

Achieving Robust Digital Preservation at a Large Academic Library: Successes and Failures

Michael Boock, michael.boock@oregonstate.edu

Hui Zhang, hui.zhang@oregonstate.edu

Open Repositories 2022

<https://beav.es/iG4>



Oregon State
University



Agenda

- Background
- Repository Level Preservation
- Digital Preservation of Electronic Thesis and Dissertation
- Challenges
- Lessons Learned

Formal Commitments to Digital Preservation

[OSU Digital Preservation Policy](#)

- “make accessible and hold in trust for future use”

[ScholarsArchive@OSU Preservation Policy](#)

- Commit to bit-level preservation of all digital content contained within the repository
- Guarantee that objects submitted to ScholarsArchive@OSU *using recommended file formats* will be retrievable in the future.

[Digital preservation plan](#)

- Current state of digital preservation efforts
- Recommends next steps

Strategic Plans

2012-2017

- “Creation of robust and flexible digital preservation and curation infrastructure.”
- “Long-term preservation system for university scholarship and digital collections.”

2018-2023

- “Develop, demonstrate and support best practices for data management and preservation of digital objects within OSULP.”
- “Implement OSULP’s **digital preservation plan**.”

MetaArchive

- Distributed Digital Preservation
- Cooperative, not a vendor
- Member-owned hardware and software.
- Membership fees and storage fees go to a central pool of support for network operations and members' co-op activities.
- Monthly community calls to share projects and for topic-based discussions.
- Includes 15 secure, closed-access preservation nodes and more than 200 TB of content.
- Geographically-dispersed
- For content that is of greatest historical value to the university community

Backup Retention Policy

- All the backups are saved in Amazon Simple Storage Service (AWS S3)
 - Database backups saved in Infrequent tier
- Server backup
 - Retention: 30 Days
 - Schedule: Monthly full, daily incremental
- App backups
 - Retention: indefinite
 - Schedule: Nightly rolling snapshots
- Database backups
 - Retention: 1 year
 - Schedule: Nightly full backup

Repositories

Samvera/Fedora: 2015-present

- Minimal preservation functions:
 - File integrity (Checksums)
 - Scheduled and ongoing fixity checks
 - Version control
 - File characterization
 - Persistent URL
- Not yet implemented:
 - PREMIS preservation metadata storage
 - Virus checking
 - Format migration

Preservation of Electronic Theses and Dissertations

1. Export ETDs of the previous academic year from ScholarsArchive@OSU in the BAG format.
2. Upload ETD BAGs to an OSULP S3 bucket acting as the “staging server” for MetaArchive.
3. Work with MetaArchive staff to coordinate the harvesting of ETD BAGs and ingestion to the LOCKSS repository.

Challenges

- Stretched library workforce where no staff is dedicated for digital preservation
- Lack of integration between institutional repository platforms and preservation networks
- Insufficient infrastructure to transfer large files over distributed network

Lessons Learned

- Outreach to admin to achieve buy-in
- Library staff knowledge-sharing and involvement
- Prioritization, because you may not be able to do everything you want to do

Related Readings

- Boock, M., & Davis, B. E. (2017). Next steps for building a flexible and robust digital preservation infrastructure at Oregon State University Libraries & Press.
https://ir.library.oregonstate.edu/concern/technical_reports/73666904w
- Boock, M., Key, C., Llebot, C., Mellinger, M., & Van Tuyl, S. (2019). ScholarsArchive@OSU Repository Core Trust Seal Self Assessment: June 2019. : Oregon State University.
https://ir.library.oregonstate.edu/concern/technical_reports/zp38wk30g
- Oregon State University Libraries Digital Preservation Policy.
<https://wiki.library.oregonstate.edu/confluence/x/eSvVWAw>
- ScholarsArchive@OSU Policies.
<https://wiki.library.oregonstate.edu/confluence/x/bivVWAw>

THANK YOU



Oregon State
University