

# DIGITAL LIBRARY INITIATIVES: TECHNOLOGICAL ADVANCES

**Sarasvathy, P, Swamy, D. and Umesha, M. B.**

## **1 Introduction**

A digital library is a collection of documents in organized electronic form, available on the Internet or on CD-ROM (compact-disk read-only memory) disks. As digital libraries evolve, they begin to realize the vast opportunities available by materials that are in digital format and ways to exploit these opportunities. Thus the digital library gets strengthened and adds value by its ability to integrate materials in digital formats that may not be as well-represented, easy-to-access or effectively usable in traditional library environments.

The Benefits that Digital Libraries Bring to a Library Patron include:

1. Access to the varied collections of libraries across geographies through automated catalogs.
2. Ability to locate both digitized as well as physical versions of scholarly text and books.
3. Ability to search the content in an efficient manner whether within the library, through the Internet or through other commercial databases.
4. Ability to instantly access the relevant content from the search results by clicking through to the digitized content or find additional items of relevant interest.
5. Ability to conduct more detailed research by being able to save search results, as well as
6. Use different criteria and parameters to narrow down their search results.

## **2. Digital Library Initiative**

### **1. Digital Library Initiative – Phase 1**

1990s: rapid growth component of U.S. National Information Infrastructure spurred by Digital Library Initiative (DLI1) and equivalent programs elsewhere domestic and international conferences established, journals founded creation of a new field attempts at community building. Major Influences on DLs.

#### **1. Six projects are supported:**

1. University of California at Berkeley: Environmental Planning and Geographic Information Systems
2. University of California at Santa Barbara: The Alexandria Project: Spatially-Referenced

#### Map Information

3. Carnegie Mellon University: Informedia Digital Video Library
4. University of Illinois at Urbana-Champaign: Federating Repositories of Scientific Literature
5. University of Michigan: Intelligent Agents for Information Location
6. Stanford University: Interoperation Mechanisms Among Heterogeneous Services

### **Digital Library Initiative - Phase 2**

Digital Libraries Initiative - Phase 2 Sponsored by:

- Ø NSF
- Ø DARPA
- Ø NASA
- Ø NLM
- Ø Library of Congress
- Ø National Endowment for the Humanities

Innovative digital libraries research and applications will be jointly supported by the National Science Foundation (NSF), the Defense Advanced Research Projects Agency (DARPA), the National Library of Medicine (NLM), the Library of Congress (LoC), the National Aeronautics and Space Administration (NASA), the National Endowment for the Humanities (NEH) and others. This announcement describes the goals and features of Digital Libraries Initiative - Phase 2 (DLI-2), with particular attention on NSF interests and requirements. More detailed information on the domain-specific interests of the partnering agencies may be obtained from them. Within NSF, DLI-2 is administered by the Division of Information and Intelligent Systems (IIS) of the Directorate for Computer and Information Science and Engineering (CISE). Supporting Directorates include the Directorate for Education and Human Resources and the Directorate for Social, Behavioral and Economic Sciences. Contacts for these and related activities at other agencies are referenced at the end of this announcement.

The current effort extends the joint NSF/DARPA/NASA "Research on Digital Libraries Initiative". Since announcement of that initiative, digital libraries research and applications efforts have proliferated; new communities of researchers, information providers and users have become engaged; the definition of a digital library has evolved; technologies have advanced; stores of digital content have increased dramatically; and new research directions have emerged. These advances point to a future in which vast amounts of digital information will be easily accessible to and usable by large segments of the world's population.

To help to achieve this, the Digital Libraries Initiative - Phase 2 plans to:

- Selectively build on and extend research and testbed activities in promising digital libraries areas;
- Accelerate development, management and accessibility of digital content and collections;
- Create new capabilities and opportunities for digital libraries to serve existing and new user communities, including all levels of education;
- Encourage the study of interactions between humans and digital libraries in various

social and organizational contexts.

### **3 Program Goals**

The primary purposes of this initiative are to provide leadership in research fundamental to the development of the next generation of digital libraries, to advance the use and usability of globally distributed, networked information resources, and to encourage existing and new communities to focus on innovative applications areas. Since digital libraries can serve as intellectual infrastructure, this Initiative looks to stimulate partnering arrangements necessary to create next-generation operational systems in such areas as education, engineering and design, earth and space sciences, biosciences, geography, economics, and the arts and humanities. It will address the digital libraries life cycle from information creation, access and use, to archiving and preservation. Research to gain a better understanding of the long term social, behavioral and economic implications of and effects of new digital libraries capabilities in such areas of human activity as research, education, commerce, defense, health services and recreation is an important part of this initiative.

Collaboration between academic, industry, non-profit and other organizations is strongly encouraged to establish better linkages between fundamental science and technologies development and use, through partnerships among researchers, applications developers and users.

The sponsoring agencies have the following special interests in this initiative:

#### **I. Research**

Research areas are organized into three areas for ease of exposition. The categories and topical areas selected are illustrative, and topics may cut across several areas. No priority is implied in the ordering of the areas or subareas.

- Human-Centered Research Human-centered digital libraries research seeks to further understanding of the impacts and potential of digital libraries to enhance human activities in creating, seeking, and using information and to promote technical research designed to achieve this.

#### **Example topics are:**

Methods, algorithms, and software leading to wide-spectrum information discovery, search, retrieval, manipulation and presentation capabilities.

- Software tools and toolkits , Browsing and navigation software for large and diverse information spaces, Intelligent search of image/video types by content, structure and context ,Semantic search and retrieval theories and models, Multilingual information access and cross-lingual data services, Advanced software for searching, filtering, abstracting and summarizing large volumes of data, imagery, and other kinds of information
  - \* Intelligent user interfaces

- User/system learning and adaptation processes associated with interactive use  
Autonomous intelligent agents to support human needs, Information presentation and visualization, Collaboration technologies and tools, User and usability studies, including human-computer interaction, human mediated communication and users and institutions with special needs; , Use in education, learning and capacity building, especially in new and naive user communities; , Economic and social implications, Social science research and humanities research applied to distributed networked information environments and contexts, Social informatics, Nature and services of libraries, universities, schools and other institutions in the transition to widespread use of digital media, Knowledge acquisition, organization, dissemination and use practiced by individuals and user communities, Factors determining usage, public acceptance and investment in digital libraries, Means and media for advancing scholarly communication

\* Content and Collections-Based Research

Content and collection-centered digital libraries research focuses on better understanding of and advancing access to novel digital content and collections. Research focusing on content from many disciplinary areas and knowledge domains is appropriate. Proposals that focus on content from scientific and non-scientific knowledge domains are encouraged as are proposals drawing on existing public domain data from Federal agencies and other established data collecting organizations. In most cases support will not be provided for routine digitizing or conversion of existing collections.

**Example topics are:**

Efficient data capture, representation, preservation and archiving

- Novel digital representations of text and non-text media and derivatives, Intelligent systems and algorithms for indexing, abstracting, interpreting, classifying and cataloging, Content-based image recognition, analysis and classification, Intelligent text processing and document management; natural language analysis for data extraction and for structure and topical segmentation, Cost-effective methods for creating and converting digital objects

\* Metadata

- Means and methods for preserving and presenting context for data elements and collections
- Metadata types and standards development

Technologies, methods and processes for addressing societal, economic and legal issues associated with the creation and use of digital collections

- Intellectual property and rights management
- Privacy and security
- Publishing in a digital environment
- Charging mechanisms for copyrighted documents
- Authentication and copyright protection
- New economic and business models corresponding to new electronic media
- Development and access to educational materials and approaches including:

- New resources for science, mathematics and engineering education at all levels
- Interactive educational tools and interfaces appropriate for different groups of users
- Creation of learning environments

#### **4 Systems-Centered Research**

Systems-centered digital libraries research focuses on component technologies and integration to realize information environments that are dynamic and flexible; responsive at the level of individual, group, and institution; and capable of adapting large, amorphous, continually growing bodies of data to user-defined structure and scale.

##### **Example topics are:**

- Open, networked architectures for new information environments capable of supporting complex information access and analysis and collaborative work, Systems scalability, federation, extensibility and composability, Networking, communications and middleware research topics relevant to digital libraries including new approaches and protocols for high bandwidth applications; metadata services; reliability and integrity of services; quality of service and payment models and issues, Advanced multimedia information capture, representation and digitization, Systems evaluation and performance studies

#### **II. Testbeds and Applications**

This focuses on development of digital libraries testbeds for technology testing, demonstration and validation and as prototype resources for domain communities - technical and non-technical. Support will be provided for development and implementation of digital libraries applications which demonstrate new technologies and are sufficiently robust and stable to serve identifiable communities and encourage collaborative work environments. Applications projects are expected to result in enduring information environments for research, learning, and advancing public use in creative ways.

##### **Example activities are:**

- Integration of functional components into useful systems to serve specific domain communities and identifying unique information requirements, technical and design issues, and metrics of performance and utility, Applications that enhance the general functionality of existing and future digital libraries by providing new concepts and tools for (e.g.) document markup, image and video management, semantic encoding, metadata, intelligent search and retrieval, and federation of existing and new digital collections, Specialized digital libraries applications designed for specific knowledge domains and communities (defense, geosciences, physical sciences, biological sciences, medicine, social sciences, arts & humanities, etc.), Improving processes which support education, learning, scholarly communication and collaboration.
- New types of digital collections
- Electronic journals, textbooks, catalogs
- New means for gathering, aggregating and establishing relationships among knowledge sources

- High-risk, "breakthrough" applications capable of providing new conceptual paradigms for information technologies and altering social and work practices on a grand scale
- Distributed knowledge-work environments
- Online educational and cultural resources in the form of virtual classrooms museums, concert halls, theaters, galleries, studios suited for a broad audiences
- Multilingual, global-scale knowledge repositories
- Multimodal access supporting information needs of mobile individuals whose primary attention is directed elsewhere

### **III. Planning Testbeds and Applications for Undergraduate Education.**

To explore the linking of digital library research efforts and testbeds for undergraduate education, NSF's Division of Undergraduate Education will provide a total of \$500,000 for planning and study projects in FY1998. Successful applicants are expected to demonstrate high potential to advance undergraduate science, mathematics, engineering and technology (SMET) education. Three types of proposals are of interest: practical digital library applications for SMET education, technical studies of digital library capabilities, and general policy studies.

#### **Research Initiatives**

The Initiative's focus is to dramatically advance the means to collect, store, and organize information in digital forms, and make it available for searching, retrieval, and processing via communication networks all in user-friendly ways. Digital Libraries basically store materials in electronic format and manipulate large collections of those materials effectively. Research into digital libraries is research into network information systems, concentrating on how to develop the necessary infrastructure to effectively mass-manipulate the information on the Net.. The key technological issues are how to search and display desired selections from and across large collections.

In May 1996, a special issue of *Computer* focused specifically on a major new US government initiative the Digital Libraries Initiative (DLI) funded by the NSF, DARPA, and NASA. The six major projects supported by the DLI each had a survey paper at this halfway point in the initiative. This issue focuses on practical outcomes from research projects major research test beds and fundamental research technologies that show what the large-scale future infrastructure might become. The papers are split between DLI and non-DLI projects. Digital libraries have become far more important nationally and internationally in 1999 than in 1996. This is largely due to the exponential growth of information in the World Wide Web, which Web searchers are increasingly failing to handle successfully. This is a special case of the increasing dependence of modern society on information technology and the increasing failure of fundamental infrastructure due to the absence of fundamental new technology. The just-released PITAC report (President's Information Technology Advisory Committee) makes this point clearly. In this report, the leaders of the US information technology research community concluded that "the current Federal program is inadequate to start necessary new centers and research programs.... The end result is that critical problems are going unsolved and we are endangering the flow of ideas that have fueled the information economy." The committee went on to recommend that "the Federal budget for the year 2000 should include a commitment to sustained growth in IT research, along with a new management system designed to foster



innovative research." Digital Libraries Initiative-Phase 2 (DLI-2) is an NSF-led initiative that builds on the successes of DLI-1 and presages the even bigger efforts recommended in the PITAC report. DLI-2 has made the initial awards for multiyear projects that will support a broader range of activities than DLI-1, including smaller projects and topics in medicine and humanities. There will be an even stronger emphasis on testbeds with real users and real collections. Many federal agencies are contributing to this initiative namely NSF, DARPA, NASA, National Library of Medicine (NLM), Library of Congress, and the National Endowment for the Humanities. The "Funding Agencies" sidebar includes a contribution from the NSF program officer discussing DLI-2, as well as contributions from the lead agencies DARPA and NLM describing their agencies' other efforts to support digital library research. The importance of digital library research is spread.

## **5 Funding Agencies**

Digital Libraries: The View from NSF Stephen Griffin, National Science Foundation the Internet and WWW have demonstrated that scholars, students of all ages, and the general public have a boundless appetite for information of all types. Millions now regularly use the Web as a primary source of information, and as an inventive medium for communicating and sharing knowledge, enabling new relationships, collaborations, and intellectual communities. The Digital Libraries Initiative (DLI), funded by NSF, DARPA, and NASA from 1994 to 1998, supported pioneering exploration into issues of organization, access, security, and use of distributed information resources.. The six DLI projects addressed a broad range of fundamental research: new document models, video capture and indexing, geographic data spaces, image retrieval, concept spaces, agent-based synthetic global economies, and new tools for classroom education, to name a few. The Digital Libraries Initiative-Phase 2 (DLI-2) supported by NSF, DARPA, NLM, LoC, NEH, NASA, and other agency partners will address a refined technology research agenda, and look to support new areas in the digital libraries information life cycle, including content creation, access, use and usability, preservation, and archiving. DLI-2 will look to create domain applications and operational infrastructure, and understand their use and usability in various organizational, economic, social, and international contexts. In short, DLI-2 will investigate digital libraries as human-centered systems. DLI-2 involvement will extend far beyond computing and communications specialty communities to engage scholars, practitioners, and learners in not only science and engineering but also arts and humanities. DLI-2 recognizes that knowledge access is inherently international and will actively promote activities and processes that bridge political and language boundaries, including sponsoring projects through a new program in International Digital Libraries Collaborative Research. Many of the most important research questions regarding systems and use are bound into the process of building and using real-world operational systems.

DLI was characterized by a single project model addressing a broad, technology-centered research agenda and building technology testbeds. Content was of secondary concern and acquired primarily through donations. DLI research illuminated the complexity and difficulty of fundamental issues of functionality, scalability, interoperability, reliability, and usability. Investigations into these and related technologically grounded questions will continue in DLI-2, but until large-scale distributed systems are built, instrumented, filled with content of value, and open to use by large and diverse populations, many important questions will go unanswered. The DLI projects reached their most potent stage as research enterprises toward the end of their funded term as the test beds matured and became heavily used..

## 6 Global Digital Library

The world's national libraries share a common mission of collecting, recording, organizing, storing, preserving in perpetuity, and providing access to their nation's cultural patrimony and intellectual output in documentary form.

- The Conference of Directors of National Libraries (CDNL) is an independent association of the chief executives of the world's national libraries, established to facilitate discussion and promote understanding and cooperation on matters of common interest to national libraries worldwide.
- Our long-term vision is the development of a global distributed digital library - comprehensive, open, seamlessly-connected, and universally accessible on the internet - giving ready access to library materials in the collections of all the national libraries of the world in the interests of scholarly research, education and lifelong learning, innovation and economic development, and the promotion of international understanding. To this end we must:
  - a) Promote, encourage, and support the development of collaborative 'digital libraries' of all types.
  - b) Promote and support the connecting of the digital collections of national libraries, and in dialogue with other libraries and archives, to provide a window to the cultures of the world and to open up our rich and diverse holdings.
  - c) support cultural and linguistic diversity and multilingualism, and also respect indigenous cultures and cultural property
  - d) Promote the development of digital tools, products and services and common standards through ICABS and IIPC, and with a particular focus on critical issues for national libraries such as long -term digital preservation, authenticity, and understanding the needs of our users.
  - e) Ensure that learnings on digital library developments and best practice are shared between national libraries.
  - f) Increase digital capability and support digital developments in the national libraries of the developing world.
  - g) Work for widespread appreciation of the importance of intellectual property issues in the digital age.
  - h) Advocate for solutions to intellectual property issues which strike the appropriate balance between the crucial public interest in ensuring access to information and ideas and also the rights of creators to be recognized and rewarded for their work.
  - i) Advocate for the development of the global digital library and for regulatory support (particularly through legal deposit and intellectual property regimes) and financial support for digital developments.
  - j) Strengthen our strategic collaboration with other cultural heritage institutions, with national and international government organizations and NGOs, publishers and information providers, and other private sector organizations.

## 7 International Activities

- a) Europe: National and EU funding
- b) NSF-EU specify five working groups
  - a. Interoperability
  - b. Metadata
  - c. Intellectual Property Rights



- d. Resource Indexing
- e. Multilingual Information Access
- c) Global connections: International Federation of Library Associations and Institutions
  - a. 1995: Libraries from 70 countries are connected
  - b. 1998: Libraries from 100 countries are connected
- d) Challenges: Connectivity & Training

## **8 New Technological Advances**

To effectively meet the information needs of their clients, digital libraries need to use a combination of technological advances and have the ability to design, Construct, Manage, and use global electronic networks. They must be able to adapt rapidly to dynamic changes in technology and the information available through them.

Many technological advances information production, management and distribution are responsible for enabling digital libraries. They are too numerous to describe in detail but include such things as advances in:

- a) Storage Media
- b) Digitization or information capturing techniques (example OCR Technology)
- c) Automatic indexing and organizing of large volumes of information
- d) Computing speed
- e) Network technology (including data compression)
- f) Content-based search and retrieval
- g) Feature- based or texture-based search and retrieval
- h) Full-text indexing
- i) Resources or knowledge discovery
- j) Multimedia and hypertext
- k) Standard (example-standardized general Mark-up Language (SGML) and Hypertext Mark-up Language (HTML), and Z39.50):
- l) Object –oriented techniques
- m) Improvements in user-interface design and visualization

## **Some Technological Requirements for Digital Libraries**

In order to accomplish the function that we have seen above, digital libraries need some typical technical requirements such as a multitude of meta-data sets, full-text retrieval, and compatibility with the Open Archives Initiative standards.

- Digital libraries are special information systems, due to their efforts to create added-value (like traditional libraries do) by describing the documents in the collections with more or less standardized sets of attributes ('fields'), called 'meta-data', which are in fact indeed very similar to bibliographic structures of the classic library approach.
- Another typical requirement for digital libraries is the capability to allow retrieval of documents based on full-text, meaning not only the added meta-data act as entry-points for searching and retrieval, but all words within text-documents also are indexed and act as search keys.
- Technical compatibility with OAI-PMH is an important characteristic of digital libraries. The Open Archives Initiative develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content. The work of OAI-PMH has

expanded to promote broad access to digital resources for e-Scholarship, e-Learning, and e-Science.

## Conclusion

Technology is moving on to make digital libraries of sound, image and moving pictures easy and common, including 3-D representations of objects. Virtual reality makes tourism "better than being there" with field trips for information. The data storage takes over from written text and the libraries might be marginalized with only old stuff. Small libraries will be a prey for large libraries. Most information will be in digital format and user has to pay for copyrighted materials.

## Reference:

- Schatz, Bruce & Chen, Hsinchun (1999). Digital libraries: Technological advances and social impacts. IEEE, doi:0018-9162/99/\$10.00
- Brezovnik, Janez & Ojstersek, Milan (2011). Advanced features of digital library of university of maribor. International Journal Of Education And Information Technologies Issue, 5(1), 31-41.
- Ranjita, Khaidem & Sophiarani, Soubam (2008). Towards development of digital library in India: An observation. 6th convention planner - 2008.
- Kumbargondar, Praveenkumar et al (2000). Development of Digital Libraries in India: Social Economic problems and prospects. Presented in PLANNER 2006, Nov 9-10, Aizwal: Mizoram University
- Giral, Angela (2000). Digital image libraries: Technological advancement and social impact on the teaching of art and architectural history. INSPEL, 34(1), 9-21.

