



HARVARD UNIVERSITY LIBRARY

## *Library Digital Initiative*

### **Review of Year One: 1998–1999**

Harvard University launched the Library Digital Initiative (LDI) in July 1998 as a five-year program to develop the University's capacity to manage digital information by:

- creating the technical infrastructure to support the acquisition, organization, delivery, and archiving of digital library materials;
- providing a team of specialists to advise librarians and others in the University community on key issues in the digital environment;
- providing librarians and staff with experience in a wide range of technologies and digital materials; and
- enriching the Harvard University Library collections with a significant set of digital resources.

This report outlines LDI activities in the first year, and highlights progress in the development of systems and services, the establishment of a grant program, related activities and collaboration. Since LDI has already improved Harvard's ability to manage and make available digital collections, the program shows promise for effective progress in the future.

<http://hul.harvard.edu/ldi/>

## Highlights 1998-1999

The Library Digital Initiative will build a common infrastructure for the University comprised of union catalogs and a set of storage and delivery services that enable Harvard to effectively manage its distributed networked collections. New systems and services developed this year contribute to this infrastructure, while the LDI grant-funded projects test their functionality.

*Critical to the success of an effective digital library are public interfaces that help users discover the variety of available resources. Catalogs, a primary means of access to many traditional library resources, also fulfill this role in the digital environment. Two new web-based union catalogs were launched in FY 1999, with a third in planning as part of LDI:*

### VIA

Visual Information Access is a catalog describing the rich variety of photographs, prints, drawings, paintings, and other visual resources held by Harvard libraries, archives, and museums. VIA allows researchers to search within one particular collection or across a number of the University's collections. The catalog is designed primarily for cultural collections, which cannot be adequately described in HOLLIS. The initial contributors to the catalog are the visual collection departments of the Fine Arts Library, the Frances Loeb Library of the Graduate School of Design, the Schlesinger Library, the Harvard University Art Museums (including the Fogg, Sackler, and Busch-Reisinger Museums), and the Peabody Museum of Archaeology and Ethnology. A related development is OLIVIA, a system for creating and maintaining visual cataloging records for display in VIA.

### OASIS

Online Archival Search Information System is a catalog of finding aids, providing access to information about the University's archival collections. Current contributors include Houghton Library, the Andover-Harvard Theological Library, Baker Library at the Harvard Business School, the Botany Libraries, the Harvard Law School Library, the Peabody Museum Archives, and the Schlesinger Library.

### Geodesy

Geodesy is a new system that will provide access to geo-spatial data. It will support searching by "geo-reference," that is by location on the face of the earth. Functional and technical planning for Geodesy began in this year, with initial implementation scheduled for next year. Key participants in the Geodesy project include the Harvard College Library Social Sciences Program and the Harvard Graduate School of Design.

The proliferation of Harvard-wide catalogs as well as other electronic resources (primarily commercial ones including indexing and abstracting databases, directories, electronic journals and reference works such as dictionaries and encyclopedias) is creating an increasingly complex and confusing landscape for library users. To address this complexity, LDI has begun planning to replace HOLLIS Plus with a more comprehensive and powerful system for the Harvard community. Frequently called "portal pages," such systems present a single, organized view of web accessible resources and services.

*Another major area of concern for digital libraries is digital resource storage and access. In response to this need, LDI has two new systems under development:*

### Digital Repository

Repositories are systems that enable the long-term management of and access to large bodies of digital objects of all types (including images, textual documents, sound files, and multimedia objects). A central repository will offer storage and access services for digital materials to the Harvard community that parallel the services offered by the Harvard Depository for traditional materials. Development of this service was initiated in early 1999 with initial implementation scheduled for the fall.

### Naming Services

Naming provides digital objects with persistent, location-independent identifiers to ensure long-term access regardless of the object's physical location, which may change over time. With the current rate of organizational and technological change, naming provides critical support to the multiple layers of linking made possible by the web. The development of a Naming service is well underway, with availability scheduled for early 2000.

## Highlights 1998-1999 — continued

*One of the key aims of LDI is to provide expertise and technical assistance to libraries, archives, museums, and research projects involved in collecting or creating digital resources throughout the University.*

### Digital Acquisition

Acquiring commercial digital materials is dramatically more complex than traditional library acquisitions. The **Digital Acquisitions Coordinator** assists Harvard librarians in addressing issues related to the licensing and acquisition of digital resources including product evaluation, negotiating licenses, and vendor management. Accomplishments in the first year include the following:

- over 250 individual resources, including 226 new electronic journals, were licensed and added to HOLLIS Plus;
- policies and procedures were developed for initiating and evaluating resources, and a distributed stewardship model has been implemented to monitor existing high-cost, high-impact resources;
- financial management of cooperative purchases was dramatically improved, including a new model for allocating costs among faculty libraries based on usage; and
- a set of licensing guidelines are in draft and will be available in winter 1999.

### Metadata

Digital library systems are driven by metadata, the information that makes it possible to find, access, use, and manage digital resources. The **Metadata Analyst** provides expert advice and design assistance for projects involving a wide variety of digital materials. Accomplishments in the first year include:

- design of a descriptive metadata model for visual resources that accommodates the different practices used by libraries, archives, and museums (developed for VIA and OLIVIA, see above);
- participation in an effort to define national technical metadata standards for the long-term preservation of image files as part of a national standards effort; and
- design of a metadata model for describing the complex structure of traditional published works in electronic form (initiated as part of the Harvard/Radcliffe Archives grant project, described in the next section).

### Reformatting

Some of Harvard's digital collection will be created by digitally reformatting existing collection materials. Reformatting involves document preparation, metadata creation, scanning, image processing and quality control. The **LDI Reformatting Advisor** advises curators on appropriate reformatting strategies, helps managers locate skilled vendors, and negotiates for products and services that conform to Harvard's guidelines. Accomplishments in the first year include:

- developing local web-based guidelines and information resources to help curators in decision-making about scanning print and film collections;
- contracting vended conversion services for scanning 35mm color film and for optical character recognition (OCR) to create searchable text from scanned page images; and
- planning for a scanning facility, Harvard College Library Digital Imaging Group (HCL-DIG), that will offer services to Harvard departments and projects on a cost-recovery basis. The facility will be operational in the summer of 1999, and has been designed to ensure safe handling, tracking, and treatment of all source material; create persistent digital images of consistent quality; and integrate the facility's systems with the digital repository to ensure reliable transfer of images and associated metadata.

## Internal Challenge Grant Program

The LDI grant program was established to encourage broad University participation in LDI, serve as a catalyst for infrastructure development, integrate management of digital materials with traditional materials throughout the libraries, and contribute meaningful digital resources to the University's collections. The grant projects have proven to be extremely useful in setting priorities for systems development, providing developers with "real world" applications during design and testing, and providing work experience in the digital environment to staff throughout the University.

*In the first round of the grant program, nineteen proposals were submitted to the Grant Review Committee and the following five were awarded in January 1999:*

### **Online Historical Reference Shelf**

is a joint project of the Harvard University Archives and Radcliffe College Archives to provide access to the Presidents' Annual Reports for 1825-1995. The reports will be converted to digital image format, and the underlying text will be converted via OCR for searching. The original source material, estimated at 95,600 pages, will first undergo disbinding, scanning and OCR.

### **Nineteenth-Century American Tradecards**

is a project to provide access through VIA to catalog records and images for 1,000 nineteenth-century trade cards selected from a collection in Baker Library.

### **The Hedda Morrison Photographs of China**

is a project of the Harvard-Yenching Library to provide access to images and descriptive catalog records for 4,800 photographs made in China from 1933 to 1946 by the important German photographer Hedda Morrison.

### **The Harvard College Library Environmental Resources and Map Collections**

are leading a collaborative effort with the Harvard Graduate School of Design to catalog and provide access to a variety of geo-spatial data. This project is driving the development of the Geodesy catalog and a set of Geographic Information Systems (GIS) services.

### **Asian Art Images**

is a collaborative project between Harvard University Art Museums (HUAM) and Fine Arts Library (FAL) to provide access to 3,600 Asian art images from transparencies (including slides and lantern slides), glass plate and film negatives in their collections. These materials will be the first application for a new digital imaging facility currently being planned in HUAM.

*Of the five preliminary proposals submitted for Round 2 in April 1999, full proposals were requested for the following two projects, with funding expected to be awarded in October 1999:*

### **Eda Kuhn Loeb Music Library**

*Music from the Archive: A New Model of Access to Rare and Unique Sound Recordings*

### **Arnold Arboretum Library of Harvard University**

*Western China and Tibet: Hotspot of Diversity*

## Related Programs

### E-Reserves

This year, the Lamont Library completed a four semester pilot project demonstrating the use of traditional course library reserves in an online environment. The project entailed both scanning existing paper materials and linking to Internet-based resources via a restricted web site for participating courses in the Faculty of Arts and Sciences. Next year, this project will be expanded to include the Divinity School as well as eight units in the Harvard College Library.

### Digital Contents Pilot Project (DCPP)

This pilot project improves access to remotely stored serial publications that are not commercially indexed. Tables of contents for serials from Widener Library that are stored offsite in the Harvard Depository are scanned, linked to HOLLIS records, and delivered over the WWW.

### Virtual Data Center (VDC)

The Harvard University Library and the Harvard-MIT Data Center are the recipients of a National Science Foundation (NSF) Digital Libraries Initiative award to support the development of a social sciences digital data library—The Virtual Data Center (VDC). The VDC will be a free, sharable software system to manage and share numerical social science data across multiple institutions, and will be based on a prototype data server developed by the Harvard-MIT Data Center.

### Copyright Guidelines

LDI has been working with the Office of the General Counsel (OGC) on issues of intellectual property related to digital materials. *Copyright Guidelines for the Library Digital Initiative* was prepared by OGC and made available to assist the Harvard community in addressing copyright issues in the digital environment.

## LDI Systems and Services in Use 1998–1999

**VIA** is a web-based union catalog of Harvard's visual resources. VIA contains 98,107 records and 19,094 thumbnail images (with many linking to larger resolution images) contributed by Harvard libraries and museums.

<http://via.harvard.edu:748/html/VIA.html>

**OASIS** is a web-based union catalog of archival Finding Aids, providing access to information about the University's archival collections. There are 182 finding aids contributed by libraries, museums, and academic departments.

<http://oasis.harvard.edu/>

**HOLLIS Plus** provides web-based access to 869 networked resources including such significant assets as ProQuest Research Library, Lexis-Nexis Academic Universe, and other databases from the Congressional

Information Service, Dissertation Abstracts/Digital Dissertations, ISI Web of Science Indices, EIU Country Reports, MEDLINE, MLA Bibliography, PsycINFO. Also included are electronic journals from major publishers such as Springer-Verlag, American Chemical Society, Oxford University Press, Academic Press, and JSTOR. Over 250 new resources were added this year.  
<http://hplus.harvard.edu/>

**DCPP** provides web accessible tables of contents for remotely stored serials publications. Tables of contents for more than 200 European local history journals were scanned and made accessible this year through links to HOLLIS records.

**E-Reserves** provides online access to course reserves materials. In FY 1999 a pilot project provided electronic reserves for twenty-two courses.

## Collaboration

Digital libraries require cooperation and collaboration on both the technical and organizational level. They are composed of many interdependent systems, created and operated by independent players inside and outside the University. A significant number of people and organizations are involved in developing and supporting the underlying technical infrastructure, as well as creating and providing digital research resources.

The Harvard University Library is establishing relationships through the Library Digital Initiative that go beyond the boundaries of the libraries within the University and extend contacts outside of the University in a diversity of fields. The process of creating digital libraries is iterative and interactive and entails a dialogue among technical experts, administrators, faculty, curators, librarians, other academic institutions, and commercial developers. Harvard faculty involvement in LDI occurs on a number of different levels including participation in technical, review, and oversight committees.

All technical and functional LDI designs are reviewed biweekly by a technical review team composed of staff from the Harvard University Library and from other departments. Participants include Scott Bradner and Charles Maxson of University Information Systems; Paul Bergen of the Faculty of Arts and Sciences Instructional Computing Group; John Howard and Tom Parris of Harvard College Library; and Micah Altman of the Government Data Center.

Harvard University Library participates actively in the North East Research Libraries Consortium (NERL), an organization comprising seventeen of the region's leading research institutions. NERL was founded in 1996 to contain the mushrooming costs of electronic purchases through joint licensing agreements. Nearly one-third of the resources cooperatively purchased by Harvard Libraries are acquired through NERL.

This year the LDI staff spent a day exchanging information with Carl Lagoze and Sandy Payette of the Cornell Digital Library Research Group (CDLRG). CDLRG investigates the technical infrastructure of digital libraries with the goal of defining core components and interactions to make them succeed.

Harvard is very active in the Digital Library Federation (DLF), a national leadership organization composed of twelve university research libraries and the Library of Congress that promotes the collection and dissemination of digital works. LDI has collaborated with DLF on many levels. Dale Flecker, Associate Director for the

University Library for Planning and Systems, chairs the DLF Architecture Committee. Flecker also helped organize and presented at a workshop co-sponsored by DLF, the National Information Standards Organization (NISO), the Association of Scholarly Publishers, and the National Federation of Abstracting and Indexing Services on issues of building electronic links from citations to electronic journal literature. Harvard, Columbia, and Oxford Universities with the University of Michigan as the lead institution were jointly awarded a DLF grant to develop an architecture and prototype for providing access to archival finding aids distributed across collections and institutions.

The Instructional Computing Group (ICG) of the Faculty of Arts and Sciences Computer Services provides support to faculty and teaching staff using computers in their coursework. The Library Digital Initiative works with ICG to contribute digital library materials to course web pages and in the development of the E-reserves system.

Robin Wendler, Metadata Analyst, actively participates in the development of the Dublin Core, an international metadata standard.

In June, Nancy M. Cline, Roy E. Larsen Librarian of Harvard College, and Dale Flecker participated in the Second International Virtual Libraries Conference sponsored by the New York Public Library and Bibliotheque Nationale de France.

# Library Digital Initiative Committees and Staffing

## Steering Committee

Nancy M. Cline  
*Roy E. Larsen Librarian of  
Harvard College (Chair)*

Gary Anderson  
*Professor of Hebrew Bible,  
Harvard Divinity School*

Allan M. Brandt  
*Professor of the History of Science  
in the Faculty of Arts and Sciences  
and Amalie Moses Kass Professor  
of the History of Medicine,  
Harvard Medical School*

Dale P. Flecker  
*Associate Director for Planning  
and Systems in the Harvard  
University Library*

Barbara S. Graham  
*Associate Director for Administration  
and Programs in the Harvard  
University Library*

Harry S. Martin III  
*Professor of Law and Librarian of  
the Law Library, Harvard Law  
School*

Judith Messerle  
*Countway Librarian for the Harvard  
Medical and Boston Medical  
Libraries*

Thomas J. Michalak  
*Executive Director, Baker Library  
Harvard Business School*

Hugh Wilburn  
*Librarian and Assistant Dean for  
Information Services, Frances Loeb  
Library, Graduate School of Design*

## Grant Review Committee

Mary Maples Dunn  
*Carl and Lily Pforzheimer  
Foundation Director,  
Schlesinger Library on History  
of Women in America, (Chair)*

Peter K. Bol  
*Professor of Chinese History in  
the Faculty of Arts and Sciences*

Robert L. Buckwalter  
*Associate Librarian,  
Law School Library*

Dale P. Flecker  
*Associate Director for Planning  
and Systems in the Harvard  
University Library*

John B. Howard  
*Librarian for Information  
Technology in the Harvard  
College Library*

Judith Messerle  
*Countway Librarian for the  
Harvard Medical and Boston  
Medical Libraries*

Donald H. Pfister  
*Asa Gray Professor of Systematic  
Botany and Curator of the  
Farlow Library and Herbarium*

## Staff

Nine new staff members were  
hired in the first year to work  
on the Library Digital  
Initiative:

Stephen Abrams  
*Digital Library Software  
Engineer*

Ivy Anderson  
*Coordinator for Digital  
Acquisitions*

Stephen Chapman  
*Preservation Librarian for  
Digital Initiatives, Harvard  
University Library Preservation  
Center*

James Coleman  
*Digital Library Projects Manager*

Wendy Gogel  
*Digital Library Projects Liaison*

Lee Mandell  
*Programming Analyst*

Edson Pereira  
*Unix Systems Administrator/  
Programmer*

MacKenzie Smith  
*Digital Library Projects Manager*

Robin Wendler  
*Metadata Analyst*

July 1999

Library Digital Initiative: <http://hul.harvard.edu/ldi/>

Dale P. Flecker  
Harvard University Library  
Office for Information Systems  
1280 Massachusetts Avenue, Suite 404  
Cambridge, MA 02138 617.495.3724