Mississippi Department of Archives and History

Recommended Practices for Digitizing Public Records

These practices are intended to assist custodians of public records, including state agencies, counties, municipalities, school districts, and other entities of the state of Mississippi, planning to digitize public records with the intent of disposing of the original records. The comprehensive management of a digital imaging system is crucial to its functional success as well as the quality, integrity, and authenticity of the imaged records. System planning, design, budgeting, procurement, procedure formulation, training, and testing all require thoughtful deliberation and patience.

The following is an overview of recommended practices in several categories of system management and operation.

1. Documentation

Comprehensive procedural and system documentation should be maintained to ensure that the operation continues to function effectively over time. The documentation should include:

a. Hardware and software specifications, brand names, versions, and dates of installation, upgrade, replacement, and conversion.

b. An overview of system purposes and uses.

c. Policies and procedures for all aspects of system operation and maintenance, including procurement, file and document preparation for scanning, data entry, quality control, indexing, corrections, expungement, redaction, back-ups, security, migration, application of safeguards to prevent tampering and unauthorized access, and printing.

d. Data structure and content, including file layout and data dictionaries.

e. Enhancement algorithms, or techniques for processing the image so that the result is visually clearer than the original image. Imaging systems should not be capable of altering a record as scanned, except for standard computer-enhancement routines used to improve legibility.

f. Documentation for providing audit trails, for establishing legal admissibility of images, and for use by future system operators as staffs change.

It is the responsibility of the system's administrators, not the vendor, to establish and maintain the above documentation.

2. Quality Control

a. To ensure the integrity and legibility of scanned images, there should be in place established procedures for quality control.   Visual quality inspection of each image is necessary and should be performed initially by the staff member scanning and then by a second staff member.

b. The accuracy of the index must also be verified through visual inspection by a second staff member of each index entry following either entry of terms or creation through optical or intelligent character recognition.

c. The system should also include the ability to rescan and to correct indexing errors before the image and/or index is written to optical media.

d. Quality control issues should be raised with vendors during the selection process and be considered when planning for time and staff budgeting. Since original records may be destroyed once reformatted, the importance of image and index quality control must not be underestimated.

No records should be destroyed until all quality checks have been completed.

3. Indexing

a. Complete, appropriate and accurate indexing capability is essential. Indexing and information retrieval needs must be assessed during system planning and design. Migration and long-term usability planning must also include consideration of continuing information retrieval requirements.

b. The importance of indexing requires that vendor claims be validated through demonstration and testing.

4. Migration

a. A comprehensive plan for refreshing data and for migrating images, indexes and related data through successive versions of hardware and software is essential for ensuring long-term access to imaged records. Not only should plans be established for the migration of images and related data, but structural data relationships should be preserved under migration. The strategy should facilitate the movement of records from one generation of technology to another and should take into consideration vendor stability and dependability, system obsolescence, and media longevity.

b. The reality of obsolescence requires that agencies and entities keep pace with constant developments and improvements. Technology trends must be monitored.

The technology choices made when systems are developed or upgraded may determine the ease of migration.

c. Systems should consist of hardware and software that conform to non-proprietary standards and should be constructed in an open-system architecture.

d. Budgeting and planning should include consideration of the costs of technology upgrades and data migration.

5. Back-up, Disaster Recovery, and Security Copies

a. Back-up procedures and disaster recovery plans should be in place with specified provisions for the imaging system. Detailed information on back-ups and disaster recovery should be obtained from vendors. Back-up expense and complexity can vary depending on the type of media and the amount of data to be stored and must be considered during the planning and selection process.

b. A regular schedule of back-ups should be instituted for all data on the system, including indexes.

c. Security copies should be labeled with information to include date, system, and software used, and any existing restrictions on access, keeping in mind that it is impossible to determine content merely by looking at a disk or tape. It is preferable that security copies be stored off-site, in an area with stable environmental conditions and with adherence to the manufacturer's specifications for the storage of the media, whether magnetic or optical.

6. Expungement/Redaction/Encryption Capabilities

a. Agencies and entities should have in place a strategy to guarantee that material exempted from disclosure is not made available to the public. Imaging systems should have the capability to expunge images and index entries and to redact confidential portions of images or indexes when required by law. System administrators may also wish to further ensure privacy of their data through the use of an encryption technique by which data is scrambled before transmission and then unscrambled (decrypted) by the receiver.

b. The potential need for expungement, redaction and encryption capabilities must be assessed and discussed with vendors when planning for long-term usability of an imaging system.

7. Legality

a. The legal admissibility of reproductions of state and county records is addressed by *Mississippi Code Annotated (MCA)* §§ 25-59-29 and 19-15-3, respectively.

b. Requirements for the legal acceptance of records are outlined in ANSI/AIIM's TR-31-2004, Legal Acceptance of Records Produced by Information Technology Systems, and the Mississippi Rules of Evidence.

8. System Selection

a. Agencies and entities should conduct a thorough survey of document and paper types, sizes, colors, and contrasts within their records and collect examples of potential problems or obstacles, such as browned and fragile papers, pencil and pen handwriting, bound volumes, photographs, and oversized items. With oversized items such as maps, plats, and architectural drawings, it is often more practical to outsource digitization to a vendor. Before selection, a scanner should demonstrate the ability to handle the job. The potential need for flatbed scanning capability rather than automatic feed alone must also be assessed, both for immediate and future needs.

b. The selection of a vendor is perhaps the most important single decision impacting an imaging system's success. A vendor's stability, accessibility, and long-term viability must be assessed when procuring a system heavily dependent on vendor support.