

**PROPOSAL TO
INSTITUTE OF MUSEUM AND LIBRARY SERVICES
2001 NATIONAL LEADERSHIP GRANTS FOR LIBRARIES**

FACE SHEET

Applicant Organization: Washington Research Library Consortium (WRLC)
Address: 901 Commerce Drive
City/State/Zip: Upper Marlboro, MD 20774

Web address: <http://www.wrlc.org>

Project Director: Bruce Hulse
Director of Library Services
WRLC
901 Commerce Drive
Upper Marlboro, MD 20774
301-390-2033 (business phone)
301-390-2020 (fax)
hulse@wrlc.org (email)

Authorizing Official: Elizabeth A. Payne
Executive Director, WRLC
301-390-2031 (business phone)

Governing control: Private nonprofit

Type of organization: 09 Library consortium

Employer ID: 52-1559828

Type of project: 03 Preservation or Digitization

Project Title: A Collaborative Digital Collections Production Center

Grant amount requested: \$233,420

Amount of matching funds: \$165,709

Grant Period: 10/01/01 -- 09/30/03 (two years)

Partners: The Catholic University of America (Mullen Library)
The George Washington University (Gelman Library)

ABSTRACT

A Collaborative Digital Collections Production Center

The Washington Research Library Consortium (WRLC), a consortium of seven university libraries in the Washington, DC metropolitan area, proposes to develop a collaborative digital collections production center (DCPC) to promote the development of digital collections among its member libraries. WRLC already provides these libraries with some of the supporting digital library technologies, including digital imaging software and equipment to support image creation, file management, and image and metadata delivery. With this project, WRLC proposes to expand its services to provide the technical and organizational infrastructure necessary to create individual digital collections and to integrate them into the WRLC digital library structure. The DCPC will provide staff and systems organized to plan and manage digitizing projects, scan materials according to specifications, and design and enter metadata developed in conjunction with library staff. The goal of this project is to fill the gap in the "digitization chain" by creating a shared facility to consolidate project management, information technology, and digital conversion experience in a production environment.

The DCPC will be designed to support conversion of the most common types of archival material: single-leaf flat text, graphic materials, and black-and-white or sepia photographs. This project will use several special collections from two different WRLC member libraries as prototypes to develop DCPC services:

- Terence Vincent Powderly papers (Catholic University). Records from the Grand Master of the Knights of Labor relating to the rise of labor unions, immigration policy, and political patronage in late nineteenth and early twentieth century America. Text and photographs.
- Fenian Brotherhood records (Catholic University). Materials describing the Irish nationalist movement in America from the 1860s-1880s. Primarily text.
- Clifford Berryman political cartoons (George Washington University). Original pen and ink drawings from 1899-1949 by a political cartoonist for the Washington Post and Washington Star. Graphic materials.
- Civil War letters of Timothy Vedder (George Washington University). Letters reflecting Civil War actions in and around Washington, DC. Manuscript.

These materials were chosen not only because they meet the characteristics outlined above, but also because they represent materials to which the libraries have already given some priority through the development of finding aids and which would benefit from digitization of some or all of the source material. Ideally other collections will be offered for conversion during the course of the project, proving the utility of the shared facility.

WRLC's DCPC will be a model service to demonstrate the value of a collaborative facility to spur development of digital collections from multiple independent library special collections.

NARRATIVE

NATIONAL IMPACT

All academic libraries contain primary research material in the form of special collections and archives which would be of significant interest to students and scholars if the existence of those materials were more widely known. The digital library technologies which are becoming increasingly available to academic libraries will go a long way to making such information more accessible, but there are significant obstacles preventing libraries from taking full advantage of these technologies. While many libraries are experimenting with digital collections on an ad hoc basis (as WRLC has), it is primarily the largest research libraries with their superior funding and staffing which have been able to develop systematic organizational approaches to support ongoing conversion of special collections material to digital form.

For most libraries, the lack of time and technological expertise among the library staff is a major obstacle to implementing digital collections, even if the supporting digital library technologies are available to them. Due to the inherently unique nature of special collections materials, where each individual page can require different treatment, digital conversion of special collections is a labor-intensive process requiring a variety of skills across its spectrum of activities. This combination of skills is rarely present in any one person and all of the skills require training and continuous application to retain. Library digitizing projects are often funded through individual grants aimed at the content of specific collections, and it is difficult for a given library to maintain the necessary level of skill and staff availability between projects.

WRLC proposes to develop a collaborative digital collections production center (DCPC) to provide the technical and organizational infrastructure necessary to promote the development of digital collections among its member libraries.

ADAPTABILITY

The organizational structure, procedures, technical and workflow guidelines, and partner relationships which WRLC seeks to define through this project should serve as models for other library consortia or distributed library systems which seek to collaborate on digitizing projects. This approach should be especially useful for smaller libraries which do not individually have the resources to support in-house digitizing projects, but which nonetheless own meaningful collections and prefer to work with a library-oriented facility rather than a commercial imaging vendor. WRLC's DCPC itself could provide a useful service to many other libraries in the Washington, DC area which are not members of the consortium.

In addition, the resulting digital collections and digital library structure are designed for maximum flexibility and utility within a variety of contexts. By their nature, consortium libraries participate in multiple overlapping systems. For instance, the WRLC member libraries belong to the shared WRLC digital library system, but they also maintain their own library websites or portals to highlight their individual collections and services, and they also provide content to their university's campus-wide course-support systems and

portals. It is important to make the libraries' digital collections available across multiple environments, by using accepted standards and widely-available technologies.

DESIGN

WRLC provides information technology services to its member libraries, and has implemented a small number of pilot digital imaging projects. With this project, WRLC aims to expand its services to include the staff and equipment necessary to plan and manage digitizing projects, promulgate standards and guidelines, scan materials according to specifications, design and enter metadata, and integrate the resulting digital collections into the WRLC digital library structure.

The DCPC will be designed to support conversion of the most common types of archival material: individual single-leaf flat text, graphic materials, and black-and-white or sepia photographs. Items of this kind can appropriately be handled by a central facility and constitute a critical mass of materials needing conversion. Other materials such as microfilms and bound volumes are more likely candidates for outsourcing to commercial imaging vendors because of the need for specialized equipment or batch processing.

The project will use special collections materials from two of the WRLC member libraries as prototypes to develop DCPC services:

- Terence Vincent Powderly papers (Catholic University). Records from the Grand Master of the Knights of Labor, detailing the organization and development of labor, immigration policy, and political patronage in late nineteenth and early twentieth century America. (<http://libraries.cua.edu/manuL-Z.html#POWDERLY>)
- Fenian Brotherhood records (Catholic University). Letters, ledgers, rosters, speeches, pamphlets, and other materials from the 1860s-1880s describing the Irish nationalist movement in America. (<http://libraries.cua.edu/manuA-K.html#FENIAN>)
- Clifford Berryman political cartoons (George Washington University). Original pen and ink drawings from 1899-1949 by a political cartoonist for the Washington Post and Washington Star. (<http://www.gwu.edu/gelman/spec/collections/manuscript/berryman.html>)
- Civil War letters of Timothy Vedder (George Washington University). Letters reflecting Civil War actions in and around Washington, DC.

These materials were chosen because they are interesting collections of manageable size with characteristics representative of the kinds of materials the DCPC will be designed to support. They will provide a proof-of-concept for the digital conversion service. These materials already have collection records in the WRLC online catalog (a shared Endeavor Voyager system) and finding aids (which in some cases are online in HTML form). Thus they represent materials to which the libraries have already given some priority and which would gain the maximum benefit from digitization of some or all of the source material. The goal is to create a linked digital collection so that users can search the text of the collection record, finding aid, or image metadata and from there link to the related images.

Project activities

1. Finalize WRLC digital library architecture and guidelines

WRLC operates a multi-platform digital library system (known as ALADIN) from a dedicated computing center with a wide-area private data network supporting more than 600 workstations in the member libraries plus hundreds of additional Internet-connected users. The primary software systems in use are Endeavor Voyager for the library automation system and shared OPAC, and OCLC SiteSearch supporting local digital collections and a catalog of Web resources (in development). WRLC staff have developed and/or integrated a variety of other software products to support patron-initiated borrowing, a patron portal (myALADIN), electronic reserves, and access management and proxying (using the Voyager patron file).

WRLC plans to enhance its current digital library structure by incorporating the necessary new technologies and data standards required to support conversion of archival collections to digital formats in a production mode. While current menu structures, database systems, and search engines are expected to support this project, new metadata formats will be incorporated to ensure interoperability with external systems, and a review of potential alternatives to the current software products is underway. See attachment 2 for a diagram of WRLC's current server system configuration which conveys the scope of WRLC's information technology services.

It is envisioned that the digital collections will be stored and managed within the WRLC digital library system as five components:

- Collection record: MARC record in the Endeavor Voyager OPAC
- Finding aid: Encoded Archival Description (EAD), stored as XML, displayed via HTML
- Collection web header: HTML page to provide an addressable entry point for each collection, linked to metadata search engine (OCLC SiteSearch)
- Image metadata: Based on Dublin Core, enhanced as needed for individual collections, stored in OCLC SiteSearch database. May include OCR'ed full text as appropriate.
- Images: Stored as image files in formats appropriate to the material (see below).

Naming conventions will be finalized to insure accessibility and maintainability of these collection components.

Access will be provided through ALADIN, WRLC's digital library system (<http://www.aladin.wrlc.org>). Metadata will also be made available to external search systems through protocols like Z39.50 (supported in SiteSearch) and the Open Archives Initiative (OAI) Protocol for Metadata Harvesting.

See the Specifications for Projects Involving Digitization for further information.

2. Establish the digital collections production facility

WRLC plans to establish the digitizing facility in a currently-unused 900 square foot room in WRLC's existing building. The facility would include workspace and equipment for three staff: a digital collections manager (project manager), metadata indexer, and scanning technician. The facility would be equipped with two scanners:

- a Fujitsu M3096GX flatbed scanner (already owned), supporting 256 greyscale up to 400 dpi.
- Color duplex flatbed scanner (to be purchased), for 24-bit color up to 400 dpi or greyscale up to 1200 dpi.

Each staff member will have an image-capable workstation with a 21" color monitor to facilitate viewing images and metadata simultaneously. The scanning technician and digital collections manager will also have Adobe Photoshop software for image editing. The facility is large enough to allow for a staging area for materials with sufficient room for additional staff and/or equipment in the future.

3. Design sub-projects for each individual digital collection identified for the project.

One of the most important activities of the overall project is to formalize the project planning steps necessary for each conversion project, and to document the process and supporting information for future projects in different libraries. Project staff will draw on the best practices of other services such as the Harvard University Library Digital Imaging Group and the Higher Education Digitisation Service (U.K) to formulate these procedures.

Conversion planning activities include:

- Define selection criteria for items within that collection to fit design goals (and, for this project, time available)
- Define workflow to support scanning and metadata creation (i.e. appropriate sequence for document preparation, scanning, finding aid conversion, metadata creation, and metadata entry) within the context of that library
- Define quality control specifications appropriate for the material
- Develop sub-project schedule.

Work on separate collections may occur sequentially or concurrently.

4. Create image and metadata files

Some of these steps may be performed in a different sequence depending on the collection or other project requirements.

- At the library, select individual documents for conversion within the scope and selection criteria
- Working with special collections staff, design the administrative and descriptive metadata record structure for each collection and material type, within the context of WRLC digital library guidelines

- Scan images according to specifications appropriate for that document type.
- Enter administrative and descriptive metadata
- Create collection web header.

5. Create finding aids database

- Implement a searchable finding aids database to support EADs for storage and indexing, and web delivery via HTML
- Create EADs for finding aids related to the project collections
- Convert other existing HTML-based finding aids to EAD as time allows (Note: This project is not attempting to convert a large number of finding aids during the project period, rather it aims to develop a structure and process for subsequent conversion projects).
- Link to completed image collections from finding aids.

WRLC will most likely implement the finding aids database using existing software products such as SiteSearch or the Java application server and XML tools.

6. Develop WRLC Digital Collections Production Center Guidebook

Develop a guidebook to document procedures and guidelines to aid WRLC and participating libraries in planning for subsequent DCPC conversion projects. The guidebook would include information such as:

- Schematic diagram of WRLC digital library systems architecture
- Baseline technical standards for center-produced collections (for text, graphics, photos, and corresponding metadata)
- Recommended or typical workflow between library staff and production center staff
- Sample templates for metadata for various material types.

7. Develop recommendations for future services

This project is intended to provide baseline imaging services for some types of commonly-occurring archival materials. It is likely that the project will reveal the need for additional kinds of equipment, software, or procedures that would improve the service to be delivered to participating libraries.

Results of the project

Specific results which are expected from this project are:

- At least three digital collections from at least 2 different libraries added to the WRLC digital library

- Guidebook to provide specific guidance for subsequent digitizing projects among the WRLC libraries.
- Cost analysis of individual conversion projects to assist with future project planning and potential cost-recovery pricing
- Lessons learned and criteria for prioritizing future collection conversions. The member libraries' existing finding aids -- especially those that are already online - - provide a natural inventory of collections which are candidates for future digital conversion projects.

MANAGEMENT PLAN AND PERSONNEL

The Project Director will be Bruce Hulse, Director of Library Services at WRLC, who will devote 5% of his time over the two-year project. He will hire project staff and manage implementation of the project to insure that goals and timetables are met. Lizanne Payne, WRLC's Executive Director, and Don Gourley, WRLC's Director of Information Technology, will also participate in this project, investing 3% of their time.

WRLC will provide information technology support from the Manager of Systems (3%) and a computer programmer (25%). The IT staff will be responsible for integrating the new collections into WRLC's Sitemsearch system and/or implementing any other necessary database systems (i.e. XML support). As necessary they will implement new metadata formats, customize user interfaces, and provide access management (if required) to link the collections to the main WRLC digital library menu.

Staff to be hired for the project will include:

- A digital collections manager
- A metadata indexer
- A scanning technician.

The Digital Collections Manager will work with the participating libraries to orient staff to the WRLC digital collections architecture (i.e. database relationships, information delivery options), establish conversion schedules, and recommend guidelines for metadata and scanning specifications for the specific collections being converted. S/he will work with WRLC's IT staff to set up the necessary system files and databases, and will supervise the metadata indexer and the scanning technician to insure quality control. The Digital Collections Manager will also prepare the standards, guidelines, recommendations, and examples for the Digital Collections Production Center Guidebook.

The metadata indexer will enter metadata for each document based on its contents and the template defined for that material type. The scanning technician will scan individual documents according to the specifications identified for that material type. The scanning technician will be responsible for adhering to quality control requirements, under the oversight of the Digital Collections Manager.

BUDGET

The total project budget is \$399,128. WRLC will contribute \$165,709 (42%), and seeks \$233,420 (58%) from IMLS. Please see the budget section for details.

CONTRIBUTIONS

For this project, WRLC will provide management and information technology staff, equipment and furnishings, facility space, and computer systems support.

WRLC's partner libraries at Catholic University and George Washington University will provide library staff time to prioritize and select collection materials and to work with project staff to create and review metadata. This library staff time is being donated by the member libraries and has not been calculated as part of the project budget. These libraries have been listed as official partners in this application in order to indicate their agreement to participate in the project.

PROJECT EVALUATION

As part of its formal evaluation of the project, WRLC will survey library directors and heads of special collections in each of its seven member libraries, those which participated in the project and those which did not, to determine their assessment of the service which was developed, their plans for using the service in the future, and their recommendations for improvement.

The Digital Collections Production Center created by this project is expected to act as a catalyst to encourage additional conversion projects. The workflow designs and planning materials developed over the course of this project should facilitate future projects and render them more efficient. Success at meeting this goal could best be exemplified by a prioritized list for conversion of additional WRLC library collections not originally identified as part of the grant project.

DISSEMINATION

WRLC plans to maintain a project website and will make available a web version of the resulting Digital Collections Production Center Guidebook, along with a summary report of the project and articles describing DCPC operations. WRLC staff actively participate in professional associations and will describe the project and the center in presentations and articles. This project might be of special interest to fellow members of the International Coalition of Library Consortia (ICOLC), who share the goal of providing central services to independent members.

SUSTAINABILITY

WRLC is well-positioned to sustain both the Digital Collections Production Center and the individual digital collections created by and for WRLC libraries.

WRLC has twelve years of experience providing information technology services to its member libraries, with a staff of 18 FTE and a dedicated computing facility. Throughout that period, WRLC and the libraries have incorporated new services as needs and

technology evolved. The funding for continued support of WRLC's information systems and services is stable. Most of WRLC's services are funded by the member institutions according to a cost-sharing formula, and the members have made long-term financial commitments to the consortium in order to provide continuity for the organization and its programs. The success of this project will demonstrate the value of an ongoing digital conversion service and will encourage integration of the DCPC into WRLC's core services and operating budget.

Similarly, WRLC's digital library system will continue to provide access to these collections. As the central library automation facility for the member universities, WRLC is the natural host for digital collections archiving, and has begun to develop plans to support such a role, taking advantage of the existing computer systems architecture. WRLC will closely follow developments and best practices deriving from other institutions such as Cornell University's digital archiving plan.

SCHEDULE OF COMPLETION

Key to project activities

Initiate project (October through December 2001)

- Hire staff
- Acquire equipment
- Acquire/configure furniture

Configure and support systems (October 2001 through March 2002)

- Finalize naming conventions
- Implement SGML/XML support for finding aids database
- Define collection databases and DTDs as necessary

Plan conversions (January through March 2002)

- Confirm selection and scope of prototype collections
- Orient library project staff to WRLC digital library structure (database relationships and delivery options)
- Establish conversion schedules for each collection
- Define metadata contents and scanning specifications

Convert collections (March 2002 through September 2003, and ongoing)

For each project collection, projects can be sequential or concurrent:

- Select and prepare documents
- Scan documents according to specifications
- Create image metadata
- Convert finding aid to EAD
- Create collection web header
- Link to online "exhibit" if present

Develop guidebook (ongoing throughout, especially June through September 2003)

- Document planning issues such as collection priorities, selection criteria
- Develop recommended workflows and division of labor between library staff and DCPC staff
- Document metadata standards and recommended templates
- Document WRLC digital library structure
- Outline areas for further analysis
- Prepare final web version of Guidebook

Evaluate project (June through September 2003)

- Survey member libraries for evaluation of DCPC services
- Analyze collection conversion costs
- Prepare final report for IMLS

BUDGET NOTES

Salaries and Wages

WRLC seeks funding from IMLS for 3 positions to staff the Digital Collections Production Center:

- a digital collections manager to direct the individual digitizing projects and provide consultation to the participating libraries about scanning and metadata requirements
- scanning technician
- metadata indexer

These positions will be filled early in the first year and will be continued through the second year of the project.

All other positions (WRLC management and information technology staff) will be provided by WRLC as part of its cost-sharing.

Travel

We are requesting from IMLS \$2,000 per year for travel as suggested by the guidelines. It would be advantageous to visit a similar facility such as the Harvard Digital Imaging Group.

Materials, supplies and equipment

WRLC will provide the necessary equipment as part of its cost-share.

This project will require the purchase of 3 image-capable workstations for the DCPC staff (i.e. 21" color monitors, fast processor speed, at least 128 MB RAM). The intended room is already wired for connection to WRLC's computer systems and network, so no additional networking equipment will be required.

The budget includes funds to purchase a color duplex flatbed scanner and a laser printer. WRLC also will provide a greyscale flatbed scanner which is already owned (not reflected in the budget at all). Because of the speed of change for equipment, we have not provided detailed specifications for equipment items at this time, but will acquire the most appropriate components available upon grant award.

Other

The budget includes WRLC cost-share funds to augment and reconfigure modular furniture to provide workspaces for the three new staff. WRLC already owns a few sets of this furniture which can be used as the basis for these new work areas, but they will need to be reassembled into the proper configuration and will probably need additional sections.

Indirect Costs

WRLC's indirect costs are included in its cost-sharing.

ORGANIZATIONAL PROFILE

The **Washington Research Library Consortium (WRLC)** is a regional library consortium established to support the information and research needs of seven universities in the Washington, D.C. metropolitan area:

- American University
- The Catholic University of America
- Gallaudet University
- George Mason University
- The George Washington University
- Marymount University
- The University of the District of Columbia

The member institutions are five private and two public universities serving a total of over 70,000 students (the largest has over 20,000, the smallest about 2,000). The libraries maintain a combined total of over 5 million volumes (about 1.5 million in the largest, about 175,000 in the smallest). WRLC's primary programs are a digital library system and a specialized high-density offsite book storage facility, plus other services to support resource-sharing among the member universities.

WRLC is a not-for-profit corporation established in the District of Columbia in 1987. It is governed by a Board of Directors comprised of the presidents of the member universities, plus the president of the Consortium of Universities of the Washington Metropolitan Area, a member university treasurer, and a member university library director. Operational oversight is provided by the directors of the member libraries, supported by advisory committees of library staff.

The member universities have executed a Participant Agreement among themselves and the WRLC corporation, a legal contract establishing a relationship "in perpetuity" with provisions for withdrawal if necessary. Members share the annual WRLC operating expenses based on a Board-approved funding formula and budget. WRLC's facility was funded by a grant from the U.S. Department of Education, and built on land donated by Prince George's County, Maryland.

Direct management of this project will be provided by Bruce Hulse, Director of Library Services.

Organizational Profile -- Partner Organizations

The Catholic University of America (Mullen Library)

The Catholic University of America, located in Washington, DC, was established by the American bishops as a graduate and research center for the study of "all branches of literature and science, both sacred and profane." In 1887 Pope Leo XIII approved the plan to found CUA, the only national Catholic center of learning in the United States with a papal charter. The university began offering undergraduate programs in 1904. As the national university of the Catholic Church in the United States, CUA is the only American university with ecclesiastical faculties granting canonical degrees in theology, philosophy, and canon law.

The library system of The Catholic University of America contains about 1.5 million volumes, with primary collections in the John K. Mullen of Denver Memorial Library, plus several other campus libraries serving individual programs.

Participation in this project will be directed by Adele Chwalek, Director of Libraries, with the primary activities performed by the Department of Archives, Manuscripts, and Museum Collections under the direction of Timothy J. Meagher, University Archivist.

The George Washington University (Gelman Library System)

The George Washington University was created by an Act of Congress in 1821, and today is the largest institution of higher education in the nation's capital. The University offers comprehensive programs of undergraduate and graduate liberal arts study as well as degree programs in medicine, law, engineering, education, business/public management and international affairs. GW enrolls a diverse population of about 19,000 undergraduates, graduate, and professional students.

The Gelman Library System consists of three campus libraries serving the primary graduate and undergraduate programs (except for law and medicine which have separate libraries).

Participation in this project will be directed by Jack Siggins, University Librarian, with the primary activities performed by the Special Collections Department, led by Francine Henderson, Head of Special Collections.

Key personnel

Bruce Hulse, Director of Library Services (Project Director)

Bruce Hulse is responsible for planning, management and operation of WRLC resource-sharing programs provided to the member libraries. Bruce has been with WRLC since 1990, and serves as part of the senior management team. He supervises a staff of seven providing services in these program areas:

- Functional and user-oriented aspects of the WRLC digital library system (software applications)
- Consortial database licensing
- Offsite storage
- Intercampus library delivery service.

Over the past year he has been leading an initiative with the member libraries to define the desired contents and structure of the WRLC digital library, to design the user interface (menu design), and to develop specifications and designs for needed features. Bruce and his staff worked with member libraries to produce WRLC's initial digital collection pilot projects.

Don Gourley, Director of Information Technology

Don is responsible for overall planning, management and operation of all WRLC computing systems and networks. Along with the Director of Library Services and the Executive Director, he serves as part of the senior management team for WRLC.

Don plans and acquires computing equipment and services, manages systems and network performance, and performs application systems development and integration. He and his staff have developed for WRLC a patron portal (myALADIN), patron-initiated borrowing requests, web-delivery of articles borrowed from library collections, and an "electronic journal title finder".

Lizanne Payne, Executive Director

As Executive Director, Lizanne Payne has overall responsibility for strategic planning, program management, and budgeting for all WRLC programs. She has been with WRLC since 1989, and originally was responsible for implementation of the consortium's first shared library automation system and OPAC. She has over 20 years experience with library automation systems at WRLC and previous positions at OCLC Local Systems and the National Library of Medicine.