

COPPUL Digital Stewardship Network Community of Practice Mixer

Jan. 25, 2024 @ 12pm Pacific / 1pm Mountain / 2pm Central

Hosted by the [CDSN Community of Practice](#)

Quick links:

[Sign up for the mailing list](#)

Introductions

(In breakout rooms)

Breakout room 1: Preservation workflows and policy

Facilitated by Andréa Tarnawsky, supported by Kate Cawthorn and other committee members

Guiding questions:

- Do you have a policy or guiding documentation for your digital preservation program?
Where are you in this process?
- What are folks' sticking points in developing this documentation?
- How do you make workflow connections with others?
 - Are there bottlenecks in terms of your workflow depending on other divisions/systems/etc.?
- Would you consider submitting your workflows to the CoP document library?
- Is there a genre of documentation that DSN CoP can host an event on?

Breakout room 2: Preservation of research data and RDM

Facilitated by Robyn Stobbs, supported by Y.G. Rancourt and other committee members

Guiding questions:

- What questions do you have about preserving research data?
- Do you have a repository or service for curation and/or preservation?
- What workflows or tools do you use?
 - Do you have appraisal/selection criteria for data?

Breakout room 3: Preservation metadata, systems, and platforms

Facilitated by Melissa Cuthill, supported by Amber Saundry and other committee members

Guiding questions:

- What questions do you have about preservation metadata, systems, and platforms?
- What preservation systems & platforms are you currently using, or are interested in exploring?
 - Do you have any integrations and/or workflows with your DAM system/repository?
- Do you capture preservation metadata?
 - How/where?
- Are there particular systems/platforms you'd be interested in having a focused CDSN CoP session on?

Roundtable sharing

Share breakout room highlights with the larger group, time permitting

UPCOMING EVENTS

- CoP Mixer: April 18th, 2024 @ 12pm Pacific / 1pm Mountain / 2pm Central

Community notes



Breakout room 1: Preservation workflows and policy

- Some folks have fulsome policy for digital preservation at an organizational level, others are just starting out in the process
- Examples:
 - <https://www.gov.br/ibict/pt-br/aceso-a-informacao/documentos-oficiais/IBICTDIGITALPRESERVATIONPOLICY.pdf>
 - <https://www.gov.br/ibict/pt-br/aceso-a-informacao/documentos-oficiais/IBICTDIGITALPRESERVATIONPOLICY.pdf>
- Documentation development of procedure and workflows
- What is the scope of policies - unit specific or organizational level?
- SFU has org level digital preservation framework and then unit level action plans for digital preservation
 - <https://www.lib.sfu.ca/help/publish/digitization/digital-framework>
- Levels of preservation for different types of content or assets included in these policies
 - For example EDTs
- How do workflows documentation connect to policies or framework
- Places we might be stuck
 - How to transition workflow documentation to other team members - for example, fixity checks
 - U of A doing integrity checking twice annually - has three copies to allow for self-healing of preserved digital assets as needed
 - PREMIS metadata workflows?
 - Fixity tools and software that has been shared:
 - [Teracopy](#)
 - [Fixitypro](#)
 - a list of tools is here: <https://coptr.digipres.org/index.php/Fixity>
 - Environmental impact of Fixity checks - computationally intensive
 - Less can be more, and can be part of procedure or policy
 - Bi-annually? Annually? Will depend on institutional needs and risk

- Workflow bottlenecks or handoff with other units and how do you address this within documentation?
 - U of A integrated metadata team and preservation team to help prevent bottlenecks - workflows are planned with metadata team involvement
 - U of C - digital archivist is a team of one, can only do so much
 - Keeping everyone updated about how software works and training
- Acquisition workflows or tools?
 - SFU has file transfer protocol built on old software so tried to create a USB version to bring to private donors and then bring back to the archives
 - Oral histories collection - complicated for making policy and workflows because community needs to lead what stewardship looks like and where preservation copies are stored and how and for how long - affordances in place to easily delete records as needed
- Workflow or other documentation creation processes or resources that we can share?
 - [BPMNs](#)
 - Process mapping
 - Visio
 - A free tool: <https://app.diagrams.net/>
 - [Digital Preservation Toolkit](#):
 - some examples of workflows contributed by various practitioners: https://coptr.digipres.org/index.php/Category:COW_Workflows
 - there're also some workflow webinars put on by the Digital Preservation Coalition (though open to all!) coming up soon: <https://www.dpconline.org/events/eventdetail/254/-/digital-preservation-workflow-webinars-2024>

Breakout room 2: Preservation of research data and RDM

- Participants include academic librarians and other specialists / related fields
- Experience with research data management (RDM), [Borealis](#) (SP Dataverse), [LOCKSS](#), digital services, repositories like [Arca](#), copyright and scholarly communications
- Folks from the west coast and beyond!
- Thinking about repositories / platforms and RDM, advising faculty / admin on preservation and RDM strategies, options, sometimes more questions than answers
- Anticipating that there will be more of a need for RDM support / expertise, managing expectations a little as sometimes folks expect librarians will have all the answers
- Smaller orgs / libraries (7-8 librarians), folks doing multiple types of work / subject areas, often working closely with other staff, research assistants, “just in time” approach
- Research in general may be new for some institutions, just had a research ethics board formed, not yet a crush of demand, but still thinking ahead to the future, relation to IT and tools available to researchers, thinking about training / best practices (i.e. naming conventions, etc.)
- Repository platforms
 - Have folks had experience with [Zenodo](#)? [FRDR](#) another option, though can be mixed opinion (FRDR delay with assigning DOI*, data has to be reviewed before it gets published)
 - *Note re DOIs: This is not quite accurate, see [FRDR's FAQ](#) which states "As soon as you start a new deposit, the system will assign a DOI to your dataset. This DOI is unique to your deposit, and it will not change. You are welcome to share the DOI ahead of data publication, e.g., if you would like to include it in an associated manuscript, however, the DOI will not resolve (or be functional) until your dataset is published."
 - Something else folks have used is [OSF](#)
 - Borealis not launched yet at one institution
 - “Put your data where you found your data” – when researchers ask about disciplinary repositories
 - Use of institutional repositories like [clRcle](#) (UBC)
 - Often recommending Borealis, FRDR, Zenodo is not off the table though
 - Importance of getting people to think about what to do with their data when the research wraps up

- Connection / overlap with archivists at universities – big gap in policy system, substantial guidance and recognizing that all of this takes money, capacity
- Wonder how many institutions have the capacity to develop workflows for OSF, Borealis has more existing info and documentation
- [Archivematica](#) integrated with Borealis
 - By default all datasets in Borealis will have bit level preservation (level 1)
 - [Permafrost](#) is another service offered by Scholars Portal for providing Archivematica as a service
 - SP looking at developing reporting around formats in Borealis
- Assumption sometimes that the library will just “take care” of everything, sometimes researchers aren’t always prepared to take part in the process, though others are super prepared, researchers across the whole spectrum
- How are folks advising researchers on preserving sensitive data / preserving long term?
 - [Controlled Access Management](#) project
 - Library folks really interested in tool kits and guidance (ex. [Sensitive Data Guidance](#), scroll down the page)
 - Concern around data access / privacy, Microsoft, etc.
- No official support for research in IT department, policies view this as “scope creep,” don’t know where to refer people to, researchers must pay for / bring their own data storage, hoping that things are secure
 - An approach that is making some progress is building a strong relationship with people working in research ethics as they are advising researchers at the beginning of their research with an eye to preserving privacy and addressing other ethical issues
 - Having conversations with ethics folks and IT and trying to bring everyone together is useful, slowly getting everyone on the same page so that researchers get the same directions / information from multiple sources (ie. ethics and IT and library), though in a larger institution there can be complications when working with larger teams
 - Librarians as non-official members of IRB, but may not be sustainable as librarians may be stretched thin, in addition not all types of research deal with the same ethics boards
 - [DMP Assistant](#) / data management plans can also be a good tool

- myLaminin is being used at Queen's U: <https://www.mylaminin.net/>
- [GDPR](#) in EU
 - Data lifecycle management, University of Geneva - <https://www.dlcm.ch/about-us/dlcm-project>
- On premise servers vs. running everything in cloud
 - Researchers have to come with their own funding to participate
- Has anyone had to figure out how to store their curation logs?
 - FYI: [Data Curation Activities](#) (scroll down for curation log definition), [Data Curator Resources](#), [Dataverse Curation Guide](#) (Canadian resource)
 - Current solution, calling them administrative data and working them into the institution's records retention policy, putting them on SharePoint
 - Subject to [FOIP](#)?
 - Should it go into a README file that goes with the data / be a section of the README, not in a separate place?
 - In a museum context, the museum database would have all of this information connected to the objects
 - Questions around should curation logs be public facing or not, could be sensitive information or just things people might not want shared
 - At some institutions they are not keeping copies of data that have not been approved / curated, only keeping the finalized or published form
 - When does the curation log "start" ?
 - Restricted files as an option in Borealis: <https://learn.scholarsportal.info/all-guides/borealis/files/#Entering-Metadata-and-Restricting-Access-for-Files>
 - Terms of Service, no confidential information on the service
 - No blanket consent, REB only signs off in certain situations

Breakout room 3: Preservation metadata, systems, and platforms

- Participants include archivists & librarians, and US participants
 - [Archivematica](#); [LOCKSS](#); cataloguing & metadata; special collections; institutional repositories
- What preservation systems & platforms are you currently using, or are interested in exploring?
 - Archivematica - open-source, Canadian
 - [Preservica](#) - another open-source system
 - MacEwan looked at this, but staffing & resourcing considerations
 - [Libnova \(Spain\)](#)
 - [Arkivum](#) (UK)
- Do you have any integrations and/or workflows with your DAM system/repository?
 - UBC: IR's DSpace with automated integration of Archivematica, looking to explore distributed storage (Westvault) for subset of content (theses & dissertations). There are multiple pipelines for different content types at UBC.
 - KPU: Content through [Arca](#)
 - WestVault linked up but nothing in yet (smooth process - check a box)
- Do you capture preservation metadata? How/where?
 - In Archivematica - metadata stored in METS file, contains [PREMIS](#)
- What kind of preservation conferences have folks found helpful?
 - [iPRES](#)
 - [National Digital Stewardship Alliance](#) (NDSA)
 - [Code4Lib](#)
 - [Coalition for Networked Information](#) (CNI) usually has a strong DP track
 - Materials from the [Fall 2023 CNI meeting](#)
 - [CNI Youtube channel](#)
 - Best Practices Exchange has some digital preservation facets (not as much as some of the other ones mentioned)

- [ArchivematicaCons](#) are conferences for users (not vendor driven or organized)
- Are there particular systems/platforms you'd be interested in having a focused CDSN CoP session on?
 - Archivematica - panel of users/use cases
 - Discussion about future of WestVault
- Discussion about what should be preserved, and how
 - UBC IR - we run everything through an Archivematica pipeline, given the nature of the content and the IR's mission
 - MacEwan - needs to make decisions about processing configurations for types of content; a lot more hands-on decisions will be needed
 - Switched to SharePoint about 4 years ago, where records are being created - how do we get them out & ingested into Archivematica?
 - KPU - RDM has challenges in the variety of file types
 - SFU's [Format Policy Registry](#)
- Discussion about policies & documentation! This cannot be separated from metadata & systems/platforms 😊
 - Forming a policy can be challenging
 - "Policy" vs. "policy" vs. guidelines vs. strong recommendations!
 - IRs seeing a shift to more non-textual content
 - VIU: Seeing more accompanying files with theses
 - Even if we can't provide as strong assurance of file usability in the future, we can preserve at bit-level etc.
- LOCKSS (Lots Of Copies Keep Stuff Safe) - open source DP system, used to build DP community
 - Historically of web materials, but can be applied to anything - does what you tell it to do; many different purposes & uses
 - Focuses on preservation layer
 - Effort to increase documentation about how it can fit into IR systems/work etc.
 - [LOCKSS networks](#)

- COPPUL had a private LOCKSS network (primarily ETDs/OJS journals); merged into WestVault later
- 2 (3?) other LOCKSS-based networks in Canada
 - Canadian Government Information (CGI) Private LOCKSS Network (PLN)
 - [PKP Preservation Network](#)
 - Synergies - not active at the moment
- Archivemata can connect to LOCKSS
 - Configure the Storage Service accordingly
 - Mark Jordan at SFU was developing LOCKSS-A-Matic - unsure of current state or usage