



Creating a Digital Collection: The Basics  
TAM-CMC Archives Workshop  
Waco, TX :: 10/28/11



Eric S. Ames  
Digital Collections Consultant  
Baylor University

# A Little About Me

---

- ▶ BA: Public Relations (TTU)
- ▶ MA: Museum Studies (Baylor)
- ▶ Digitization Collections Consultant
- ▶ Proud daddy

[Photo: Kelly Hosch Photography, Temple, TX]





# Pre-Planning

Section I

# Pre-Planning: Assessment

---

- ▶ **Assessing your collections**

- ▶ What kind of materials will you be digitizing?
  - ▶ Archival: manuscripts, letters, photographs, maps
  - ▶ Library: books, bound items
  - ▶ Newspapers
  - ▶ Video/audio



# Pre-Planning: Outcomes

---

- ▶ **Scanning for preservation**
  - ▶ Higher resolution
  - ▶ Larger files
  - ▶ Long-term responsibility
- ▶ **Scanning for access**
  - ▶ Lower resolution
  - ▶ Smaller files
  - ▶ Shorter retention schedule
- ▶ **Both**
  - ▶ More complex
  - ▶ Most likely



# Pre-Planning: Staff

---

- ▶ Who scans it?
  - ▶ Staff
    - ▶ As primary responsibility?
    - ▶ An “other assigned duty”?
  - ▶ Volunteers
    - ▶ Primary focus?
    - ▶ Training concerns/quality of work
  - ▶ Contract staff / temps
    - ▶ Short-term solution
    - ▶ One-off or specialized jobs



# Pre-Planning: Organization

---

- ▶ **How are the materials organized?**
  - ▶ Chronological
  - ▶ By accession number
  - ▶ By serial number/publication information
  - ▶ By name/subject/title
  - ▶ Other



# Pre-Planning: File Management

---

## ▶ Identifiers

- ▶ Human-readable best
  - ▶ 2003\_03\_11.tif (trinomial)
  - ▶ Tribune\_October1899-03

## ▶ Checksums/File fixity

- ▶ “File fingerprints” for tracking
- ▶ FastSum

## ▶ Storage

- ▶ Multiple redundancies: scanning machine, external storage, off-site





# Pre-Planning: Timeline

---

- ▶ **Have a target end date**
  - ▶ “1<sup>st</sup> quarter of 2012”
- ▶ **Be reasonable**
  - ▶ It will take longer than you think
- ▶ **Under-promise, over-deliver**
  - ▶ Who are your stakeholders?
  - ▶ How much do they understand about the process?
  - ▶ What are their expectations?



# Equipment and Software

Section 2

# Equipment: Minimum Specs

---

- ▶ **Compatible with your computer**
  - ▶ Mac or PC
- ▶ **USB 2.0 or firewire connectivity**
  - ▶ Listed on box, almost all are USB
- ▶ **At least 300 dpi, 24-bit color**
  - ▶ Grayscale also useful
- ▶ **8.5" x 11" scanning surface or larger**
  - ▶ Perfect for photographic prints, letters, standard size paper
  - ▶ If an option, look for something larger



# Equipment: Software

---

- ▶ **Scanner-based**
  - ▶ Acquisition of images (EPSON scan)
  - ▶ Basic image editing
- ▶ **Photoshop (or equivalent)**
  - ▶ GIMP (GNU Image Manipulation Program)
- ▶ **Google docs/Excel**
  - ▶ Workflow tracking





# Material Selection

Section 3

# Materials: Copyright

---

- ▶ **Public domain**
  - ▶ 1923 is considered “bulletproof”
- ▶ **“Orphaned works”**
  - ▶ Author/artist is known, but current owner is not
- ▶ **Under copyright**
  - ▶ Request permission or metadata only
- ▶ **Permissions-granted works**
  - ▶ Make sure it's covered in the paperwork



# Materials: Rarity

---

- ▶ **One-of-a-kind**
  - ▶ The only known example of X in existence
- ▶ **Most-used rationale**
- ▶ **Do your homework**
  - ▶ Make sure it really is rare



# Materials: Fragility

---

- ▶ Determine a threshold
- ▶ Conserve/preserve first
- ▶ Be aware of risks during scanning
- ▶ Make plans for rehousing after scanning





# Materials: Other

---

- ▶ **Most requested**
  - ▶ Protection of originals
  - ▶ Access via Internet
- ▶ **Largest holdings**
  - ▶ Most complete set of historic city photographs





# Digitization Basics

Section 4



# Digitization: Basics

---

- ▶ Choose an appropriate dpi and color setting
- ▶ Always do a preview scan
- ▶ To color bar or not to color bar?
- ▶ Borders
  - ▶ 1-2mm around images



# Digitization: Basics, File Management

---

## ▶ Storage locations

- ▶ At least three: scanning computer, an external hard drive, off-site
- ▶ Spinning discs the rule

## ▶ Checksums

- ▶ Assign to files as you save them to external source

## ▶ Backups

- ▶ Further steps include backing up to tape

## ▶ Derivatives

- ▶ Match filenames to preservation copies
- ▶ Keep or discard
  - ▶ External media like a CD-R, DVD-R, flash drive





# Metadata

Section 5

# Metadata: Basics

---

- ▶ “Data about data”
- ▶ Think catalogue info
- ▶ Organized according to schemas
- ▶ Dublin Core is one of most common



# Metadata: Dublin Core

---

## Fields

- ▶ Title
- ▶ Creator
- ▶ Subject
- ▶ Description
- ▶ Publisher
- ▶ Contributor
- ▶ Date
- ▶ Type
- ▶ Format
- ▶ Identifier
- ▶ Source
- ▶ Language
- ▶ Relation
- ▶ Coverage
- ▶ Rights

For more info:  
<http://dublincore.org>





# Metadata: Considerations

---

- ▶ How much is enough?
- ▶ Embedded URLs to reference collection
  - ▶ Source: [www.ourcollection.com/texas](http://www.ourcollection.com/texas), for example
- ▶ Copyright clearance statements
- ▶ Disclaimers
  - ▶ No known copyright restrictions
  - ▶ These materials are representative of the time in which they were created and may be considered offensive by current standards. User discretion is advised.





# Online Presentation

Section 6

# Online Presentation: Basics

---

- ▶ **Free, purchased, or “near-free”**
  - ▶ Free: Flickr, Picasa
  - ▶ Purchased: Virtual Exhibit plugin for PastPerfect, CONTENTdm
  - ▶ “Near-free”: Flickr Pro
- ▶ **Level of access**
  - ▶ High-quality vs. low
  - ▶ Worldwide vs. controlled
- ▶ **Social media capabilities**
  - ▶ Easy to share?
- ▶ **Downloads?**
  - ▶ High-resolution or low



# Summary / Q&A / Contact info

Section 7

# Contact Info

---

[eric\\_ames@baylor.edu](mailto:eric_ames@baylor.edu)

(254) 710-1576

## Helpful links

- ▶ <http://contentdm.baylor.edu>
- ▶ <http://digitalcollections.baylor.edu>
- ▶ [www.flickr.com](http://www.flickr.com)
- ▶ <http://docs.google.com>
- ▶ <http://dublincore.org>
- ▶ <http://goo.gl/FKWGN> : Course materials, including this PowerPoint, templates, etc.

