LOTS OF COPIES KEEP STUFF SAFE

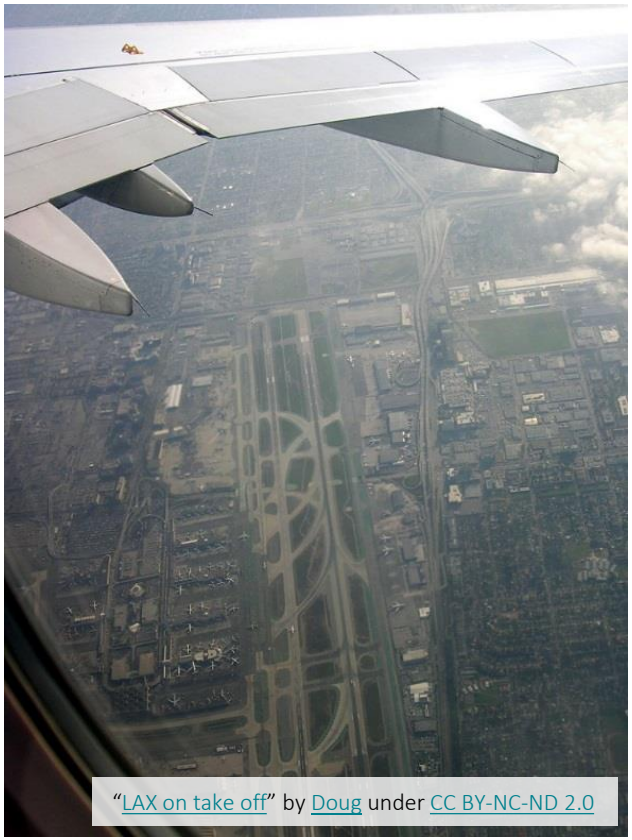
LOCKSS for COPPUL, and Lots of LOCKSS Keeping Stuff Safe

Nicholas Taylor (@nullhandle)
Program Manager, LOCKSS and Web Archiving
Stanford University Libraries

COPPUL Webinar
14 March 2017

overview

- how does LOCKSS align w/ COPPUL's digital preservation goals?
- LOCKSS ties to diverse technology ecosystems (and why that matters)
- growth + maturation of LOCKSS community + networks



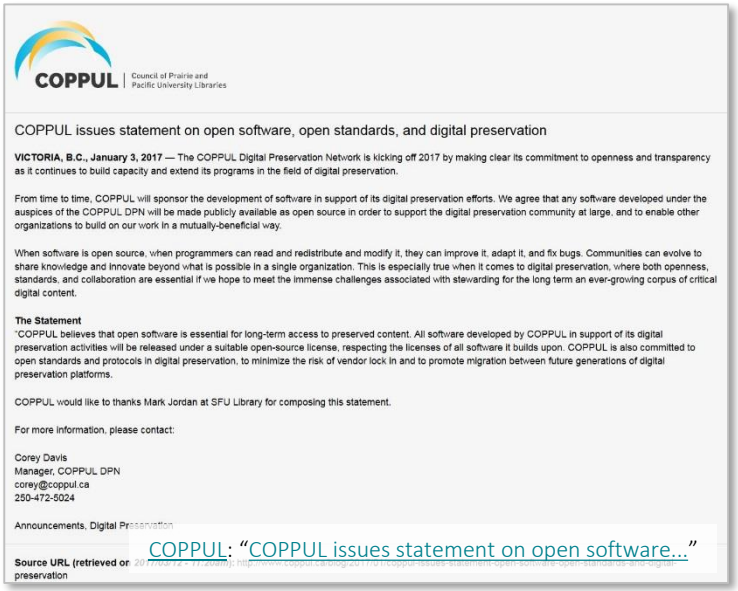
COPPUL Strategic Directions 2016-2018: Digital Preservation

- COPPUL will **develop digital preservation options for COPPUL members** and will assume a **leadership role in national discussions on digital preservation**.
- Strategies:
  - Create a **sustainable governance and funding model** for the COPPUL Digital Preservation Network (DPN).
  - Create online resources and provide educational and professional development opportunities to **increase the capacity of COPPUL members to undertake digital preservation**.
  - **Modernize the COPPUL Private LOCKSS Network** (PLN) and provide other preservation storage and service options.
  - Pursue agreements with allied organizations and programs to **align COPPUL's preservation activities with other regional, national, and international initiatives**.

COPPUL: "Strategic Directions 2016-2018"

COPPUL statement on open software, open standards, + digital preservation

*"COPPUL believes that **open software is essential for long-term access to preserved content**…COPPUL is also **committed to open standards and protocols** in digital preservation, to **minimize the risk of vendor lock in** and to **promote migration** between future generations of digital preservation platforms."*



COPPUL | Council of Prairie and Pacific University Libraries

COPPUL issues statement on open software, open standards, and digital preservation

VICTORIA, B.C., January 3, 2017 — The COPPUL Digital Preservation Network is kicking off 2017 by making clear its commitment to openness and transparency as it continues to build capacity and extend its programs in the field of digital preservation.

From time to time, COPPUL will sponsor the development of software in support of its digital preservation efforts. We agree that any software developed under the auspices of the COPPUL DPN will be made publicly available as open source in order to support the digital preservation community at large, and to enable other organizations to build on our work in a mutually-beneficial way.

When software is open source, when programmers can read and redistribute and modify it, they can improve it, adapt it, and fix bugs. Communities can evolve to share knowledge and innovate beyond what is possible in a single organization. This is especially true when it comes to digital preservation, where both openness, standards, and collaboration are essential if we hope to meet the immense challenges associated with stewarding for the long term an ever-growing corpus of critical digital content.

The Statement

"COPPUL believes that open software is essential for long-term access to preserved content. All software developed by COPPUL in support of its digital preservation activities will be released under a suitable open-source license, respecting the licenses of all software it builds upon. COPPUL is also committed to open standards and protocols in digital preservation, to minimize the risk of vendor lock in and to promote migration between future generations of digital preservation platforms.

COPPUL would like to thanks Mark Jordan at SFU Library for composing this statement.

For more information, please contact:

Corey Davis
Manager, COPPUL DPN
corey@coppul.ca
250-472-5024

Announcements, Digital Preservation

Source URL (retrieved on 2017-01-17 11:20am**):** http://www.coppul.ca/blog/2017/01/coppul-issues-statement-open-software-open-standards-and-digital-preservation

COPPUL: "COPPUL issues statement on open software…"

LOCKSS

COPPUL digital preservation principles

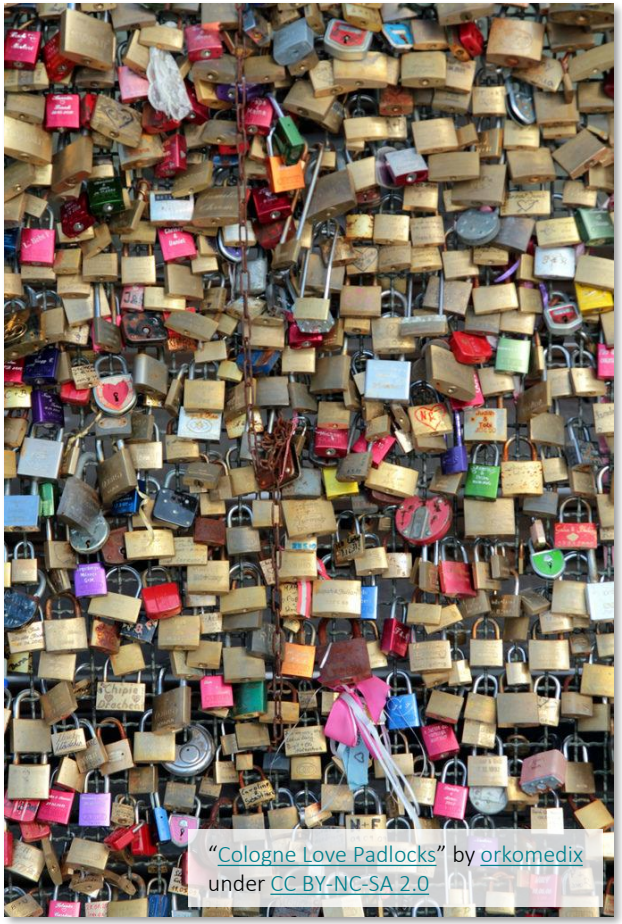
- be deliberate about business model
- value the network
- partner w/ other communities
- apply the best available tools
- commit to open technologies

LOCKSS

LOCKSS Alignment

lots of LOCKSS

- LOCKSS (principle)
- LOCKSS (program)
- LOCKSS (software)
- Global LOCKSS Network
- Private LOCKSS Networks
- CLOCKSS



community-centric

- preservation is an **active** community effort
- lots of **communities** keep stuff safe
- enable preservation of the **content your community cares about**
- **enable** libraries to be **digital libraries**

LOCKSS

open-source software

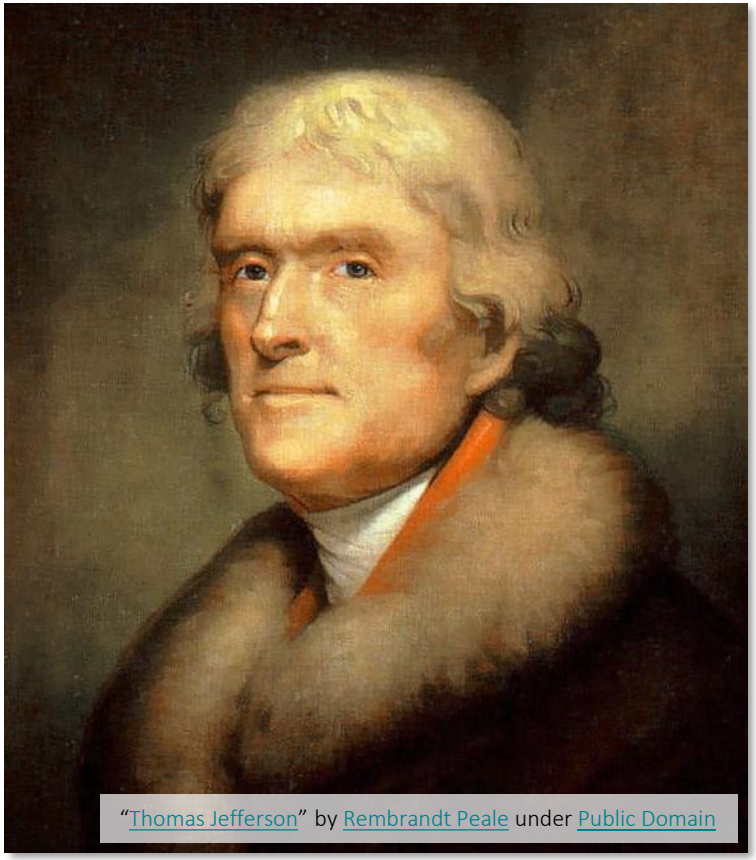
- complements digital preservation
- mitigates lock-in
- facilitates data portability
- builds on open standards
- enables collaboration
- enhances security
- empowers adopters

"Teaching Open Source Practices, Version 4.0 (high res)" by opensource.com under CC BY-SA 2.0

lots of copies

"The lost cannot be recovered; but let us save what remains: not by vaults and locks which fence them from the public eye and use, in consigning them to the waste of time, but by such a multiplication of copies, as shall place them beyond the reach of accident."

National Archives: "from Thomas Jefferson to Ebenezer Hazard, 18 February 1791"



"Thomas Jefferson" by Rembrandt Peale under Public Domain

LOCKSS

decentralized copies

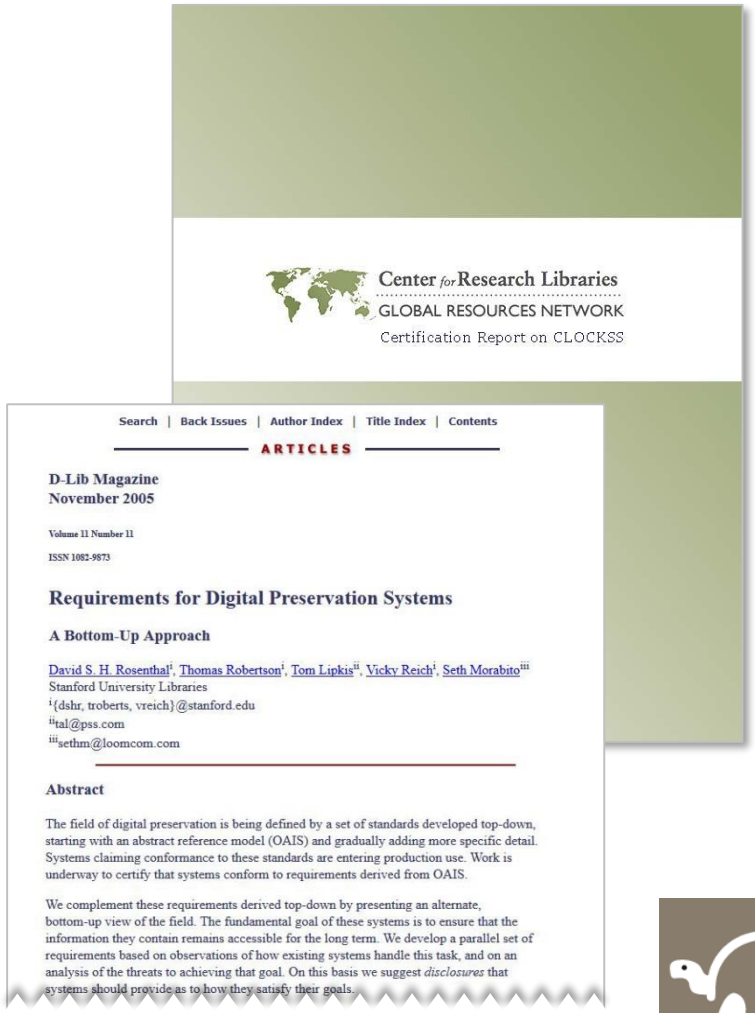
- no monopoly on copy-making
- independent, de-correlated copies
- no central point of failure or vulnerability
- local custody, self-determination

articulated threat model

- long-term **bit integrity** is a hard problem
- more (correlated) copies **doesn't necessarily** keep stuff safe
- don't underestimate:
  - people making **mistakes**
  - **attacks** on information
  - organizational **failure**

community-validated

- built upon peer-reviewed research
- successfully operating for almost 20 years
- CRL TRAC assessment of CLOCKSS
  - overall score matching previous best
  - only perfect technology score awarded to date



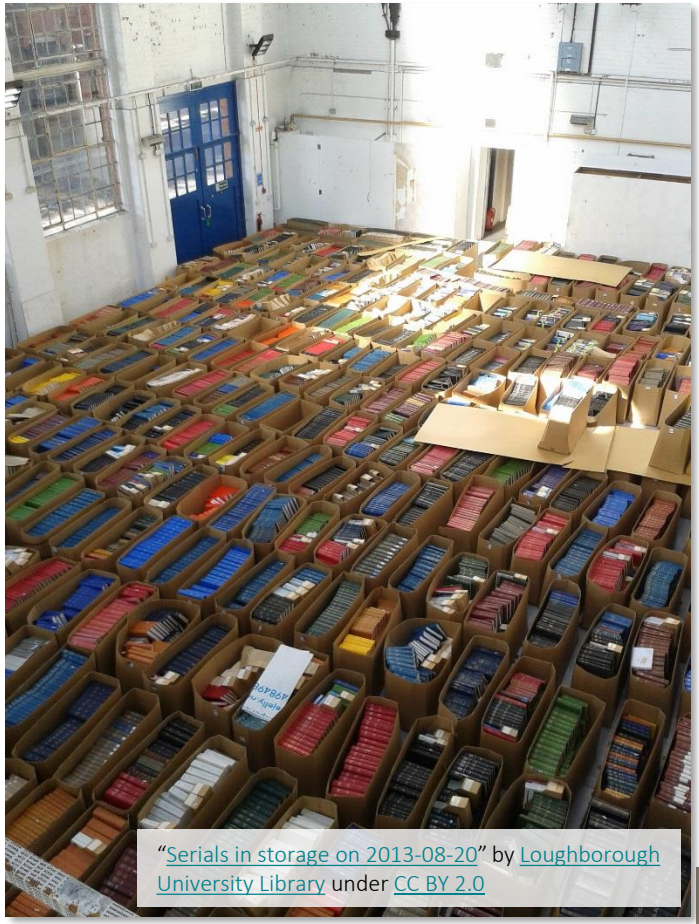
LOCKSS

affiliate, collaborate, interoperate

- ties to:
  - publishing
  - digital cultural heritage
  - web archiving
- cross-pollinate best practices + technology
- reduce costs by sharing
- better keep pace w/ evolving technology

publishing

- original use case
- still a core capability
- software, workflows to ingest wide range of genres from 1000+ publishers
- workflows handle content received via FTP, web harvest, bags
  - multiple approaches needed to support range of publishers: small, large, local, international, subscription, open access

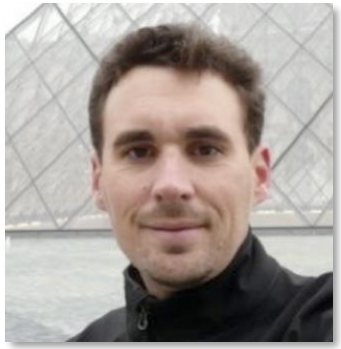


digital cultural heritage

- area of growing application
- all kinds of digitized + born-digital materials
- interconnection opportunities:
  - Archivematica
  - Fedora
  - LOCKSS-O-Matic
  - Mirador

LOCKSS + DLSS administrivia

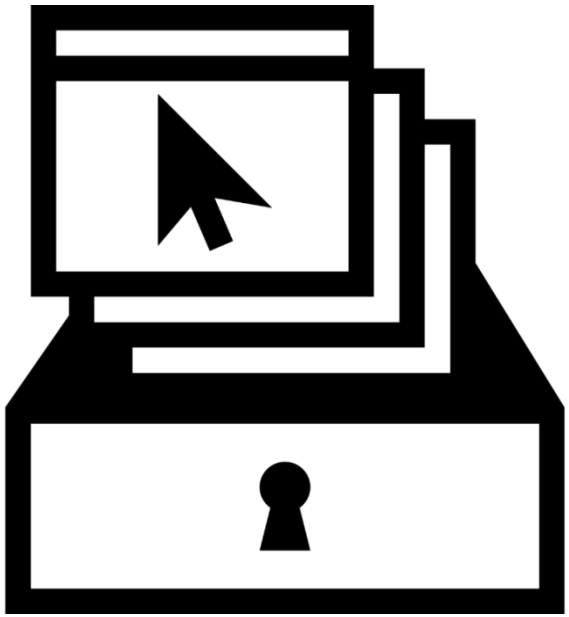
- LOCKSS has joined SUL Digital Library Systems & Services (DLSS)
- led by Tom Cramer, Director & Assistant University Librarian
- LOCKSS + SUL Web Archiving, under Nicholas Taylor



"SPO.101514.SLIDERlathrop.jpg" by Michael Hong

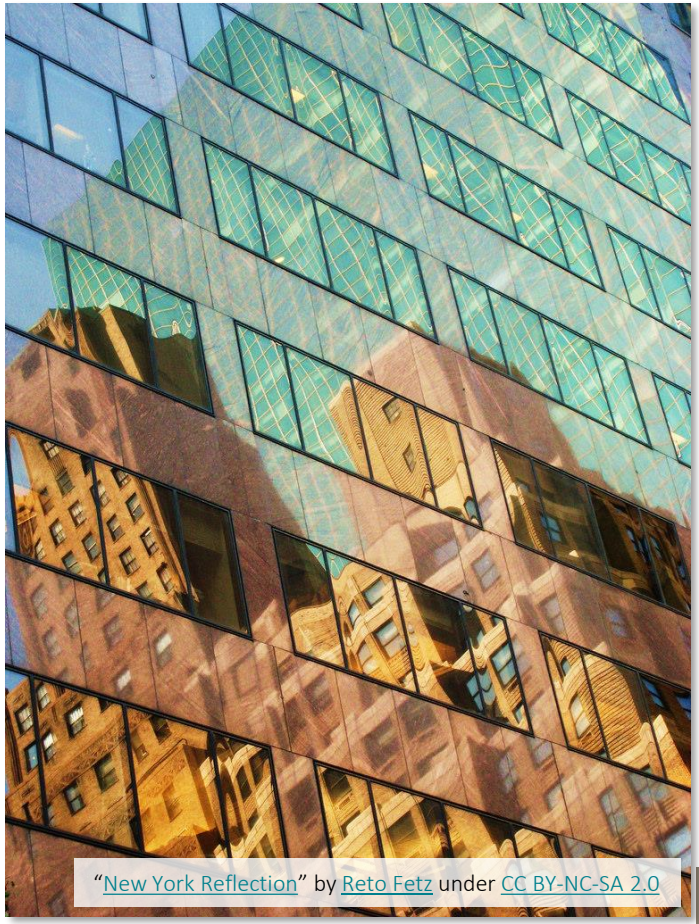
web archiving

- at its core, a web archiving system
- aligning w/ web archiving mainstream
- collaborating w/ Internet Archive on WASAPI
  - data transfer APIs
  - other candidate APIs
  - community framework?

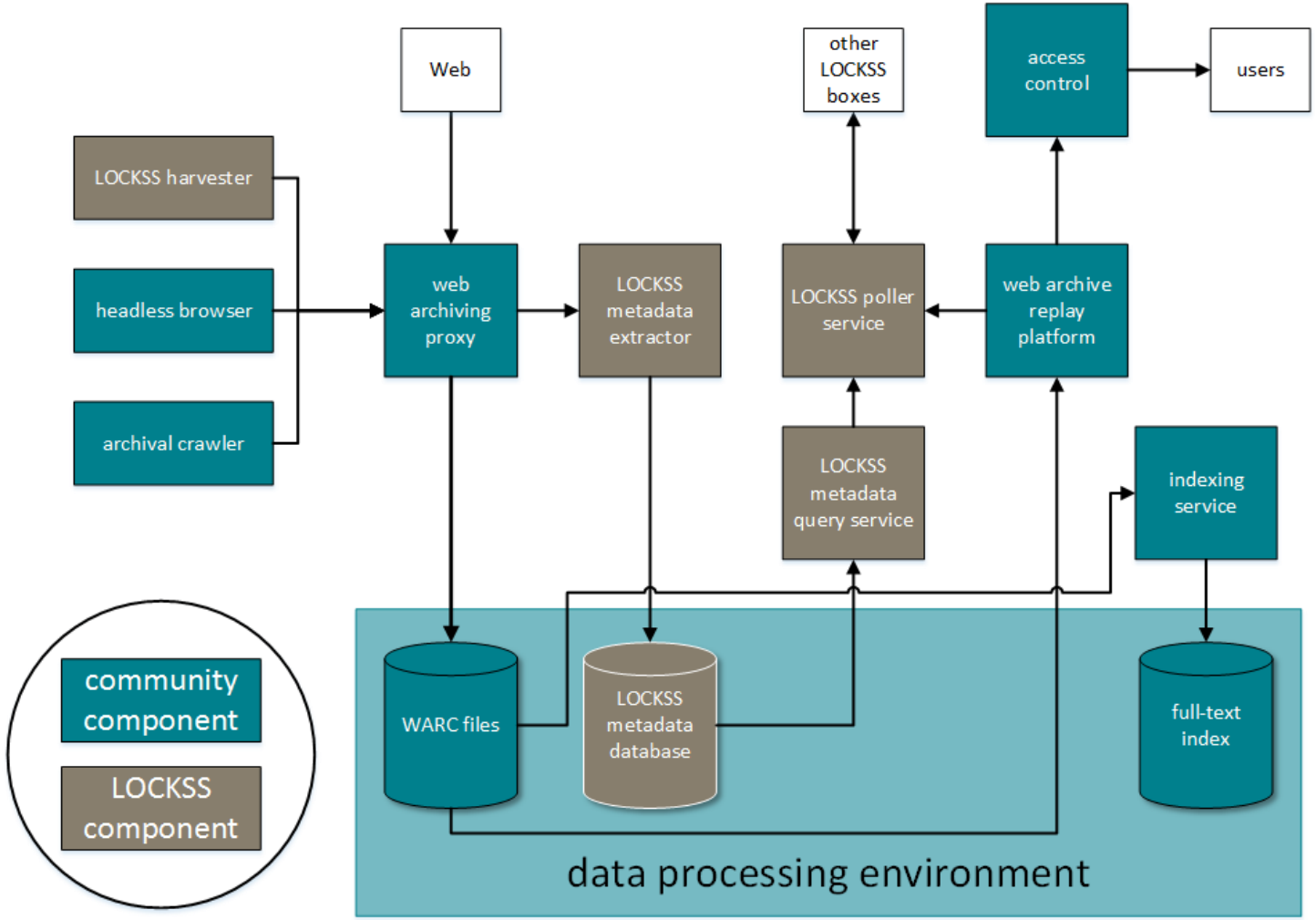


software re-architecture

- reduce support + operations costs
- de-silo components + enable external integration
- prepare to evolve w/ the Web



leveraging community components



lots of LOCKSS networks

- tens of networks
- hundreds of institutions
- all types of content
- preservation via diverse:
  - technologies
  - institutions
  - networks

LOCKSS

maturing networks

- large regional networks
- “super-node” setup
- integrating w/ other systems + workflows
- organizing to share best practices + technology
- tuning business + service models

LOCKSS

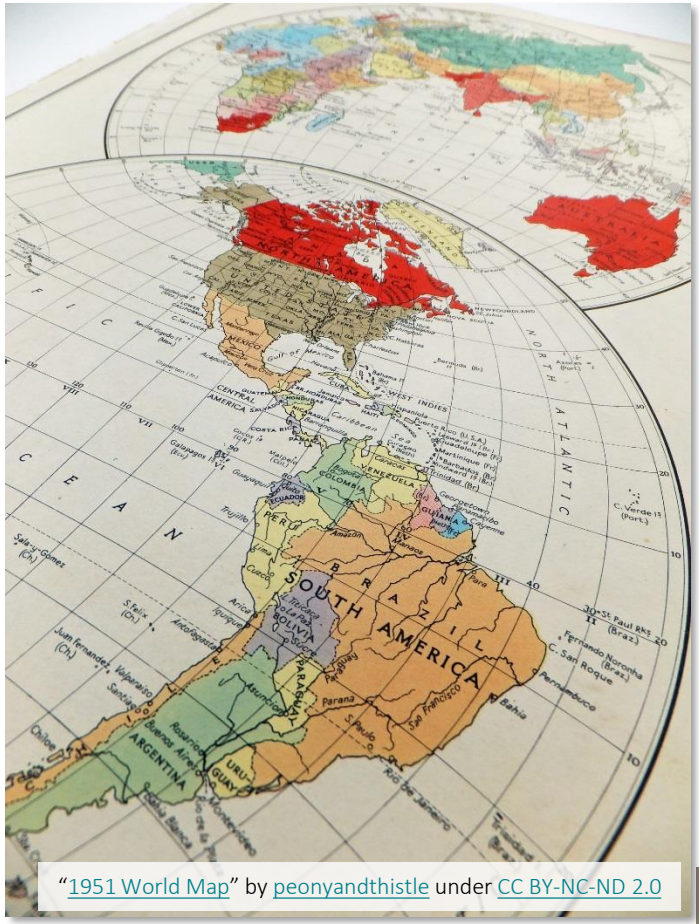
new LOCKSS networks

- Ivy+ Preservation Network
- Perma.cc
- Software Preservation Network
- national networks

LOCKSS

national networks

- KeepSafe Europe
- also Mexico, Russia, U.K.
- motivated by:
  - a more uncertain global political climate
  - local content, local preservation
- providing:
  - in-country preservation
  - perpetual access
  - non-consumptive use
  - collaboration framework for long-tail content



improving network support

- create **documentation**
- enable **self-setup**
- evolve **business model** for LOCKSS Program support
- support **community collaboration**
- hire a Partnerships Manager

connecting w/ new communities

- streamline system←→network **data exchange**
- promote **API-oriented** architectures
- **contribute upstream** to shared tools
- broaden, diversify **community outreach**

LOCKSS

takeaways

- LOCKSS value proposition is **open-source software** for **distributed**, **resilient**, **local** digital preservation
- LOCKSS technical development informed by historical ties + **new opportunities for interconnection**
- LOCKSS programmatic focus shifting to better serve a **growing, diversifying community** of networks + institutions

questions for you

- **how can LOCKSS support COPPUL**'s strategic directions around digital preservation?
  - how can we **ensure our business models are complementary** + mutually sustainable?
  - **what technical capabilities does COPPUL need** to provide flexible + robust digital preservation services?
  - how can we **foster distributed digital preservation capacity** + knowledge throughout the network?
  - **what communities could we be reaching out to** + collaborating with more actively?