



California Digital Library

Report to the Digital Library Federation

October, 2004

Volume 5, Number 1. Fall 2004

http://www.diglib.org/pubs/news05_01/

TABLE OF CONTENTS

- i. Collections, services, and systems
- ii. Projects and programs
- iii. Specific digital library challenges
- iv. Digital library publications, policies, working papers, and other documents

I. Collections, services, and systems

A. Collections

Shared Collections

With collaborative leadership and core support provided by the California Digital Library (CDL), the UC libraries' shared digital collections have grown to provide faculty, students, and staff from all UC campuses with access to about 10,000 journal titles, 250 reference databases, 240,000 images, and more than 7,000 finding aids. These shared digital collections not only provide the UC community with access to a wealth of materials that individual campuses might not have been able to afford independently, but they also make information equally accessible to all UC students and faculty at any time, regardless of location. Through shared licensing and cataloging, the campus libraries avoid costs of roughly \$34 million per year.

The new UC Libraries Shared Print Program has been made possible in part by services such as the Melvyl Catalog, Request, and desktop delivery, which are all hosted by the CDL. The program is working to define and develop user access mechanisms that will integrate broadly with CDL services, ensuring that the development of digital and paper

repositories are complementary, and that digital access mechanisms are available to support efficient access to shared print materials. To date, the Shared Print Program has developed "dim archives" of online content from Elsevier's Science Direct and ACM. Currently, we are receiving print archival copies from several other online publishers, and are working with JSTOR to develop a dim repository of JSTOR digitized journals.

Counting California

Counting California provides users with a single interface for accessing a variety of data and statistics about California from local, state, and federal government agencies. New content has recently been added to the web site covering topics such as population, business, crime, education, and the economy. In addition, an improved search page allows users to select from a list of topics or agencies and search by keyword, table title, subject, agency, and publication title.

<http://countingcalifornia.cdlib.org/>

Image Collections

The CDL's Image Services project provides access to more than 240,000 digital images from licensed and freely-available collections, including Museums and the Online Archive of California, Art Museum Image Consortium, Saskia Art & Architecture, the David Rumsey Map Collection, and others. A library of UC images from the visual resource collections on UC campuses is being added. Through the use of Insight software from Luna Imaging, users can zoom in on image details, create groups and presentations, and view descriptive data for each image.

<http://www.cdlib.org/inside/projects/image/>

B. Services

Layered Service Model

In order to more effectively manage and deliver essential ongoing library services, the CDL and UC libraries are developing more ways to leverage their collective resources. Through the use of a shared infrastructure of building block collections and services, the campus libraries will be able to create enriched sets of collections and services customized to local needs. The following tools are components of this shared infrastructure.

eXtensible Text Framework (XTF)

The eXtensible Text Framework (XTF) is a flexible indexing and query tool that supports searching across collections of heterogeneous data and presents results in a highly configurable manner. The CDL will use XTF as a building block for new services such as the digital preservation repository, and to replace a number of aging systems.

<http://www.cdlib.org/inside/projects/xtf/>

Interface Customization Tools

The CDL is developing a digital object repository to support the collections it manages, which currently contain more than 150,000 images, texts, and other materials encoded in the Metadata and Encoding Transmission Standard (METS). The interface customization tools developed by the CDL allow institutions who contribute digital objects to access their collections through a customized interface.

<http://www.cdlib.org/inside/diglib/repository/customize/>

Metasearch Infrastructure

The metasearch infrastructure project provides tools and software that UC campus libraries will be able to use to craft search portals tailored to specific audiences and needs. In addition, services will be built to harvest metadata from remote repositories and crawl selected web sites. The resulting indexes will enrich the resources available for searching within specific disciplines. As part of this project, the CDL is collaborating with the Interactive University at UC Berkeley. Their Scholar's Box concept will guide the addition of tools that scholars can use with UC portals to gather raw materials for personal collections of resources for research and teaching.

<http://www.cdlib.org/inside/projects/metasearch/>

II. Projects and programs

A. Projects

New Project Announcements

The Web at Risk: A Distributed Approach to Preserving our Political Cultural Heritage

Through a grant awarded by the National Digital Information and Infrastructure Preservation Program at the Library of Congress, the CDL is partnering with New York University and the University of North Texas to develop web archiving tools that will be used by libraries to capture, curate, and preserve collections of web-based government and political information. Although it is anticipated that the technology will be useful in the general capture and persistent management of web-based information, work will initially focus on the web-based information produced by U.S. state and federal governments and by local political activities, such as the California gubernatorial recall election of 2003. Other partners include the UC libraries, San Diego Supercomputer Center, Stanford University Computer Science Department, and Sun Microsystems Inc., among others.

<http://www.cdlib.org/inside/projects/preservation/>

Update on Existing Projects

Bibliographic Services Infrastructure

The CDL is identifying and building components of a new bibliographic services infrastructure. Current sub-projects include researching options for an electronic resources management system and decision support tools for collaborative collection development. The goal is to create efficiencies in the creation and management of bibliographic data and the services that rely upon it. The project draws upon extensive experience with the various projects and programs hosted by the CDL, including shared cataloging, link resolution systems, persistent IDs, and the Melvyl Catalog (which was successfully migrated to a new platform in 2003). <http://www.cdlib.org/inside/projects/>

Documenting the American West

With the support of a three-year William and Flora Hewlett Foundation grant, the CDL is assembling a virtual collection of the American West that draws from the resources of major research institutions. A range of tools will support extensive re-configuration, integration with online learning environments, and continued growth through the addition of relevant research and teaching materials produced in the course of its use. Based upon needs discovered during user assessment workshops in July, the CDL will develop tools to configure and integrate these virtual collections with local, personalized content. <http://www.cdlib.org/inside/projects/amwest/>

B. Programs

eScholarship

The eScholarship program facilitates innovation and supports experimentation in the production and dissemination of scholarship. Through its partnerships across the university community and the use of innovative technology, the program is working to improve all aspects of scholarly communication, including its creation, peer review, management, dissemination, and preservation.

<http://osc.universityofcalifornia.edu/responses/escholarship.html>

eScholarship Repository

The eScholarship Repository offers UC faculty a central, online location for depositing working papers, technical reports, research results, and conference proceedings from a wide range of disciplines. Two new features have recently expanded the eScholarship Repository's capabilities. Peer-review tools support the publication of articles, journals, monographs, and edited volumes, providing UC faculty with an alternative to publishing their research in for-profit journals. Another new feature is the seminar series service, which UC faculty can use to give their seminars, lecture series, and colloquia a lasting and highly visible presence on the Internet.

<http://repositories.cdlib.org/escholarship/>

UC Office of Scholarly Communication

In response to the continuing economic crises in scholarly communication and the broadening concern within UC about the impacts those crises are having on academic practices, the CDL has established an Office of Scholarly Communication for the university. The new program will seize opportunities to leverage technology and assist scholars in the use of new and sustainable means of scholarly communication.

<http://osc.universityofcalifornia.edu/>

III. Specific Digital Library Challenges

Utilizing Shared Infrastructure

The UC libraries have a history of successfully building shared resources and services. Looking ahead, new practices will be needed to adequately deal with library services that are collaborative and multi-institutional, and that provide access to a host of materials in a variety of formats, much of which are not acquired or owned by any single library. A new plan for systemwide strategic directions for libraries and scholarly information at UC addresses these key issues: <http://libraries.universityofcalifornia.edu/planning/>

The advances in technology that have rapidly expanded the world of information available to library users have also made it possible to design new ways to manage and provide access to those resources. "Layered" service designs offer the potential for each UC library to develop innovative and customized services without sacrificing the economies that are traditionally associated with centralized and commercial services. In a layered model, a library can develop services that use core shared infrastructure, which can include (but is not limited to) collection development tools for shared collection decisions, digital object repository services, preservation utilities, bibliographic and metadata management utilities, access customization tools, digital rights and other policy frameworks, and publishing utilities.

With leadership provided by the CDL, and through consultation with systemwide groups, the UC libraries are addressing the challenges posed by the pursuit of building and sharing a core infrastructure. Examples of these challenges include:

- i. Developing principles and guidelines for the identification, evaluation, selection, and implementation of online tools and services for sharing, accessing, and integrating scholarly content in all forms.
- ii. Developing guidelines for the type and frequency of the evaluation of tools and services, both before and after they are implemented.
- iii. Identifying and developing the resources, expertise, and staffing required for the local adoption, extension, and integration of the core infrastructure.
- iv. Identifying the potential impacts of and contributions to national developments in digital library standards vis a vis CDL layered services.

- v. Creating sustainability and co-investment strategies, including grants and other funding opportunities.

http://libraries.universityofcalifornia.edu/about/success_stories.html

IV. Digital library publications, policies, working papers, and other documents

A wide selection of guidelines, technical reports, working papers and news bulletins is available at: <http://www.cdlib.org/news/index.html>