

Proceedings of the 9th Annual Federal Depository Library Conference

October 22 - 25, 2000

OCLC Content Management Services

John A. Hearty, OCLC
Dublin, OH

Agenda

- Dimensions of Digital Archiving
- Obstacles to Digital Archiving
- OCLC's Digital Archiving Experience
- OCLC Vision
- OCLC Strategy
- OCLC Content Management Programs
- OCLC/GPO Content Management Program
- Implementation Sequence and Schedule

Digital Archiving Dilemma

"The explosive growth of information in digital form is stretching the ability of traditional archives to ensure its accessibility and usability to future users. In addition, technology evolution is causing hardware and software systems to become obsolete in a few years, resulting in inaccessible media formats and data structures."

From the Digital Archive Directions Workshop, College Park, MD, June 1998

Dimensions of an Ideal Digital Archive

- Security
- Access
- Economy
- Openness
- Library Needs
- Publisher Needs
- End-user Needs

Dimensions: Security

- Destruction due to natural hazard

- Deterioration of physical media over time
- Technological obsolescence
- Business failure

Dimensions: Access

- Standards and tools for descriptive, structural, and archival metadata
- Provisions for sustaining accessibility over time and through stages of migration
- Contract protection for access arrangements

Dimensions: Economy

- Digital archiving must be lower than existing costs
- Opening up of shelf space/reduction in capital expense
- Greatly reduced "circulation" costs
- Elimination of duplicate content costs (physical and electronic)
- Cost recovery options for content owners and service providers

Dimensions: Openness

- Facilitates broader access
- Open linking with other front-end services and archives—expands access opportunities/reach
- Ease of adaptation as better technologies appear

Dimensions: Library Needs

- Assurance of long term access
- Low and predictable costs
- Leverage existing investments
- Choices in access, archiving, and other content management facilities

Dimensions: Publisher Needs

- Provision for compliance with intellectual property rights
- Credible protection against unauthorized use
- Subscription/usage fee models which reduce publisher risks

Dimensions: End-User Needs

- Ease of searching and browsing archival contents
- Content in context
- Integration with other information sources
- Immediate access to content from around the world

Obstacles to Implementing Digital Archives

- Standards and tools for content capture and creation (including metadata)

- Rights, permissions, and access management
- Long term preservation and access
- Library trust
- Stewardship of the cultural, historical, and scientific record is a cornerstone of librarianship
- Few organizations trusted to assume this critical role for the library community.

Standards and Tools for Content Capture and Creation

- Diverse content structures require flexibility in workflow and metadata creation facilities
- Digital content lends itself to automation of otherwise labor-intensive processes
- Standards and best practices not yet well established

OCLC Response

- OCLC CORC—Cooperative Online Resource Catalog.
 - Web-based
 - Support for MARC 21 and Dublin Core
 - Flexible resource description options
 - Descriptive and pathfinder metadata tools
 - To be upgraded to adapt to digital content management application
- Cooperative digital publishing network

Rights, Permissions, and Access Management

- Balancing desire to expand access with rights and needs of content owners
- Support for distributed, linked access environment, with a high level of object migration expected.
- Contracts to support more distant time horizon than the community is accustomed to

OCLC Response

- Independent third party with track record of working as a liaison between publishing and library communities
- 30 years of experience managing complex access control scenarios
- PURL technology to manage object migration challenge
- New Naming Service to enhance inter-service linking

Long term preservation and access

- Many unique concerns:
 - Deterioration of physical media over time
 - Technological obsolescence—data, document formats, display technologies
 - Business failure

OCLC Response

- Guarantee a process, not an outcome.
 - Establish principles to provide maximum probability of perpetual access
 - Consistent with DLF position and OAIS framework
- "Perfect solutions are the enemy of getting anything done" Clifford Lynch
- "Establish some archival repositories," with two requirements:
- Participants will adopt OAIS framework
 - Agreement on minimum requirements for a digital archival repository

From DLF Web site: <<http://www.clir.org/diglib/preserve/presjour.htm>>

Library Trust

"When faced with the reality that no vendor can provide "guaranteed forever archiving," library decision-makers are pragmatic in their assessment about this issue (e.g., they need to realize a certain "comfort level" with a vendor and simply acclimate to this realization) and are highly supportive of, and receptive to, the guiding principles OCLC has outlined regarding this issue."

Carl Hendrickson, Market Measurement

OCLC Archiving Principles

- OCLC will dedicate staff to monitor all aspects of related technology
- OCLC will escrow a percentage of service fees to fund future technology migration
- All materials archived by OCLC will be placed in safe and redundant storage facilities and routinely refreshed
- If necessary, OCLC will microfilm and store information and data deemed important by the membership or by the content owner
- OCLC will establish three advisory groups to monitor OCLC's archiving program and provide direction (1 - Computing and Storage, 2 - Software and Data Formats, and 3 - Information Technology)
- OCLC will embrace the OAIS framework, work collaboratively with the DLF and other key payers, and continue to play an active role on all International Standards Groups that influence any aspect of archiving
- OCLC will guarantee to provide to the content provider, subscriber and/or copyright owner on a usable transport and in an appropriate format all content if OCLC is unable to continue maintaining the archive

OCLC Response

- Fill the vacuum. Continue OCLC's traditional role of applying technology to address library needs to the digital archiving problem
- 68% of libraries surveyed believe OCLC will do the "best job" fulfilling this role

OCLC Digital Archiving Experience

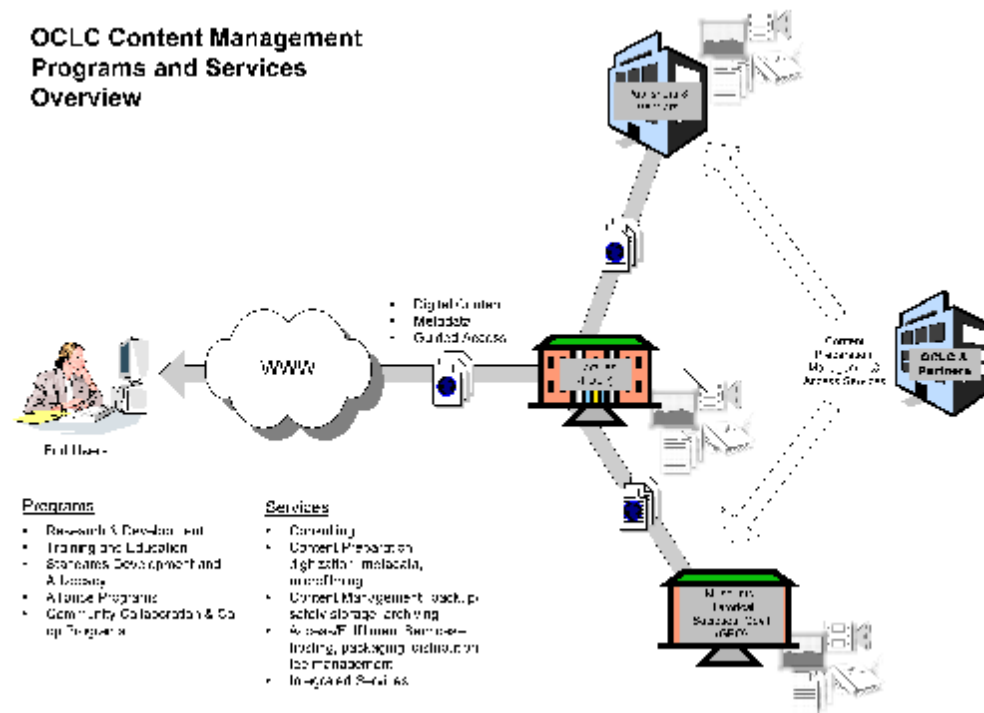
- WorldCat Backup
- RLG Collaboration
- JSTOR Safety Storage
- netLibrary Escrow Services
- ECO Archive
- Electronic Archive Pilot
- FDLP/ERIC Digital Library Pilot Project

OCLC Electronic Archiving Pilot Project: FDLP/ERIC Digital Library Pilot Project

- Partnership between GPO, NLE, and OCLC
- Two-year file of ERIC Research Reports
- Each report is approximately 133 pages of TIFF images
- Memorandum of Understanding signed
- OCLC acknowledged as GPO/FDLP partner
- Great potential for Depository Libraries

OCLC Vision for Digital Content Management Services

Help libraries and other institutions fulfill and expand their role as stewards and providers of local information, and become premier, highly sought-after destinations on the global Web.



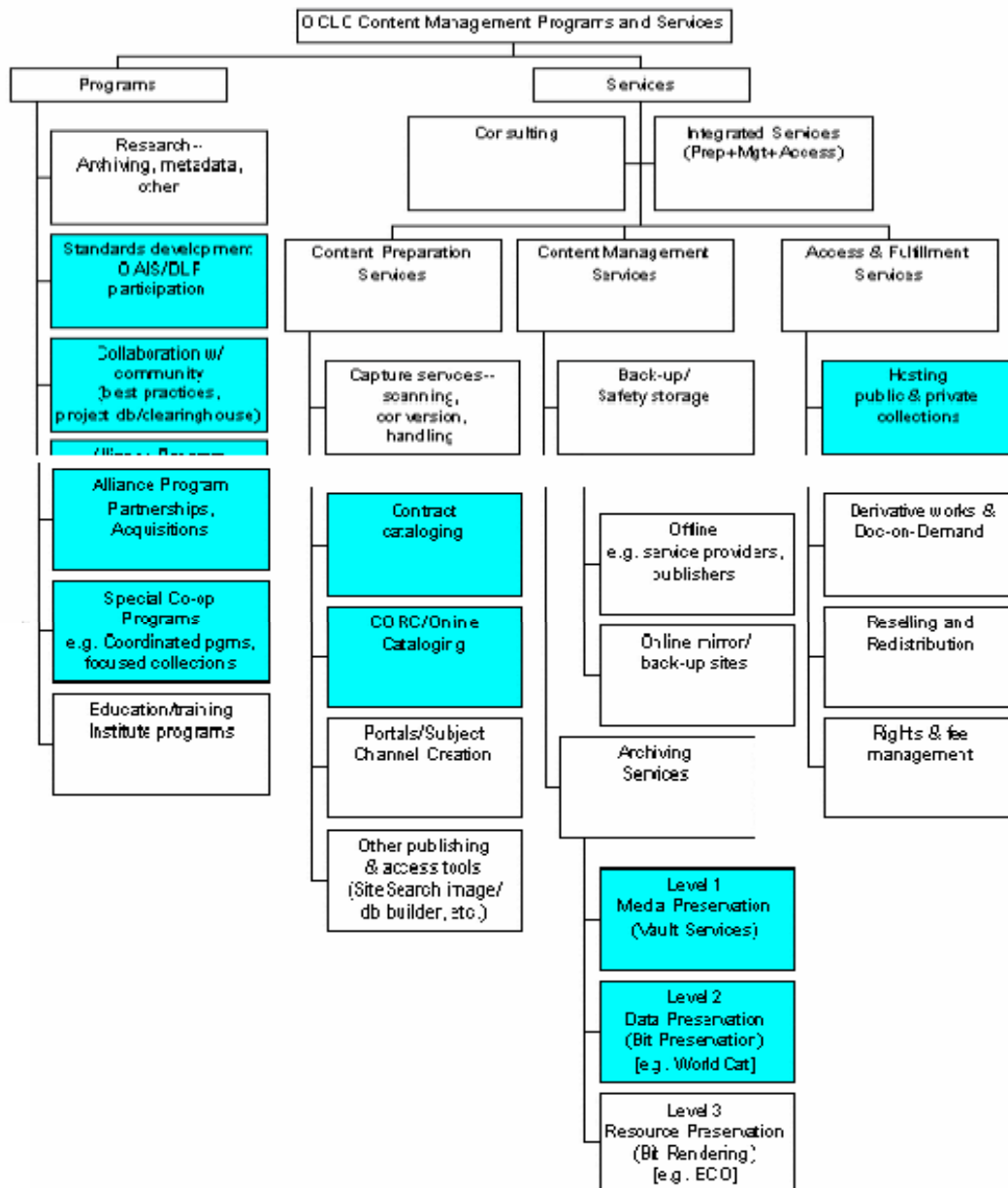


Digital Content Management Strategies

- Make digital publishing of locally controlled content easy and cost effective for all interested libraries and other like-minded institutions
- Become a major node on a global network of trusted electronic archives
- Employ a partnership-intensive program to achieve scale and tap into needed talent and other resources
- Actively promote OCLC's interest and commitment in this area

OCLC Digital Content Management Program

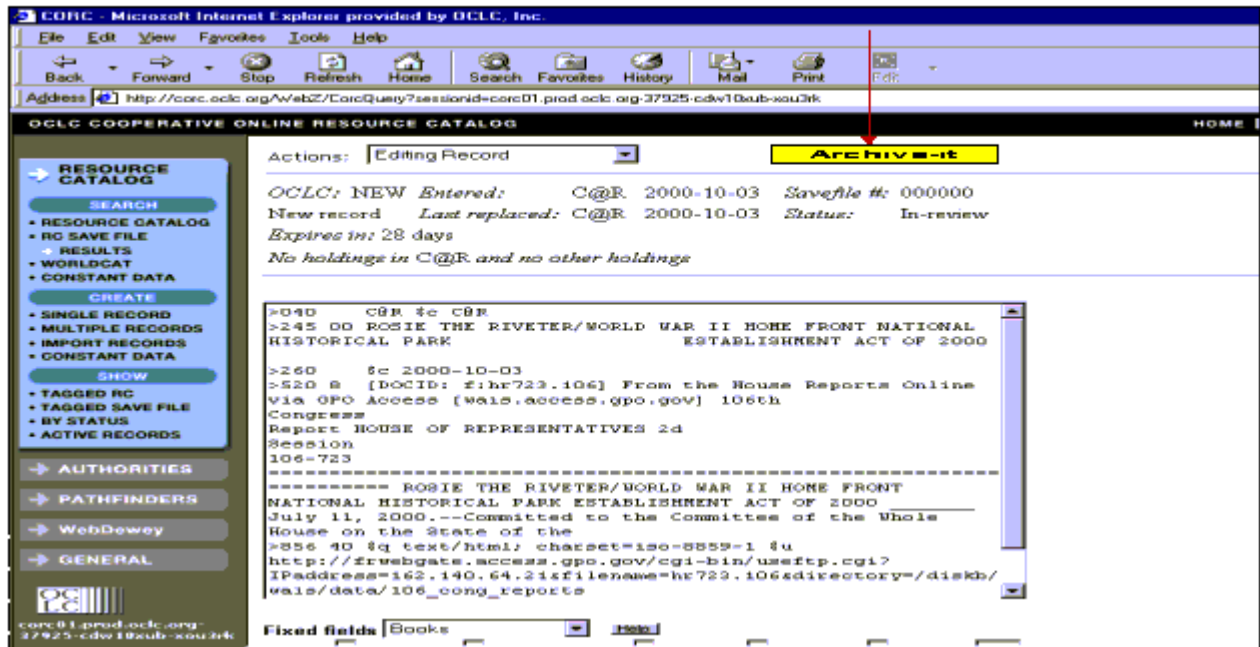
OCCL Digital Content Management Program



- Content Preparation
 - Discovery
 - Selection
 - Cataloging
 - Workflow Management

Archive-it" button on CORC interface

Archive-it" button on CORC interface



GPO/OCLC Digital Content Management Program

- Content Preparation
 - Ingest Function
 - Archival Storage
 - Data Migration & Refreshing
 - Administration
- PURL Maintenance
 - Backup/Redirection Functionality
 - Validation
 - Migration Tracking
- Access
 - GPO Access
 - WorldCat/FirstSearch
 - Library OPAC

Implementation Sequence & Schedule

3-6 months	CORC Archiving Interface with button to archive at GPO. Existing CORC selection and workflow tools minimally enhanced. PURLs manually inserted.	Archiving System development	At the end of 6 months, GPO uses the CORC-based tools to archive to GPO's archive using TeleportPro.
3-6 months	CORC Archiving Interface	Archiving System	
6-9 months	CORC development continues. OCLC PURL server enhanced.	Archive development continues.	GPO continues to use tools to archive to their archive.
9-12 months	Enhanced CORC-based features including selection, harvesting, and reporting. Archiving button expanded to include the options of archiving at OCLC. Minimal descriptive metadata records exported from CORC. Retrieval of documents from the OCLC archive.	Creation of archive metadata, storage and preservation, archive maintenance.	GPO pilot is merged into the general Electronic Archiving solution. Available to all libraries.