2014

University
Employee
Guide to:

Document Imaging and Management of University Records

Questions:
University Records Officer
recmgmt@library.wisc.edu
Voice: 608-262-3284

The capture and imaging of paper records, is an increasingly popular strategy for Academic and Administrative units and offices looking to manage their records electronically. When implemented appropriately, digitization of records can dramatically improve workflow through access and retrieval of university records; while at the same time reduce the need for physical records storage space.

However, digitized records introduce new responsibilities, costs, and challenges not present with the management of paper records. This guideline is intended to assist UW-Madison units/departments in determining if digitization of their records is an appropriate course of action and if so, to assist them in developing a secure and robust digital imaging program with associated processes.

This document covers the following:

- **BENEFITS AND QUESTIONS ON DOCUMENT IMAGING**
- **NEEDS ASSESSMENT AND COST ANALYSIS**
- **SYSTEM SPECIFICATIONS AND SELECTION**
- **TECHNICAL SPECIFICATIONS**
- **INFORMATION SECURITY AND TRAINING**
- **THE DOCUMENT SCANNING PROCESS**
- **INDEXING AND METATDATA**
- **STORAGE AND MIGRATION**
- **LEGAL ISSUES AND RECORDS MANAGEMENT**
- **DEFINITIONS**

University Records that are scanned and image are required to be authentic, complete and accessible and relied on as evidence.

> *Document imaging or Digitization is the conversion of paper-based documents to digital images, making them readily accessible, thereby enhancing the business processes and workflows for university departments and units on campus.*

Authentic: The product of routine, authorized copying and registration processes.

Complete: An accurate, legible reproduction of the original, without substantive changes or deletions of content.

Accessible: Available and readable to all those with the rights to access it, for as long as it is required.

Careful analysis and planning is essential for the successful development and implementation of an imaging project and ongoing management of information in the system. The University Records Officer can assist in this.

Points to consider:

- *Scanning is not a substitute for development of good records management practices. The problems caused by poor record and filing systems will not be improved in a digital format and will get worse without a plan.*

- *Scanning should be implemented as part of a larger information management strategy which should be addressed by the campus or departments /unit.*

Questions to ask before beginning document imaging and scanning:
- Is there a need for imaging?
- Do benefits outweigh the cost?
- Is imaging a "nice to have" or a "must" system?
- Will imaging system make the department more efficient and effective to faculty, staff and students?
- Does current staff have the time and resources to implement the imaging system? *(see Imaging Limitations below)*

Benefits of Document Imaging and Capture:
- Concurrent access -two or more staff retrieving same file
- Rapid file retrieval -high retrieval of active documents, better customer service
- Image Enhancement
- Workflow -automate and control certain business processes (schedule processing, route documents, track document status)
- Physical space savings
- Preservation and protection if the records have archival value

Limitations to Document Imaging:
- Costs for -document preparation, indexing, quality assurance, hardware, software, maintenance, administration
- Risk of hardware and software obsolescence
- Need a uniform Taxonomy and classification system for images
- Need for a migration plan for long term or archival records

The cost of an imaging system includes not only the initial purchase of the hardware, software and technical support, but also on-going costs. The university has a campus wide license for imaging software. This is the recommended solution for document imaging of university records.
_____

To begin evaluating whether scanning records will be beneficial to your academic or administrative unit review the following steps.

STEP 1: NEEDS ASSESSMENT AND COST ANALYSIS

Before embarking on a digitization program, units/departments need to assess whether digitization is an appropriate solution for their records management needs. In general, digitization is best implemented in cases where quick, simultaneous, and/or distributed access to records is necessary for fulfillment of a

department's responsibilities, or when such access would significantly increase the efficiency of a department's operations.

Digitization is <u>NOT</u> an appropriate solution for reducing storage costs for paper records, as the cost of hardware, software, server space, training, and maintenance of systems can greatly exceed the cost for storing those same records in paper form.

Records that have a <u>high reference value</u> and <u>short retention</u> are best to image.

Some of the factors that must be evaluated during the initial needs assessment include:

- **Program Purpose:** Why is the digitization being undertaken? Will digitized records be maintained in lieu of or in addition to paper records?

- **Business Process Evaluation**: How will the digitization program improve the unit/departments business processes or workflow?

- **Information Security:** Who will have departmental access to images? How will images be secured from unauthorized access? There should be a documented process in place for access to sensitive information.

- **Amount and Accumulation:** How many records are to be digitized? What is the estimated rate of addition per year to these records?

- **File Format:** What type of document (textual, photograph, map, etc.) is to be captured and imaged? In what file format will the record be maintained in? (jpg. pdf, tiff/ etc.)

- **Records Retention:** How long must digitized records be retained? Has the appropriate record series been identified from the University's Records Schedules? No public records should be digitized without a current records schedule in place.

For a cost analysis, some of the factors to consider include:

- **System Hardware and Software**
- **Image Management Application**
- **Facilities Upgrades/Site Preparation (**including additional storage space)
- **Project Management**
- **Training**
- **Staffing**
- **Ongoing Maintenance, Support, and *Upgrade*** *(generally about 10-20% annually of initial implementation cost)*

If a cost analysis does not yield a net benefit for a digitization program, consider retaining files in paper form in your office, or explore off-site storage at the State Records Center.

STEP 2: SYSTEM SPECIFICATIONS AND SELECTION

Due to the broad scope of most imaging projects, departments are strongly encouraged to undertake document imaging if done through an enterprise-wide document management system. UW-Madison has a campus-wide license for a document imaging/content management system for use by university departments. If your department will not be using the campus wide license then the following system specification should be requested in a RFP.

Request for Proposals for a Digital Imaging Systems should include, at minimum, the following requirements:

- **Open systems architecture**, including non-proprietary compression standards. This type of architecture allows the system to be upgraded over time without a significant risk of records loss. It also supports the importing and exporting of digital images to and from other sources. If proprietary standards or architecture are unavoidable, the vendor should provide a bridge to systems with non-proprietary configurations and/or license the software beyond the length of the contract.

- **Specifications for hardware/software that will require vendors to support and maintain their product(s).**

- **Controls and system auditing tools.** Effective audit trails can automatically detect who had access to the system, whether staff followed existing procedures, or whether fraud or unauthorized acts occurred or are suspected.

- **Image authenticity/integrity tools.** The system should ensure that the images are protected from accidental or intentional modification. Equipment should also conform to methodology for media error detection and correction.

- **Records management system integration.** System records should be linked to approved RRDAs and retention periods and include provisions for automatically or manually purging records beyond their scheduled retention dates. See Records Management, below.

- **Appropriate document scanning resolution.** Consider data storage requirements, document scanning rates, and the accurate reproduction of the image. See Technical Specifications, below.

- **Access to records.** Systems should use an indexing system database that provides for efficient retrieval, ease of use, and up-to-date information on the scanned images stored in the system. The index storage method should be based on standard relational database technologies with access using standard SQL queries. See Metadata Integration, below.

- **Appropriate levels of security.** Systems should ensure that only authorized personnel are able to create copy, modify, or use scanned images within the system. Different types of scanned records may include different security requirements. See Information Security, below.

STEP 3: TECHNICAL SPECIFICATIONS

Digitization technologies allow offices to control the resolution, size, color, bit-depth and other qualities that affect how the image appears on a computer screen or is output to a printer. Furthermore, once captured, a digital image can be saved in numerous file formats that may or may not include compression technologies that reduce the file size of the file. Choices offices make in these areas need to be cost-effective while still producing an accessible, accurate, authentic, reliable, legible, and readable record throughout its life cycle.

Convenience copies, those that are not used for preservation but to be used in the office, may be of more diverse formats and resolutions in order to best fit the needs of the office. The office, for example may wish to create JPG or PDFs files from the source files that are of lesser resolution and are compressed for day to day use.

STEP 4: INFORMATION SECURITY AND TRAINING

Departments should appoint a staff member, preferably one with systems administration experience, as the administrator of any digital imaging system. This administrator should be responsible for overall project management, and the development and maintenance of written system documentation which describes the requirements, capabilities, limitations, design, operation, and maintenance of the digital imaging system. All other personnel to be given access to the system should undergo comprehensive training on the system before being granted privileges to add or dispose of records. For security reasons, only personnel who require access to digitized records for their daily job duties should have access to the digital imaging system. Departments should consult with the CIO Office of Campus Information Security. http://www.cio.wisc.edu/security-risk.aspx

STEP 5: THE DOCUMENT SCANNING PROCESS

Document Arrangement:
Prior to scanning, documents to be imaged should be arranged in such a way that the organization of those documents is clearly discernible.

Document Preparation:
Office personnel in charge of scanning or coordination with scanning vendors should also prepare documents for efficient processing (remove staples, unfold paper, remove extraneous documents, etc.). For the sake of consistency and security, employees responsible for scanning should be specially designated and trained for this purpose.

Once scanning has completed, digital images should be inspected by the system administrator or other responsible party to ensure the accuracy, legibility, and readability of the documents. In cases of scanning projects with very large numbers of documents, a visual quality evaluation of a sample of documents may be appropriate.

Document Scanning Resolution

When determining document-scanning resolution, consider data storage requirements, document scanning throughput rates, and the accurate reproduction of the image.

- 300dpi –Use a scanning density of 300 dpi for good quality images in imaging modern office records

- 600 dpi or higher scanning density is appropriate for deteriorating documents, and documents with a visual element such as engineering drawings, maps, or documents with background detail.

The display resolution of the inspection/verification monitor and printer should match the scanning density of the document scanner. When scanning continuous tone images, such as photographs, maps, and illustrations, use gray scale or color imaging technology.

Usually Grayscale is the most commonly recommended format. Black and White will give you the smallest file type but will also produce the lowest quality scan.

File Format or Scheme

a. In most instances you will want to work with the PDF file format, there are some document Management systems that prefer TIFF files.

- PDF files allow you to create multipage files that are ideal for most large scale document scanning projects.

- They can become completely searchable, with use of an OCR program, allowing you to quickly jump to relevant information within the document.

- Acrobat Reader is a free program that allows you to open and read PDF files and one that can be downloaded by anyone.

- With recent advances in the PDF/A file format you can expect PDF files to be around for many years to come.

Image Types	Bit Depth	Color Mode	Resolution (ppi/spi)	Scale	File Format	Compression
B&W Text	1-bit	B&W (bitonal)	600 ppi/spi	100%	TIFF	CCITT Group 4
B& W Text with Illustrations (charts, artwork, graphs, photos)	8-bit	Grayscale	400 ppi/spi		TIFF	None
Color Illustrations with Text	24-bit	RGB	400 ppi/spi		TIFF	None

(From National Archives and Records Administration' Technical Guidelines – pp.32-36)

Next perform a visual quality control evaluation of each scanned image and related index data.

Perform a Quality Check of Images:
In imaging review the source materials to be imaged:
- how old they are,
- what type of substrate they're on, if they are handwritten or typed, if they include multi-part forms on colored paper,
- how worn/used the materials are,
- if there are two-sided sheets mixed in the materials or if there are any NCR-type originals or odd sized things, such as legal sized or receipts/note pages.

These documents all have different requirement for getting a good quality scanned image. Assess your source material by taking a sample set and then image them at a variety of resolutions and view. These should be viewed on the scanning workstations and also work stations where they would be most viewed. Not all monitors are the same. You can also gauge the files sized for images scanned in the various resolution and can make a good estimate of the volume of storage that will be needed to store the images. To help ensure the integrity of long-term and archival records stored on the system, staff members should perform a visual quality evaluation of each index entry and scanned image. Overall system quality control is best when the scanned image is temporarily stored on magnetic media, permitting corrections through rescans as needed. Depending on the system configuration, corrections may be performed at the scanner capture station or at designated inspection/rescan workstations. Training and supervision of the operations staff is a key factor in maintaining acceptable image and index quality as well as user satisfaction with the system.

STEP 6: INDEXING AND METADATA

Use an indexing database that provides for efficient retrieval, ease of use, and up-to-date information about the digital images stored in the system. The indexing database should be selected after an analysis of unit operations and user needs. Reliable access to scanned images depends on an accurate, up-to-date index database. Indexing a digital image involves linking descriptive image information with header file information. Normally, index data is manually key-entered using the original documents or the scanned images, either at the time of image capture or later in the production process. Index data verification, in which database entries are compared with the original source documents for completeness and accuracy, is crucial because an erroneous index term may result in the inability to retrieve related images.   A barcode page with metadata can also be used which contains customized information and came be indexed immediately with a barcode indexer.

Without proper indexing, information contained in an image file will not be easily located and could be irretrievable lost.   Metadata is information associated with digital images that describes the content and structure of the digital image and its context of creation. Unit/Department must capture metadata that will allow the records to be identified, organized, searched and preserved. To increase the accessibility and ease of retrieval of digitized files, metadata should be included within records or linked to them for this purpose.

Metadata is "structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource." (NISO, Understanding Metadata)

METADATA ELEMENTS	OBLIGATION
Identifer	Mandatory
Title	Mandatory
Subject	Optional
Description	Optional
Creator	Mandatory
Date	Mandatory
Addressee	Mandatory for email, Optional for other records
Record Type	Mandatory, where applicable
Relation	Mandatory, where applicable
Function	Optional, but highly recommended
Aggregation	Mandatory
Language	Mandatory
Locaton	Optional
Security & Access	Mandatory
Disposal	Mandatory
Format	Mandatory
Preservation	Optional

For more information on metadata Standards:
http://dublincore.org/documents/2008/01/14/dces/

Chart 1 Metadata Elements from the Dublin core.

Consistent data entry practices (names and date formats, etc. and controlled vocabularies are recommended to increase accessibility. A workflow to add this metadata to records at the point of

digitization should also be created. Where appropriate, records creators should use controlled vocabularies to increase accessibility and group like documents together.

Taxonomies provide a structure for describing and organizing records based on their function while Metadata describes the information about the records, such as data, origin and document type. A file plan identifies categories to aid in the folder arrangement and can aid in the classification of the scanned images which correlate back to the department or unit filing system. Implement this filing system into your electronic recordkeeping system.

***Begin imaging according to the original physical arrangement of the records* and require unique identifiers or metadata elements to facilitate their management.** The more metadata captured the more complete the record will be.

 Chose a naming convention that allows you to quickly identify the file that you want to open:
   i.   Example: If you are a doctor's office, a naming format that closely represents the format used with the paper documents often works best. (LastName_FirstName_DateOfBirth.pdf)

   ii. For large scale projects chose a name that can be quickly identified within the paper documents. If the information you need resides on one page that is located anywhere within the first 10 pages of the file you may find yourself wasting valuable time trying to name each file when another name could be used instead and is much easier to identify.

The key is consistency in using naming conventions for access points.

STEP 7: STORAGE AND MIGRATION
What kind of reliable long-term storage methods are available the will both maintain the authenticity of the records as evident and also provide a means to search and retrieve the records for access. The file format and size of the image will assist in determining which storage method to use and what means of tracking, retrieving and viewing the scanned image. Regular maintenance is required and a migration plan should be documented when the storage method becomes outdated as required in Wisconsin Administrative Rule 12.

It is recommended that source files of digitized records be stored on a network server or as part of an enterprise-wide document management system and NOT on removable media.   Appropriate campus IT staff must be notified if the official copy of any public record is to be stored on campus servers. If, for whatever reason, it is not possible to preserve source records on a server, it is recommended that these records be stored on removable WORM (write once, read many: e.g. CD-R, DVD-R) discs only. Rewritable media are generally unacceptable storage media for digitized records because they do not preserve the authenticity (fixity) of records. At least 2 copies of each disk should be made and kept in separate, secure, locations. Be sure to label external storage media with particular care since it is impossible to determine content merely by looking at a disk or tape.

Although the term "open systems architecture" is defined in various ways, Campus departments/units should follow a system design approach that permits future component upgrades with minimal degradation of system functions. This open system architecture allows the system to be upgraded over time without a significant risk of records loss. It also supports the importing and exporting of digital images to and from other sources.

One key factor in achieving open systems architecture is the adoption of non-proprietary standards. **The flexibility of an open systems architecture helps enable long-term records to be accessed and transferred from one hardware or software platform to another.** Digital Imaging Systems which contain records with retention periods longer than 5 years should include provision for migrating records, i.e. converting images and indexes to newer file formats or storage media to prevent hardware or software obsolescence. A migration strategy will document how a department will transfer long-term and archival records from one generation of hardware and software to another generation without losing system functionality. The strategy should be written and available with current system documentation.

STEP 8: LEGAL ISSUES AND RECORDS MANAGEMENT
Records maintained within a digital imaging system are subject to open records requests under **Wis. Stat. 19.31-19.39**, and must be producible on short notice as required.

Digitized copies of paper records may be considered the copy of record under Wisconsin Administrative Policy ADM 12 if the records in question meet the following criteria:

1. **Accessible:** The records can be retrieved for reference or access within a reasonable period of time.

2. **Accurate:** The retrieved file correctly reflects the original record.

3. **Legible:** The letters, numbers, and symbols in the document are uniquely identifiable.

4. **Readable:** The records can be opened on an accessible program and easily read by any and all users.

5. **Reliable:** The electronic record reflects the initial record each and every time it is accessed.

6. **Authentic:** The electronic record correctly reflects the input of creators and editors and can be substantiated.

If all six of these Administrative 12 criteria are satisfied throughout the retention period specified by a record's retention schedule, the electronic copy is considered the official record,

*and the paper input may be discarded. Please keep in mind that these are the criteria for records concerns **only**; if your records contain student information or other confidential information, you must also consider requirements to keep these confidential records secure.*

Records stored within digital imaging systems must be managed according to Wisconsin Stat. 16.61, which requires Records Retention Schedules to be in place for any and all records series created and maintained in the course of University business. Records Retention Schedules define a record series, provide a retention period for that series, and gives instructions for records disposition (destroy, destroy confidentially, or transfer to Archives). It is recommended that records with temporary disposition be destroyed upon or shortly after expiration, in order to protect records integrity in the event of a security breach.

A readability and retrieval check must be done at least every 3 years and transfer to new storage media at least every 5 years.

Privacy and Security Standards
Departments are to ensure that systems and practices used for managing and maintaining imaging systems and data comply with applicable state and federal regulations and standards, including, but not limited to:

- Family Educational Rights and Privacy Act (FERPA) (12 U.S.C. 1232g and implementing regulations)
- HIPAA - Medical Privacy for confidentiality of medical information
- Protection of Personally Identifiable Information such as Social Security Numbers.

Other Records Management issues to consider as part of the imaging process:

File Maintenance
Departments must have in place procedures for maintaining and filing source and imaged documents throughout the required and approved retention period.

Back Up
Design backup procedures to create security copies of digitized images and their related index records. System component reliability is critical to system success. Prolonged or repetitive downtime can seriously affect office operations. Creating a duplicate copy of records in another format or another system is an effective method of ensuring access to long-term information. Backup copies also support system integrity and legal admissibility requirements. The Department/Unit may select the backup storage media (optical, magnetic, paper, or microform) that best meets the office's records requirements. Security copies of the records should be stored in an offsite, environmentally controlled location. Retention of these copies should not be kept longer than the original scanned image.

Departments must have in place procedures for regularly backing up imaged documents and associated indexes. The backup copy is to be available if the record copy becomes inaccessible due to disaster, vandalism, theft, malfunction, or any other cause.

Disposal
Departments must have procedures for properly eliminating or disposing of imaged records that exceed retention periods. NOTE: Just deleting the pointer in the index is not sufficient. The actual record must be deleted from the media.

Archival Records
Departments are responsible for ensuring the safekeeping of all records, including those identified as "archival" or "potentially archival." Archival documents are retained in accordance with the applicable retention schedule and transferred to UW-Archives.

Final Recommendations:
It is recommended that university departments/unit interest in digital imaging utilize the Campus-wide License for Digital Imaging and capture system.

Refer to the DoIT Knowledge Base for Imaging - **https://kb.wisc.edu/imaging/**

For all Departmental/Unit imaging projects contact the University Records Officer for Records Management Consultation to begin the process of developing a records management plan for conversion of these records.

Contacts:
Peg Eusch, CRM University Records Manager   meusch@library.wisc.edu

----

Definitions:

ASCII text file: A file that contains characters represented by the American Standard Code for Information Interchange. ASCII text files, which can be produced by many word processors and other computer program, are typically devoid of formatting information.

Archiving: The process of moving copies of less-frequently accessed electronic records from primary on-line storage devices to secondary storage media for long-term storage. Archiving may take place automatically by a predetermined time period, Predetermined level of file activity on the primary on-line storage device, or at the user's request.

Backup: The process of duplicating information, primarily for protection against loss or damage.

Digital Document: Computer-process able records that have both the following characteristics: (1) they are created by computer programs for purposes that would otherwise be served by

paper or photographic documents. If a digital Document did not exist, the same information would and could be created and maintained in non electronic form. (2) They can be printed to produce paper of photographic document of comparable content, appearance and functionality.

Electronic Imaging: The process of capturing, storing, and retrieving documents, or data, regardless of its original format, using micrographics or optical disk technologies.

File Plan: A systematic categorization scheme that groups documents pertaining to a given matter. A file plan defines topical or other categories into which documents will be grouped, and a folder is established for each category. Within a given folder, digital documents are stored and individually labeled computer files. Also know as file taxonomy.

Index Database: A database created to keep track of digital documents that related to a given person, account, case claim, subject, or other matter.

Metadata: data about data, or information about records (e.g. which describes the record's context, content and structure, and their management through time).

Office of record: usually the office which originates the record or which is directly and primarily responsible to take action on a task or transaction.

Source documents: paper documents that contain information to be converted to electronic records. A record – hard copy or original- on which and original transactions was captured.

Workflow: A process that automates the routing of documents among designated recipients according to user-defined rules and relationships.

FOR MORE INFORMATION:
Records Management: http://archives.library.wisc.edu/records/Index.html
Information Security: http://www.cio.wisc.edu/security.aspx
Office of Legal Affairs: http://legal.wisc.edu/
DoIT Knowledge Base for Imaging - **https://kb.wisc.edu/imaging/**

Sources:

University of Wisconsin System Administration. Office of General Counsel. Digitization Systems: Overview and Caution on the Imaging of Public Records.

University of Wisconsin System Administration. Office of the General Counsel. Records Management Guidelines for Information Technology Systems.

University of Wisconsin – Milwaukee Information Technology Policy on Digitization of Department Records No. S-63 4-24-2013

Harvard University "Guidelines for Scanning University Records"