K, started out with a concerted focus on society related applications beyond pure technology innovation. This international dimension allowed the NSF led DLI-2 to reset its priorities, making the transition to a human-centered focus along with its research partners worldwide.

Collective Accomplishments

Although there is not much formal evaluation of the technical achievements associated with the joint-agency digital libraries initiatives, we can provide a qualitative statement about the collective work by the research community at large, including all the performers in or outside the U.S. and not just those directly funded under the program. Among the key accomplishments are:

- Advances in scientific knowledge for content creation, store, search and retrieval. This includes new ideas and methodologies that area applicable across a wide range of media technologies, such as text, image, voice, as well as novel user interfaces.
- The development of scaleable testbeds and software tools for experimental research in a variety of knowledge domains. These are improvements in DL infrastructures that would be hard to build otherwise without the joint-agency initiatives.
- Technology innovation and integration. This includes many novel ideas and techniques developed by individual projects or research groups that offer business opportunities if the right financial arrangements are brought forth. Less innovative but equally important is the integration of technologies, especially those involving legacy software and databases.
- Role in the development of hundreds of operational digital libraries worldwide. Many of these operations involve institutions far beyond traditional local or national libraries that house art collections or museums, and the online or chive publications for professional organizations, etc.
- A hand in the transformation of traditional library science schools into programs that offer modern digital curricula for training the future librarians as knowledge workers.
- Role in the creation of many DL byproducts, including many technical publications and literally countless workshops and conferences held every year to report on the latest advances.

Digital Challenges facing the Information Society

As digital libraries research makes its inroads along the technology path, there is a lack of visible progress on the side of economic and social benefits often envisioned by the research community. First, while, in ten years, digital library has become a household name, its impacts on both the economy and the society have been limited. It is true that DL research has generated excitement and occasional successes associated with the media industry (e.g., Google, Inc. in the search engine arena), by and large, it has not led to a "billion-dollar" industry as in many other computer and communications businesses (e.g., database, graphics, object-oriented languages, RISC processors, etc.), after similar research investment. This economic outcome, however, is likely to materialize if a combination of public and private research investments, coupled with sound government policies continues in the next decade. Previous studies [10] have shown that a 20 to 30 year time span is needed for a new technology to make it to the main-stream business from its initial conception. Digital libraries will have its turn at the right time.

On the other hand, digital libraries as a field of research must now ask the question where it is likely to go in relationship to society of the information age. A clear, or even perceived, lack of tangible society benefits would certainly dampen the shared vision as we have articulated all along during the course of past ten years. More importantly, at least in the United States, this will certainly weaken the arguments for more public investment in further research. Further, the information society faces many social challenges that technology advances alone would not be sufficient, unless they are properly channeled to bring in simultaneous social benefits. Examples of excessive or unused technology abound. All we need is witness the telecommunications collapse as of late in many parts of the world. And computer and communications technologies are not the only ones that pose such challenges to the society.

Take "information glut" as another example. Recent studies have shown that we are producing significantly far more bits than we are consuming in the digital world. Lyman and Varian [3] have shown that we collectively produce a grand total of approximately 2 million terabytes (from all media) in a year (1999 data). However, an average American household could only consume (by reading, watching TV, listening to music, etc.) about 3 terabytes a year, a small fraction of that the grand total produced. What does this all mean? This means in part we are becoming increasingly more "illiterate" as we fail to find and use digital information resources for our daily life. This is a serious challenge to our society, as we face the problem of "digital literacy" that goes beyond the traditionally required ability to read and write. In the information society, citizens are expected to find information quickly and efficiently and to make the judgment about the information they receive. In theory, computer and network tools such as search engines would help, but their indexed web coverage represents only a small fraction of the information resources available. This is one of the most pressing challenges in digital library research and a huge opportunity for it to be a part of the information-glut solution.

Digital literary is one that only sits at the top of a long list of many critical challenges that face the information society. Closely related are the problem of distance and life-long learning and education in the networked environment. The increasing disconnect of certain population groups deprived of the opportunities of access to computing, network and information resources – the problem of "digital divide" – created by the disparity in wealth, racial backgrounds, language and culture backgrounds, physical limitations, etc. Networked health care information and service delivery is another critical challenge. Also rising on the societal calendar is the access to government information and doing business and making transactions with government agencies at all levels on the network, and allowing citizens to take part in government business and policy making more directly and effectively. Ultimately, we are looking at the challenges of knowledge communities, both regionally and on a global basis. These are all areas that digital library research should have its agenda for making a serious contribution to in the next ten years. Table

2 summarizes a view of what the technical and policy issues are and where DL can seize as the new opportunities for research.

Table 2.   Challenges, Issues, and Opportunities for Future DLs

Challenges	Key Issues	Research Opportunities
Digital Literacy	Fluency in digital skills: content search, evolution, judgment	Digital libraries; Search engines; Information resources discovery; Knowledge assembly; Metadata Tools
Digital Divide	Disconnect between the haves and have-nots; universal access to digital resources for all	Multi-channel distribution of resources; Multi-modality communication media
e - Government	Universal access to government information, services, products; Improved citizen's participation in government; collaboration between public and private sectors	Friendly human-computer interfaces; economy of scale and scope across knowledge domains or agency missions
e - Education	Networking schools; Training of Next generation IT workforce; Distance and life-long learning	Learning technologies; online safety and protection of digital contents; integration of learning and entertainment
Networking health care	Networked medical services; home and self health care; deploying and improving IT technologies in medical research and applications	Online sensing, imaging and other diagnostic tools; In silica and virtual experiments; Network availability (bandwidth, latency), ubiquity
e - Commerce	Privacy; Trust; Reliability; Mobility; Convenience; Internet Taxation and regulations	Cyber security; Integrating with legacy and existing systems; e-payment
e- Community	Work and play; Group and public Infrastructure; development of social capital; virtual "third places"; resource-sharing	Collaboration technology; Public policy; Internet as "public space"
Global Reach	Internet governance and control; Rules for collaboration vs. competition; Interaction across national and cultural borders	Cross-culture content development; Multi-language interfaces

Research Agenda for the transformation

Meeting these digital challenges would certainly require a concerted effort on the part of many players, including the researchers, industries, Government at all levels, and most importantly the societal communities. This effort needs to include a rethinking of how to channel the technological advances, energies and resources of DL research into tangible society benefits – a transformational process that enables change in the everyday activity of people and organizations in the information society. Cliff Lynch in a recent EDUCAUSE Review article [7] talks about this concept. He describes the forty years of libraries and information technology as a period of progression "from automation to transformation". While Lunch's focus is on the transformational context of the libraries themselves, the concept of transformation applies broadly to connections of digital libraries to societal contexts. This concept is further illustrated by the U.S. President's Information Technology Advisory Committee in a series of reports on strategies

of furthering the research investment in digital libraries and information technologies [8,9]. The underlying argument for the investment strategy is indeed to accelerate the "transformation" of society in virtually every aspect of life: from the way people communicate, learn, work and play, do commerce, get health care, access government information and services, and so on. Table 3 summarizes the ten most essential areas of life in which such a transformation would be enabled and the exemplary roles digital libraries and information technologies could play in these transformations.

Table 3. Digital Libraries and Information Technology in the Transformation Process

Transforming The Way We	Technical Challenges for Digital Libraries Research (Examples)	Benefits from Digital Libraries Technologies (Examples)
1. Communicate	• Improved human interface w/ computers • Global DL connections	• Users of Internet can communicate in any language worldwide
2. Deal with Information	• Improve data access methods • Policy in dealing w/ e-distribution of information	• Ubiquitous access to information from different networked media: books, journals, TV broadcast
3. Learn	• Scaleable computing infrastructure and data repositories • Software for developing educational material and learning technology	• School children world wide can access the Library of Congress, British Museums, or the Ethnology Museums in Osaka
4. Conduct Commerce	• Security, privacy and reliability for e-transactions improved • Integrating information infrastructure and physical infrastructure	• Consumers make e-transactions w/ confidence and convenience • DL infrastructures accessible or sharable for consumers use; as models adaptable for commercial transactions
5. Work	• Establish network and infrastructure for combining work w/ training • Software for group interaction and collaboration	• Workers can access job-related or training material online • A progressive workplace is synonymous to digital library
6. Practice Health Care	• Ensuring privacy of information repositories • Develop networked medical information and treatment systems	• Patients access biomedical information systems for better self assessment and making informed medical decisions • Medical procedures can be demonstrated at a distance
7. Design and Build Things	• Develop software for computer aided design • Virtual reality technologies for design and manufacture	• e-access to all engineering design archives and manufacture techniques • Cross-culture interchange becomes a key to creativity
8. Conduct Research	• Develop collaboration technology • Create globe computational grid	• Shared data and other resources over networks and across time, location, and national borders

		• All journals are available online
9. Deal with the Environment	• Computational models for climate and environmental research • Networked data systems for environmental monitoring	• DLs become critical resources for energy and environmental management • Help keep policy makers better informed
10. Conduct Government	• Improved systems for delivering sensitive data over networks • Provide education and training programs for citizens' participation in e-government	• Government services and information are available to all regardless of location, computer literacy, physical limitations • Enhanced crisis management

Source: Adapted from PITAC Reports [8,9].

Whether we succeed in making these transformations for the society is course to a large degree a function of how we direct our future research efforts as a discipline. There are many ways we can set the goal posts for future DL research, but the following three address specially the many digital challenges that face the information society we discussed.

- First, future content and media based research must take advantage of the advances in computing and networking technologies. Trends and progress in supercomputing, grid computing, network middleware, embedded systems, information appliances and the like should be factored into the research agenda of digital libraries. These advances could dramatically alter how information is stored, processed, accessed, and otherwise used and have a significant impact on DL core technologies of the future.

- Second, government and private research programs should place increased emphasis on standards and tools that serve as the "middleware" across disparate domains for knowledge access, discovery, and use. Research efforts such as those in metadata (e.g., Dublin Core [6] and other similar initiatives) should receive serious investment in both capital and human resources. Interagency support and partnerships of government and industry may be the best program model where multiple agency missions and business applications may intersect and can justify their investment more effectively.

- Third, the next generation digital library research must consider large-scale projects that involve global reach. This means, for example, content collections and software tools that cultivate cross-culture, multilingual applications (for example, see [4,5]). Further, as in most emerging areas, digital library work must also leverage local and regional resources where contents, people, experiences are likely to be richly concentrated. Leveraging all three levels of financial, physical, and people resources is a key to the creation of successful research partnerships.

Conclusions

A decade of research efforts, by government design and natural evolution, has put digital library on the science and technology map in many parts of the world. The collective impacts have been significant as well as substantial, as evidenced by the advances reported in publications and conferences that have multiplied manifolds in just a few years. There have been also some modest business successes, but the real progress in that is yet to materialize. To continue a path of growth, future digital library research must be able to demonstrate a real connection to society

benefits as originally envisioned. This paper outlines such a role by addressing the challenges that face the information society and where digital library fits in meeting these challenges.

References

1. Y. T. Chien, Digital Libraries, Knowledge Networks, and Human-centered Information Systems, Proc. of the International Symposium on Digital Library ISDL'97, Tsukuba, Japan, Nov. 1997.
2. S. Griffin, The Digital Libraries Initiative: A U.S. Federal Program of Research and Applications, Proc. of the Digital Library Workshop, Tsukuba, 2000; www.dli2.nsf.gov.
3. P. Lyman and H. Varian, How much information? An ongoing and living document which is available online. http://www.sims.berkeley.edu/research/projects/how-much-info/.
4. S. Sugimoto, W. Lee, Q. Zhao, M. Nagamori, T. Sakaguchi, and K. Tabata, An Experimental Study on a Collaborative Development of a Multilingual Subject Gateway for Public Library Users, Proc. of the International Conference on Digital Library: IT Opportunities and Challenges in the New Millennium, Beijing, China, July 2002.
5. S. Sugimoto, Helping Information Access across Languages using Simple Tools – multilingual projects at ULIS and lessons learned, to appear in Proc. of the International Conference on Asian Digital Libraries, Singapore, December 2002.
6. S. Weibel and T. Koch, The Dublin Core Metadata Initiative: mission, current activities, and future directions, DiLib Magazine, 6, 12, December 2000.
7. C. Lynch, From Automation to Transformation: forty years of libraries and information technology in higher education, EDUCAUSE Review, January/February 2000.
8. Report of the U.S. President's Information Technology Advisory Committee (PITAC), Information Technology Research: Investing in our Future, February 1999.
9. Report of PITAC's panel on digital libraries, Digital Libraries: Universal Access to Human Knowledge, February 2001.
10. Report of the National Research Council, Information Technology Research, Innovation, and E-Government, Computer Science and Telecommunications Board, 2002.