
Private LOCKSS Networks: Overview and Working Examples

Aaron Trehub, Auburn University Libraries
Andrew Waller, University of Calgary

Lyrasis
Staying On TRAC Post-Workshop Webinars
March 10, 2011

Digital preservation: a definition

“Digital preservation combines policies, strategies and actions that ensure access to digital content over time.” (Source: American Library Association, Association for Library Collections & Technical Services, Preservation and Reformatting Section, Working Group on Defining Digital Preservation)

Digital preservation: another definition

Digital preservation is the flipside of digital collection-building.

You can't have one without the other.

(Or you shouldn't.)

Why is digital preservation necessary?

Here are a few reasons...



Floods (University of Hawaii Library, 2004)



Hurricanes (Hurricane Katrina, 2005)



Tornadoes (Enterprise, Alabama, 2007)



Floods, again (Cedar Rapids Public Library, Iowa, 2008)



Freezing and flooding (University of Calgary, Calgary, Alberta, 2011)



Digital preservation is necessary because...

- Stuff (hurricanes, 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
-

One solution: Distributed Digital Preservation (DDP)

- Multiple copies of digital content...
 - At geographically dispersed locations...
 - With 75-125 miles between them...
 - Preferably out of hurricane pathways...
 - And across different power grids...
 - With different system administrators...
 - On live media, with audits for “bit rot”...
 - And in at least 3 copies (LOCKSS = 6 copies)
-

DDP + LOCKSS = PLN

- LOCKSS = “Lots of Copies Keep Stuff Safe”
 - Open-source software, developed at Stanford University
 - Originally designed for e-journals
 - Now being used to preserve locally created content in Private LOCKSS Networks (PLNs)
 - List of member libraries at <http://www.lockss.org/lockss/Libraries>
-

Partial list of PLNs

- MetaArchive Cooperative (USA-Brazil-UK)
 - Council of Prairie and Pacific University Libraries (COPPUL: Canada)
 - Persistent Digital Archives and Library System (PeDALS: USA)
 - Data Preservation Alliance for the Social Sciences (Data-PASS: USA)
 - LOCKSS-und-KOPAL-Infrastruktur-und-Interoperabilität (LuKII: Germany)
-



DDP/PLN in Alabama: ADPNet



THE PROJECT

In 2006, the Institute of Museum and Library Services (IMLS) awarded a two-year National Leadership Academic Libraries (NAAL) and seven Alabama institutions. These seven institutions are building a preservation network for the state of Alabama using open-source LOCKSS ("Lots of Copies Keep Stuff Digital Preservation Network (ADPNet) will preserve locally created digital assets and will demonstrate a solution for academic institutions, state agencies, and community cultural heritage organizations. Aaron Technology at Auburn University, is the project director. The participating institutions are the Alabama Department of Archives and History, Auburn University, Spring Hill College, Troy University, the University of Alabama, the University of Alabama at Birmingham, and the University of North Alabama.

By the end of the two-year project, institutions of higher education, libraries, state agencies, and other cultural heritage institutions in Alabama should have an increased awareness of the importance of preserving significant digital assets. The project will also show that a LOCKSS preservation network can support the digital preservation needs of different types of institutions. A representative body will govern the network and assure the long-term operation of ADPNet.

Building local digital collections with a LOCKSS box provides libraries with mission-critical assets that grow over time. How can libraries make the most of this valuable asset? How can they best influence how these assets grow or influence what publishers and genres are included? How can they take advantage of knowledge and best practices gained through real-world experience of others?

Having multiple copies of our content on a geographically dispersed network gives us more confidence that it will still be around in five years, ten years, or - given weather conditions down here - next week.

- Aaron Trehub, Director of Library Technology, Auburn University, Alabama

Join the LOCKSS Alliance.

Participating Institutions: Alabama Department of Archives and History, Auburn University, Spring Hill College, Troy University, the University of Alabama, the University of Alabama at Birmingham, and the University of North Alabama.

NAVIGATION

- Home
- Participating Libraries
- Publishers and Titles
- How It Works
- For Libraries
- For Publishers
- LOCKSS Alliance
- Talks
- Publications
- Installing LOCKSS

“...let us save what remains: not by vaults and locks which fence



ADPNet's mission

To create a low-cost, low-maintenance, geographically distributed digital preservation network for libraries, archives, museums, and other cultural memory organizations in Alabama.



Background

- Inspired by Auburn's experience with the NDIIPP MetaArchive Cooperative
- Two-year IMLS grant: September 2006 through September 2008
- Grant administered by Network of Alabama Academic Libraries (NAAL) in Montgomery
- Project director at Auburn University Libraries
- Commitments from seven Alabama institutions



Member institutions

1. Alabama Department of Archives & History
2. Auburn University
3. Spring Hill College
4. Troy University
5. University of Alabama
6. University of Alabama at Birmingham
7. University of North Alabama
8. ?



Basic info about ADPNet

- ADPNet is a statewide PLN for Alabama
- Closed or “dark” archive: for preservation ONLY, not access or display
- LOCKSS servers (“nodes”) at all seven member institutions
- Each institution contributes content for harvesting and archiving by the network
- Self-supporting since 2008



Basic info, continued

- ADPNet currently contains 200 digital collections (“archival units”) from all seven member institutions
- Almost 3 terabytes harvested to date
- Network capacity: 8 terabytes at each node, expandable to 16 terabytes
- Plenty of room for more collections
- More collections on the way, including audio and video files



Membership

- Membership is open to all libraries, archives, and museums in Alabama that have publicly available digital content
- Four membership levels:
 1. Anchor (base fee: \$4,800/year—1.5TB)
 2. Host (base fee: \$2,400/year—500GB)
 3. Participant-Large (base fee: \$700/year—1.5GB)
 4. Participant-Small (base fee: \$300/year—500MB)
- ...Plus fees for additional storage space



Governance

- ADPNet is part of Network of Alabama Academic Libraries (NAAL)
- Two committees:
 1. ADPNet Steering Committee (policy)
 2. ADPNet Technical Committee (hardware/software)
- Member institutions are represented on the ADPNet Steering Committee
- Governance policy and technical specs available at <http://adpn.org/>



Costs and fees

- ADPNet membership fee or LOCKSS Alliance Fee
- Servers: LOCKSS server and Web server (for making content available to the network)
- Storage upgrades (as network grows)
- Staff time: typically several hours per month
- Communication: e-mail listserv, monthly conference calls, annual business meeting
- Travel: mostly in-state



The future

- Add more content to the network
 - Recruit more member institutions in Alabama, including...
 1. Public libraries
 2. Museums
 3. Other cultural heritage organizations
 - Swap servers with PLNs in other states and countries (COPPUL in Canada?)
 - Promote ADPNet model
-

DDP/PLN in Canada: COPPUL

- COPPUL=Council of Prairie and Pacific University Libraries
 - Regional university library consortium in Western Canada
 - 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
 - www.coppul.ca
-

COPPUL PLN - History

- Came out of the COPPUL Collections group in spring 2006
- Task group investigated options
- LOCKSS was selected
- Two-year pilot with eight initial partners
- COPPUL paid for some of the LOCKSS Alliance membership costs
- 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 – Basic info

- Much the same as ADPNet
 - 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 – 916GB, 432 archival units – but much more is on deck
-

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 - 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
 - Also update the COPPUL Collections and Electronic Resources group
 - Work closely with ADPNet and the Synergies PLN
-

COPPUL PLN – Costs and fees

- 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 – same sort of thing as ADPNet
 - Communication: Near-monthly Skype calls involving steering committee
 - Essentially no travel costs (conferences)
-

COPPUL PLN – The future

- Add more content + keep plug-in development going
 - Server swap – University of Alberta will be hosting an ADPNet box
 - New members – maybe one?
 - Relationship with COPPUL Digital Curation and Preservation Task Group
 - PLN meeting in Vancouver, Oct. 24-25, 2011
-

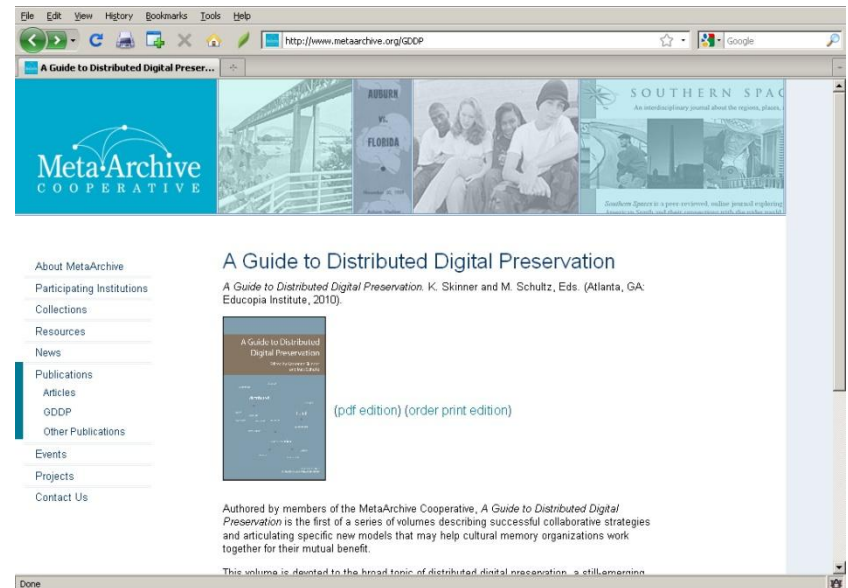
DDP/PLN: Will it work for you?

What you'll need:

- Quorum of at least six institutions
 - Hardware (LOCKSS boxes)
 - In-house IT support
 - Governance structure
 - Enough \$\$\$ to cover membership and/or LOCKSS fees (your institution might already be a member!)
 - Support from your directors (or equivalent)
 - Commitment of a few years to get things going
 - Enough sweat equity to keep it running
-

Before you get started...

...Check out
MetaArchive's
"A Guide to Distributed
Digital Preservation"



Guiding principles

- Keep it simple (if you can)
 - Keep it cheap (ditto)
 - Go with a lightweight governance structure
 - Aim for low maintenance and low administrative overhead
 - Take advantage of existing structures and relationships (in Alabama: NAAL; in Canada: COPPUL)—and learn from them
-

Closing thought

Digital preservation isn't free, but it's a lot cheaper than having to rebuild your digital collections.

For more information...

ADPNet: <http://adpn.org/>

COPPUL PLN: <http://bit.ly/gVh6uu>

LOCKSS: <http://www.lockss.org/>

MetaArchive Cooperative: <http://www.metaarchive.org/>

MetaArchive GDDP: <http://www.metaarchive.org/gddp>

Aaron Trehub
Auburn University Libraries
trehuaj@auburn.edu
334-844-1716

Andrew Waller
University of Calgary
waller@ucalgary.ca
403-220-8133
