

COPPUL Digital Stewardship Network Community of Practice Mixer

October 31, 2024 @ 12pm Pacific / 1pm Mountain / 2pm Central

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

Quick links:

[Sign up for the mailing list](#)

[The Planning Committee is looking for new members!](#)

Suggested topics for future mixers:

- Sharing lessons learned from @Risk North 3
- How folks have set up web archiving programs at their institutions
- Multi-modal / multi-format options for ETDs

INTRODUCTIONS

- Name, Organization
- Optional ice breaker question: Pumpkin pie or pumpkin spice latte? (or neither!)

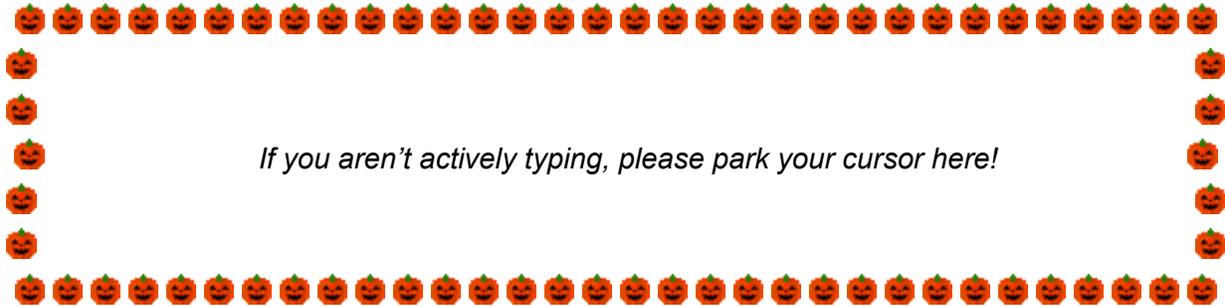
ROUNDTABLE

- Share descriptions of current work, or ask any questions about digital preservation

UPCOMING EVENTS

- November 28, 2024
- January 23, 2025
- April 24, 2025

Community notes



Spooky Preservation Stories

- Digital stewardship story time!
 - Collection of an electro-acoustic composer who worked a lot with early computer music
 - Items donated on magnetic reel, titled “moldy reels” and “possibly moldy reels”!
 - Didn't realize that yes, there was in fact dried mold over everything!
 - Some of these were the only recordings available
 - 6 months of mold remediation, and the reels are now digitized
 - Some of the reels have spatial requirements for how they are played, very cool!
- Digital preservation of theses
 - One of the technicians realized that some files were missing PDFs, about 300!
 - Versions available online, but they hadn't been included with the preservation files
 - Back then we weren't doing checksums either, but now we are!
- More of a confession than a horror story
 - When Archive-It went dark a few weeks ago, I realized that we hadn't backed up our WARC's recently – it has now moved up my to-do list!
 - There's not a great alternative to Internet Archive infrastructure out there right now
 - Comment: for anyone who had crawling activities scheduled for that week (i.e. for the recent BC elections) it highlighted the issues with relying on just one platform / service provider. Luckily we did a bunch of crawling

ahead of time, and have since been able to do the post-election crawls, but that's still a good question of where to put the WARCs

- An example I use to try and encourage folks to back up their data
 - Deduce was a popular service back in the day, lots of students and researchers lost all their grant data, analysis, etc.
 - The company now allows you to download copies of your projects
 - Now I ask: at what stage in your research do you need to back up stable copies?
 - Using AWS, but it's all active storage, no guarantee that this will be around forever
 - Issues for people who need to preserve their data but can't actually share it (i.e. publishing in a repository)
- A lesson learned about changing practices
 - Preservation of electronic theses and dissertations (ETDs)
 - Long-tail issues
 - Creative arts theses, went fully digital in the past but these theses did not
 - Instead putting things on CDs and DVDs, getting shipped off to the library
 - Decided to upload in a closed collection
 - Some of the files were not loaded correctly onto the media in the past or since became corrupted
 - Needing to use older workstations and programs to access media (disk drives, iTunes, etc.)
 - Media from the 90s, unsure what to do with it, sitting on a shelf
 - Accepting non-textual supplementary ETDs for awhile, but now we're seeing a lot more of these "multi-file" ETDs (no longer called supplementary)
 - We have file format guidelines (https://wiki.ubc.ca/Library:Circle/File_Format_Guidelines), but we don't look at everything, we have to rely on the students and their advisors to follow their recommendations or not, and sometimes you just do get non-preferred formats
 - No definition of what an ETD can entail, in terms of policies, but our Grad School encourages emerging ETDs for PhDs

<https://www.grad.ubc.ca/current-students/dissertation-thesis-preparation/emerging-dissertation-approaches-designs>)

- We've had people wanting to deposit PDFs with embedded media, and when we transform that to PDF/A that media gets lost
- FYI Spooky data podcast about data vampires:
https://techwontsave.us/episode/241_data_vampires_going_hyperscale_episode_1
- Difficulties in trying to acquire DAT machine and other A/V tech / hardware (that works!)
- Working with a large image collection
 - Organization had hired students to digitize the entirety of the collection
 - Everything had been scanned with the office Xerox machine and put on USB sticks, unlabeled and undescribed
 - The physical media was also in varying conditions, some scrapbooks, newspaper clippings, different adhesives, etc. all mixed together
 - Community organizations and interns doing their best with what they have
- Special collections research appointment
 - Preservation scans of a publication
 - Loaded up the disc, and all the files had been named by issue date, month, day, year
 - Not very easy to find things as a result
 - The time you take before starting a project is so important, being able to plan out what you're going to do

Not a Spooky Preservation Story

- Update from the Scholaris Preservation Expert Group
 - Recently sent out a survey about institutional repositories
 - Working on another survey of digital preservation practices / gathering information from websites and online policies
 - Look out for more info from this group in future