

National Archives and Records Administration

NARA 1550, Supplement 3

November 18, 2019

SUBJECT: Digitization Process for Internal Digitization Projects

NARA digitization processes differ based on a variety of factors and NARA staff should follow all Office and unit-specific digitization standard operating procedures (SOPs) when digitizing. This Supplement sets forth major considerations and agency-wide guidance to be followed for all in-house digitization including systematic digitization. If certain steps do not apply to a custodial unit, they may be skipped. Generally, though, this Supplement applies to all custodial units. For targeted digitization in support of exhibits, education or social media efforts consult Supplement 2, Digitization Process for Exhibits, Education, and Social Media.

Before starting any digitization at NARA, staff must obtain approval from their supervisor and management chain, which includes the Office Digitization Liaison and their Office Executive. All digitization activities must be included in the Office's required annual digitization plan submitted to the Digitization Governance Board. (See Section NARA 1550.5.a.). All digitization projects must comply with NARA handling requirements for locating, pulling, delivering, and refiling of holdings.

1. Considerations that should be addressed throughout the lifecycle of a digitization project:

a. Digitization Planning

- (1) Holdings Evaluation
- (2) Staffing and Resource Evaluation
- (3) Schedule Development
- (4) Project Plan Development

b. Digitization Project Execution

- (1) Equipment Selection
- (2) Digital File Format and Technical Specifications Standards

c. Post-Digitization

- (1) Auditing, Quality Assurance, and Quality Control
- (2) Catalog Submission and Metadata

(3) Long-term Storage

2. Digitization Project Planning – Holdings Evaluation.

Offices should consider the following questions in relation to the holdings proposed for digitization. These factors impact the approach, the equipment selected, and the time and resources required to complete the work. Answers to these questions should impact selection of records for digitization and should be addressed within the Office digitization plans.

a. Analog format:

- (1) Are the holdings textual? Are the textual materials loose sheets or bound volumes?
- (2) Is it a photographic print, negative, transparency, or aperture card or other still image? Are the photographs/still images mounted or otherwise affixed?
- (3) Is it microfilm? Is it accessioned microfilm or a microfilm publication?
- (4) Is it audio? Is the original media approaching obsolescence and therefore the digitized version will become the record copy?
- (5) Is it video? Is the original media approaching obsolescence and therefore the digitized version will become the record copy?
- (6) Is it motion picture film? Is the original media unable to be preserved photo-chemically and therefore the digitized version will become the record copy?
- (7) Do the materials require special equipment or hardware and/or software to digitize?

b. Physical condition:

- (1) Can the materials be safely handled for digitization?
- (2) Is special handling assistance required?

c. Size:

- (1) Do the holdings require a large format scanner, camera or other special equipment?
- (2) Do the holdings require custom book cradles or other custom supports?

- d. Restriction status: After reviewing content per NARA 1601: Screening Federal Records for Information Covered by FOIA Exemptions, including, “Screening Guidelines for Record Series That Are Likely to Contain Some National Security-Classified or Other Exempted Information” consider the following:
 - (1) Is there personally identifiable information (PII) in the materials?
 - (2) Are there copyright or other intellectual property rights considerations?
 - (3) Will there be material that becomes open for public release after initial digitization and inclusion in the Catalog and how will that change be managed?
- e. Post-production activities:
 - (1) What post-production activities are needed to render the digital format usable by the public?
 - (2) Will the materials be enhanced? If so, metadata should be captured to indicate the digital format was enhanced during digitization or post-production and both enhanced and original digital files should be provided with a way to maintain a link between the digital files.

3. Digitization Project Planning - Staffing and Resource Evaluation.

Offices should consider the following staffing and resource questions when considering digitization projects including proposing a new project to the Office’s annual digitization plan. Staff who wish to propose new work and/or a continuation of and/or extensions to ongoing digitization projects should discuss the best approach for completing the work with their supervisor.

- a. Staffing
 - (1) NARA staff
 - (a) Is the digitization project on the unit’s annual work plan?
 - (b) Does the unit have sufficient and appropriate staff to support the digitization project?
 - (2) NARA volunteers (including Citizen Archivist Scanners, see Supplement 1 for guidance on Citizen Archivist Scanner Projects).

- (a) Has the project been coordinated with the Volunteer Coordinator? Has the project been approved for Citizen Archivist Scanners? If yes, consult NARA 1550, Supplement 1.
- (b) Will there be sufficient oversight for volunteers doing digitization work including training in scanning, metadata creation, and quality control?
- (c) Does the project require that volunteers receive records handling training from Preservation Programs?

b. Resource Evaluation

- (1) Have the required standards, equipment, and special requests been identified? Does the custodial unit have the equipment, services, technology, and knowledge (training) to support the project? If additional technology is needed, has that equipment and technology been requested and approved? Or does the digitization need to take place at a NARA Digitization Lab or a vendor?
- (2) Is the digitization taking place at the custodial unit location or at the NARA Digitization Lab or the Motion Picture or Audio Visual Preservation Lab? If at a NARA Lab, has the project been scheduled with the appropriate Digitization Lab or the Motion Picture or Audio Visual Preservation Lab following their procedures?

c. Contracting

- (1) What need is the contract seeking to fill?
- (2) Does a current contract in place at NARA meet this need? If not, have requirements been established and enough time allowed to enter into a new contract? Are there any special considerations for handling the records, for digitization planning, or for the deliverables to be provided by the contractor?

4. Digitization Project Planning - Schedule Development.

All digitization projects should have a schedule with resource time estimates and completion dates. Questions to consider include:

- a. How much archival preparation do the materials need?
 - (1) How long will archival preparation take with the available number of staff?
 - (2) What information must be gathered to assure reassembly (if required) and refiling?

- b. How long will metadata creation take?
 - (1) Is there an existing file or item list?
 - (2) Does a file or item list need to be created from scratch?
- c. How long will digitization and quality control take?
- d. Is the digitization project associated with any outreach (such as education, exhibit, or social media activities)? Is there a hard deadline?

5. Digitization Project Planning - Project Plan Development.

Project plans help systematically address all of the factors pertinent to a digitization project, establish a common set of expectations, and create baseline expectations to permit tracking and accountability. Plans must be approved at the management level designated by each Office. All approved digitization project plans should be included in the Office annual digitization plan submitted to the Digitization Governance Board. Digitization project plans should include the following information:

- a. Designated archival project manager.
- b. Justification for digitization.
- c. Digitization project number.
- d. Project scope:
 - (1) Citation(s) for series/collections to be digitized. Summary of work to be completed for non-series or collection level digitization including responses to FOIA requests.
 - (2) Volume of materials.
 - (3) Estimated item count.
- e. Condition and format of the original holdings.
- f. Conservation assessment.
- g. Archival preparation requirements.
- h. Type of equipment to be used for digitization.
- i. Location of digitization.

- j. Unit doing the digitization:
 - (1) Archival unit.
 - (2) Digitization Lab.
- k. Number of scanning stations/digital workstations.
- l. Number of staff.
- m. Estimated start and end dates.
- n. Metadata plan:
 - (1) At what level will holdings be described?
 - (2) File naming conventions.
- o. Digital file specifications.
- p. Quality control plan for files, content, and metadata.
- q. Estimated data volume to be created.
- r. Digital storage, backup, and recovery plan.
- s. Cost and time estimates.
- t. Reporting plan.
- u. Digitization SOP.

6. Project Execution - Equipment Selection.

The type of digitization equipment used should be determined by a number of factors including, but not limited to, the format/size of the original materials, available funding, space, approval by Information Services for new hardware/software, staff expertise with the equipment, and Office Executive approval.

A list of recommended equipment is available at [NARA@work](#).

Consult with Preservation Programs for further guidance on ensuring the appropriate match between material type, condition, and equipment.

7. Technical Specifications for Digitization Projects including Scanning and File Standards.

NARA uses the [Federal Agencies Digital Guideline Initiative \(FADGI\)](#) 3-star standards described in “[Technical Guidelines for Digitizing Cultural Heritage Materials \(2016\)](#)” as the minimum standard for digitization specifications to create digital files and to capture technical metadata. Note that the standards provided here are not exhaustive and only cover the most common digitization projects and formats. Digital file specifications may differ from project to project and will depend on the type, quality, and the material being digitized.

For specific questions about technical specifications and file formats for a digitization project not covered in NARA 1550 or included in FADGI guidance please contact:

- Digitization Division digitization@nara.gov for questions about textual, photographic, microfilm, and large-format holdings.
- Audio-Visual Preservation Lab AVPresLab@nara.gov for questions about audio-visual material.
- Motion Picture Film Preservation Lab motionpicturelabinquiries@nara.gov for questions about films, movies, and other motion picture material.

a. General digital image guidance:

- (1) Holdings with image content in color (e.g., photographs, graphs, charts, etc.) should be digitized in color. Justifications for exceptions to color scanning should be included in the digitization project plan.
- (2) Holdings must be digitized using the standards described herein.
- (3) All page edges of the holdings should be visible in the scanned image. Holdings should be digitized with a slight border around the edges to show that the entire record has been captured in the image. Borders must be created during the capture process and cannot be added in post-processing.
- (4) No part of the original record should be cut off or cropped off in the scanned image. All borders or mounting boards should be included. Spatial resolution may have to be adjusted for the actual image size.
- (5) Explanatory quality targets, or “slugs,” may be inserted on each image when the digital image is the result of a poor quality original, but is the best copy possible. Options for language in quality target/slugs:

- (a) “Best Possible Image” would indicate that the digital image is the best possible image that is safely able to be taken of the record if, even with conservation treatment, the record is in poor condition. For example, there is faded ink, smeared media, large losses, or information in a bound volume is lost in the gutter.
- (b) “This image contains all information on the document” would indicate that one or more edges of the record are bleeding out of the image area and the camera/scanner operator can confirm that no information is present on these edges.
- (c) “All unique information in this document is visible in this image” would indicate that part of the text is not visible in the digital image, usually due to a fastener or an attachment, but this hidden text is not unique (e.g. found elsewhere on records or printed form text).

b. Textual, Cartographic, and Graphic Materials.

(1) **File Type:**

- (a) JPEG maximum quality level.
- (b) PDF (option for textual materials).

(2) **Spatial Resolution:**

- (a) 300 ppi as the minimum standard for digitization resolution for originals sized up to 34” x 55”,
- (b) For originals with dimensions over 34” x 55”; ensure pixel array measures 3000 pixels, at minimum, across the long dimension of the image area (on the long edge).

(3) **Color Mode:** sRGB, or grayscale.

(4) **Bit Depth:** 8-bit grayscale, 8, 16 or 24-bit color.

c. Photographic Prints.

(1) **File Type:** JPEG maximum quality level.

(2) **Spatial Resolution:**

- (a) 400 ppi as the minimum standard for digitization resolution,

- (b) For all originals, especially for those under 5" x 7", ensure pixel array measures 4000 pixels, at minimum, across the long dimension of image area, excluding mounts and borders.
- (3) **Bit Depth:** 8 or 16-bit.
- (4) **Color Mode:** sRGB, or grayscale.
- d. Photographic Negatives and Color Transparencies.
 - (1) 35mm to 4" x 5":
 - (a) **Format type:** JPEG maximum quality level.
 - (b) **Spatial Resolution:** 6000 pixels on the long axis/3000 ppi.
 - (c) **Bit Depth:**
 - (i) **Color:** 16-bit (24 or 48-bit for selected).
 - (ii) **Black and White:** 16 bit.
 - (d) **Color Mode:** sRGB, or grayscale.
 - (2) Larger than 4" x 5":
 - (a) **Format type:** JPEG maximum quality level.
 - (b) **Spatial Resolution:** 1500 ppi.
 - (c) **Bit Depth:**
 - (i) **Color:** 16 bit.
 - (ii) **Black and White:** 16 bit.
 - (d) **Color Mode:** sRGB, or grayscale.

e. Dynamic Media

Dynamic media requires specific SOPs and training tailored for each digitization project. The following guidelines are suggested for dynamic media digital file formats and represent the current preferred formats for digitization. Adjustments should be made when necessary for digital file surrogates created for access and reference purposes.

- (1) Audio Recordings.
 - (a) **Format type:** WAV.
 - (b) **Resolution:** Uncompressed 96kHz/24-bit.
 - (2) Video Recordings.
 - (a) **Format type:** AVI.
 - (b) **Resolution:** SD Uncompressed 720x486, 10-bit.
 - (3) Film and Motion Pictures. Follow NARA guidance on [digitizing moving images](#).
- f. If your digitization project plan calls for the creation of preservation master files, follow the standards below. Preservation master files are a color and tonal corrected scan (digital image) for preservation that can then be used to create derivative files for access:
- (1) Include imaged dimensional and/or color reference targets,
 - (2) Create an uncropped image,
 - (3) Use 8-bit uncompressed grayscale TIFF or 8, 16, or 24-bit RGB color TIFF,
 - (4) Create both _us file and positive image of a negative,
 - (5) Spatial resolution standards:
 - (a) Textual holdings (4000 pixels longest side):
 - (i) Minimum 400 ppi at original size up to 34" x 55" originals.
 - (ii) Minimum 300 ppi at original size for originals between 34" x 55" and 46" x 73".
 - (iii) Option for less than 300 ppi at original size for originals larger than 46" x 73".
 - (b) Photographic Prints:
 - (i) Minimum 800 ppi 4" x 5" and smaller.
 - (ii) Minimum 600 ppi larger than 4" x 5".

- (c) Film originals: Maintain original file dimensions including extent, frame rate, aspect ratio. Consult NARA guidance on [digitizing moving images](#).

8. Project Execution - Digitizing Holdings.

Once NARA staff are ready to start digitizing materials, they should complete the following steps to ensure that the materials will be safely and accurately digitized. While each step outlined below may not be necessary for every project, these steps represent the basic tasks that should be taken into consideration when digitizing non-dynamic media. These steps should be part of unit digitization SOPs.

- a. Perfect the arrangement of all selected holdings as they are to be digitized.
- b. Confirm that SOPs for digitization are in place, that technical specifications are programmed into machines, where possible, and that the destination for digital files is prepared.
- c. Remove fasteners when unique information is obscured or if the record cannot be safely handled or digitized otherwise and include related slug:
 - (1) Separate all holdings that are stapled together prior to placing on scanner plate or under camera,
 - (2) Remove and address sticky (i.e. Post-It) notes and other attachments or inserts following unit SOPs.
- d. Ensure that the original record is properly placed for the selected scanning equipment:
 - (1) For flatbed scanners, ensure that original record is placed away from the sides of the glass edge. For non-edge based scanners (overhead, camera, etc.) make sure all edges of the image are visible in the scanning field and in the resulting digital capture.
 - (2) Unfold the record prior to digitizing if a fold is obscuring the text on a document.
 - (3) Request appropriate conservation treatment, if deemed necessary, if a tear obscures text.
 - (4) Ensure that all original information - both front and back - in the record is captured.

- e. Digitize all contents of a file folder or all identified selected material. Staff are not required to digitize exact duplicates (e.g. 2 copies of the same report).
- f. Insert a declassification slug on each image if the holdings are associated with a declassification project number.
- g. Conduct spot checks throughout the digitization process to ensure that all information is being captured correctly and image file names are accurate.

9. Post-Digitization - Quality Control.

This section describes quality control for imaging of non-dynamic media. Quality control checks are important tasks to ensure the accuracy and completeness of the digital images. Each project will differ but NARA staff should check either a pre-determined sample or percentage of the digital images or the entire set of digital images (a 100% check) prior to finalizing the project and submitting the digital files and metadata to the National Archives Catalog and to ERA 2.0 for long-term storage and preservation. Quality control checks include:

- a. Image quality check:
 - (1) Ensure that digital images meet specifications as outlined in NARA 1550.
 - (2) Ensure that the digital image matches the analog original and that all text (including handwritten text) is legible.
- b. Content quality check:
 - (1) Ensure that all original materials have been digitized.
 - (2) Gaps or missing items/pages identified during the content quality control check should be resolved as soon as possible prior to submission to the National Archives Catalog and ERA 2.0.
- c. Metadata quality check:
 - (1) Ensure the accurate completion of descriptive metadata at the appropriate archival level (e.g., file unit or item).
 - (2) Ensure parent National Archives Identifier (NAID) is included, if available.
 - (3) Ensure the accurate capture and completion of technical metadata.

10. Post-Digitization - Catalog Submission.

Staff must submit all digital files and associated metadata to Innovation staff for upload to the National Archives Catalog for public access. These steps are outlined in detail in Adding Descriptions and Digital Objects to the National Archives Catalog.

11. Post-Digitization – Long-term Storage.

All digital files and relevant metadata must be placed in the Digital Object Repository of ERA 2.0 for long-term storage and preservation. Once the files added to the Catalog and stored in ERA 2.0, the Office that conducted the digitization must delete the files from local storage, including shared drives, hard drives, area networks, and other non-cloud storage.