

Effective Tools for Digital Object Management

University of North Texas Libraries
Digital Projects Unit

Jeremy D. Moore
Lab Manager

Sarah Lynn Fisher
Digital Imaging Technician

About the UNT Digital Projects Unit:

- Supports the UNT Libraries with guidance and digital services including imaging, archival storage of electronic files, and metadata development.
- Pursues research opportunities in digital preservation and access, in partnership with libraries and historical societies across the state.
- Collaborative efforts such as the [Portal to Texas HistorySM](#) and the [UNT Libraries' Digital Collections](#) freely provide digital content to a worldwide audience.

Overview of DPU Digitization Workflow

- ❑ Organize materials to be digitized
 - Track using book flags, post-its, and carts
- ❑ List physical objects on an internal Wiki project page
- ❑ Move digital objects through the workflow in folders named for each step of the process
 - Folder descriptions are matched to the internal Wiki
- ❑ Manage data folders and digital objects using consistent identifiers
- ❑ Employ MagickNumbering system when large numbers of files make up one digital object

Physical Organization

- Organization of physical objects in a logical and traceable way ensures efficient production, i.e. the correct object is scanned only once!

DPU Physical Organization Methods:

- Carts of varying sizes
- Book flags
- Post-its to track scanning progress

Carts



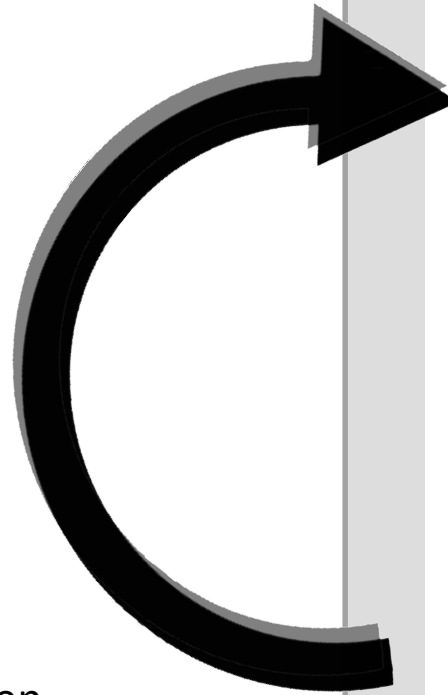
Small carts for micro-grants and smaller scale projects.



Large carts for large scale digitization projects.

Book Flag

- Inserted in each object to be scanned



- Flag is rotated 180° when scanning is complete

SCANNED

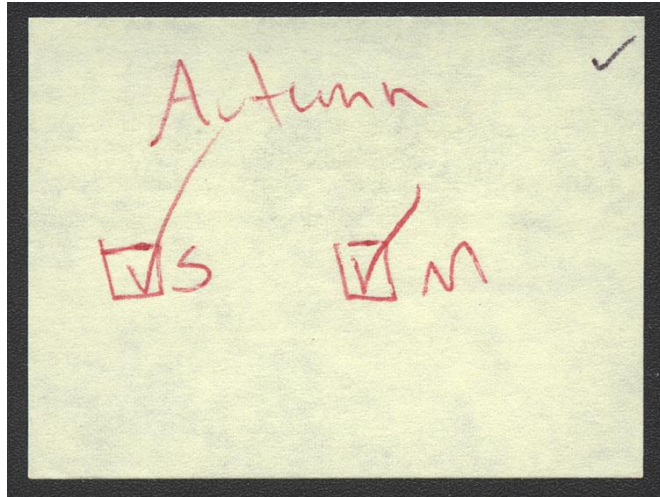
by: Leah

date: 5/14/09

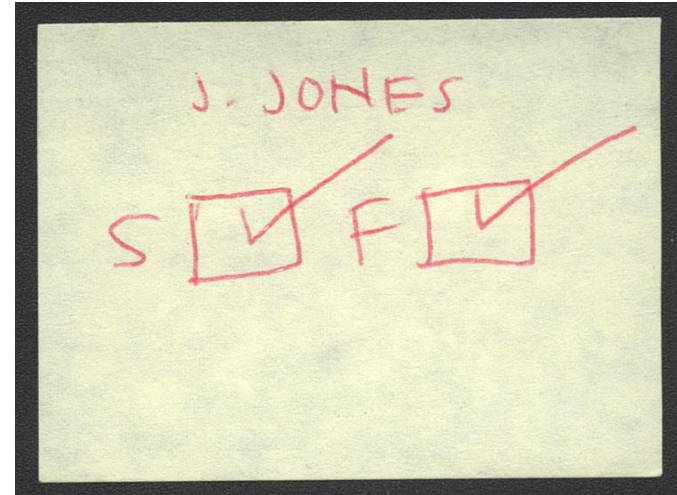
notes:
files saved here:
P:\Digi Proj\scanned_for_UNT

NOT SCANNED

Post-It Flags



"Scanned" and "Metadata" check boxes for items with metadata. Includes name of student assistant for QC tracking.



"Started" and "Finished" check boxes for items delivered in folders. Will be changed to "Began" and "Finished" to avoid confusion with the "Scanned" check box for metadata items.

Internal Wiki

- Each project is given a separate page detailing the materials to be digitized, the project's specific digitization workflow, imaging instructions, and information for metadata creation.
- Project pages also track physical location of materials
- Multiple users can easily edit content to update progress; Wiki will therefore display current project status

Examples of Internal Wiki Project Pages

Internal Serial Project

Collaborative Project

Instructions

[\[edit\]](#)

1. Read before scanning:

- [How To Scan Yearbooks](#)

2. Place your name in parentheses next to the book you are currently scanning. Once you have finished a yearbook COMPLETELY, cut and paste the asterisk and the year into the "ToQC" section. After the year add the computer station # (taped to the top of your computer), your name, and the date you finished.

Example:

- 1917 (Computer #4, Jeremy Moore, June 13, 2008)

Important: Please Read

Note any scans where text is cut off on the edge of the page in a text document (scan_problems).

When magic numbering, start with object number 3 (in most cases). (The front cover and front inside cover are object numbers 1 and 2, respectively.)

All black and white pages should be scanned in bitonal, 600 dpi.

Scan page in color if there is any color ink on the page (excluding added handwriting) or if the color of the page is different from the white pages in the book.

When you are finished scanning, save files in a folder labeled by the yearbook year (e.g. 1999) in the 1.ToQC folder here: P:\DigiProj\scanned_for_library_yearbooks

[\[edit\]](#)

Yucca/Wings/Aerie

ToScan

[\[edit\]](#)

- 1915 **M**
- 1919 **M**
- 1922 **M**
- 1926 **M**
- 1928 **M**
- 1935 **M**
- 1943 **M**
- 1944 (No yearbook for this year)
- 1949 **M**
- 1958 **M**

Identifiers

- Identifiers catalog a digital object
- A unique identifier is used throughout the digitization process for filenames and folder structures
- Identifiers link digital objects to their corresponding metadata records
- Effective identifiers ensure that digital objects will not be lost
- A good identifier is:
 - reasonably unique
 - inexpensive to generate

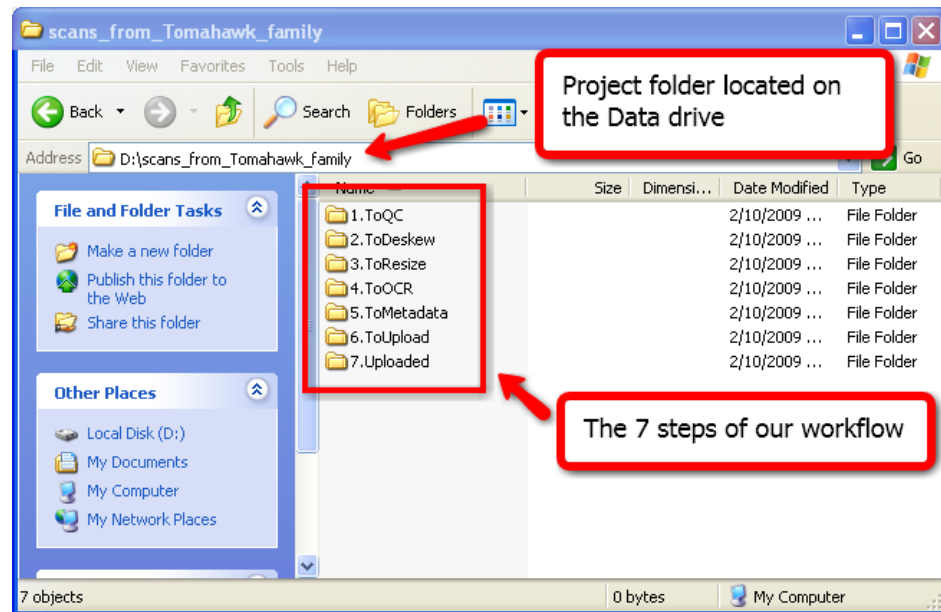
Folder Management

- Main folder is named “scanned_for_xxx,” where “xxx” is the name of the project. Example: “scanned_for_Amon_Carter” is a project funded by the Amon Carter Foundation.
- Folders within the main folder are named for each step of the digitization workflow:

- Example:

1. ToQC
2. ToDeskew
3. ToResize
4. ToOCR
5. ToMetadata
6. ToUpload
7. Uploaded

- These steps may vary from project to project.



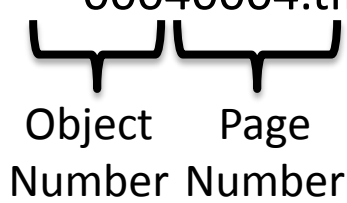
Folder Management

- Digital objects move through the workflow folders as each step of the process is completed
- Within the workflow folders 1 Object + 1 Metadata Record = 1 Folder
 - Object may be a single photograph or a 600-page book
 - Metadata record (.xml files) is always grouped with the objects it describes as the object folder moves through the workflow

Naming Object Folders and Files

- Use only A-Z, a-z, 0-9, underscore (“_”) when naming folders and files
- Match the name of the file to the object folder
- Separate more than one digital image with a suffix (_01)
- Use the three letter extension for a file: tif, jpg, png, txt
- Objects made of many digital images require special naming consideration

MagickNumbering!

- Internal naming schema for books logically handles both object-order and pagination.
- Particularly useful during the Quality Control process
- MagickNumbers consist of 8 digits:
 - 00010001.tif
 - 00020002.tif
 - 00030003.tif
 - 00040004.tif

Object Number Page Number
- Roman numerals and letter codes allow for different content to be entered in the page number section:
 - fc = front cover
 - fi = front inside
 - tp = title page
 - i = roman numeral “i” or “I”

Summary

- Organization starts with the physical objects.
- Consistent tracking is key for physical and digital objects.
- Establish naming/identifier conventions before digitization begins.

Questions?