

Converting Paper Records to Digital

Several program offices in APHIS have expressed interest in implementing scanning operations to convert paper records into digital images. Benefits include reduced storage costs and space, improved accessibility, and reduced risk of damage to records. When combined with effective indexing, imaging the files can shorten information retrieval time and allow access to materials for multiple users at various locations.

Image files come in many different types of software-dependant formats, such as .gif, .jpg, and .tif. Most formats are proprietary and may not be supported long term by manufacturers and may vary from vendor to vendor. Many file formats use compression to force more data into less storage space and speed image storage and transmission. Images converted to some file formats, such as JPEG (.jpg), don't necessarily look the same after compression due to data loss.

For this reason it is recommended that images be saved in a standardized format, such as Adobe Portable Document Format (PDF), whenever it's feasible. In addition to no data loss, the format is approved by the National Archives and Records Administration (NARA) for permanent records. For more information about formats approved by NARA see Appendix A.

Search and retrieval normally depends on some form of indexing, which assigns specific metadata to each document, such as author, recipient, date, title, and content keywords. This index, or metadata, can be simple or sophisticated, and is typically an electronic database that is linked to the images. Useful indexing requires careful planning and forethought before any actual imaging begins.

Recordkeeping requirements for imaged records

A plan must be developed to protect the information in the records for the entire retention period. When imaging records, document the procedures and provide audit trails to serve as the record that the images were created properly and validated. For records with long retention periods, a migration plan may be needed to ensure that the information in the images can be accessed throughout the entire retention period of the records.

Scheduling Records

Temporary Records

Imaged copies of records already scheduled as temporary do not need to be scheduled if the nature and content of the records remain identical to the description in the schedule.

- Apply the disposition authority approved by NARA for the paper records to the image files.
- Retain the paper copies of temporary records that have been imaged only when there is a compelling business reason.

Retention schedule items for temporary textual records must be added or revised whenever the nature or content of records changes.

- Reengineered work processes or other changes may result in creation of a different series of records.
- When the nature or content of a records series changes, both the paper and imaged copies of the records must be scheduled, even if all of the records are temporary.

Permanent Records

When paper records that are scheduled as permanent are imaged, the imaged files must also be scheduled. The agency may not dispose of the paper copies until NARA has approved a new schedule for them. The schedule should provide the disposition of both the paper and imaged copies and specify which version is the recordkeeping copy. See appendix A for more information.

Do indexes created for imaged records need to be scheduled?

- Indexes and other finding aids for temporary records are disposable under [General Records Schedule 23, item 9](#), unless the finding aid contains an abstract or other information that can be used independently of the related records.
- Indexes and finding aids for permanent records that have been imaged must be scheduled.

Appendix A

Expanding Acceptable Transfer Requirements: Transfer Instructions for Existing Permanent Electronic Records Scanned Images of Textual Records

1.0 PREFACE

As part of the National Archives and Records Administration's (NARA) electronic government (e-gov) initiative, and in cooperation with other Federal agencies, NARA is issuing guidance to supplement current requirements in 36 CFR Part 1228.270 for transferring permanent electronic records to NARA. This guidance expands currently acceptable formats to enable the transfer of existing scanned images of textual records ¹ to NARA. The guidance does not apply to "born digital" images or scanned images of primarily non-textual records (e.g., maps and photographs).

2.0 BACKGROUND

Scanned images is a priority electronic records format identified by NARA and partner agencies as part of the Electronic Records Management (ERM) initiative, one of the twenty-four e-gov initiatives under the President's Management Agenda. A major goal of this initiative is to provide the tools for agencies to access electronic records for as long as required and to transfer permanent electronic records to NARA for preservation and future use by government and citizens. Scanned images is the second of three formats examined under this ERM e-gov initiative; the other priority formats include email with attachments and electronic records in Adobe's Portable Document Format (PDF).

The following sections provide transfer requirements specifically for existing scanned images that have been appraised or scheduled for permanent retention in electronic (i.e., scanned) form. These requirements address acceptable scanned image file formats, image quality specifications, transfer documentation, and related information on how to transfer these records to NARA. If an agency has existing permanent scanned images that do not meet the image quality specifications in this guidance, contact the NARA appraisal archivist (see section 6.0).

Additional information about the ERM e-gov initiative can be found on NARA's web site at: <https://www.archives.gov/news/articles/archives-guides-agencies-on-managing-electronic-records>

3.0 TRANSFER REQUIREMENTS FOR SCANNED IMAGES OF TEXTUAL RECORDS

Effective upon promulgation of these requirements, NARA will accept transfers of scanned images that have been scheduled as permanent records on a SF 115, Request for Records Disposition Authority. To facilitate preservation processing and future access to these records, agencies must comply with the following minimum requirements.

3.1 Image File Format

3.1.1 NARA will accept scanned images in the following file formats:

3.1.1.1 **Tagged Image File Format (TIFF)**, in 'II' format, ² versions 4.0 (April 1987), 5.0 (October 1988), and 6.0 (June 1992).

3.1.1.2 **Graphics Interchange Format (GIF)**, versions 87a (June 1987), and 89a (July 1990).

3.1.1.3 **Basic Image Interchange Format (BIIF)**, International Standards Organization (ISO) Standard 12087-5, Part 5 (December 1998).

3.1.1.4 **Portable Network Graphics (PNG)**, version 1.0 (October 1996).

3.1.2 NARA will periodically update this set of acceptable formats and versions.

3.1.3 NARA prefers uncompressed images; with prior consultation between NARA and the transferring agency, NARA may accept compressed images created using a lossless compression method (e.g., ITU Group 4, LZW) when it is impractical or technically impossible to transfer uncompressed files (e.g., due to limitations on bandwidth or storage media capacities). NARA will not accept images created using lossy compression techniques because the potential information loss that can result from compression can make such images unsuitable for archival preservation.

3.2 Acceptable Image Quality Specifications

Scanned images of textual records transferred to NARA must meet the following minimum requirements for scanning resolution and pixel (bit) depth to support archival preservation and continued use.

3.2.1 *Bitonal (1-bit) scanned at 300-600 ppi.*

This is appropriate for documents that consist exclusively of clean printed type possessing high inherent contrast (e.g., laser printed or typeset on a white background). Scanning at 600 ppi is recommended.

3.2.2 *Gray scale (8-bit) scanned at 300-400 ppi.*

This is appropriate for textual documents of poor legibility because of low inherent contrast, staining or fading (e.g., carbon copies, thermofax, or documents with handwritten annotations or other markings), or that contain halftone illustrations or photographs. Scanning at 400 ppi is recommended.

3.2.3 *Color (24-bit RGB [Red, Green, Blue]) scanned at 300-400 ppi.*

Color mode (if technically available) is appropriate for text containing color information important to interpretation or content. Scanning at 400 ppi is recommended.

3.3 Transfer Documentation

This guidance supplements transfer documentation requirements in 36 CFR 1228.270(e) to ensure that transfers of scanned images of textual records are clearly identified and described. Agencies must also submit a signed *Standard Form 258, Agreement to Transfer Records to the National Archives of the United States (SF 258)*, as required by 36 CFR 1228.272.

3.3.1 For each transfer, agencies must supply documentation that identifies:

- 3.3.1.1 Imaging system and version(s) (i.e., application software and storage system)
- 3.3.1.2 Operating system and version(s)
- 3.3.1.3 Records Management Application (if applicable) and version(s)
- 3.3.1.4 Image file format(s) and version(s)
- 3.3.1.5 Image quality specifications (i.e., resolution, pixel (bit) depth, compression technique)
- 3.3.1.6 Total number of documents in the transfer
- 3.3.1.7 Total number of images in the transfer
- 3.3.1.8 Arrangement of files on the transfer media
- 3.3.1.9 Structure of image header (if applicable)

3.3.2 In addition, agencies must supply the following:

- 3.3.2.1 Control information necessary to enable retrieval of each file, including whether the file is a single image or multiple images. Such information can be conveyed through metadata (e.g., a thumbnail index), captured for each document either in the image header or in accompanying documentation.
- 3.3.2.2 Finding aids, indexes, and other information used to retrieve the records also must be transferred to NARA to support future access to the records.
- 3.3.2.3 The Optical Character Recognition (OCR) versions of the images (i.e., additional files enabling full text searches linked to the appropriate scanned image), if available.

3.3.3 Agencies also must supply a description of the quality control inspection performed as part of the imaging process and a report on the results of the last inspection performed on the images and the date of that inspection.

3.3.4 For inspections conducted after the effective date of this guidance, agencies must use one of the quality sampling methods presented in ANSI/AIIM, *Sampling Procedures for Inspection by Attributes of Images in Electronic Image Management and Micrographic Systems (ANSI/AIIM TR34-1996)*.

4.0 TRANSFER MECHANISMS

4.1 Agencies must transfer scanned images as uncompressed file(s).

4.2 Agencies must transfer scanned images as unaggregated file(s).

4.3 In addition to the approved transfer media listed in 36 CFR 1228.270(c), agencies may transfer records on Digital Linear Tape (DLT) Type IV or through File Transfer Protocol (FTP). FTP transfers currently are limited to a maximum size of five Gigabytes (5 GB). For FTP transfers, the offering agency must provide the following information to NARA:

- IP address for the Agency FTP site,
- A temporary user name and temporary password for NARA to use in accessing the Agency FTP site,
- A designated time or time span for NARA to access the records and download a copy to NARA.

5.0 LEVELS OF ACCESS

NARA will provide access to the creating agency and to all researchers requesting scanned images accessioned from Federal agencies, subject to review of content for FOIA exemptions as is feasible. While compliance with these requirements will improve future access to accessioned scanned images, NARA's ability to provide access to certain records will vary according to their hardware and software dependencies. At the present time, NARA provides users with a copy of fully releasable electronic record files on any of the media currently approved by NARA. For scanned images transferred to NARA the user will be responsible for obtaining the necessary hardware and software to view the records.

6.0 CONTACT INFORMATION

For assistance in scheduling scanned images, or to discuss how to handle permanent scanned images that do not meet the specifications in section 3.0, contact the APHIS records officer.

For technical assistance in transferring scanned image records to NARA, contact the Electronic and Special Media Records Services Division (NWME), 8601 Adelphi Road, College Park, MD 20740. The general telephone number is 301-837-3420.