



Todd Grappone (UCLA)
Mary Elings (UC Berkeley)
Edson Smith (UCLA)

Building a Digital Preservation Strategy Across a Broad University System

CNI Fall 2019

INTRODUCTION TO THE PROCESS

- Overview of the University of California's Libraries Advisory Structure (UCLAS)
- CoUL states plans and priorities annually.
- Maximize long-term access to digital content has been a priority for a decade. Lots of work led to this point.
- DOC recognized the importance and marrying IT and library preservation activity.

Digital Preservation Working Group

Investigate Internally

- Look at the current and planned UC activity

Develop a Baseline

- Use OAIS as a baseline to develop a high-level overview

Investigate Externally

- Perform a review of external preservation service providers, e.g., CLOCKSS, DPN, HathiTrust and Portico

Phase One Keys

- Articulate gaps between existing UC Libraries digital preservation capabilities and practices compared to current best practices and building blocks.
- The phase one team is *not* charged to write a prescriptive document that outlines a single approach to digital preservation for the UC Libraries. Rather, the group will develop both a high-level snapshot of current UC Libraries practices and capabilities, as well as an overview of the building blocks an academic research institution the size of the University of California should consider related to digital preservation practices, policy, capabilities, expended resources and potential service providers.
- From these parallel investigations, the phase one working group will build the foundation for subsequent phases as outlined below.

Phase One Focus

- Look Outward
- Look Inward
- Look Across
- Phase Two Charge



Timeline

Scope

Methodology



Exemplars (12)

- CLOCKSS
- DPN
- HathiTrust
- Portico
- LOCKSS (Stanford)
- Preservica
- Rosetta (Ex Libris)
- Chronopolis (UCSD)
- Merritt
- University of Michigan
- University of Illinois
- Internet Archive

UCs (11)

- UC Berkeley
- UC Davis
- UC Irvine
- UC Los Angeles
- UC Merced
- UC Riverside
- UC San Diego
- UC San Francisco
- UC Santa Barbara
- UC Santa Cruz
- CDL

Topics Covered



- Organization
- Mission
- Business Model
- Succession
- Rights Management/IP
- Integrity/Fixity
- Sustainability
- Property
- Architecture
- Ingest
- Metadata
- Access
- Roles/Responsibilities
- Storage/Replication

Exemplar Matrix

	Business model	Architecture	Ingest Req's	Metadata Req's	Storage and Replication	Access	Succession
Chronopolis	University	Multiple Nodes/OAIS	Limited	Agnostic	Distributed/Multiple Copies	Dark	Evolving
CLOCKSS	Consortium	Multiple Nodes/OAIS	Strict	Agnostic	Distributed/Many Copies	Dark	Strong
DPN	Consortium	Multiple Nodes	None	Minimal (DC)	Distributed/Amazon	Dark	By Contract
Hathi Trust	Consortium	Multiple Nodes/OAIS	Strict	Strict	Semi-Local/Exploring Amazon	Dark/Light	Strong
Internet Archive	Independent	Multiple Nodes	Limited	Agnostic	Distributed/Multiple Copies	Light/Dark possible	Relies on community
Illinois	University	Local/OAIS	Limited	Minimal (MODS)	Local/Exploring Amazon	Dark/Light	None
LOCKSS (Stanford Libraries)	Vendor	Multiple Nodes/OAIS	Limited	Agnostic	Distributed/Four Copies	Dark	Evolving
Merritt (CDL)	Consortium/University	Multiple Nodes	Limited	Agnostic	Distributed/Multiple Copies	Light/Dark possible	Strong
Michigan	University	Local	None	Minimal	Local/Amazon	Dark/Light	None
Portico	Consortium	Multiple Nodes/OAIS	Limited	Agnostic	Distributed/Multiple Copies	Dark	Loose
Preservica	Vendor	Local/OAIS	None	Minimal/XIP	Distributed/Amazon	Dark/Light	Customer copies
Rosetta (Ex Libris)	Vendor	Local/OAIS	None	Minimal (DC)	Configured by implementer	Dark/Light	Strong

UC Matrix

	Business model	Architecture	Ingest Req's	Metadata Req's	Storage and Replication	Access
Berkeley	University	Tind/Merritt	Limited	Basic	Distributed/Amazon	DAMS, CDL
Davis	University	Fedora/Archivematica	Limited	Minimal	Distributed/Amazon	DAMS
Irvine	University	Nuxeo/Merritt	Limited	Basic	Distributed	CDL, Dark/Light
Los Angeles	University	Various	Limited	Basic	Distributed/Amazon	DAMS, Dark/Light
Merced	University	Nuxeo/Merritt	Limited	Basic	Distributed	CDL, Dark/Light
Riverside	University	Nuxeo/Merritt	Limited	Basic (DC)	Distributed	Light
San Diego	University	Various	Limited	Minimal	Distributed/Data Center	DAMS, Dark/Light
San Francisco	University	Nuxeo/Merritt	Limited	Basic	Distributed	CDL, Dark/Light
Santa Barbara	University	Samvera/Local/Merritt	Limited	Basic	Distributed/Data Center	DAMS
Santa Cruz	University	Samvera/Merritt	Limited	Minimal	Distributed	DAMS

Phase One Report – April 2019

UC Digital Preservation Strategy Working Group

Phase One Report

April 10, 2019

Edson Smith (Chair), UCLA
John Chodacki, California Digital Library
Mary Elings, UC Berkeley
Todd Grappone (DOC), UCLA
Greg Janée, UC Santa Barbara
Charlie Macquarie, UC San Francisco
Rice Majors (DOC Liaison), UC Davis
Kevin Miller, UC Davis
Erik Mitchell (CoUL Liaison), UC San Diego
Adrian Petrisor, UC Irvine
Chrissy Rissmeyer, UC Santa Barbara
Roger Smith, UC San Diego

1



Working Group Findings

UNIVERSITY OF CALIFORNIA UC



UNIVERSITY OF CALIFORNIA UC



Three Tiers of UC Digital Preservation

The Top Tier

- California Digital Library (CDL)
- Chronopolis at UC San Diego

The Bottom Tier

- CDL clients

The Independent Middle Tier

- Digital preservation is largely aspirational
- No one is working together
- Legacy systems continue to be in use
- Staff and economic resources are limited
- Focus is primarily on DAMS

Technology Findings

There's nothing magical about a preservation repository

- All the issues are well understood
- We have the tools and reference models to build on
- We have standards in place, and they're keeping up with technology
- Data integrity is no longer a significant problem

Storage

- Almost all storage in the middle tier is local
- Storage costs are a universal concern

Other Findings

Non-technology findings

- The OAIS model is still sound
- There is general agreement about Best Practices
- Certification is essential
- Staff limitations
- Prioritization

Working Group conclusions are supposed to be out-of-scope, but...

- Technology should not be our focus
- System-wide resources are limited and activities not coordinated
- Each campus does not need it's own repository
- Our challenges are to **build collaboration** and to develop a **well-articulated governance model**.

“While research universities and cultural heritage institutions are innately long-running, they operate on that implicitly rather than by making explicit long-term plans.”

-- DPN sunset announcement, Dec 2018

Phase 2

Lots of preservation activity in the system

- Nearly all of the UC libraries using a TRAC Certified repository
- That preservation activity is disjointed and not complete
- Limited or no digital preservation unit or staff in most UC Libraries

Charge specifics

- Develop a high-level taxonomy of content types and associated preservation requirements
- Identify those content types that are common to the majority of UC libraries
- For the identified content type(s), develop an estimate of the amount of content held
- Throughout the process, provide opportunities for campus library staff to participate in activities that build expertise in techniques for assessment of local practices, workflow development, and approaches to content appraisal and selection for preservation decision-making.

Closing Thoughts

Digital Preservation Strategy Working Group, Phase One

Todd Grappone

Associate University Librarian for
Research and Development, UCLA
grappone@library.ucla.edu

Mary Elings

Assistant Director and Head of
Technical Services at The Bancroft
Library, UC Berkeley
melings@library.berkeley.edu

Edson Smith

Data Management Strategist, UCLA
edson@library.ucla.edu

John Chodacki, California Digital Library
Mary Elings, UC Berkeley
Todd Grappone (DOC), UCLA
Greg Janée, UC Santa Barbara
Charlie Macquarie, UC San Francisco
Rice Majors (DOC Liaison), UC Davis
Kevin Miller, UC Davis
Erik Mitchell (CoUL Liaison), UC San Diego
Adrian Petrisor, UC Irvine
Chrissy Rissmeyer, UC Santa Barbara
Edson Smith (Chair), UCLA
Roger Smith, UC San Diego
Perry Willett, California Digital Library

Phase One Report

[https://libraries.universityofcalifornia.edu/groups/files/doc/docs/
DPS_Phase_One_Report_20190410.pdf](https://libraries.universityofcalifornia.edu/groups/files/doc/docs/DPS_Phase_One_Report_20190410.pdf)