

Preservation is our Mission

Linda Newman

University of Cincinnati

Linda.newman@uc.edu

Why do this?

File corruption does occur.

Political bad actors exist.

Local and regional natural disasters happen.



aptrust.org

- Bit level preservation – 6 copies in Amazon – 3 in S3 (east coast), 3 in Glacier (west coast), checksums confirmed quarterly
- Modified LC bagIt format bags
- Dark archive
- Option to replicate in another DPN node.
- Cincinnati first partner to load production content

So – is bit level preservation solved?

No – challenges include:

- Getting deposit agreement thru local bureaucracy
- Length of time to export from repositories
- Local space needs for export
- Scripts break with large collections
- No clear method for keeping track, and avoiding overwriting good content with bad.

What would make it better?

- Drag and drop or easier submit method requiring less local file manipulation, less local commitment of storage when preserving entire collections.

What would make it better?

- Communication between Samvera and storage layer (whether Fedora or other), to verify –
 - When a work (metadata + files) has been exported for preservation
 - When a work (metadata or files) has legitimately changed
 - When a work (metadata or files) has illegitimately changed

Are things solved now?

- No: no checksum algorithm for preservation use case – as our dark archives grow, our frequency of checksum processing will diminish without a better algorithm.
- No: no one knows how many copies it takes to really keep stuff safe (1 Petabyte for a century still guesswork).(<http://blog.dshr.org/2007/06/petabyte-for-century.html>)
- No: We're relying on cloud storage when our local computer centers are now going there too; or we are relying on distributed yet homogeneous systems with shared vulnerabilities.
- No: no global reach to our preservation dark archives – yet.

MY QUESTIONS:

- What preservation actions and event tracking are appropriate for shared incorporation into our Samvera solutions (i.e. Hyrax)?
- What preservation action and events are so implementation specific that they should not be incorporated into shared Samvera solutions? (But we can share info.)
- What leadership role does the Samvera community have in furthering digital preservation in general?

YOUR QUESTIONS?



linda.newman@uc.edu