



**The 5 W's of Planning an Imaging Project – Part 2
(Who and How)**





State and Local Records Management Division | Texas State Library and Archives Commission
(512) 463-7610 | slrminfo@tsl.texas.gov | <https://www.tsl.texas.gov/slrsm>

Welcome to our Webinar

To hide or unhide your Control Panel
Click the arrow key at the top of the Grab Tab 

Trouble hearing with headset or computer speakers?
Click the [Audio Setup](#) link in the Audio Pane for source selection and/or adjust the volume settings on your speakers or computer

Have a question or comment?
Type in the blank box in the Questions Pane
Click "Send" 

Course Credit 

✓ To obtain a certificate of completion:

1. Go to the online course; <http://bit.ly/imagingpt2>
2. Create an account or log in.
3. **Click the "Enroll me in this course" link!!!**
4. Score a 70% or higher on the 5-question quiz.
5. Save or print your certificate.

HELP: Email us at slrminfo@tsl.texas.gov

✓ An archived recording and all supplemental materials will be available at the link above (<http://bit.ly/imagingpt2>)

About Us

- Records Management Assistance unit
 - 6 Government Information Analysts
 - Consulting and Training
 - Retention, destruction, imaging, managing email, disaster preparedness and recovery...
 - Retention Schedule Reviews/Development



Imaging Part 1: Recap

- ✓ Understand importance of metadata / indexing
- ✓ Scan selectively: take the time to outline the expected benefits and conduct feasibility studies
- ✓ Destroying the source document: know the rules, understand the challenges

<http://bit.ly/imagingpt1>

Objectives of Part 2

- ✓ Know the pros and cons of in-house projects vs. outsourcing
- ✓ Understand the processes involved with imaging
- ✓ Be better equipped to make decisions regarding an imaging project

WHO?

**DO IT YOURSELF?
HIRE A VENDOR?**

Questions to ask

- How much material is there?
- Timeframe?
- Confidentiality or security concerns?
- Condition and format of the materials?



In-House: Pros

- “Learn by doing”
- Make adjustments along the way
- More control over quality and security
- Ensure that project meets requirements



In-house: Cons

- Larger organizational investment
 - Difficult to determine per-item cost
- Limited production capabilities and facilities
- Staffing considerations



Vendor: Pros

- Costs are more predictable
- Expertise, training, technology obsolescence costs absorbed by vendor
- Can handle large volume and high throughput
- Can do parts or all of a project



Vendor: Cons

- Security, handling, and transportation issues
- Potential communication challenges
- Vulnerability due to vendor instability
- Organization is removed from imaging functions



Who to involve?

- Consider using a systems integrator
- Involve technical staff early and often
- Consider your users' needs
- Get support from upper management

Imaging Contract

- New contract with multiple pre-approved vendors
- Three categories of services offered:
 - Project-based
 - Flat rate
 - Preservation
- Contact Nan Pfiester @ 512-463-5477 for more information
- Check Council on Competitive Government website for contract documents, pricing, user's guide, sample statement of work, and vendors' websites
http://ccg.state.tx.us/contracts/doc_image.php

HOW?

COMPONENTS OF AN IMAGING PROJECT

Components of imaging project

- Inventory & workflow analysis to help plan project
- Staff training
- Select materials to be imaged
- Document prep
- Create/choose standardized metadata
- Quality control and document processing
- Long-term maintenance and preservation

Components of an Imaging Project



<http://www.library.cornell.edu/preservation/tutorial/management/management-01.html>

What to decide

- In-house or vendor?
- Hardware selection – scanners and storage
- Software
- Bit depth and DPI (dots per inch)
- OCR (optical character recognition)
- File formats
- Compression

Hardware selection

- Hardware serves three functions:
 - Captures, stores, and retrieves images

- Hardware includes:
 - Computers
 - Scanners
 - Storage



- When selecting hardware, plan according to anticipated peak workload

Hardware - Scanners

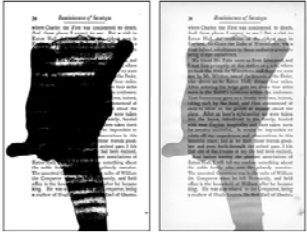
- Feed methods
 - Flatbed
 - Automatic document feeders (ADFs)
 - Overhead
 - Simplex (one-sided) and duplex (two-sided)
 - High speed
- Throughput
 - The speed at which a scanner processes a page
 - 60 ppm = 60 pages per minute

Bit Depth

- Bit depth is the number of bits used to define each pixel
- The higher the bit depth, the greater the number of tones represented

Bit depth	Displays	Recommended for
1-bit or "bi-tonal"	black and white	Typewritten documents
8-bit grayscale	256 shades of gray	Black and white photographs, half-tone illustrations, handwriting
24-bit color	Approximately 16 million colors	Color graphics and text, color photographs, art, drawings, maps

Bit Depth

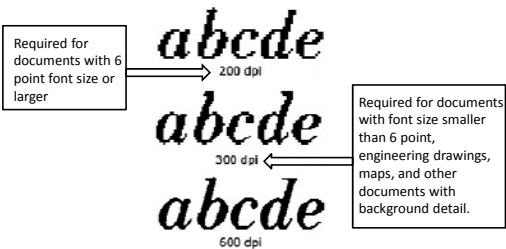


Scanning Text: Compare bitonal (left) and grayscale (right) scanning of a slanted text page.

DPI

- DPI is how resolution of image is measured
- Dots per inch
- You may also see ppi (pixels per inch) or lpi (lines per inch)
- The higher the number of dpi, the higher the legibility of reproduced image
- Higher dpi means larger file size, takes longer to scan

DPI



Hardware – Storage

- Network backed storage is ideal
- Choose storage media that is widely popular and accepted
- Buy enough storage to last you a few years (and for backups too)



Hardware - Storage

- Criteria for evaluating storage media:
 - Speed (read/write, data transfer)
 - Capacity
 - Reliability (stability, redundancy)
 - Standardization
 - Cost
 - Fitness to task

QUESTIONS?

Software

- Software performs tasks such as:
 - Scanning
 - Optical Character Recognition (OCR)
 - Indexing
 - Database Management System (DBMS)
 - Image Retrieval



File Formats



File formats

- Directly affects the quality and file size of images
- Consider:
 - What kind of materials you're imaging
 - Retention periods
 - Frequency and nature of use

File Formats

- Generally, open-source and non-proprietary is ideal
- Otherwise: widely available and accepted formats, and formats that have become industry standards
- Formats should be stable, well-supported, and well-documented

Compression

- First, decide if file compression is acceptable
- Lossless – Resulting file looks identical to uncompressed file (TIFF)
- Lossy – lose data during compression process (JPEG)



Effects of Lossy Compression on Text: Close-up comparison of a section from a map saved in lossless GIF (left) and lossy JPEG (right).

QUESTIONS?

What to do

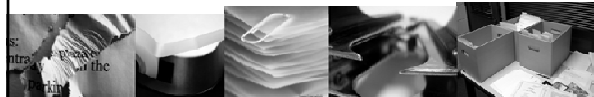
- Staff selection & training
- Document prep
- Adding metadata
- Indexing
- Quality control
- File maintenance



Prepare Files

Make documents “scanner ready”

- Remove documents from folders/files (unfold or unroll if necessary)
- Arrange files in order, correct misfiles, and remove duplicates
- Conduct audit or inventory
- Remove all staples, paper clips, post-its, flags, etc.
- Mend torn pages



Metadata

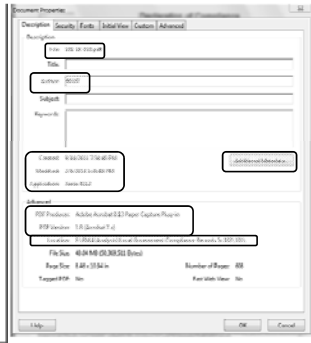
- Metadata helps describe, track, organize, and/or maintain an image
- An indexing database will help ensure efficient and accurate retrieval of records
- Important to capture image metadata at or near time of creation



Metadata – what to capture

- Title
- Creator/Author
- Date (Date of document creation/creation of image)
- Unique identifier (such as a coding system for different document types)
- Format (OS, software configurations, versions)
- Definition
- Rights/security
- Data type (written document, photograph, etc.)
- Keywords
- Comments

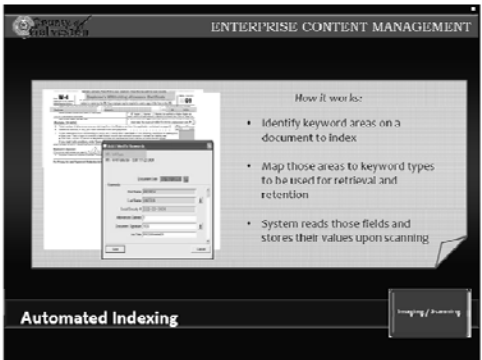
Metadata



Indexing Examples

- Employment Applications
 - Date, name, position, hiring department
- Purchase Orders
 - Date, purchase order #, requisition #, vendor name





Automated Indexing


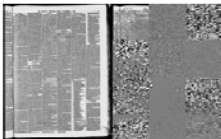
Enterprise Content Management: E-Mail, Paper, Applications, and Shared Drives:
Covering all the Bases of Hard-to-Manage Records

How it works:

- Identify keyword areas on a document to index
- Map those areas to keyword types to be used for retrieval and retention
- System reads those fields and stores their values upon scanning


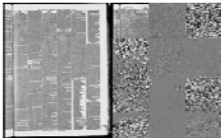
Quality Control During Imaging

- Periodically test or adjust scanner settings
- Document QC procedures
- Back up your scans



Quality Control During Imaging

- Perform a visual quality control check on every image
- Ensure no information is lost
- Verify index is correct and user-friendly



Document Processing

- Rotating
- Deleting blank pages
- Renaming

OCR

- Optical Character Recognition
- Text in a document becomes machine readable text
- Allows for full-text searching or editing
- Errors are inevitable
 - Scanning at higher resolutions results in less errors

Maintenance

- Try to store media in optimal environmental conditions
- Ensure security of media
- Refresh or migrate media periodically
- ...and don't forget about retention and disposition!
- Digital Preservation Webinars:
<http://bit.ly/DigPres-One> <http://bit.ly/DigPres-Two>

CONCLUSION

Imaging Part 2: Summary

- Pros and cons of imaging in-house or using a vendor
- What to decide
 - Hardware, software, bit depth, DPI, file formats, compression
- What to do
 - Document prep, metadata, indexing, quality control, maintenance

Case studies

Annual e-records conference:
<https://www.tsl.texas.gov/slrn/training/conferences/index.html>
– 2012: Galveston County
– 2011: TSLAC Records Management Assistance, University of North Texas



Thank you for attending!

- Take the 5-question quiz: <http://bit.ly/imagingpt2>
- Questions for Marianna:
msymeonides@tsl.texas.gov or (512) 463-5448
- To exit:
 - Unhide the Control Panel (click red arrow key)
 - File / Exit – Leave Webinar
 - Please fill out the survey that pops up! ☺

Q&A
