Digital Preservation In Academic Libraries: Recent Canadian Initiatives

Andrew Waller

Wayne Johnston
What We’re Going To Be Talking About - Andrew

• What is digital preservation?
• Why is digital preservation necessary?
• The COPPUL Private LOCKSS Network (PLN)
• Other PLNS, including Synergies and the Alabama Digital Preservation Network (ADPNet)
What We’re Going To Be Talking About
- Wayne

• Trusted Digital Repositories
  • Overview
  • Scholars Portal TDR
  • Library and Archives Canada TDR

• Research data management
  • Background
  • the role for libraries
  • Agri-Environmental Research Data Repository
What Is Digital Preservation?

• Short definition: “Digital preservation combines policies, strategies and actions that ensure access to digital content over time.”

• Slightly longer definition: “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.”
Why Is Digital Preservation Necessary?

- Stuff (hurricanes, floods, etc.) happens
- Hardware breaks
- Software crashes
- Files get corrupted
- Publishers fold, merge, disappear
- Somebody hits the wrong button
- Somebody drops something
- Burning files to CDs and DVDs isn’t enough
- It costs time/money/hassle to restore lost collections
It Can Happen To You!
(Freezing and flooding, University of Calgary, Calgary, 2011)
COPPUL

• COPPUL=Council of Prairie and Pacific University Libraries
• www.coppul.ca
• Regional university library consortium in Western Canada; same sort of thing as OCUL
• 22 members in British Columbia, Alberta, Saskatchewan, and Manitoba, plus some associate members within the COPPUL area and elsewhere in Canada
• Almost all publicly-funded institutions
COPPUL PLN - History

• Came out of the COPPUL Collections group in spring 2006
• Task group investigated options
• LOCKSS was selected – a distributed digital preservation solution
• Two-year pilot with eight initial partners
• COPPUL paid for some of the LOCKSS Alliance membership costs initially
• Has been an official COPPUL program since start of 2010
• One of two PLNs in Canada
• Learned/borrowed much from ADPNet
COPPUL PLN – What Do We Preserve?

• Collections of local or regional interest that would not be preserved elsewhere

• This includes:
  – Locally-hosted OA journals, esp. those using OJS
  – Locally-digitized collections
  – Small university press publications
  – Journals with a western or local focus
  – Websites and online resources that form an important part of a member’s collection
COPPUL PLN – The Basics

• Closed or “dark” archive
• LOCKSS nodes at all nine member institutions
• Each institution contributes and archives content
• Self-supporting since 2010
• Not too big yet - 521 archival units – but much more is on deck

• [http://www.coppul.ca/pln.html](http://www.coppul.ca/pln.html)
• [http://coppullockssgroup.pbworks.com](http://coppullockssgroup.pbworks.com)

OLA Super Conference 2012
COPPUL PLN - Governance

- The COPPUL PLN falls under COPPUL’s umbrella – is an official program of COPPUL
- Steering committee – 1 official rep from each member institution; others also at meetings
  - Chair is Andrew Waller, University of Calgary
- Technical committee – drawn from member institutions; reports to steering committee
  - Chair is Mark Jordan, Simon Fraser University
- Report twice a year to the COPPUL directors
- Work closely with ADPNet and the Synergies PLN
COPPUL PLN - Membership

• Open to all full COPPUL members
  – Athabasca University*
  – Simon Fraser University*
  – University of Alberta*
  – University of British Columbia*
  – University of Calgary*
  – University of Manitoba*
  – University of Saskatchewan*
  – University of Victoria
  – University of Winnipeg*

*=pilot phase participant
COPPUL PLN – Costs

• No PLN fees, just need to be able to pay LOCKSS Alliance fee (no financial support from COPPUL beyond the pilot phase)
• Local hardware, software, and staff costs
• Communication: Near-monthly Skype calls involving steering committee
• Essentially no travel costs (some conferences)
COPPUL PLN – Some Issues

• Adding more content
• Keep plug-in development going e.g. CONTENTdm, DSpace
• New members – maybe one or two?
• Relationship with COPPUL Digital Curation and Preservation Task Group
COPPUL PLN - Summary

• A success – PLN is established, functional, stable
• Still can grow – there’s lots of content that meets the criteria that can be added
• Still can develop
• Job is not done (as if it ever will be)
Other PLNs

• Synergies: http://www.synergiesprairies.ca/preservation

• Alabama Digital Preservation Network (ADPNet): http://adpn.org/

• PLN list (partial): http://www.lockss.org/lockss/Private_LOCKSS_Networks
Some Resources (part 1)

- LOCKSS: [www.lockss.org](http://www.lockss.org)
- Synergies Prairie Node: [http://www.synergiesprairies.ca/](http://www.synergiesprairies.ca/)
Some Resources (part 2)

OAIS compliance

• Organizational infrastructure
• Digital object management
• Technical infrastructure and security risk management
trusted digital repository

scholars portal is in the process of developing a robust long-term archiving platform for all its collections, beginning with the 23,000,000+ articles housed on the portal. working from concepts set out in the open archival information model and codified in the trustworthy repositories accuses (trac) checklist, scholars portal seeks to create a digital preservation repository that will ensure the long-term integrity and usability of the objects in its collection.

our commitment

we commit to ensure the integrity, authenticity and usability of the objects in its collection now and in the future.

digital preservation

our way: ocul’s trusted digital repository

stephen marks, digital preservation librarian, scholars portal; karl nilsen, digital preservation assistant; arianne steinman, digital preservation assistant; scholar’s portal

discuss the process by which scholars portal has moved toward its accreditation as a trusted digital repository seeking to join the ranks of other well-known repositories such as portico and hathitrust. this presentation will center around tdr requirements and how scholars portal had changed to meet those requirements, through both the adoption and rejection of common tools and protocols. convenor: mary claire vandenburg, queens university
Digital Initiatives at LAC

Digital Projects

LAC Trusted Digital Repository

Abstract

In all domains the amount of digital information is increasing at a rapid rate, this raises crucial questions of preservation. Our intellectual capital, as laid down in educational, scientific, public, cultural and other intellectual resources, is increasingly at risk by the volatile character of digital objects and the rapid developments in information technology.

The growing need for adequate management and preservation of digital information is recognized in many stakeholder communities. The Library and Archives Canada (LAC) is committed to being a Trusted Digital Repository (TDR), to provide reliable and long-term access to the digital documentary heritage of Canada.
Research data management

- Accelerates scientific progress
- Increases the visibility and impact of research
- Ensures compliance with funding agency policies
- Avoids duplication of research
- Enables replication and verification of research results
- Enhances collaboration

Research Data: Unseen Opportunities
Canadian Association of Research Libraries, 2009
OECD Principles and Guidelines for Access to Research Data from Public Funding
Canadian Institutes of Health Research
www.cihr-irsc.gc.ca

Natural Sciences and Engineering Research Council of Canada
www.nserc-crsng.gc.ca

Social Sciences and Humanities Research Council
www.sshrc-crsh.gc.ca

OLA Super Conference 2012
At the Watershed: Preparing for Research Data Management and Stewardship at the University of Minnesota Libraries
Leslie M. Delserone

“Like many research universities, the University of Minnesota (UMN), and its Libraries, attempts to understand the nature and intensity of data produced by its researchers, and address the management and stewardship of its institutional research output. The recent activities of the Libraries, in support of and in conjunction with campus-wide efforts, illustrate a set of approaches that show the Libraries to be at the watershed, integrated into and contributing leadership to the growing river of campus-wide exploration and planning of cyberinfrastructure needs.”
All too often the results of expensive and time-consuming research as represented by rich data sets are lost due to the absence of sound data management plans. Redundant research is undertaken because the previous research data is no longer available. Opportunities for analysis of data across time are lost along with the historical data sets. Even when data has been properly stored and preserved it benefits no one if it isn't easily discovered, retrieved and repurposed.
Agri-environmental Research Data Repository: work plan

- needs assessment
- data audit
- environmental scan
- standards, policies, protocols
- data capture and migration
- data discovery and retrieval
- data citation and linking
- research data technician
- pilot the system
- move into production
- promotions and outreach
- sustainability strategy
- service level agreement
Agri-environmental Research Data Repository: needs assessment, data audit

- Interviews with key stakeholders
- Email broadcast to all researchers
- Public education component

“When I read your letter I got pretty excited as the lack of dissemination and use of data generated in the 'public' sector has always worried me. [...] If the Final Report is already stored at the Library, and there is already a system of cataloguing and searching it, why not entrust the same institution to store the data.”
Agri-environmental Research Data Repository: early data sets

- Agrometeorology
- Capacity development
- Environmental geology
- Food processing
- Food safety
- Soil physics
The Dataverse Network
An Open-Source Application for Publishing, Citing and Discovering Research Data
Created at IQSS at Harvard University

you and your Library
Agri-environmental Research Data Repository

Agri-environmental research data repository Dataverse

OLA Super Conference 2012
Metals in the Human Environment Strategic Network (MITHE-SN)

Search within this community and its collections:  
Go  
Advanced Search

Principal Investigator, Dr. Beverley Hale, University of Guelph

The MITHE Strategic Network is a collaboration of academia, government, and industry. The Network conducts research in support of science-based environmental and human health risk assessments for metals in water, soil and food, within well articulated and planned inter-disciplinary research themes.

The research program is strongly linked across Canada, with field sites in the Maritimes, Quebec, Ontario, The Prairies, British Columbia, the Yukon and Nunavut. The Network also features strong links among academic and government scientists. Implicit in the MITHE-SN approach is a commitment to joint, interactive, centralized planning, project accountability for both intellectual and financial objectives, and regular reporting of research progress beyond the usual publication in scientific journals.

Collections in this community

- Aquatic Ecosystems [9]
- Foods and Ingested Particles [7]
- Publications [1]
- Soils and Plants [6]
Questions? Discussion?

Andrew Waller
Licensing and Negotiation Librarian
Open Access Librarian
Libraries and Cultural Resources
University of Calgary
waller@ucalgary.ca
(403) 220-8133 voice
(403) 284-2109 fax

Wayne Johnston
Head, Research Enterprise and Scholarly Communication
University of Guelph
wajohnston@uguelph.ca
(519) 824-4120 Ext. 56900