Mississippi Department of Archives and History

Reformatting Standards

STANDARDS FOR ARCHIVAL IMAGING

The Mississippi Department of Archives and History (MDAH) approves the following standards for creating preservation copies of original paper records that have been scheduled for permanent retention and disposal after imaging. Prior to disposal of any records, they must have a retention schedule approved by the State Records Committee or Local Government Records Committee as directed in *Mississippi Code Annotated (MCA)* §§ 25-59-1 et. seq. and *MCA* §§ 25-60-1. Born-digital records (like MS Word documents and Excel spreadsheets) scheduled for permanent retention should be preserved in their original formats.

1. OPEN SYSTEM ARCHITECTURE

   The design of the system shall permit future system upgrades with minimal effect on system operation. The system architecture shall allow flexibility in exporting and importing data to other non-proprietary systems.

2. STORAGE

   Prior to disposal of original paper records, images of long-term or permanent records must be stored either on non-rewritable storage media or in a system with a backup/redundancy routine that includes multiple copies of the images with at least one copy stored off-site. If the records being imaged are scheduled for permanent retention at MDAH, the originals can be disposed of once the images have been received and verified by MDAH.

   Non-rewritable media include write once read many (WORM), compact disc read-only memory (CD-ROM), digital versatile disc ± recordable (DVD±R), and microfilm produced in accordance with standards approved by MDAH. Digital optical discs are required to have a guaranteed minimum shelf life of five years and a minimum post-write life of twenty years, and a plan to refresh and migrate as necessary should be in place. Information regarding optical media storage may be found in ISO 18925:2013, "ImagingMedia - Optical Disc Media - Storage." ANSI/AIIM MS59-1996, Standard Recommended Practice for Media Error Monitoring and Reporting Techniques for Verification of the Information Stored in Optical Digital Data Discs, is the standard for reporting the error rate data to the operating system for user evaluation.

3. NON-PROPRIETARY IMAGE FILE FORMAT

   TIFF is the approved file format for alphanumeric and graphic paper documents that are scheduled for archival retention but will be discarded after imaging. TIFF, JPEG 2000, PDF/A-1, and PNG are recommended formats for documents that will not.

4. COMPRESSION ALGORITHMS

   Uncompressed files are preferred for archival preservation, but the ITU-T (formerly CCITT) Group 3 and 4 compression algorithms shall be used if storage limitations require compression of files.

5. COLOR SPACE AND BIT DEPTH

Color (24-bit RGB) scans are recommended for most documents, especially those with color, graphical elements, or poor contrast between text and its background. Grayscale (8-bit) is acceptable for documents in which color content does not exist or is not considered important. Bitonal (1-bit black-and-white) images should be reserved for typed documents with good contrast when preservation of the text is the only requirement.

6. SCANNING RESOLUTION

300 pixels per inch (ppi) or greater is acceptable for original documents. 300 ppi is recommended for oversized documents. 600 ppi is recommended for average-sized documents, especially for bitonal scans. For digitization of micrographics, 400 ppi relative to the original dimensions of the documents should provide good results. If Optical Character Recognition (OCR) is used, a minimum of 400 ppi is suggested for average-sized documents, and the OCR results must be compared with the original text to determine if the resolution used was high enough to create an accurate copy.

7. NAMING CONVENTIONS

File and folder names should include only alphanumeric characters. Special characters that are used by operating systems ought to be avoided (i.e. / > < + = ' " ^ | \ { } [ ] # , ; ? ! $ * &). A valid three- or four-character file extension for the file type (e.g. .tif, .docx) should appear at the end of each file name. File names should contain enough descriptive information so that they can be identified outside their storage location but be short enough that they migrate well.

8. INDEXING SYSTEM

The indexing system used to retrieve images shall be a relational database. Alternative indexing systems must be approved in writing by MDAH. Information on establishing index fields in an electronic imaging system may be found in ANSI/AIIM TR40-1995, "Suggested Index Fields for Documents in Electronic Image Management (EIM) Environments."

9. QUALITY CONTROL

Information regarding the establishment and use of procedures for the ongoing control of quality within an electronic imaging system may be found in the following standards:

- ISO 14524:2009, "Electronic Still-Picture Cameras - Methods for Measuring Opto-Electronic Conversion Functions (OECFs)"
- ISO 15739:2017, "Electronic Still-Picture Imaging - Noise Measurements"
- ISO 16067-1:2003, "Spatial Resolution Measurements of Electronic Scanners for Photographic Images - Part 1: Scanners for Reflective Media"
- ISO 16067-2:2004, "Electronic Scanners for Photographic Images - Spatial Resolution Requirements - Part 2: Film Scanners"
- ISO 21550:2004, "Electronic Scanners for Photographic Images - Dynamic Range Measurements"
- ISO 22028-1:2016, "Extended Colour Encodings for Digital Image Storage, Manipulation and Interchange - Part 1: Architecture Requirements."

10. TEST CHARTS AND PATTERNS

Information regarding appropriate use of test charts and patterns in document imaging applications may be found in ANSI/AIIM TR38-1996, "Compilation of Test Targets for Document Imaging Systems"; ISO 16067-1:2003, "Spatial Resolution Measurements of Electronic Scanners for Photographic Images - Part 1: Scanners for Reflective Media"; and ISO 16067-2:2004, "Electronic Scanners for Photographic Images - Spatial Resolution Requirements - Part 2: Film Scanners."

11. EXPUNGEMENT/REDACTION/ENCRYPTION CAPABILITIES

Explanation of procedures for expunging information on WORM optical systems may be found in ANSI/AIIM TR28-1991, "The Expungement of Information Recorded on Optical Write-Once-Read-Many (WORM) Systems."

12. LEGALITY

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

b. Requirements for the legal acceptance of records are outlined in ANSI/AIIM's TR31-2004, "Legal Acceptance of Records Produced by Information Technology Systems," and the Mississippi Rules of Evidence.

STANDARDS FOR MICROFORMS

MDAH approves the following standards for reproduction of public records on microforms:

- ANSI/AIIM MS18-1992 (R1998), "Splices for Imaged Film - Dimensions and OperationalConstraints"
- ANSI/AIIM MS19-1993, "Recommended Practice for Identification of Microforms"
- ANSI/AIIM MS23-2004, "Recommended Practice - Production, Inspection, and Quality Assurance of First-Generation, Silver Microforms of Documents"
- ANSI/AIIM MS26-1990, "35mm Planetary Cameras (Top-Light) - Procedures for Determining Illumination Uniformity of Microfilming Engineering Drawings"
- ANSI/AIIM MS42-1989, "Recommended Practice for the Expungement, Deletion,Correction, or Amendment of Records on Microforms"
- ANSI/AIIM MS43-1998, "Recommended Practice for Operational Procedures/Inspection and Quality Control of Duplicate Microforms of Documents and from COM"
- ANSI/AIIM MS45-1990, "Recommended Practice for the Inspection of Stored Silver-Gelatin Microforms for Evidence of Deterioration"
- ANSI/AIIM MS111-1994, "Recommended Practice for Microfilming Printed Newspapers on 35mm Roll Microfilm"
- ISO 18901:2010, "Processed Silver-Gelatin-Type Black-and-White Films - Specifications for Stability"

DEFINITIONS

1. AIIM - the Association for Information and Image Management, a standards-setting body affiliated with the American National Standards Institute (ANSI), which is the principal developer of standards for microforms and information storage technologies involving images, such as optical discs and scanners.

2. ANSI - the American National Standards Institute, a private national standards organization in the United States, which coordinates the development and maintenance of various industry standards.

3. Bit Depth - the number of bits used to represent each pixel in an image. For example, a typical color image using 8 bits per channel is often referred to as a 24-bit color image (8 bits each for the three red, green, and blue light channels).

4. Bitonal - an image consisting only of black and white pixels.

5. CCITT - International Telegraph and Telephone Consultative Committee - predecessor of the International Telecommunication Union (ITU).

6. CD-ROM (Compact Disc - Read Only Memory) - a data storage system using compact discs as the medium.

7. Color Space - an organization of colors that supports reproducible representations of color across viewing media.

8. Grayscale - a range of gray shades from white to black. A grayscale image is made up of pixels with values of light intensity (light to dark).

9. IEC - International Electrotechnical Commission.

10. ISO - The International Organization for Standardization, which coordinates nations' standards.

11. ITU-T - Telecommunication Standardization Sector of the InternationalTelecommunication Union.

12. JPEG 2000 - an image-encoding system from the Joint Photographic Experts Group that uses wavelet compression for color and grayscale images.

13. Open System - a system with characteristics that comply with specified, publicly maintained, readily available standards and that therefore can be connected to other systems that comply with those same standards.

14. Optical Character Recognition (OCR) - the conversion of images of text into machine-encoded text that can be indexed and searched using computer software.

15. Optical Disc - an electronic data storage medium that can be written to and read using a low-powered laser beam.

16. Original Record - a public record in the format as created or received, or in the format as reproduced in accordance with standards approved by the Mississippi Department of Archives and History.

17. PDF/A-1 (Portable Document Format / Archival) - open standard version of the file format used to capture, distribute, and store formatted, page-oriented documents containing fonts, graphics, and images uncompressed.

18. PNG (Portable Network Graphics) - a patent- and license-free file format that uses lossless compression.

19. PPI (Pixels Per Inch) - the number of pixels displayed in an image. Pixel is short for picture element - a single point in a graphic image.

20. Public Records - all documents, papers, letters, maps, books, tapes, photographs, films, sound recordings or other materials regardless of physical form or characteristics made or received pursuant to law or ordinance or in connection with the transaction of official business by any agency or by any appointed or elected official. Books, periodicals and other published material normally found in a library are excluded from this definition (as defined in *Mississippi Code Annotated* §25-59-3).

21. TIFF (Tagged Image File Format) - a family of bitmap file formats for describing and storing color and grayscale images.

22. WORM (Write Once, Read Many) - storage media (usually recordable CD-ROM or optical disc) that is not rewritable. Information can only be written to the disc once. It is permanently stored on the disc.