

---

---

# Digital Preservation – Policies and Practices

Eric Johnson, Numeric and Spatial Data Services Librarian  
Miami University Libraries

---

---

# Digital Preservation

How much do you know about digital preservation?

Has any form of digital preservation been implemented at your institution?

Do you have institutional support for digital preservation?

---

---

---

# Digital Evaporation

Something can be gone and you would never know it

Digital files are much more vulnerable than paper or film

Bit rot can destroy a file, must examine the file to verify

---

---

# Miami University's situation

Scattered files

CONTENTdm

DSpace repository

miscellaneous files on servers and individual computers

Our first attempt - 2010  
DVDs and CDs

What was the digital preservation field like in 2010?

What we tried

---

Why we failed

---

# Current Practices at MUL

Locally created content held on a variety of servers and hard drives

Preservation currently consists of periodic backups of some content

Use of CONTENTdm and DSpace for collections access

---

---

# Digital Preservation Committee 2.0

Content and Files/projects lost

New management prompted revisiting our current practices

After five year gap, decision to start from scratch

---

---

# Task Force created

Digital Preservation committee membership included representation from a variety of departments

Preservation

Digital Scholarship

Library Systems (IT)

Technical Services

Branch Libraries

---

Special Collections and Archives

---

# Current Mandate

In 2015, our dean mandated that we make “recommendations as needed for short and long term planning, maintenance, updates, and changes to digital preservation policy and practice”

Determine the steps needed

Set priorities

---



---

# Inventory of Digital Holdings

- What types of files do we have?
  - Where did they currently live?
  - How much do we have?
  - What are they currently used for?
  - How are they currently accessed and by whom?
-

---

# We Found: Types of file content

Commercial databases, eBooks

Scanned images, oral histories, finding aids, manuals

Scholarly creations (published papers, slide decks,  
spreadsheets)

Emails

Born digital materials (donated materials; archives)

Locally created web sites

---

---

# Three main types to preserve

- Purchased content (eBooks, commercial databases)
  - Digitized content (scans of artifacts)
  - Scholarly creative content (published papers)
-

---

# We Found: Where do they live?

- On vendors' web sites
  - University and library servers (CONTENTdm, DSpace, Omeka, WordPress)
  - Network and cloud storage (Production drive, Google docs)
  - Removable hard drives, flash/thumb drives
  - CD-roms and DVDs
  - Under people's desks (on personal work computers)
-

---

# We Found: How much do we have?

- CONTENTdm: 700 GB
  - Digital collections archival masters: 2 TB
  - DSpace: 130 GB
  - Other: 1 TB
-

---

# We Found: Current Use

Commercial content - in place of print copies, patron research

Public display and access of digitized special collections and archives content

Digitization masters/archival storage

3D printer models - first creation, reprint, modifying

Scholarly repository/access

---

---

# We Found: Current Access

Patrons through web interfaces

Clients mediated by staff

Staff projects

Staff - collections management and maintenance

---

---

# Library Needs and Limitations

- The committee considered what local skills would be needed
  - How stable is our funding, how broad is our skill base?
  - Decided to go with a solution that provided more technical support rather than requiring in-house library technical expertise
-



---

# Desired Solution Characteristics

- Listed out all characteristics
    - access (end user)
    - access (staff)
    - preservation tools
    - technology requirements
    - support
    - succession planning
    - costs
  
  - Each member ranked characteristics based on importance - average ranking used to determine priorities
-

---

# Compared potential solutions

- Developed list of available solutions
  - Ranked software on each of the features
  - Decision matrix helped guide our discussions
-

---

# Available Solutions

- Archives Direct
  - ArchivesSpace
  - Bepress digital commons
  - BitCurator
  - Box.com (A.K.A. Box.net)
  - CLOCKSS
  - CONTENTdm
  - DSpace (locally hosted)
  - DSpace Direct
  - DuraCloud
  - HathiTrust
  - Local Gluster storage
  - LOCKSS
  - MetaArchive
  - OCLC Digital Archive
  - Preservica
  - Portico
  - Rosetta
  - Widen Media Collective
  - Zetta.net
-

---

# Top Solutions: A Closer Look

Talked to vendors

Talked to user groups and product support teams (and read blogs & comments)

Asked for more information and demonstrations

---

---

# Recommendations

- We summarized our findings
  - Presented our recommendations to the Dean
  - Purchased content - Portico or LOCKSS
  - Local content - Preservica or MetaArchive
-

---

# Moving Ahead, Holding Back

- Recommendation of Portico accepted and implemented right away
    - no real competition
    - simple to implement
    - didn't require any change in library workflows
  - Acceptance of Preservica took about a year
    - competing products sounded similar and overlapping
    - additional research
    - comparison to tools from existing contracts (CONTENTdm/OCLC)
    - increased staff support and buy-in
    - further costs analysis
-

---

# Preparation for Implementation

- Metadata cleanup
  - Preservation Policies
  - Local best practices
  - Works flows
  - Staff roles and responsibilities
-

—

# Questions?