COLLABORATIVE, SELF-GOVERNED, SECURE DIGITAL PRESERVATION WITH LOCKSS

Use the chat box at the right of the screen to tell us who you are, where you're from, and who is participating with you today.

(To open the chat window, click on the CHAT icon in the upper right corner.)

Connect to the audio portion of the webinar through your phone line or through VoIP.

INSTITUTE *of* **Museum** *and* **Library** SERVICES

CoSA

SERI
State Electronic
Records Initiative

COUNCIL OF STATE ARCHIVISTS

WELCOME!

Spring 2019 State Electronic Records Initiative webinars

- April: Blockchains for Public Recordkeeping

- May: IQ and You: South Carolina's approach to Governor's Records in a Proprietary File Format

- June: Bit by Bit: Alabama State Electronic Records Project Update

Kathryn Baringer
Maryland
State Archives

INSTITUTE of **Museum**and**Library** SERVICES

CoSA

SERI
COUNCIL OF STATE ARCHIVISTS
State Electronic Records Initiative

SERI ONLINE RECORDINGS

- **Managing Digital Content Over Time**

  - Identify
  - Select
  - Store
  - Protect
  - Manage
  - Provide access

- **https://archives.utah.gov/community/SERI/SERI-resources.html**

TODAY'S PRESENTER

Nicholas Taylor

Stanford Libraries

WIRED SUBSCRIBE

LILY HAY NEWMAN SECURITY 03.30.18 09:18 AM

THE RANSOMWARE THAT HOBBLED ATLANTA WILL STRIKE AGAIN

ZDNet MENU US

MUST READ: The Industrial Internet of Things: A guide to deployments, vendors and platforms

Hackers wipe US servers of email provider VFEmail

Hackers did not ask for a ransom. VFEmail described the incident as 'attack and destroy.'

By Catalin Cimpanu for Zero Day | February 12, 2019 -- 10:59 GMT (02:59 PST) | Topic: Security

Recommended Content:
White Papers: Guide to Antivirus (AV) Replacement: What You Need to Know

Amazon cloud crash wipes out customer data; Will users be compensated?

By **Molly McHugh** — Posted on April 28, 2011 10:57 am

gt

SECURITY

Ohio's Recent Spate of Cyberattacks Is Indicative of the National Trend

Local governments are increasingly having to face that they are not exempt from damaging and costly ransomware attacks.

BY WILL GARBE, DAYTON DAILY NEWS / JUNE 4, 2018

FOX NEWS Login Watch TV

STORY · Published January 24, 2008 · Last Update January 13, 2015

Angry Employee Deletes All of Company's Data

By | Fox News

Vox

FUTURE PERFECT   EXPLAINERS   MORE

"Climate change" and "global warming" are disappearing from government websites

The deletions follow a pattern of policy changes on climate change under the Trump administration.

By Umair Irfan | Updated Jan 11, 2018, 12:30pm EST

SHARE

Museum and Library SERVICES
CoSA State Electronic Records Initiative

UNDERSTAND + MITIGATE THREATS

long-term **data integrity** is hard

needs architecture informed by **actual leading threats** to data

don't underestimate:

- people making **mistakes**
- **attacks** on information
- organizational **failure**

"Fragile" by Garrett Coakley under CC BY-NC 2.0

WHAT IS LOCKSS?

a widely-accepted **principle** to protect data

a digital library-focused **program** of Stanford Libraries

a research-informed **software** app for p2p digital preservation

an international **community** of institutions + networks

INSTITUTE *of* **Museum**and**Library** SERVICES

CoSA

SERI State Electronic Records Initiative

COUNCIL OF STATE ARCHIVISTS

MORE THAN LOTS OF COPIES

lots of copies is **necessary but not sufficient**

central points of failure **can undermine all copies at once**

LOCKSS provides:

- continual **integrity checking + repair**
- b/t mutually-distrusting, **independent peers**
- on a network that **your community controls**

"Domino's" by david pacey under CC BY 2.0

ROUTINE AUDIT + REPAIR

ensuring **long-term data integrity**

- must **read data** to know it's good
- easier to **repair data sooner**

network nodes **conduct polls** to validate integrity of distributed copies

more nodes = **more security**

- more nodes can be **down**
- more copies can be **corrupted**
- …and polls will **still conclude**

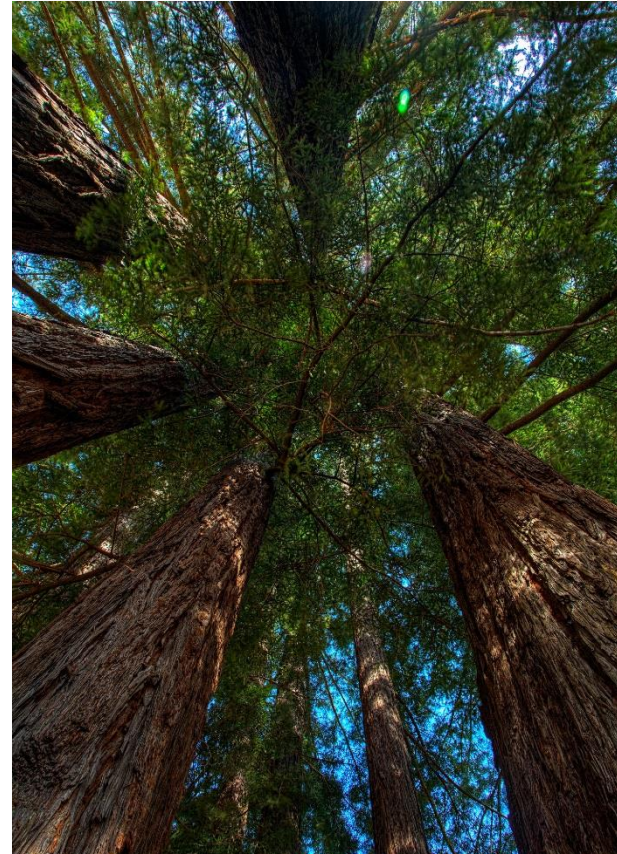"DSC_4346" by Dennis Jarvis under BY-SA 2.0

COMMUNITY + DIGITAL PRESERVATION

communities **complement** LOCKSS:

- **resilience** against organizational failure
- native **heterogeneity**

preservation is an **active** community effort

sustain **local, distributed** capacity

lots of **communities** keep stuff safe

"Redwood Canopy" by Floyd Stewart under CC BY-NC-SA 2.0

HOW LOCKSS WORKS

DISTRIBUTED DIGITAL PRESERVATION

align w/ **best practice**

achieve **resilience** not possible w/ centralized solution

for use either:

- as **dedicated** preservation solution
- to **supplement** local preservation (e.g., for most important materials)

particular to LOCKSS among service providers:

- strongly **research-based**
- articulated **threat model**
- supports **local custody**

"Stone stacks" by Jack Malvern under CC BY-NC 2.0

INSTITUTE of **Museum** and **Library** SERVICES

CoSA

SERI State Electronic Records Initiative

COUNCIL OF STATE ARCHIVISTS

CONTENT LIFECYCLE

ingest content

- web harvest, OAI-PMH, direct interconnect, or drag-and-drop via LOCKSS-O-Matic

manage content

- web-accessible GUI to monitor preservation activity (+ select new content for archiving, in some networks)

preserve content

- each node retrieves content independently
- once stored, audit + repair takes place automatically, on ongoing basis

deliver content
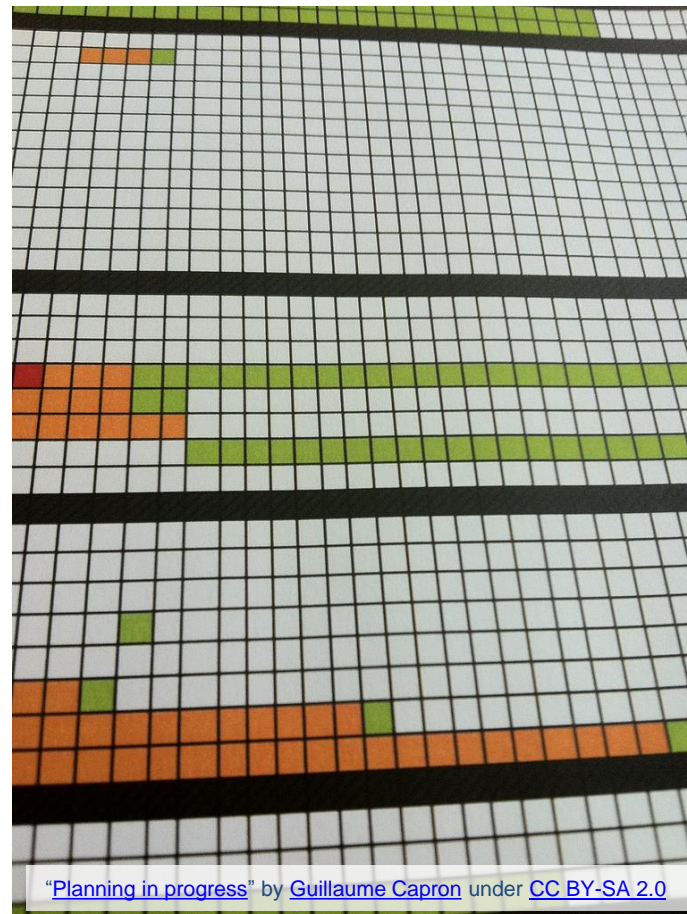
- proxy server, web server, OpenURL

SETUP, SUPPORT, COSTS

organize your community around content of shared concern

we will **consult on fit**, technical requirements, workflow integration

pilot implementation w/ subset of nodes to validate workflows

production implementation

ongoing support

participants are asked to join the LOCKSS Alliance (annual fee)

"Planning in progress" by Guillaume Capron under CC BY-SA 2.0

INSTITUTE of **Museum**and**Library** SERVICES

CoSA

SERI State Electronic Records Initiative

COUNCIL OF STATE ARCHIVISTS
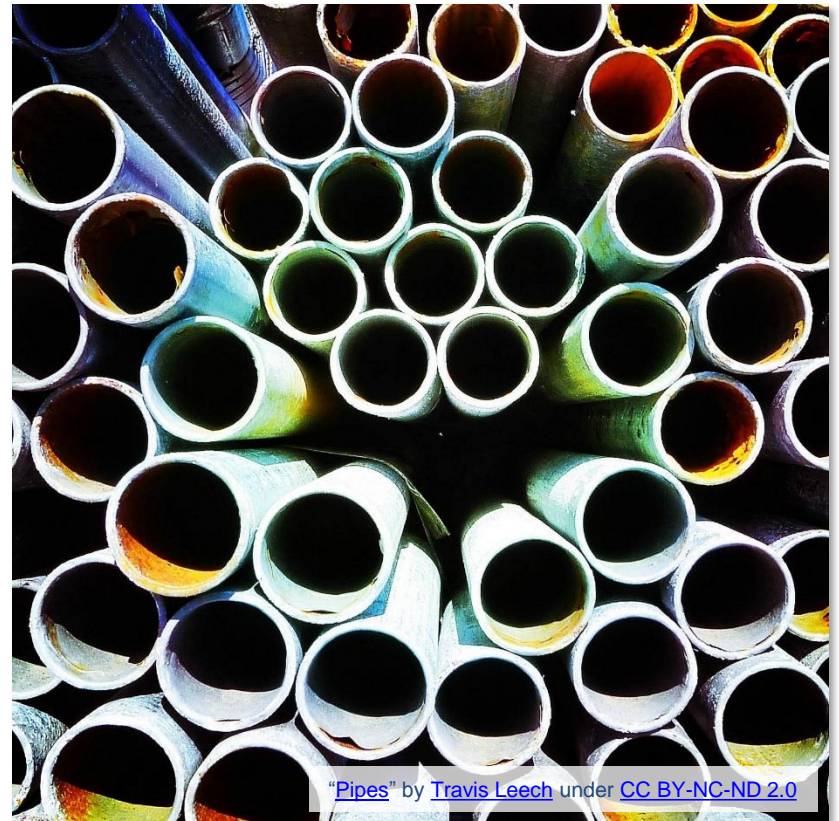
START NEW OR JOIN EXISTING NETWORK

[start a new network](#)

- [recommend 4+ copies](#)
- we can host node(s)

also an option to join an [existing network](#)

- reach out to network communities directly

we are exploring how to **better support needs of individual orgs** that aren't aligned w/ a logical community

"[Pipes](#)" by [Travis Leech](#) under [CC BY-NC-ND 2.0](#)

USE CASES

SERVICES

CoSA **State Electronic Records Initiative**

POST-CANCELLATION ACCESS FOR E-RESOURCES

networks:

- [Global LOCKSS Network](#)

restore **best features of print journal holdings** lost w/ online publishing transition:

- **local custody** (vice contingent access)
- lots of **decentralized copies** (vice fewer, centralized copies)

to better assure:

- preservation of **scholarly record**
- continuing **library role as steward**

GOVERNMENT INFORMATION

networks:

- Canadian Government Information
- Digital Federal Depository Library Program

partnership (or hedge) to **ensure permanent access** to government information

save + **re-decentralize** government information

CGI DPN

WHY LOCKSS FOR PRESERVING GOVINFO?

for **public/publishable** content:

- leverage **community** of collaborators
- **open**, extensible networks

for **private/protected** content:

- local IT control
- jurisdictional determination of storage

high standards for **data integrity assurance**

CGI DPN

End of Term Web Archive
US Federal Web Domain at Presidential Transitions

PEGI PROJECT
PRESERVATION OF ELECTRONIC GOVERNMENT INFORMATION

DataRefuge

Endangered Data Week

EDGI Environmental Data & Governance Initiative

DATAMIRROR

INSTITUTE of Museum and Library SERVICES

CoSA

SERI State Electronic Records Initiative

COUNCIL OF STATE ARCHIVISTS

VARIED DIGITIZED + BORN-DIGITAL CONTENT

networks:

- Alabama Digital Preservation Network
- MetaArchive Cooperative
- WestVault

all types of content

- including special collections + records

service models:

- **all depositors** also host infrastructure
- **subset of orgs** hosts infrastructure but serves whole community

governance + infrastructure both **community-based**

WEB ARCHIVE CO-PRESERVATION

networks:

- Ivy Plus Libraries Confederation Preservation Network

growing relative importance of web archives for **collection development**

- especially for **online-only resources** previously in print

decentralized local custody + preservation to **complement Archive-It**

- Archive-It also participates in network

WRAP UP

CoSA State Electronic Records Initiative

TAKEAWAYS

many threats to long-term persistence of digital content

- **don't discount** attacks, mistakes, org failure

LOCKSS is used to preserve **all types of content**, including govinfo

learn more at our **new website**: lockss.org

please contact us with any questions, or ideas on how we can work together

"Partir" by Chiara Conticelli under Fair Use

QUESTIONS + FEEDBACK

what content do you care to preserve?

what threats are you concerned about?

who also cares to preserve content you care about?

what are **limits on dissemination** of content you care to preserve?

what could you **preserve collaboratively**?

"Satellite" by Martin Davidsson under CC BY-NC 2.0

WEBINAR EVALUATION

We appreciate your feedback!

After you exit the webinar, you will automatically be taken to an online webinar evaluation.

Please take a couple of minutes to complete the survey and help us plan future webinars.

STAY CONNECTED & INFORMED

CoSA Website
http://www.statearchivists.org

CoSA Twitter Handle
@StateArchivists

CoSA Facebook Page
www.facebook.com/CouncilOfStateArchivists