Mississippi State University

Scholars Junction

Policies and Procedures   University Libraries

2019

DPAU Technical Guidelines.2019

Jenifer Ishee

Follow this and additional works at: https://scholarsjunction.msstate.edu/libpolicies

Digital Preservation and Access Unit (DPAU)

Technical Guidelines

for Digitization of Cultural Heritage Materials

Rev. August, 2019

I.   PURPOSE:

This document sets forth the Digital Preservation and Access Unit's (DPAU) guidelines for digitizing cultural heritage materials. It is DPAU's goal to adhere to best practices and nationally recognized standards in the creation of master image files for preservation purposes and to comply with the FADGI 4-star quality level of imaging.

II.   FOUNDATION:

These guidelines are based on the following standards:

- Federal Agencies Digital Guidelines Initiative (FADGI) Technical Guidelines for Digitizing Cultural Heritage Materials (2016)
- Library of Congress Recommended Format Statements (2018-2019)
- The Association for Library Collections and Technical Services Preservation and Reformatting Section Minimum Digitization Capture Recommendations (2013)
- NISO Standard ANSI/NISO Z39.87-2006 (R2017) Data Dictionary – Technical Metadata for Digital Still Images
- PREMIS (Preservation Metadata: Implementation Strategies) Data Dictionary version 3.0 (June 2015)
- WCAG 2.1 (AA level) Standards (June 2018)

III.   DEFINITIONS:

A. **Master Archival files**: The guidelines in section V are for the creation of Archival Master files. Due to the stress of digitizing cultural heritage materials, a digital master file should be generated for every object created. The Master Archival file will have long tonal scale, wide color gamut, and be minimally adjusted to be use-neutral.

B. **Production Master files**: These are files produced by processing the content in one or more Master Archival files. For example, an image of a large map can be produced by stitching together several segments of the map to create one whole image.

C. **Derivative files**: These are created from the Master Archival and Product Master files and can have the ppi reduced for viewing online (generally 200-400 ppi), depending on the format. [NOTE: CONTENTdm, the content management platform utilized by DPAU, archives JPEG files as JPEG2000].

D. PDF:

i. <u>PDF/A</u>: The primary purpose for the PDF/A format is to represent electronic documents in a manner that preserves their static visual appearance over time, independent of the tools and systems used for creating, storing or rendering the files. To this end, PDF/A attempts to maximize device independence, self-containment, and self-documentation. Recommended by LOC for compliance with IOS standards (ISO 19005. Document management - Electronic document file format for long-term preservation).

ii. <u>PDF/UA</u>: This framework allows assistive technology, PDF applications and the PDF file format to work together optimally so that as many people as possible have access to PDF documents with sound technical support. The purpose of the PDF/UA is to create a *universally accessible* PDF document. PDF's can be validated and tested for compliance with Acrobat Pro, Window's NVDA, and Access for All's open source PDF Accessibility Checker.

"PDF/UA-compliant PDF documents are universally accessible from a technical standpoint, which means they allow improved communication for people with disabilities in particular. Many legislative bodies refer to WCAG 2.0, the international guidelines for accessible web content. When it comes to PDF files, these guidelines can be met by complying with PDF/UA's specific, unambiguous requirements." (From *PDF/UA in a Nutshell*, Olaf Drümmer and Bettina Chang)

*Digitized materials should be both PDF/A and PDF/UA compliant for digital preservation goals and accessibility standards.

iii. <u>OCR</u>: PDF files should be OCR'd before converting to PDF/A for maximum searchability. This can be accomplished through Adobe Pro. [NOTE: CONTENTdm supports integrated OCR functionality through the OCR Extension. Using the OCR Extension, full text can be generated from JPEG2000, JPEG, PNG, GIF and TIFF files.

IV. TESTING

Testing will be routinely performed to ensure compliance with DPAU digitalization guidelines. The following tools will be utilized to assist in testing:

a. OpenDICE, a FADGI conformance measurement and analysis tool which has been developed as an open source alternative to the DICE (Golden Thread) system. OpenDICE will provide an analysis of the imaging variables that comprise the FADGI star ratings.

b. AutoSFR, a program developed to assist imaging professionals determine the actual resolution present in images. This information assists in the determination of the appropriate scanning resolution for specific collections.

c. As mentioned in Section III.D.ii above, DPAU will utilize software to test for PDF/UA compliance.

V. GUIDELINES:

The Master Archival files should adhere to the following guidelines:

A. Bound Volumes: Rare and Special Materials:

Performance Level: 4 Star

Master File Format	TIFF, JPEG 2000, PDF/A
Access File Formats	All
Resolution	400 ppi
Bit Depth	16
Color Space	Adobe 1998, ProPhoto, ECIRGBv2
Color	Color

B. Documents (Unbound): Manuscripts and Other Rare and Special Materials

Performance Level: 4 Star

Master File Format	TIFF, JPEG 2000, PDF/A
Access File Formats	All
Resolution	400 ppi
Bit Depth	16
Color Space	Adobe 1998, ProPhoto, ECIRGBv2
Color	Color

C. Oversize Items: Maps, Posters, and Other Materials

Performance Level: 4 Star

Master File Format	TIFF, JPEG 2000
Access File Formats	All
Resolution	400 ppi
Bit Depth	8 or 16
Color Space [1]	Grey Gamma 2.2, SRGB Adobe 1998 ProPhoto, ECIRGBv2
Color [2]	Grayscale or Color

D. Newspapers

Performance Level: 4 Star

Master File Format	TIFF, JPEG 2000, PDF/A
Access File Formats	All
Resolution	400 ppi
Bit Depth	8
Color Space	Grey Gamma 2.2 SRGB
Color	Color

E. Prints and Photographs

Performance Level: 4 Star

Master File Format	TIFF
Access File Formats	All
Resolution	600 ppi[1]
Bit Depth	16
Color Space	Adobe 1998 ProPhoto, ECIRGBv2
Color	Color

F. Photographic Transparencies: 35mm to 4"x5"

Performance Level: 4 Star

Master File Format	TIFF
Access File Formats	All
Resolution	4000 ppi[1]
Bit Depth	16
Color Space	Grey Gamma 2.2; Adobe 1998 ProPhoto ECIRGBv2
Color	Grayscale or Color as appropriate

G. Photographic Transparencies Larger than 4" x 5"

Performance Level: 4 Star

Master File Format	TIFF
Access File Formats	All
Resolution	2000 ppi[1]
Bit Depth	16
Color Space	Grey Gamma 2.2; Adobe 1998 ProPhoto; ECIRGBv2
Color	Grayscale or Color as appropriate

H. Photographic Negatives: 35mm to 4"x5"

Performance Level: 4 Star

Master File Format	TIFF
Access File Formats	All
Resolution	4000 ppi
Bit Depth	16
Color Space	Gray Gamma 2.2; Adobe 1998 ProPhoto ECIRGBv2
Color	Grayscale or Color as appropriate

I. Photographic Negatives Larger than 4" x 5"

Performance Level: 4 Star

Master File Format	TIFF
Access File Formats	All
Resolution	2000 ppi[1]
Bit Depth	16
Color Space	Gray Gamma 2.2; Adobe 1998 ProPhoto ECIRGBv2
Color	Greyscale or Color as appropriate

J. Paintings and Other Two-Dimensional Art (Other Than Prints)

Performance Level: 4 Star

Master File Format	TIFF
Access File Formats	All
Resolution	12,000 Pixels on long dimension or Minimum 600 ppi of physical object
Bit Depth	16
Color Space	Adobe 1998 ProPhoto ECIRGBv2
Color	Color

K. Printed Matter, Manuscripts, and Other Documents on Microfilm

Performance Level:4 Star

Master File Format	TIFF, JPEG2000
Access File Formats	All
Resolution*	4000 ppi
Bit Depth	8
Color Space	Gray Gamma 2.2
Color	Grayscale

VI.   METADATA

DPAU will utilize the following metadata schemes for creating of descriptive and technical metadata:

- Dublin Core Metadata Element Set
- Anglo-American Cataloging Rules (AACR2)
- Library of Congress Authorities to guide the creation and structure of descriptive fields
- Cataloguing Cultural Objects (CCO, 2005)
- PREMIS (Preservation Metadata: Implementation Strategies) Data Dictionary version 3.0 (June 2015)
- NISO Standard ANSI/NISO Z39.87-2006 (R2017) Data Dictionary – Technical Metadata for Digital Still Images

VII.   FILE NAMING CONVENTIONS

Good filenames will ensure efficient management (storage and backups) of large numbers of digital files, and ensure the names are compatible with most software applications and file systems now and in the future, including compression utilities, Web protocols (URLs), optical storage file systems, and different operating systems. These conventions will mirror the directory structure as indexed in the archives drive.

File naming conventions should adhere to the following format:

<institution>_<department>_<collection#>_<date>_<short item description>_<sequence description>

- NOTE: Only use sequence description for digital objects consisting of more than 1 file
- NOTE: Date should follow international formats:
    - xxx-xx-xx (year, month, date)
    - xxxx-xx (year, month)
    - xxxx (year)

Examples:

MFM_MSS_65_1975-03-17_MCLphoto.tif

MFM_MSS_65_1975-03-17_MCLphoto.0001.tif

Special Collections Departments Identifiers:

MSS: Manuscripts
UA: University Archives
CPRC: Congressional and Political Research Center
FVW: Frank and Virginia Williams Collection of Lincolniana
USG: Grant Collection
Rare Books: Will be identified by catalog call number