Implementation of OCLC's Digital Archive in North Carolina

Montana Statewide Digital Preservation Planning Meeting

May 22, 2009

Jennifer Ricker, Digital Collections Manager

State Library of North Carolina

Background Information: Organizational Structure

- Two divisions within the Department of Cultural Resources
  - The State Library of North Carolina
    - Targets a variety of statewide needs (Library Development, LBPH, NC ECHO, G&H)
    - The Government and Heritage Library provides direct services to state government employees and researchers
      - State Publications Clearinghouse collects and processes state agency created publications
      - Digital Information Management Program develops recommendations and tests solutions that support the identification, collection, cataloging, storage, and preservation of state government information for permanent public access
  - The North Carolina State Archives
    - Provides services to state government and public
      - Designated the official archival agency of the State of North Carolina
      - Public records include all formats regardless of "physical form or characteristics"

Background Information: Mandate and Responsibilities

- Separate mandates, missions, administrative structures
- Mandates overlap, often work together to solve challenges
- Statutory mandate
  - Gather, provide access to, and permanently store state agency electronic publications and records in North Carolina

	Pubs	Websites	Geospatial Data	E-mail	Databases	Other Records
Archives		√	√	√	√	√
Library	√	√			√	

Background Information: IT

- **Information Technology**
  - DCR-IT provides desktop support
  - ITS provides network support
  - No defined support for long-term management of born-digital state records at this time

Background Information: Requirements of SLNC

- **Must be hosted**
- **Must be easy to use administratively**
  - Due to size and makeup of staff paraprofessionals will be entering state publications and basic metadata
    - fit in existing workflows
    - integrate with existing tools
- **Must be standards based**
  - Must allow for data to be easily imported/exported in a standard format
- **Must be secure and reliable (preferably OAIS compliant)**
- **Must provide meaningful reporting**
- **Must ensure long-term access (preferably more than just bit-level preservation)**

Comparison of Digital Archive to SLNC Requirements

SLNC Requirements	Must be hosted	Must be easy to use administratively	Must be standards based	Must be secure and reliable	Must provide meaningful reporting	Must ensure long-term access
Digital Archive	OCLC hosts the storage in Dublin, Ohio	Content is automatically added either directly to the Digital Archive if we are using CDI or to a volume that can be shipped to OCLC and added to the Digital Archive at any time. Integrated with Connexion and CONTENTdm which are tools we already use.	The Digital Archive supports the mandatory elements of PREMIS for its metadata.	OCLC manages the storage environment by checking storage media life and file integrity (virus, fixity, format verification), backing up all files in off-site storage locations), and verifying inventories. Is OAIS compliant.	Provides a monthly accession report, health report for stored files, file integrity report, and a dissemination report.	Only provides bit-level preservation, but that is good enough for now.
	✓	✓	✓	✓	✓	✓

Storage for SLNC

- Hosted storage
  - Began using summer 2007
  - Current storage: ~100 GB
  - Add content two ways (CDI and bulk)
  - Content backed up in multiple locations
- Local storage
  - Local Storage Area Network (SAN)
  - Content backed up by ITS (off-site)
 - Limited space, funds for growth

CDI Workflow

- the WorldCat record is automatically modified to include a link to the CDM record in the 856 field

- a copy of the file is **<u>automatically</u>** passed to and stored in DA

- the WorldCat record is automatically crosswalked from to DC and the file and the new DC record are added to CDM

  - DC metadata is manually edited to include additional info. (e.g., creating/submitting Agency)

Bulk Workflow

- Files **<u>automatically</u>** stored in Volumes when loaded to CDM

- On schedule, **<u>manually</u>** move volumes to external drive (or have the volumes housed there from the beginning) and create manifest

- Before shipping external drive to OCLC (optional)
  - Validate checksums
  - Add metadata (update manifest)

- Ship external drive to OCLC so files (including metadata files) can be added to DA

Integration of Tools
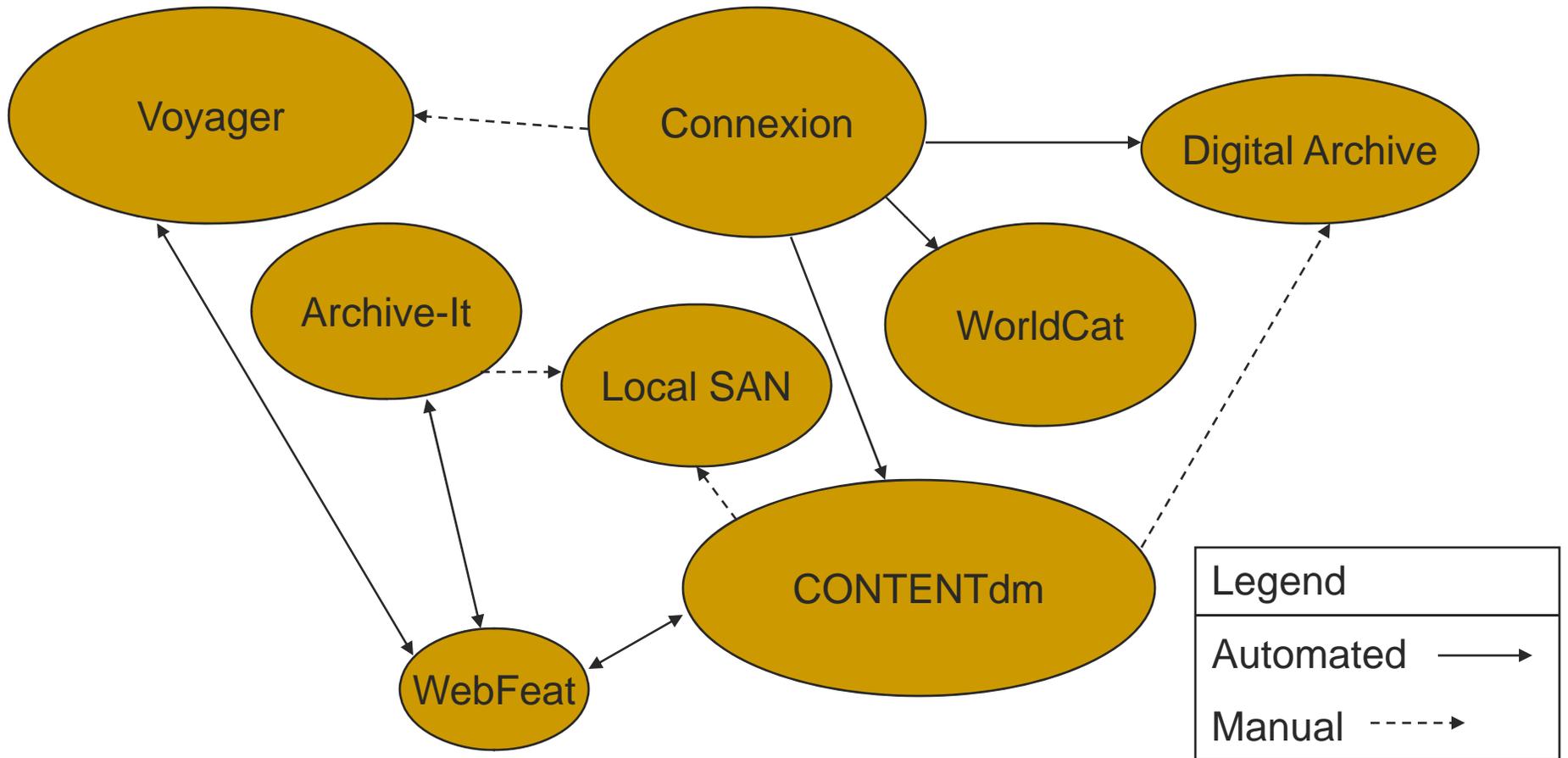
Pros and Cons

- **Pros**
  - Meets stated requirements
  - Requires no additional staffing
  - Integrates with other tools
  - Works!
- **Cons**
  - Not a joint solution with the Archives
  - Only provides bit-level preservation
  - Expensive

Other Options

- Considered by SLNC
  - CEP (not hosted)
  - ENCompass (too cumbersome, no longer available)
- Under Consideration by SLNC
  - Washington State Digital Archives
    - NDIIPP grant through Oct. 2009, extension requested
    - Establish the Washington State Digital Archives (WSDA) model as one that can be replicated in other states and/or as a viable hosting service
    - No North Carolina publications submitted yet
  - iRODS
    - NHPRC grant through May 2011
    - Build a distributed production preservation environment that meets the needs of archival repositories for trusted archival preservation services
    - Currently identifying required rule set for distributed data management environment and loading Archive-It and CDM files

Funding

- **LSTA grant funds for testing**

- **Dedicated state budget funds for subscriptions**

  - Cost share joint subscriptions with Archives

Other Considerations

- **State Agency participation**
  - ○ Could be accomplished via CDM (project clients or Admin Module)
- **DCR-wide Solution (SLNC, State Archives, and NC ECHO)**
  - ○ Ideally need a solution that works for all of us