Planning and Infrastructure for Analog to Digital Preservation Projects

Linda Tadic

Director of Operations

ARTstor

lt@artstor.org

Why digitize?

To preserve and provide access to content :

Content that is heavily used, so digitization serves to protect and preserve the original source

Content that should be used, but the original format and/or condition limits access and discovery

Digital projects are expensive ~

Every decision you make impacts costs

The budget determines what, how, and how much you can digitize

Digital project considerations BEFORE beginning:

1. Intellectual property rights
2. Care of the original source
3. File format and resolution
4. Storage and infrastructure (file formats and resolutions will determine storage requirements)
5. Digital asset management (workflow AND system)
6. Metadata structure and workflow
7. Production and QC scenarios (in-house or outsource?)

(1) Intellectual property rights

- Who owns the source being digitized? What can you do with the digital surrogate? Do legal research (due diligence)
- What are the underlying (3rd party) rights:

  still images: rights to the object displayed in the photograph; photographer's rights

  sound recordings: composition rights, public performance rights, distribution rights; interviewees' rights (right to privacy)

  moving images: script, music, etc.

- Metadata: How will you track IPR information?

(2) Care of the original source

Process: inventory and inspect (note condition of source)

Metadata should be created throughout the production workflow; use inventory as basis for production management data

Clean before digitizing; re-house into archival containers after digitizing

(3) File format, resolution, and compression

Addressed by next speaker

The selected file formats and resolutions will determine your storage requirements

(4) Storage and infrastructure

Nuts and bolts:

- Estimate storage needs for first 5 years (resolution determines file size, which determines storage needs)
- Need to store archival digital masters, plus access files. What is your archival digital master format? Will you store different “states?”
- What is your bandwidth to supply access files? What format will you provide? (JPEGs? MPEG4?)
- In-house storage (consider AC, electrical, seperate server room) or hosted (on demand) storage?

(4) Storage and infrastructure

Storage options:

Physical carriers aren't as important as the digital file, which lives in many places and moves a lot over its lifecycle.

- File delivery media (CDs, DVDs, HDDs) ($)
- Servers and SNAP drives ($$)
- Enterprise storage (SAN) (online, nearline) ($$$)
- Tape back-ups ($$)
- Staff to maintain servers ($$)

(4) Storage and infrastructure

Create a migration and digital preservation plan

- Redundancy is key
- Use metadata to track migration and preservation
- Capture technical metadata throughout production
- Migration plan: how often are files moved to new carriers?

(5) Digital Asset Management

DAM is a workflow/process, not just a system

Consider the steps that can be managed:

- Creation
- Quality control
- Version control / authenticity (what “states” will you keep?)
- Catalog (metadata added throughout lifecycle)
- Access
- Distribution
- Preservation

(5) Digital Asset Management

Digital asset management systems:

What do you want it to do?

- Manage final files
- Manage production workflow, versions, and quality control
- Hold technical and preservation data only, or also descriptive?
- Cataloging tool?
- Stand-alone system, or linked to other databases?

(5) Digital Asset Management

Digital Asset Management Systems options:

- Off-the-shelf (Excel, Access, FilemakerPro, Extensis)
- Open source digital repositories (DSpace, Fedora)
- Commercial products (ContentDM, Artesia, NorthPlains)
- Hosted (on-demand) DAMS and storage
- Build your own

If you buy, be sure the DAMS supports your platform (Mac or Windows)

(6) Metadata

What standards will you use or adapt?

What information will you track? Where will it live?

- Descriptive (content)
- Original source information
- Digital file creation information (technical metadata)
- Quality control information
- Digital preservation information
- Legal information

(7) Production Scenarios

Outsource or in-house? (Balance costs: equipment, staffing, physical environment)

Production scenarios:

- in-house (equipment, staff, room)
- outsourced to vendor
- vendor sets up shop on site
- 2-vendor production (scanned at 1st, then “post” at 2nd)
- new analog photography is created and then scanned

(7) Project Management

Schedule monthly deliverables by all parties:

- Processors (pull and prepare the collection)
- Vendor 1 (ramp up production gradually)
- Vendor 2
- Quality control
- Catalogers
- Processors (return the collection)
- IT staff (load files)

Budget estimates template