

United States National Archives and Records Administration Digitization Regulations

Michael Horsley, US National Archives and Records Administration, College Park, Maryland, USA

Abstract

The United States National Archives and Records Administration (NARA) has issued new regulations that establish standards for the digitization of US government records. The regulations are part of an effort to transition to a fully electronic government, and allow US federal agencies the authority to digitize and destroy source records and the electronic version become the recordkeeping copy. The specifications draw upon established international digitization standards such as ISO 19264, Metamorfoze, and FADGI guidelines.

By adopting the image quality specifications found in ISO 19264 and the image analysis method described by FADGI, NARA has effectively defined the minimum requirements for a digital surrogate to serve as legal and evidentiary purpose as the source record. This paper presents the records management context of digitization, as well as discussing the quality management, documentation, image and metadata specifications, and validation requirements.

Introduction

“Inspection with the aim of finding the bad ones and throwing them out is too late, ineffective, and costly. Quality comes not from inspection but from improvement of the process.”

W. Edwards Deming

The National Archives and Records Administration (NARA) is an independent agency established in 1934 to identify, protect, preserve, and make publicly available the historically valuable records of all three branches of the Federal government. NARA manages the Federal government’s archives, administers a system of Presidential Libraries, operates museums, conducts education and public programs, provides oversight of government-wide records management activities, and provides temporary storage of other agencies’ records on their behalf.

The Records Management Policy and Standards Team develops records management policies and guidance related to the creation, management, and transfer of electronic records and participates in the development of international records and document management standards. Key records management guidance maintained by the team are: “2015-04: *Metadata Guidance for the Transfer of Permanent Electronic Records*” [1], and “2014-04: *Revised Format Guidance for the Transfer of Permanent Electronic Records*” [2]. Specifically, we are responsible for creating regulations.

Regulatory Context

The Federal Records Act requires that NARA issue regulations and provide guidance to agencies to ensure proper records management. Through a NARA reviewed scheduling and appraisal

process agencies determine the value of their Federal records, how long temporary records must be retained before destruction, and which records are permanently valuable and must be transferred to NARA for preservation.

When Congress enacts a Federal law, or statute, the law often does not include all the details that explain how people, businesses, or government organizations should follow the law. To develop technical, operational, and legal details that make the laws work on a day-to-day level, Congress authorizes Federal government agencies - including NARA - to create regulations, or rules. Regulations are mandatory requirements that apply to federal agencies and may include details and requirements about, what is legal and what isn't, how to meet the requirements, timelines that must be met, and definitions of special terms.

Development of the Regulation

Beginning in 2022, physical permanent federal records heading to the National Archives (NARA) must be in digital form and with few exceptions, the originals will be disposed. This means all federal government documents must be digitized, a massive undertaking. NARA has published digitization regulations 36 CFR 1236 Sub-Part E [3] in the Code of Federal Regulations (CFR) that incorporate the Federal Agencies Digital Guidelines Initiative (FADGI) [4] technical guidelines, as well as defining records management, quality control, and validation of digital surrogates that will serve as the new records copy. Because the original records will not be retained, the image quality specifications for this is quite high, requiring rigorous application of FADGI guidelines to document mass digitization.



Figure 1. 36 CFR 1236 Subpart E Digitizing Permanent Federal Records

The regulation issues recommendations that: Agencies must digitize permanent records to ensure that electronic copies are authentic, reliable, usable, and have the integrity necessary to allow for the disposal of the original records. These digitization standards for permanent records ensure that agencies can continue to use digital versions for the same business purposes as the original records, and that the digital records will be appropriate for preservation in NARA’s archival holdings. It provides agencies with guidance necessary to digitize and destroy original source permanent records.

The regulation defines the requirements for digitization as a records management activity, drawing principles from the FADGI, *Technical Guidelines for Digitizing Cultural Heritage Materials (2023)*, the *International Standards organization publication ISO 13028:2010, Implementation guidelines for digitizing records*, and the requirements described in ISO 15489-1:2016 -- Records management -- Part 1: Concepts and principles. The regulation defines digitization standards for paper, photographic print, and mixed media records as record management requirements. Digitization requirements for specific media types will be added to the regulation as they are completed.

Relationship between FADGI and NARA regulations

FADGI is a collaborative effort by federal agencies to articulate common sustainable practices and guidelines for digitized and born digital content. The FADGI standards were developed out of NARA’s 2004 digitization guidelines, but is a guidance document and not a standard, nor is it a regulation. The FADGI guidance consists of an implementation of the ISO aim points and tolerances as applied to over 14 different records material types divided into 4 different image quality.

NARA has adopted the FADGI quality metrics and Digital image conformance evaluation method. However FADGI is a guidance document intended for an international audience and does not apply to all the requirement for the digitization of permanent federal records. Federal Agencies must comply to NARA Regulations found in 36 CFR 1236. However it is the technical basis for the regulations. NARA developed a new FADGI category for modern textual records as the minimum.

The Modern Textual records category is similar to FADGI 3 star for unbound textual general collections and is designed for modern office records that have white paper and black text. NARA estimates that over 80% of existing textual records are those created after 1950, and within that group the vast majority of those records were created by word processing applications and printed on modern office paper.

The regulations contain a second category of specifications for photographic print materials and records requiring higher quality specifications. Agencies must analyze the information content of the source records and apply more precise imaging standards to capture fine detail or maintain color fidelity. For example, records such as cartographic prints, or color illustrations may require the equivalent of FADGI 4 star quality imaging.

3.5 Documents (Unbound): Modern Textual Records

Master File Format Name and Version	Acceptable Compression Codes
TIFF 6.0	Uncompressed, DEFLATE (ZIP)
JPEG2000 part 1 (ISO/IEC 15444-1:2019)	JPEG 2000 part 1 core coding system lossless compression. Federal agencies may use up to 20:1 visually lossless compression.
Portable network graphics 1.2 (PNG)	DEFLATE (ZIP)
PDF/A	DEFLATE (ZIP), JPEG 2000 part 1 core coding system lossless compression. Federal agencies may use up to 20:1 visually lossless compression.
Measurement Parameters	
Resolution (Sampling Frequency) (Units are Pixels Per Inch/ppi minus Reproduction Scale Accuracy)	≥ 204ppi (300 ppi – 2%)
Bit Depth	8 or 16
Color Space	Gray Gamma 2.2, Adobe RGB (1998), sRGB, ProfPhoto, ECHRGB_y2
Color Mode	Color or grayscale
Measurement Parameters	Performance metric values Difference from aim (applies to 20 ≤ L* ≤ 100)
Tone Response (OECF) L* (Units	

Figure 2. FADGI Documents (Unbound): Modern Textual Records

Requirements in the regulation

The regulation is composed of several sets of requirements such as: Records Management, Documentation, Quality Management, File Formats, Digitization specifications, for permanent paper and photographic prints, Metadata, and Validation.

The Records management section discusses the concepts of intellectual and physical control, including the inclusion of finding aids and descriptive information, as well as accounting for records completeness. The documentation section outlines the various documentation that must be created to manage a digitization project and provide accountability for the process.

The Quality Management (QM) section provides a deep discussion of the mandatory quality assurance and quality control. The QM approach of the regulations takes the form of a total quality management approach to preventing defects and process improvement. The Digitization specifications describe the minimum image quality and file attributes that are based on FADGI guidelines.

The metadata requirements define the minimum mandatory administrative, technical, and descriptive elements for the records. Finally the Validation step provides the high level assertion that the project meets all the requirements.

Temporary Records, Current Work, Future Work

Prior to issuing standards for the digitization of permanent records, NARA issued 36 CFR 1236 Sub-Part D that contains the standards agencies may use when digitizing temporary records. The vast majority of records created by the government are temporary and may have short or long term retention periods. To provide agencies with practical flexibility to manage these records, the standards and requirements are not as strict as those for permanent records. However agencies must maintain these records to serve as the same purpose as the source record.

We are currently working on issuing regulations for the digitization of transmissive film materials including microforms, radiographs, aerial film, black and white and color negatives, and transparency film. The technical standards will adopt FADGI 3 star as the minimum requirement, but we will follow closely the developments by ISO TC42 JWG26 work on film materials and revise the requirements as needed.

Our final product will cover dynamic media such as audio, video, and motion picture records. Currently there are no accepted international or federal standards. Also the technology is expensive and the archival formats are not well accepted. Our current position is that agencies should transfer these materials to NARA as soon as possible. We will work closely with the FADGI Audio-Visual working group and NARA special media preservation staff to develop practical records management policies.

Supplementary Products

As part of its mission our team creates supplementary products to augment the instructions found in regulations. The formal structure of regulatory language does not allow for easy explanatory text to help explain the context of some of the mandatory elements. Supplementary products such as bulletins provide more detailed explanations of NARA policy. White Papers, Guidance, and Blog posts, allow our team to go into greater detail about complicated or technical aspects of the regulations. To address issues that have emerged and to further provide detailed information we have created several supplementary products including a Digitization Quality Manage Guide, a Success Criteria White paper, a Non-Compliant Digitized Records FAQ, and a Digitization Resources Webpage.

Quality Management

NARA has authored a Digitization Quality Management Guide [5] to assist agencies meet NARA’s requirements. The guide walks through the various stages of the Quality life cycle.

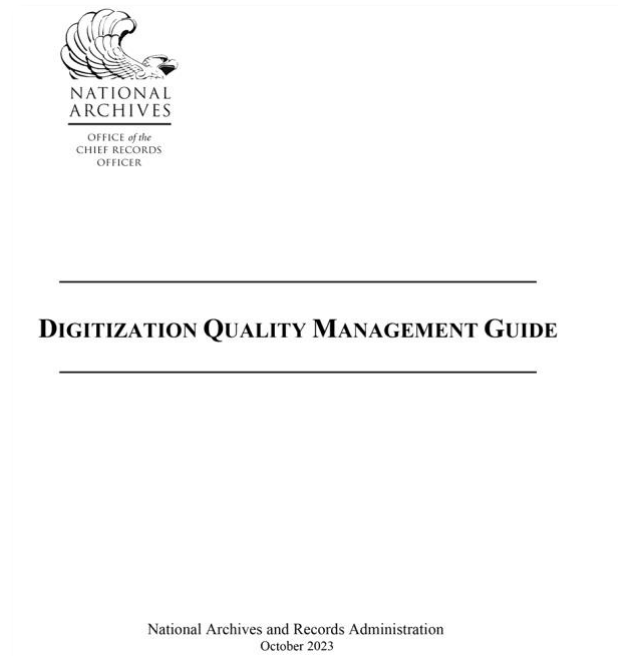


Figure 3. Digitization Quality Management Guide

The goal of quality management (QM) in a digitization project is to prevent defects before they happen. Prevention of defects is the most efficient, cost-effective, and productive activity in a digitization workflow. Digitization is an image manufacturing

process that relies upon quality assurance (QA) to establish specifications and requirements, and a quality control (QC) process to test for and inspect defects. A well-run digitization operation will be proactive rather than rely upon QC inspection to achieve quality.

NARA’s digitization regulations could be considered a total quality management document that brings together records management practices, digital imaging quality standards, quality control, inspection, and validation steps. Each phase of a digitization project affects the other parts. For example, establishing intellectual and physical control contributes to creating metadata as well as identifying missing records. The quality control and inspection steps have been optimized to rely on automated processes where possible and limit human inspection to the phases that cannot be automated. By relying on objective testing and analysis for image quality, we eliminate wasteful subjective inspection of attributes.

Success Criteria

If not properly managed, digitization can result in incomplete/insufficient digital record quality to serve all of the same business purposes as the originals. NARA has authored a success criteria white paper [6] to assist agencies in digitizing permanent records and meeting federal records management requirements in laws, regulations, and NARA-issued policy. It helps agencies adhere to the requirements in 36 CFR 1236 subpart E, and other recordkeeping requirements, to help mitigate many of the risks inherent in digitization. The white paper provides high-level success criteria for digitizing permanent records, which are organized around four key concepts: Policies, Access, Systems, and Disposition.

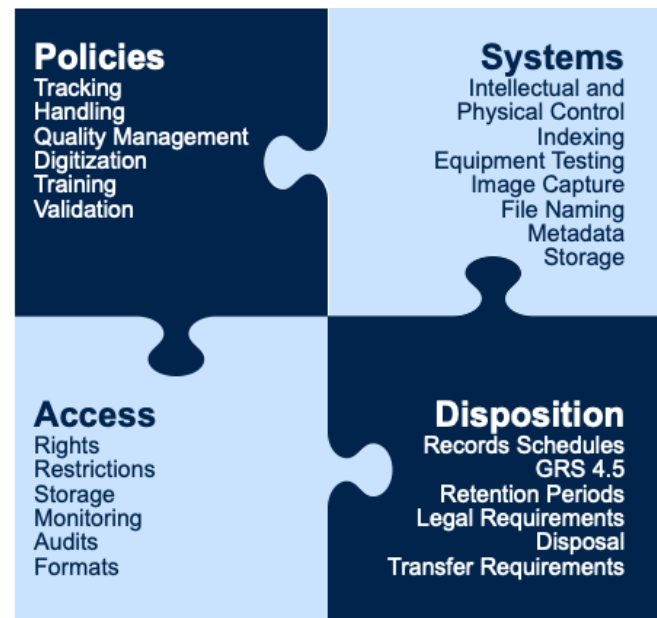


Figure 4. NARA Success Criteria

Non-Compliant Records FAQ

Non-compliant digitized records are those that do not meet the standards in the regulations. Agencies may have ongoing digitization projects that continue to generate non-compliant

digitized records. NARA recognizes that it will take agencies time to implement the digitization standards for permanent records. In order to assist federal agencies navigate through this transition, NARA has created an FAQ [7] regarding non-compliant permanent digitized records that provides answers to questions that apply to any digitization work that does not comply with the digitization standards. The FAQ answers questions about non-compliant digitized records that were created before the regulations were issued. It is reasonable to assume that most records digitized before the regulations were issued will not be compliant with all the digitization standards. If agencies validate that they have digitized permanent records according to NARA's digitization standards, then agencies can transfer the digitized records to NARA under the existing permanent records schedule, and destroy the source records.



Figure 5. NARA Non-Compliant Digitization FAQ

Digitization Resource Page

We have created a resource web page [8] to gather all the related topics and resources into one place.



Figure 6. Digitization Resources Webpage <http://www.archives.gov/records-mgmt/policy/digitization>

Conclusion

NARA's Digitization regulations define the steps necessary to reformat source records to create an electronic record copy. The standards and specifications are taken from the industry best

practices and recognized standards. The state of imaging science and technology has grown in sophistication to the extent that we can create reasonable facsimiles of the original that can serve as digital surrogate. Recognizing the experience of past imaging technologies we have taken care to delineate a process to create a new record copy that is faithful to the original but also more accessible to the communities that the records serve.

References

- [1] <https://www.archives.gov/records-mgmt/bulletins/2015/2015-04.html>
- [2] <https://www.archives.gov/records-mgmt/bulletins/2014/2014-04.html>
- [3] 36 CFR 1236 Sub-Part E in the Code of Federal Regulations (CFR)
- [4] "Technical Guidelines for Digitizing Cultural Heritage Materials 3rd ed.", Federal Agency Digital Guideline Initiative, Rieger, Tom et al editors, 2023.
- [5] "Digitization Quality Management Guide", National Archives and Records Administration, 2023 <https://www.archives.gov/files/records-mgmt/policy/digitization-quality-mgmt-guide.pdf>
- [6] "Success Criteria for Digitizing Permanent Records" National Archives and Records Administration, 2023 <https://www.archives.gov/files/records-mgmt/policy/digitization-success-criteria.pdf>
- [7] "Frequently Asked Questions (FAQ) about Non-Compliant Permanent Digitized Records" National Archives and Records Administration, 2023 <https://www.archives.gov/records-mgmt/policy/non-compliant-digitized>
- [8] Digitization Resource page, National Archives and Records Administration, 2023, <https://www.archives.gov/records-mgmt/policy/digitization>

Author Biography

Michael Horsley is a Senior Electronic File Format Analyst in the Office of the Chief Records Officer for the U.S. Government, United States National Archives and Records Administration. He is a technical editor for the 2023 FADGI Technical Guidelines 3rd ed. A member of ISO/TC42 JWG 26, TC 46, and TC 171. He has over 25 years of digital imaging, records management, and archival experience at the Smithsonian Institution, Library of Congress, and National Archives.