University of Wisconsin-Madison
A CALL TO ARMS
Archives and Records Management Services

Bulletin # 7 December 2001

University departments and units have access to a variety of information and records storage technologies. The choices they make have significant ramifications for the ongoing management of University records and information. Because of frequent technological change, including hardware and software obsolescence, University offices must take steps to manage and protect their records for as long as they are needed to meet operational needs and protect the legal, financial and historical interests in their records. For campus units in the process of deciding upon the appropriateness of digital imaging technology, this Bulletin answers several basic questions about this technology and provides a set of recommendations that address: project planning, systems specifications and selection, and system implementation.

CONTENTS

REPRODUCTION OF UNIVERSITY RECORDS USING DIGITAL IMAGING SYSTEMS

Guidelines issued by the University Archives and Records Management Service and endorsed by the Campus Records Review Group.

INTRODUCTION

University departments have a wide variety of information and records storage choices. These choices bear directly upon how well a department will be able to access their information and records now and in the future. Increasingly, the following questions are being asked: " Is it acceptable to scan University records and store them on CD? " "Will the University Archives accept CDs for archival preservation? " "If I scan my documents and burn a CD, is there a set of criteria that I should follow? " "Is a CD a legally acceptable storage medium for official records? " This Bulletin responds to these questions and provides recommendations for the development and implementation of imaging systems for University records.

The Guidelines are advisory and not intended to serve as a set of requirements; however, the degree to which they are incorporated into system design will greatly affect the long-term accessibility of the electronic records involved. National technical standards and professional best practices form the background for the recommendations. Many of them are listed at the end of this Bulletin.

BACKGROUND

What is document scanning? What is meant by burning a CD?

There are two types of digital imaging: 1) imaging which is essentially taking a picture of a document and storing the image or 2) imaging to convert the image from analog to machine readable data.

In the context of these Guidelines digital imaging refers to the electronic technology that captures an image of a document and converts it to an electronic format for storage. It is a reproduction technology. It requires a scanner for image

recognition; computer software, memory, and storage; and graphic terminals and laser printers to reproduce the stored images. Normally, some type of data compression software is used to compress the images prior to storage.

Compact Disc, or CD, technology is one of many information technologies that have become available for information storage. The technology is relatively easy to use and affordable. A CD-R is a type of compact disc technology. It is normally made of a rigid plastic and the more recent versions have high-density storage capacities. CDs are said to be "burned" because in placing data on the CD minute holes are burned into the plastic surface of the CD. CDs have become a popular storage medium because they can easily store large quantities of data and documents. The technology to scan and "burn" CDs can easily be configured into an office computer desktop. It permits offices to remove information from their computer hard drive or LAN and store it near line, keeping it accessible.

Is it acceptable to image University records and store them on CD?

ANSWER: Only if the unit responsible for the record has developed procedures to assure the CD will be readable during its required retention life.

University records regardless of their format are subject to Wisconsin Public Records laws and University rules with regard to records retention and disposition. University officials (including chairs, deans, and unit heads) are responsible for managing official records so that they remain accessible throughout their retention life. The establishment of a records retention schedule determines the retention life; see ARMS Manual, Chapter 3 (http://www.library.wisc.edu/libraries/Archives/rm/manual/ManualRDA_ch3_.html) for instructions on developing a records schedule.

Wisconsin Public Records Laws do not require that records be created, maintained, or stored in any particular format. Administrative Rule 12, however, does require that records that are maintained ONLY in electronic media meet certain criteria to insure their access, retrievability, and authenticity. See the ARMS FAQ on Administrative Rule 12 (http://www.library.wisc.edu/libraries/Archives/rm/faq_adm12.htm) or the Primer on Administrative Rule 12 (http://enterprise.state.wi.us/static/erecords/Primer.htm) for further description.

Are CDs acceptable as evidence in legal proceedings?

ANSWER: It is not the CD itself that will be legally acceptable; it is the information contained on the CD that may be utilized in a legal proceeding and therefore subject to the rules of evidence (see ARMS Bulletin #2 for a discussion of rules of evidence and electronic record keeping).

The issue of greatest concern is readability; can the information be accessed in the event of a court action? Rules about electronic records and record keeping are changing. Many State and Federal agencies have established rules and regulations about the maintenance of records. It is therefore advisable to monitor legal activities within your particular community or industry to determine what may or may not be legally acceptable.

It is also important to remember that all types of documentation can be subject to subpoena and discovery action. Courts are generally unsympathetic to a government office that cannot retrieve specific files or records and may rule unfavorably against the party who has lost the record.

What are the implications of storing University records exclusively on CD?

ANSWER: There are several implications for storing University records on CD, and University departments need to be aware of the implications of such a move and develop procedures to meet the rules.

First, it is important to remember that University records are in large part public records. State records law does not mandate that records be maintained in any particular format. Wisconsin's recent administrative rule on electronic records management does require, however, that in instances in which State agencies choose to retain their records in electronic formats ONLY, certain criteria must be met. In other words, if the

image copy will be the only record retained or the paper original will be destroyed following the imaging process, the requirements outlined in Admin Rule 12 apply. In essence, the rule requires that records remain accessible, retrievable, and authentic throughout their retention life.

Electronic records require far more care and maintenance than their hard copy counterparts. For example, the signed, original time sheets must be maintained for a period of 5 years and then can be destroyed. Throughout that 5 years, the time sheets must remain accessible and retrievable if necessary to satisfy administrative need or in the case of audits. If the time sheets are imaged and stored on a CD, the department must make sure that the CD is readable throughout that 5-year retention life. To insure readability, the office will need to retain hardware and software that allow readability.

RECOMMENDED PROCEDURES FOR UNIVERSITY DEPARTMENTS AND OTHER UNITS CONSIDERING DIGITAL IMAGING FOR RECORDS STORAGE

The text of the following recommendations are extensively patterned after the "Digital Imaging Guidelines for State Government Records" - Kansas State Historical Society. We are grateful to them for allowing us to use them in preparing this Bulletin. Some language and references to specific processes have been modified to fit those of Wisconsin State government.

A. Project Planning

Recommendation 1: Prior to selecting a digital imaging system, conduct a records and workflow analysis to determine and document existing and planned agency information needs.

The examination of existing workflow patterns and records is the crucial first step in determining the need for a digital imaging system. A records analysis assesses existing operations to determine what records are best suited for digital imaging applications. A workflow analysis assesses the processes of records creation, access, and retrieval to determine areas where re-engineering can improve operational efficiency. This reorganization of business or work processes may be simple or extensive in approach. Implementing a digital imaging system significantly impacts the current work processes because personnel create, retrieve, use, and store documents in a different way. The detail and complexity of the process of re-engineering affects the project schedule, cost justification, and the bid specifications.

The University Archives and Records Management Service can assist in analyzing a department's record keeping systems.

Recommendation 2: Prior to selecting a digital imaging system, conduct a cost benefit analysis to determine the cost justification of a system purchase and to determine the possible benefits to the agency with its implementation.

Cost justifying a digital imaging system allows a financial comparison between the current and proposed record keeping systems to help in making a procurement decision. The cost-justification goal of a digital imaging system is to offset the cost of the equipment and software by reducing personnel and storage costs or allowing the existing staff to process more work through the improvement of work processes. To determine a cost estimate, the following components should be considered: system hardware, system software, application software, communications hardware and software, system maintenance, training, project management, facilities upgrades/site preparation, staffing costs, and other miscellaneous costs. A typical cost justification includes the following major areas: a study of current operations, a proposed system architecture, equipment pricing, and financial indicators, including payback period, net present value, and rate of return.

B. System Specifications and Selection

Recommendation 3: Require "open systems architecture" for digital imaging applications or require vendors to provide a bridge to systems with non-proprietary configurations.

Although the term "open systems architecture" is defined in various ways, public officials 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 an 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.

Recommendation 4: Where data longevity or records integrity is a primary concern, use a recording media that is not rewritable.

The storage capacity of optical disks versus paper is a primary advantage to the use of digital imaging systems. However, optical disks are not the only option. Other storage solutions that can be used with digital imaging systems include output to microfiche or microfilm, digital tape, and magnetic disks. The selection of a storage media may depend on budget considerations for the unit.

When selecting optical storage media, the issues of data longevity and integrity must be considered. There are a variety of optical disks on the market today: Write Once Read Many (WORM), Rewritable, and Compact Disk-Read Only Memory (CD-ROM)/Compact Disk-Recordable (CD-R). Each has its own advantages and disadvantages. WORM and CD-R are not rewritable. Computer Output to Laser Disk (COLD) is also a non-rewritable option. These media offer a high level of data security because alteration of data is not achievable without destruction of the media itself. If a record is no longer needed, software may allow the pointer to the data to be disabled, preventing normal access. Because the data cannot truly be deleted, however, it may remain accessible by other means.

Recommendation 5: Use a non-proprietary digital image file format. If using a proprietary format, provide a bridge to a non-proprietary digital image file format.

A digital image file format is a structured container for information about each digital image and the image data. Information about the digital image file includes, but is not limited to, its name, width, length, resolution, and compression techniques. The computer requires this information to interpret the digital image. It is essential to use a non-proprietary image file format to ensure the ability to transfer successfully digital images between different systems or when a system is upgraded or modified.

American National Standards Institute (ANSI)/Association for Information and Image Management (AIIM) MS53-1993, Standard Recommended Practice - File Format for Storage and Exchange of Images - Bi-Level Image File Format: Part I details a standard definition for file formats. Despite the existence of a standard, there is not an agreed-upon, industry-wide image format standard. Many digital imaging systems use the Tagged Image File Format, or TIFF. Because different versions of TIFF exist (TIFF-5, TIFF-4, etc.), there is still no absolute guarantee that images can be transported seamlessly from one system to another. Comprehensive documentation of the digital image file format, including TIFF, is recommended.

A number of other file formats exist, such as Graphics Interchange Format (GIF), Joint Photographic Experts Group (JPEG), and Bitmap (BMP). These file formats are commonly used in conjunction with hypertext markup language (HTML) for Internet and intranet applications. Many systems or third-party graphics packages will convert images from one to another, although often with unpredictable results.

Recommendation 6: Use International Telecommunications Union (ITU) Group 3 and Group 4 compression techniques or have the vendor provide a bridge to these techniques.

The large file sizes of typical scanned documents require digital image compression to support data transmission and to promote storage efficiency. Today most digital imaging systems use standard compression algorithms to "shrink" images. Standard compression techniques are instrumental in ensuring a migration strategy for records needed for long-term use. Two international standards are currently available. Using compression techniques conforming to either of these specifications will increase the likelihood that the images can be used with other technologies or migrated between systems.

Recommendation 7: When determining document-scanning resolution, consider data storage requirements, document scanning throughput rates, and the accurate reproduction of the image. Validate vendor claims using a sampling of the unit's documents.

A digitized image consists of black and white dots or picture elements (pixels) measured in dots per inch (dpi). The higher the number of dpi, the higher the legibility of the reproduced image will be. Images scanned at higher dpi rates, however, use more storage space on the disk and may require longer scanning times. The selection of scanning density involves a trade-off between image clarity, storage capacity, and speed. When selecting a scanner, ask the vendor to perform a quality test on a broad sampling of documents at various dpi settings so that an appropriate end-to-end throughput rate and resolution can be determined.

For good quality images in imaging modern office records, use a scanning density of at least 300 dpi. A higher scanning density (600 dpi or higher) 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.

Recommendation 8: Select equipment that conforms to the standard methodology for media error detection and correction. The system should provide techniques for monitoring and reporting verification of the records stored on a digital optical disk, and the system administrator should actively follow the status of the monitors.

Digital imaging technology uses two methods within the Error Detection and Correction (EDAC) system to minimize digital image recording and retrieval errors. The first method uses error correction codes to detect and correct data read errors automatically. The second employs correction code software to determine if and when the utilization of error correction codes is approaching a critical point. Monitoring the error correction status information provides an audit trail to measure the progress and degree of disk degradation. Tracking error correction trends will indicate an appropriate timetable for recopying disks.

The Association for Information and Image Management's (AIIM) Standards Committee has developed a standardized methodology for reporting the error rate data to the operating system for user evaluations. ANSI/AIIM MS 59-1996, Media Error Monitoring and Reporting Techniques for Verification of Stored Data on Optical Digital Data Disks, describes these standards.

Another precaution against losing long-term records because of defective disks is to require the use of digital optical disks with a guaranteed minimum shelf life of five years and a minimum post-write life of twenty years.

Recommendation 9: Specify that the Small Computer System Interface (SCSI) command "Write and Verify" be used when writing data to digital optical disks.

The "Write and Verify" command, available within the Small Computer System Interface (SCSI), is valuable for assessing how accurately the scanned information is transferred from the central processing unit of the computer to the digital optical disk. "Write and Verify" requires verification from the system that the digital image is correctly written to the disk and provides additional protection for continued access to long-term records.

Recommendation 10: 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.

Recommendation 11: Provide specific plans for an ongoing process of migrating long-term and archival records from older to newer hardware and software platforms.

Agencies must ensure that their long-term and archival records are continually accessible,

(Wisconsin Administrative Rule 12; see FAQ at http://www.library.wisc.edu/libraries/Archives/rm/faq/faq_adm12.htm). Systems as physical devices could be operational for ten years or more, but system technology will often be superseded within two to three years. If the system stores records with retention periods exceeding the life span of the hardware and software, it is essential that the administrator plan for future data migration. A migration strategy documents how an organization 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.

Current strategies for migrating digital imaging system records include: upgrading equipment and software as technology evolves and periodically recopying disks as required; recopying optical disks based upon projected longevity and/or periodic verification of the records; or, transferring the data from an obsolete generation of optical disks to a newly-emerging technology, in some cases bypassing the intermediate generation that is mature but at risk of becoming obsolete.

C. System Implementation

Recommendation 12: Assign a permanent staff member as systems administrator and require the vendor to provide a project director during the installation and training periods.

The assignment of a qualified staff member, preferably with systems administration experience, is critical to the effective implementation and maintenance of a digital imaging system. The systems administrator should be responsible for overall project management, and the development and maintenance of written system documentation that describes the requirements, capabilities, limitations, design, operation, and maintenance of the digital imaging system. Making a vendor representative responsible for installing the equipment and training the systems administrator and other appropriate unit staff will help to ensure successful implementation of the system.

Recommendation 13: Establish operational practices and provide technical and administrative documentation to ensure the future usability of the system, continued access to long-term records, and a sound foundation for assuring the system's legal integrity.

It is the responsibility of office administrators, rather than vendors and manufacturers, to maintain written documentation of system procedures, also called Standard Operating Procedures or SOPs, including access and security policies and procedures. Security and access policies should be developed to protect the system and the records from alteration or unauthorized use.

In regard to legal admissibility and trustworthiness, records stored on a digital imaging system should be treated no differently than records stored on magnetic disk or tape. The key is for the systems administrator to become familiar with how the rules of evidence apply to such records. Procedural controls should be established and followed to protect the integrity of the records.

These procedural controls should be documented and should reflect requirements for the legal acceptance of records as outlined in AIIM TR31-1992, Performance Guideline for the Admissibility of Records Produced by Information Technology Systems as Evidence. This AIIM performance guideline stresses the importance of specifying the processes used to create the records, demonstrating that records are produced and relied upon in the regular course of business, establishing quality control and audit procedures, conducting formal training programs, and providing written documentation for each procedure. Case histories indicate that system requirements for good archival maintenance are consistent with the requirements for the admission of records under the "rules of evidence" laws. Records administrators should be familiar with how the rules of evidence apply to Wisconsin's public records laws. Policies and procedures should be followed to protect the integrity of long-term records.

Recommendation 14: Perform a visual quality control evaluation of each scanned image and related index data. Write the scanned image to optical media only after the evaluation process is completed.

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 before writing the digital image to optical media. 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.

When the system is operational, a routine scanning quality test, as outlined in ANSI-AIIM MS44-1988 (R1993), Recommended Practice for Quality Control of Image Scanners, should be performed on a weekly or monthly basis.

Recommendation 15: 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 government office 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.

Recommendation 16: Provide adequate environmental conditions for the digital optical disks.

Even in an optimum environment, digital optical disks are susceptible to deterioration. Adverse storage conditions, especially high humidity, can cause rapid deterioration of the media. A prudent storage guideline for digital optical disks is to adhere to the temperature and humidity levels recommended for magnetic media storage. Technical specialists recommend a stable environment, with a temperature between 65 and 75 degrees, and a relative humidity between 30 and 50 per cent. Digital optical disks should never be stored in direct sunlight nor placed near sources of heat.

Dust, debris, and fingerprints affect digital optical disks. Plastic cartridges should never be removed; nor should the cartridge shutter be opened to expose the digital optical disk's recording surface. To protect disks from warping, they should not be subject to pressure and should be stored in an upright position when not in the disk drive.

Agency officials should request that the vendor supply specifications for the storage of digital optical disks and ensure that office conditions meet these specifications during installation of a system.

Recommendation 17: Budget annually between fifteen and twenty per cent of the original system acquisition cost for upgrades, training, and maintenance.

Administrative managers should be aware of the high cost of maintaining and upgrading digital imaging systems. Unless these costs are factored into the continuing support of system maintenance and improvement, the system is in danger of becoming obsolete and requiring a far greater cost outlay to restore its effectiveness. Also, records stored in an outdated system tend to be at greater risk than those in a well-supported system. Continued planning and budgeting for the migration of long-term and archival records, as discussed in Recommendation 11, is essential for the success of any digital imaging project.

APPENDIX A– A PRE-PLANNING CHECKLIST

The following checklist of questions can be used to guide your pre-planning efforts:

- Who are your current and future users?
- Will the digital images act only as "finding aids" (i.e., after locating a desired document through a search of the image database, will your users be referred to the original document for study/research)? Or, will the digital images serve as surrogates for the deteriorating originals, which will be stored or discarded?
- Have the originals already been photographed? If so, can the slides/transparencies serve as source documents for digitization? Were color/gray-scale targets used during photography? How high is the photographic quality? What is the condition of the film?
- What are your quality expectations? How do you define a "good image"? How much detail do you need to capture (e.g., do you need to capture the fine details such as the lace work around a woman's sleeve in a portrait, or fine lines in a pencil sketch?)? Do digital images need to capture the colors faithfully, or can you (and your users) tolerate a minor color shift?
- Will you provide Web access to the collection?
- What are your users' hardware/software capabilities and network configurations? Will their equipment and network support the delivery of images in a timely fashion while retaining the image quality?

The above listing is taken from:
http://www.library.cornell.edu/preservation/kodak/kodak-htm.htm#technology
Using Kodak Photo CD Technology for Preservation and Access: A Guide for Librarians, Archivists, and Curators by Anne R. Kenney and Oya Y. Rieger

APPENDIX B -- LIST OF SOURCES

CD Technology Main
http://www.disctronics.co.uk/cdref/cdmain.htm

Digital Projects Guidelines (Arizona State Library, Archives and Public Records)
http://www.lib.az.us/digital/dg_a13.html

The following ANSI/AIIM standards can be purchased through http://www.ansi.org.

ANSI/AIIM MS44 (1988, R1993): Recommended Practice for Quality Control of Image Scanners

ANSI/AIIM MS54 (1993): Standard Recommended Practice – File Format for Storage and Exchange of Images – Bi-Level Image File Format: Part 1

ANSI/AIIM MS56 (1996): Media Error Monitoring and Reporting Techniques for Verification of Stored Data on Optical Digital Data Disks

ANSI/AIIM TR31 (1992): Performance Guideline for the Admissibility of Records Produced by Information Technology Systems as Evidence

Archives and Records Management Services
Main Office
B134 Memorial Library
728 State Street
Madison, WI 53706-1494
Phone: (608) 262-5629
Fax: (608) 265-2754
Recmgmt@library.wisc.edu
www.library.wisc.edu/libraries/Archives/

ARMS is a periodical bulletin from the University of Wisconsin-Madison's Archives and Records Management Services.