The Canadian Information Network for Research in the Social Sciences and Humanities

Westell, Mary E.; Au Yeung, Tim

http://hdl.handle.net/1880/48762
Presentation

Downloaded from PRISM: https://prism.ucalgary.ca
The Canadian Information Network for Research in the Social Sciences and Humanities

http://www.synergiescanada.org

Tim Au Yeung and Mary Westell
Libraries and Cultural Resources
University of Calgary
March 5, 2010
“Synergies is a national project . . . whose intent is to provide a platform for the potential transformation of the Canadian scholarly record.”

(DEVAKOS AND TURKO, 2007)
BACKGROUND

• A national project funded by CFI, with matching contributions from provincial and other agencies

• For the Prairies, our matching comes from the Alberta government, vendor in-kind, and “in-cash” matching from the libraries

• The project is four years and began in 2007
SYNERGIES GOALS

• Build capacity for online publishing and scholarship in the Social Sciences and Humanities

• Provide access to peer-reviewed journals, data sets, theses, conference proceedings and scholarly books
SYNERGIES GOALS

• Through common standards and interoperability mechanisms
  – Provide a platform for access to publicly funded research
  – Build a technical infrastructure that will support digital preservation
SYNERGIES NATIONAL PARTNERS

- University of New Brunswick
- Universite de Montreal
- University of Toronto
- University of Calgary
- Simon Fraser University
• University of Calgary
• Athabasca University
• University of Winnipeg
• Affiliate members
### Synergies Services

<table>
<thead>
<tr>
<th>Publishing Services Workflow</th>
<th>Dissemination Services</th>
<th>Long-term Preservation and Permanent Referencing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-publishing tools (all scholarly genres and data)</td>
<td>Interoperability Link to bibliographic databases</td>
<td></td>
</tr>
<tr>
<td>General public Web interfaces Search tools, etc.</td>
<td>Digital Warehouse Storage, Distribution and Archives</td>
<td>Copyright</td>
</tr>
<tr>
<td>Research-specific Web interfaces Data mining, etc.</td>
<td>Metadata Preparation and Distribution</td>
<td>e-Commerce? Individuals or Institutions</td>
</tr>
</tbody>
</table>
DIGITAL PRESERVATION
DIGITAL PRESERVATION

“Digital preservation combines policies, strategies and actions to ensure access to reformatted and born digital content regardless of the challenges of media failure and technological change. The goal of digital preservation is the accurate rendering of authenticated content over time.”

ALA Association for Library Collections & Technical Services, Preservation and Reformatting Section, June 24, 2007
DCC DIGITAL OBJECT LIFECYCLE MODEL
ROLE OF LIBRARY (TRADITIONAL)

Create → Publish → Access → Preserve → Withdraw
ROLE OF LIBRARY (SYNERGIES)

Create → Publish → Access → Preserve → Withdraw
DIGITAL LIFECYCLE CHALLENGES

Loss of context
Authenticity
Inaccessible Formats

Create

Lost
Multiple Versions
Access Disruption
Digital Corruption

Publish

Format obsolescence
Media obsolescence

Access

Preserve

Withdraw

Loss of context
Authenticity
Loss of masters

Legacy Versions
# DIGITAL LIFECYCLE CHALLENGES

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss of context</td>
<td>Metadata</td>
</tr>
<tr>
<td>Authenticity</td>
<td>Guidance</td>
</tr>
<tr>
<td>Inaccessible Formats</td>
<td>Persistent Identification</td>
</tr>
<tr>
<td>Lost</td>
<td>Perpetual Access</td>
</tr>
<tr>
<td>Multiple Versions</td>
<td></td>
</tr>
<tr>
<td>Access Disruption</td>
<td></td>
</tr>
<tr>
<td>Digital Corruption</td>
<td></td>
</tr>
<tr>
<td>Loss of masters</td>
<td></td>
</tr>
<tr>
<td>Format obsolescence</td>
<td></td>
</tr>
<tr>
<td>Media obsolescence</td>
<td>Central Archive</td>
</tr>
<tr>
<td>Legacy Versions</td>
<td></td>
</tr>
</tbody>
</table>
METADATA

• Packaging metadata (METS)
• Technical metadata
• Provenance metadata (PREMIS)
GUIDANCE

• Best practices for creators and publishers
• Recommendations for
  – File formats
  – Creation of descriptive metadata
  – Retrospective digitization standards
• Central information website
PERSISTENT IDENTIFICATION

• Persistent identification allows
  – Distinguishing between versions of content
  – More durable location

• Requires
  – A standard means of assigning an identifier
  – An assignor authority responsible for preventing duplicates
  – A resolver authority responsible for location resolution
PERSISTENT IDENTIFICATION

- Publisher
  - Article is "published"
  - Publishing system determines local persistent location
  - Article presents permalink
  - Article local persistent location changes
  - Article is withdrawn

- Persistent ID Assignor
  - Publisher requests PID
  - PIDA returns PID
  - PIDA checks database to avoid duplicate
  - PIDA adds PID to database to ensure it is never assigned again and publisher to record

- Persistent ID Resolver
  - Publisher requests persistent location
  - PIDR updates PIDA with responsibility tag
  - PIDR generates a persistent tiny URL
  - PIDR add object information to record to support PID schema operators
  - PIDR updates access location for persistent location
  - PIDR generates tombstone page
PERPETUAL ACCESS

• Goals
  – Multiple copies synchronized across multiple geographic locations
  – Redundant modes of access in event of network / local failure
  – How? Secure storage and distributed access

• Approach
  – Private LOCKSS network (PLN)
  – Canadian administration of configuration
CENTRAL ARCHIVE

- A central archive for master files allows for:
  - Reduced resource requirements for file/media migration
  - Greater awareness of threats
  - Resource sharing for preservation activities
  - Envisioned to be a “cloud” service
OVERVIEW OF INFRASTRUCTURE

Synergies LOCKSS Network

Persistent Identification System

Publishing System (OJS, Erudit)

Archival Package (PREMIS, METS)

Central Archive

Preservation Activities

Guidance Site

Creators
Publishers
Repository Managers
PRESERVATION DELIVERABLES

– Develop a central resource for creators and institutions to support creating preservable content
– Establish a set of technical minimum standards for depositing into the Synergies platform
– Create a network of persistent identifier servers to provide long term access handles
PRESERVATION DELIVERABLES

• Establish secure storage for content
  – Create/identify plugins for common platforms like DSpace and OJS to preserve content and ancillary materials
  – Deploy a private LOCKSS network for Synergies nodes
  – Deploy archival servers for master files
PRESERVATION DELIVERABLES

• Validate the Synergies infrastructure in terms of trusted digital repository certification

• Ensure future accessibility
  – Establish a technology watch system to monitor key formats, applications and standards
  – Identify and/or develop tools to aid in place migration of content, including distributed migration
QUESTIONS?
WWW.SYNERGIESCANADA.ORG