

# DIGITAL LIBRARIES INITIATIVE - PHASE 2

## Sponsored By:



**National Science Foundation**



**Defense Advanced Research Projects Agency**



**National Library of Medicine**



**Library of Congress**



**National Aeronautics & Space Administration**



**National Endowment for the Humanities**

## In Partnership With:

**National Archives And Records Administration  
Smithsonian Institution**

PROPOSAL DUE DATES: JULY 15, 1998  
MAY 17, 1999  
Announcement Number NSF 98-63 (NEW)

# Digital Libraries Initiative - PHASE 2 (DLI-2)

## FY 1998 - FY 2002

### INTRODUCTION

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 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.

Electronic information is being created by many people and data gathering instruments in many forms and formats, stored in many repositories around the world, and becoming increasingly interconnected via electronic networks. Digital libraries research is faced with the challenge of applying increasing computational capacity and network bandwidth to manage and bring coherence, usability, and accessibility to very large amounts of distributed complex data and transform it into information and knowledge. Since digital libraries are meant to provide *intellectual access* to stores of information, research in this initiative is concerned with developing concepts, technologies and tools to gain use of the fuller knowledge and meaning inherent in digital collections. For example, for users this means intelligent search, retrieval, organization and presentation tools and interfaces; for content and collections providers this means new information types, structures, document encoding and metadata for enhancing context; for system builders this means designing hardware and software systems capable of interpreting and implementing users' requests by locating, federating and querying collections to provide the user with the structured information sought.

### 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
  - alternative document and text markup systems, structuring principles and distributed management models
  - structuring and linking of information objects and documents
  - 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
- interoperability of content and collections
- domain-specific information objects
- 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

### *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
- intelligent agents
- interoperability
- 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 FY 1998. 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.

#### **CATEGORIES OF SUPPORT**

All awards for this announcement made by NSF will be as grants or cooperative agreements to academic institutions and qualified non-profit research organizations. Partnership arrangements with other groups are encouraged, including subcontracts with the single proposing organization.

NSF expects to fund two general types of projects under this initiative:

- a) Individual investigator research grants. Awards will not exceed \$200,000 per year, for 1 to 3 years.
- b) Multi-disciplinary group research projects. Awards will not exceed \$1,200,000 per year, for 1 to 5 years.

The number of awards will depend on the quality of proposals received, the availability of funds, and considerations for creating a balanced overall program. Total support for the initiative from federal sponsors is projected to be \$40-\$50 million over the 5 year Initiative. Awards will not exceed \$1,200,000 per year, except in exceptional circumstances. Ideas for projects requiring support above this level should be discussed with the NSF program officer before proposal preparation.

#### **PROPOSAL SCHEDULES AND PREPARATION**

**Letters of Intent Are Due April 15, 1998 for the FY 1998 Competition and February 15, 1999 for the FY 1999 Competition.**

Organizations or persons considering submission of a proposal should send an electronic mail message with the following statement: "I am interested in submitting a proposal to the Digital Libraries Initiative focusing on (Research, Testbeds and Applications, Planning Testbeds and Applications for Undergraduate Education)." It should include:

- Brief Abstract of Proposed Work - one page maximum

If prospective applicants intend to apply for joint support from other agencies, a statement to that effect should be included.

- The Proposed Principal Investigators

Name  
Title  
Organization  
Mailing address  
Phone  
email address

Submit the letter of intent as an electronic mail message to dli2@nsf.gov or send a letter with the same information to:

Digital Libraries Initiative - Phase 2  
Division of Information and Intelligent Systems  
Suite 1115  
National Science Foundation  
4201 Wilson Boulevard  
Arlington, VA 22230

**Full Proposals**

**Proposal deadlines are: July 15, 1998, May 17, 1999**

The proposals must be marked DIGITAL LIBRARIES - NSF 98-63 in the top left hand box, "Program Announcement", on the cover sheet (NSF Form 1207). Proposals prepared as specified in the NSF Grant Proposal Guide (NSF-98-2) and sent to:

National Science Foundation PPU  
Digital Libraries Initiative - Phase 2  
4201 Wilson Boulevard  
Arlington, VA 22230

**Special Conditions For Projects Seeking Multiple Sponsorship From Other NSF Programs Or Other Federal Funding Agencies**

For projects seeking multiple support from other NSF programs or other Federal funding agencies the following special conditions apply:

- a) For those seeking joint support from other NSF disciplinary programs, the proposal should be written, and the budget constructed to allow separate and independent review and funding of each component, the domain-specific component and the DLI-2 component. The proposal should clearly indicate targeted NSF programs appropriate for support.

- b) For those seeking joint funding from Federal agencies with extramural programs available for support of digital libraries-related research and applications, the proposal should be written, and the budget constructed to allow independent review and funding of each specific component by each agency. Potential co-sponsors include programs in the group of agencies supporting this initiative and other Federal agencies.

An organization wishing to submit a multiple-funding proposal assumes the responsibility for establishing communication between appropriate agency personnel and for coordinating the various interactions that may be required to prepare a necessarily more complex proposal: ascertaining the nature and levels of commitment to the proposed project by each source, consideration of funding timing and review cycles; planning for managing staggered funding from different sources; and being responsive to various reporting demands.

The proposal must clearly establish the linkage between digital libraries research and the larger effort under consideration.

## **PROPOSAL EVALUATION**

Proposals will be evaluated by panels of experts representing all sponsoring agencies, with mail reviews solicited as necessary. Cost sharing is encouraged and will be considered in evaluating proposals; multi-institutional proposals are encouraged to have cost-sharing of at least 25 percent. Cost sharing must be from non-Federal funds.

Evaluation criteria applied to all NSF proposals listed in Grant Proposal Guide, NSF 98-2, are:

1. What is the intellectual merit of the proposed activity?

How important is the proposed activity to advancing knowledge and understanding within its own field or across different fields? How well qualified is the proposer (individual or team) to conduct the project? (If appropriate, the reviewer will comment on the quality of prior work.) To what extent does the proposed activity suggest and explore creative and original concepts? How well conceived and organized is the proposed activity? Is there sufficient access to resources?

2. What are the broader impacts of the proposed activity?

How well does the activity advance discovery and understanding while promoting teaching, training, and learning? How well does the proposed activity broaden the participation of underrepresented groups (e.g., gender, ethnicity, disability, geographic, etc.)? To what extent will it enhance the infrastructure for research and education, such as facilities, instrumentation, networks, and partnerships? Will the results be disseminated broadly to enhance scientific and technological understanding? What may be the benefits of the proposed activity to society?

### **Additional Criteria For This Announcement**

1. Creativity and Originality.

Proposals should represent new and innovative research, designed to open new areas of activity in the digital library field.

2. Information Utility

The commercial, educational, or scientific value of the information used should be as high as possible to as many users as possible, and support needs and goals of our society.

3. Metrics

Each project should present a plan for evaluating and measuring the effect of the work done.

### **Additional Criteria For Group Awards**

In general, group and multi-institutional proposals involving significantly large and valuable collections are expected to show potential for novel functionalities and features, identifiable user clienteles, interoperability, scaling, extensibility, and durability, and a detailed description of how these resources might eventually be made available to and used by others. These proposals should, as appropriate include active participation from client groups, technology vendors, commercial enterprises, and private or governmental archival establishments. The proposals will also be evaluated based on:

1. Survivability

Credibility of the plan for continuing the services to the user community after the expiration of research funding.

2. Impact

Potential for general impact on the development of techniques, environments or paradigms that will advance the utilization of digital libraries.

3. Cohesiveness and Management

The extent to which the group is integrated, has a common focus and the quality of management and coordination plans.

4. Education and Training

The degree to which research and education are integrated and activities involve participation and training of students.

Proposals which are initially funded at a level exceeding \$1,000,000 per year for three or more years will be evaluated based on the proposed work plan by teams of experts midway through the term of the project to determine performance levels. Funding for the balance of the project term may be revised based on this evaluation. All publications, reports, data and other output from all awards must be prepared in digital format and meet general requirements for storage, indexing, searching and retrieval.

## **PUBLIC BRIEFINGS**

Special regional briefings will be held relative to this announcement. These briefings will begin in March 1998. Locations and schedules will be posted on the sponsor announcement web sites. (see below)

## **SPONSOR WEB SITES FOR ADDITIONAL INFORMATION**

The following Web sites of the sponsoring Federal agencies provide additional information about this announcement and related activities and programs of support.

### **NATIONAL SCIENCE FOUNDATION**

<http://www.nsf.gov>

### **DEFENSE ADVANCED RESEARCH PROJECTS AGENCY**

<http://www.darpa.mil/ito>

### **NATIONAL LIBRARY OF MEDICINE**

<http://www.nlm.nih.gov/ep>

### **LIBRARY OF CONGRESS**

<http://lcweb2.loc.gov/ammem/dli2/>

### **NATIONAL AERONAUTICS AND SPACE ADMINISTRATION**

<http://www.nasa.gov>

### **NATIONAL ENDOWMENT FOR THE HUMANITIES**

<http://www.neh.gov/html/guidelin/dli2.html>

## AWARD ADMINISTRATION

NSF awards will stipulate that the awardee is responsible for contributing the specified and agreed upon amount of cost sharing during the award period. In light of this requirement, NSF funds will not be awarded until required cost sharing commitments have been assured. A letter from an institutional officer stating the amount and source of eligible cost sharing, and assuring availability and commitment of these funds during the proposed award period must accompany the signed cover sheet for an award recommendation to be processed. In addition, NSF award conditions may specify special reports, on-site inspections, or other requirements.

NSF requires prospective grantees to furnish, upon request by NSF's Division of Grants and Agreements, basic organization and management information that will assist the NSF Grant Officers in assessing their financial and managerial responsibility. These requirements are described in the NSF Grant Policy Manual (July 1995), particularly Chapters IV and V. Grants awarded as a result of this solicitation are administered in accordance with the terms and conditions of NSF GC-1, "Grant General Conditions," or FDP-III, "Federal Demonstration Partnership General Terms and Conditions," depending on the grantee organization. Any Cooperative Agreement resulting from this announcement must comply with NSF GC-1 and Cooperative Agreement General Conditions, CA-1.

## OTHER INFORMATION

The Foundation provides awards for research in the sciences and engineering. The awardee is wholly responsible for the conduct of such research and preparation of the results for publication. The Foundation, therefore, does not assume responsibility for the research findings or their interpretation.

The Foundation welcomes proposals from all qualified scientists and engineers and strongly encourages women, minorities, and persons with disabilities to compete fully in any of the research related programs described here. In accordance with federal statutes, regulations, and NSF policies, no person on grounds of race, color, age, sex, national origin, or disability shall be excluded from participation in, be denied the benefits of, or be subject to discrimination under any program or activity receiving financial assistance from the National Science Foundation.

Facilitation Awards for Scientists and Engineers with Disabilities (FASSED) provide funding for special assistance or equipment to enable persons with disabilities (investigators and other staff, including student research assistants) to work on NSF projects. See the program announcement or contact the program coordinator at (703) 306-1636.

Privacy Act and Public Burden. The information requested on proposal forms is solicited under the authority of the National Science Foundation Act of 1950, as amended. It will be used in connection with the selection of qualified proposals and may be disclosed to qualified reviewers and staff assistants as part of the review process; to applicant institutions/grantees; to provide or obtain data regarding the application review process, award decisions, or the administration of awards; to government contractors, experts, volunteers, and researchers as necessary to complete assigned work; and to other government agencies in order to coordinate programs. See Systems of Records, NSF 50, Principal Investigators/Proposal File and Associated Records, and NSF-51, 60 Federal Register 4449 (January 23, 1995). Reviewer/Proposal File and Associated Records, 59 Federal Register 8031 (February 17, 1994). Submission of the information is voluntary. Failure to provide full and complete information, however, may reduce the possibility of your receiving an award.

Public reporting burden for this collection of information is estimated to average 120 hours per response, including the time for reviewing instructions. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Gail A. McHenry,

Reports Clearance Officer, Information Dissemination Branch, National Science Foundation, 4201 Wilson Boulevard, Suite 245, Arlington, VA 22230.

The National Science Foundation has TDD (Telephonic Device for the Deaf) capability, which enables individuals with hearing impairment to communicate with the Foundation about NSF programs, employment, or general information. To access NSF TDD, dial (703) 306-0090; for FIRS, 1-800-877-8339.

Catalog of Federal Domestic Assistance Numbers:  
47.070 Computer and Information Science and Engineering  
47.075 Social, Behavioral, and Economic Sciences  
47.076 Education and Human Resources

OMB# 3145-0058  
P.T. (Program Type) and K.W. (1000000,0400000,0300000) codes from the Grants Keyword Thesaurus  
NSF 98-63

## **AGENCY CONTACTS**

### **National Science Foundation**

Stephen M. Griffin  
National Science Foundation  
Division of Information and Intelligent Systems  
4201 Wilson Boulevard  
Arlington, VA 22230

Electronic mail: sgriffin@nsf.gov  
Telephone: 703-306-1930  
Fax: 703-306-0599

Herbert H. Richtol (undergraduate education planning testbeds and applications)  
National Science Foundation  
Division of Undergraduate Education  
4201 Wilson Boulevard  
Arlington, VA 22230

Electronic mail: hrichtol@nsf.gov  
Telephone: 703-306-1666  
Fax: 703-306-0445

William Sims Bainbridge  
Directorate for Social, Behavioral and Economic Sciences  
4201 Wilson Boulevard  
Arlington, VA 22230

Electronic mail: wbainbri@nsf.gov  
Telephone: 703-306-1741  
Fax: 703-306-0495

### **Defense Advanced Research Projects Agency**

Ronald L. Larsen  
DARPA/ITO  
3701 N. Fairfax Dr.

Arlington, VA 22203

Electronic mail: rlarsen@darpa.mil  
Telephone: 703-696-2227  
Fax: 703-696-2202

## **National Library of Medicine**

Milton Corn  
Associate Director  
National Library of Medicine  
Rm 5N505, Bldg 38A  
Bethesda, MD 20894

Electronic mail: corn@nlm.nih.gov  
Telephone: 301-496-4621  
Fax: 301-402-2952

## **Library of Congress**

Laura Campbell  
Director, National Digital Library  
Library of Congress  
Washington, DC 20540

Electronic mail: lcam@loc.gov  
Telephone: 202-707-3300  
Fax: 202-707-0815

## **National Aeronautics and Space Administration**

Mark Leon  
Ames Research Center  
Mail Stop 269-3  
Moffett Field, CA 94035-1000

Electronic mail: mleon@mail.arc.nasa.gov  
Telephone: 415-604-6498  
Fax: 415-604-4036

## **National Endowment For the Humanities**

George Farr  
Director  
Division of Preservation and Access  
National Endowment for the Humanities  
Room 411  
Washington, D. C. 20506

Electronic mail: gfarr@neh.gov  
Telephone: 202-606-8570  
Fax: 202-606-8639

or

Jeffrey Field  
Deputy Director  
Division of Preservation and Access  
National Endowment for the Humanities  
Room 411  
Washington, D. C. 20506

Electronic mail: jfield@neh.gov  
Telephone: 202-606-8442  
Fax: 202-606-8639