Developing an E-Records Action Plan

New York State Archives
www.archives.nysed.gov

Outcomes for Today

- Assess the scope of your electronic records holdings
- Inventory electronic records
- Identify at risk electronic records
- Develop a plan for processing and preserving electronic records

---

The Case for Digital Preservation

The Case for Digital Preservation

- Why preserve electronic records?
  - To ensure accountability of your organization or business
  - To meet local, state and federal legal requirements
  - To preserve information assets generated by your organization
  - To ensure the authenticity of digital materials in the long term, for access, use and research

The Case for Digital Preservation

- Preservation Issues
  - Technological Changes
    - Obsolescence of technology and playback mechanisms
    - Media fragility
    - File formats and their degradation
  - Organizational Changes & Challenges
    - Growing volume of digital resources
    - New partnerships
    - New expectations

"The perfect is the enemy of the good"

-Voltaire

Your Electronic Records Inventory

The Scope of Your E-Records Inventory

- Born-digital records
- Digitized materials
- Digital files in your collection
- Your archives' operational records
- Other digital materials

Physical Media

- Floppy disks
- Zip disks
- CD/DVD-ROM
- Flash drives
- Hard drives (internal and external)
- Networked storage
- Cloud storage

File Formats

- Word processing documents
  - PDF/PDF-A
  - DOC/DOCX
  - XLS/XLSX
  - HTML/HTM
  - XML
- Databases
  - MDB
- Spreadsheets
  - XLS
- Presentations
  - PPT
- E-mail
- Audio recordings
  - WAV
  - MP3
- Video recordings
  - WMV
  - MOV
- Digital photographs
  - JPEG
  - TIFF
  - RAW
- Websites (public or private)

Your Electronic Records Inventory

- Identify collections likely to contain electronic records
- Locate media containing these records
- Gather relevant donor and collection information about your electronic records
- Use electronic records inventory form to survey your holdings

Your Electronic Records Inventory

- Media manufacturer
- Label on media (consider photographing)
- Serial number
- Note any visible damage

Your Electronic Records Inventory

- Is it part of a larger collection?
- If part of a larger collection, is it all electronic files, or are paper-based materials included?
- Is information included in the electronic files available in other formats (e.g. paper, microfilm, audio or video tape)?
- Do backups exist?

Your Electronic Records Inventory

- Series or collection title?
- Series or collection ID?
- Does information about the records exist elsewhere

Your Electronic Records Inventory

- Acquisition documents
- Loan forms
- Gift correspondence
- Exhibition contracts
- Inventory slips
- Invoices
- Purchase orders
- Releases

Your Electronic Records Inventory

- Creator name
- Date of creation
- File sizes
- Quantity
- Most common file types
- Arrangement/file structure
- Software application used to create records
- Access restrictions

Your Electronic Records Inventory

- Date of the Inventory?
- Person conducting the inventory
  - Name
  - Title
  - Department

Reviewing the Findings from Your Inventory

- How many files do you have?
- Prioritize for preservation based on file content
- Prioritize based on file type and media
- Flag most vulnerable files and media

At-Risk Records In Your Repository

At Risk E-Records

- Older/obsolete media
- Obsolete file formats
- Playback equipment unavailable

At Risk E-Records

- No accession documentation available for record
- Little or no metadata available for the record
- Small or large volume of data

Additional Risks

- Location/geographical issues
- Organizational concerns
- Legal problems and issues
- Financial issues
- Political concerns

Basic Processing and Preservation

Processing Workstation

- Laptop or desktop
- Antivirus software
- Separate from network/internet?
- Storage and back-up media
- CD/DVD/floppy disk drives?

Basic Processing and Preservation

- Start working with files on media you can read
- Plan solutions for media you cannot read
  - Hardware
  - Software
  - Vendors
  - Collaboration

NDSA Levels of Digital Preservation

Table 1: Version 1 of the Levels of Digital Preservation

	Level 1 (Protect your data)	Level 2 (Know your data)	Level 3 (Monitor your data)	Level 4 (Repair your data)
Storage and Geographic Location	- Two complete copies that are not collocated - For data on heterogeneous media (optical discs, hard drives, etc.) get the content off the medium and into your storage system	- At least three complete copies - At least one copy in a different geographic location - Document your storage system(s) and storage media and what you need to use them	- At least one copy in a geographic location with a different disaster threat - Obsolescence monitoring process for your storage system(s) and media	- At least three copies in geographic locations with different disaster threats - Have a comprehensive plan in place that will keep files and metadata on currently accessible media or systems
File Fixity and Data Integrity	- Check file fixity on ingest if it has been provided with the content - Create fixity info if it wasn't provided with the content	- Check fixity on all ingests - Use write-blockers when working with original media - Virus-check high risk content	- Check fixity of content at fixed intervals - Maintain logs of fixity info; supply audit on demand - Ability to detect corrupt data - Virus-check all content	- Check fixity of all content in response to specific events or activities - Ability to replace/repair corrupted data - Ensure no one person has write access to all copies
Information Security	- Identify who has read, write, move and delete authorization to individual files - Restrict who has these authorizations to individual files	- Document access restrictions for content	- Maintain logs of who performed what actions on files, including deletions and preservation actions	- Perform audit of logs
Metadata	- Inventory of content and its storage location - Ensure backup and non-collocation of inventory	- Store administrative metadata - Store transformative metadata and log events	- Store standard technical and descriptive metadata	- Store standard preservation metadata
File Formats	- When you can give input into the creation of digital files encourage use of a limited set of known open formats and codecs	- Inventory of file formats in use	- Monitor file format obsolescence issues	- Perform format migrations, emulation and similar activities as needed

---

NDSA Elements

- Storage
- Fixity
- Security
- Metadata
- File Formats

Storage

- Analog storage: paper, microfilm
- Near-line storage
- Offline storage
- Online storage
- Third-party storage
  - Offsite storage
  - Cloud storage

Storage: NDSA requirements

- Select appropriate storage option(s)
  - Move from mixed storage media to uniform storage system
- Minimum of two copies of files

Fixity

- Fixity
  - Checksum or hash value is a unique value based on contents of file
  - Values generated by algorithms (e.g. md5, sha256)
  - Every file has its own value
  - If value changes, file has changed

Fixity: NDSA requirements

- Check fixity information if provided
- Generate fixity information if none exists

Security: NDSA requirements

- Identify who can perform what actions on which files (delete, move, copy and/or change files)
- Restrict who has the ability to *modify* files to fewest number of people

Metadata

- Data about data
  - Descriptive
  - Administrative
  - Structural

Metadata: NDSA requirements

- Inventory of files and their location
- Store a copy of the inventory separate from files.

File Formats: Background

- Obsolete file formats
- Proprietary formats

File Formats: NDSA requirements

- Encourage your organization to use open file formats
- Encourage your organization to use a limited number of file formats

More Information

- New York State Archives documentation on electronic records management, inventorying, and preservation:
www.archives.nysed.gov/a/records/mr_erecords.shtml

- National Digital Stewardship Alliance (NDSA)
www.digitalpreservation.gov/ndsa/

- Articles and Publications by Ricky Erway - OCLC Office of Research
www.oclc.org/research/publications/reports.html

Action Plan Elements

- What you have
  - Summary of inventory results
- Plan for at risk files/media
  - Analysis of inventory results
- Processing and preservation plan
  - Workstation
  - Processing outline

Next Steps

- Advocate for electronic records/digital preservation at your organization
- Identify at-risk materials and develop a plan to make them accessible
- Decide on a long-term storage solution or service

Questions?

- Thank you!
- Contact info...