CSU Japanese American Digitization Project

Technical Reference Guide

Prepared by Steve Kutay
Edited by Maureen Burns
mailto:stephen.kutay@csun.edu
Updated December 17, 2018

Contents

Introduction

The digitization specifications and guidelines drafted here are intended to provide workflow consistency and quality control for digital objects produced across participating CSU campus libraries and partner archives as part of the CSU Japanese American Digitization Project (CSUJAD). This guide incorporates best practices for digitization as recommended by the Federal Agencies Digitization Guidelines Initiative (FADGI) created for still images and audio-visual materials revised in 2016. The FADGI guides addressed here are *Technical Guidelines for Digitizing Cultural Heritage Materials* by the Still Image Working Group, and MXF Application Specification and Broadcast Wave Metadata by the Audio-Visual Working Group.

The FADGI Guidelines

Still Images and Text. This guide incorporates, at the very least, minimum specifications for preservation scanning/reformatting as described by the Still Image Working Group of the Federal Agencies Digitization Guidelines Initiative in *Technical Guidelines for Digitizing Cultural Heritage Materials* (2016) and the California Digital Library's *CDL Digital File Format Recommendations* (2011). This guide is not explicitly intended to replace workflows and/or policies already adopted by institutions responsible for the storage and maintenance of their own preservation files, assuming such local practices are commensurate with previous FADGI/CDL or comparable guidelines. However, this guide addresses the essential factors determining the FADGI compliance for institutions that wish to conform to the revised 2016 system and establishes a minimum acceptable quality required for materials that are digitally preserved according to the commitments shared by CSU partners and the granting agencies that have funded the project.

The revised 2016 FADGI *Technical Guidelines* restructure the information into a four star system indicating levels of derivative creation and use of digitally scanned images and text (4 star – recommended, 3 star – minimal):

- 1 Star - Minimal reference. Not intended to be used as a surrogate (e.g. a thumbnail).
- 2 star - Low resolution reproduction that does not meet the minimal specifications of a surrogate.
- **3 star - A digital surrogate that contains minimal specifications of resolution and size.**
- **4 star - A digital surrogate that contains the recommended specifications of resolution and size.**

NOTE: For those interested, other preservation image quality control standards outside of FADGI are ISO 19264-1 and Metamorfoze.

Scanner and Camera Calibration - DICE

To facilitate the best practices defining the FADGI four-star system is the Digital Imaging Conformance Evaluation (DICE) to facilitate image analyses that ensure digitization equipment is capable of reproducing accurate images within a range of acceptable tolerances. OpenDICE and AutoSFR are free downloadable applications for analyzing device level target images taken with your cameras and scanners to ensure that your equipment is providing quality reproductions of source materials. This saves significant time and training that would otherwise be required to do it manually. The software provides profiles containing an

acceptable range of measurements across multiple image parameters. Should your target renderings fail to fall within acceptable measurements, your equipment will require recalibration, which depending on the type and model, can be performed manually or by the manufacturers or other qualified maintenance services. In some cases a device cannot be altered from its default settings and will require replacement altogether. For this reason, great care should be taken to purchase high quality, professional level cameras and scanners.

Device Level Targets

A FADGI compliant DICE quality control environment uses color targets to test the image reproduction accuracy of cameras and scanners used in the reproductions of 2-dimensional objects. A robust FAQ and tutorial regarding the FADGI quality controls are available from ImageAccess. Targets are expensive and often challenging to locate. Color targets can be purchased from camera/video supply vendors or select image calibration companies. *X-Rite Colorchecker SG* and *DICE Golden Thread* targets (described in the above FAQ) are recommended for use with OpenDICE and AutoSFR software.

Audio-visual materials. The current Application Specification is AS-07 which promotes the use of the Material eXchange Format (MXF) by the Society of Motion Picture and Television Engineers (SMPTE) for preserving moving image media for the *long term*. This format is capable of capturing embedded data such as timecodes used in broadcast recordings. If your institution has film or broadcast video, and possesses the knowledge and digital storage required to preserve your moving image materials for the long term, consider the MXF industry standard, if it is not already implemented locally. It is likely, however, that some, if not most partners are not yet able to deploy AS-07, in which case it is recommended that a strategy for *medium term* preservation be used. The specifications for medium term (or minimal) preservation will be referenced here and will be consistent with the minimum video capture recommendations of the California Digital Library (CDL) and the Library of Congress

Access derivative processing. No comparable set of best practices are referenced regarding the processing of access derivatives to populate the database of the CSUJAD Project. However, guided by our responsibility to comply with the mandates of the *Americans with Disabilities Act* (ADA) derivatives will be processed accordingly. Beyond this, decisions affecting derivatives are intended to produce the consistent performance of digital objects across all contributing CSUJAD institutions.

Embedded metadata. Adding core embedded XMP standard (or IPTC for XMP) metadata is recommended for preservation and access files. Applications such as Adobe Bridge, Photoshop, Acrobat and others can be used to write and read embedded metadata through *File/File Info*.

RELATED RESOURCES: Adobe XMP, XMP/Wikipedia and XMP ISO 16684.

File Outputs

FADGI defines three levels of analog-digital outputs that serve multiple purposes, from long term preservation to online surrogates.

- Archival Masters (3 or 4 star conformance rating) are captured without alteration or additional digital processing. These high resolution, uncompressed files represent the best copies intended for sustainable long-term storage.
- Production masters (3 or 4 star conformance rating) are high quality copies of archival masters intended for use as reproducible master files that may contain color adjustments or other types of adjustments, such as stitching to form a complete composite image from separate archival master images required of maps and other unusually large documents. Applied corrections should not alter the intellectual content of the document, leaving all essential characteristics intact.
- Access files (2 star conformance rating) are produced from either type of master file. Typically access files are reduced in size either through reduction in tonal range, compression, or other processes in order to facilitate file performance for end user access as a surrogate of the original. Access files should maintain the integrity of the intellectual content found in the original. File formats used for access can not be considered stable for archival use given changes in proprietary algorithms used in the rendering of these files.

NOTE: Benefits and limitations of specific file formats can be found at Sustainability of Digital Formats developed by the Library of Congress. Format recommendations are given in the document type sections of this guide.

Scanning Specifications - Still Images

Specifications presented here are for the most common still images for the CSUJAD project, which are prints and photographs, photographic transparencies, and negatives. See the FADGI Technical Guidelines for less common sizes and formats such as paintings and other two dimensional art (page 46), X-Ray film (page 48), microfilm (page 50).

Prints and Photographs

Includes photographic prints, graphic-arts prints (intaglio, lithographs, etc.), drawings, some paintings,
(e.g., watercolors), and some maps.

	Preservation Master		Access Derivative
	4 Star (recommended)	3 Star (minimum)	
Master File Format	TIFF	TIFF	-
Access File Format	-	-	JPG, JPG2000
Resolution	600 ppi	400 ppi	200 ppi
Bit Depth	16	8 or 16	8
Color Space Options	Adobe 1998 ProPhoto ECIRGBv2	Adobe 1998 ProPhoto ECIRGBv2	Grey Gamma 2.2 SRGB Adobe 1998 ProPhoto ECIRGBv2
Color	Color	Color	Grayscale or Color

For aerial photographs, double the resolution of the access derivative to ensure clarity

online. See FADGI Technical Guidelines (page 33) if assessing tolerance of

measurement parameters. **Recommended Technologies**

- Planetary scanners
  (overhead)
- Digital cameras
- Flatbed scanners

Not Recommended Technologies

- Drum scanners
- Lighting systems that raise the surface temperature of the original more than 4 degrees F (2

degrees C) in the total imaging process. Excess heat can cause bowing or cupping of the source document.

Object Level Reference Targets

The use of standard IT8 reference targets are recommended for tone and color reproduction and dimensional references and are required to conform to the Digital Image Conformance Evaluation (DICE). Unlike object level targets (see Device Level Targets), these targets are placed next to source documents during the image process. The use of targets for positive color transparencies is unnecessary as references should already be embedded. Reflective and transmissive Targets should be placed close to, but not touching the original to allow adequate space for cropping of the image during the creation of access copies. For more information on using reference targets, see [Technical Guidelines for Digitizing Cultural](#) [Heritage Materials](#) (page 72).

Cropping

Scanned preservation **images should contain a small border around the entire image (roughly 3 to 8 mm)**. Film images (such as slides) with mounts may include a border around the mount if there is desired information on the mount itself.

Access Derivatives

File formats and specifications for *access derivatives* for use online is more flexible than for preservation files, and are selected here for consistency throughout the CSUJAD project.

Format

Access derivatives of preservation still images will be saved as compressed JPEG (.jpg) or JPEG 2000 (.jp2) format.

Resolution

Still image (non-aerial) resolutions will be 200 ppi. **Aerial images will be 300 to 400 ppi** as necessary to deliver the required detail. For the best full view display of regular images on most monitors, **access images will be scaled to 1024 pixels** along the longest edge. Aerial images will be scaled to 2048 pixels or greater as needed to deliver the required detail.

Embedded Metadata

Recommended. Much of the embedded technical metadata for images is automated, however, descriptive metadata pertaining to structure (if part of a compound object), administration (management), and rights are especially useful. Use ITPC for XMP standard.

RELATED RESOURCES: [FADGI Guide Still Image Working Group, IPTC](#), [IPTC Guide](#), [IPTC4XMP](#)

Photographic Transparencies - 35mm to 4 x 5 inches

Black-and-white or color (positive) transparencies. For transparencies larger than 4" x 5" see [Technical](#) [Guidelines for Digitizing Cultural Heritage Materials](#) (page 38).

	Preservation Master		Access Derivative
	4 Star (recommended)	3 Star (minimum)	
Master File Format	TIFF	TIFF	-
Access File Format	-	-	JPG, JPG2000
Resolution	4000 ppi	3000 ppi	200 ppi
Bit Depth	16	16	8 or 16
Color Space Options	Adobe 1998 ProPhoto ECIRGBv2	Adobe 1998 ProPhoto ECIRGBv2	Grey Gamma 2.2 SRGB Adobe 1998 ProPhoto ECIRGBv2
Color	Grayscale or Color	Grayscale or Color	Grayscale or Color

See [FADGI Techinical Guidelines](#) (page 36) if assessing tolerance of measurement

parameters. **Recommended Technologies**

- Film scanners
- Planetary scanners (overhead)
- Digital cameras
- Flatbed scanners

Not Recommended Technologies

- Drum scanner

The use of the Kodak IT8 Q60 reference targets are recommended for use with transmission scans. The use of a calibrated viewing system including a 5000k light box are critical.

Cropping

Scanned preservation **images should contain a small border around the entire image (roughly 3 to 8 mm).**

Access Derivatives

File formats and specifications for *access derivatives* for use online is more flexible than for preservation files, and are selected here for consistency throughout the CSUJAD project.

Format

Access derivatives of preservation still images will be saved as compressed JPEG (.jpg) or JPEG 2000 (.jp2) format.

Resolution

Still image (non-aerial) resolutions will be 200 ppi. Aerial images will be 300 to 400 ppi as necessary to deliver the required detail. For the best full view display of regular images on most monitors, access images will be scaled to 1024 pixels along the longest edge. Aerial images will be scaled to 2048 pixels or greater as needed to deliver the required detail.

Embedded Metadata

Recommended. Much of the embedded technical metadata for images is automated, however, descriptive metadata pertaining to structure (if part of a compound object), administration (management), and rights are especially useful. Use ITPC for XMP standard.

RELATED RESOURCES: [FADGI Guide Still Image Working Group, IPTC](#), [IPTC Guide](#), [IPTC4XMP](#)

Photographic Negatives - 35mm to 4 x 5 inches

Black-and-white or color negatives. For negatives larger than 4" x 5" see Technical Guidelines for Digitizing Cultural Heritage Materials (page 43).

	Preservation Master		Access Derivative
	4 Star (recommended)	3 Star (minimum)	
Master File Format	TIFF	TIFF	-
Access File Format	-	-	JPG, JPG2000
Resolution	4000 ppi	3000 ppi	200 ppi
Bit Depth	16	16	8 or 16
Color Space Options	Gray Gamma 2.2 Adobe 1998 ProPhoto ECIRGBv2	Gray Gamma 2.2 Adobe 1998 ProPhoto ECIRGBv2	Grey Gamma 2.2 SRGB Adobe 1998 ProPhoto ECIRGBv2
Color	Grayscale or Color	Grayscale or Color	Grayscale or Color

See FADGI Technical Guidelines (page 40) if assessing tolerance of measurement

parameters. **Recommended Technologies**

• Film scanners
• Planetary scanners (overhead)
• Digital cameras
• Flatbed scanners

Not Recommended Technologies

• Drum scanner

NOTE: Given a lack of available targets on the market for negatives, it is recommended to manually adjust parameters for highlight, shadow, and mid-tone measurements prior to scanning the original image.

Cropping

Scanned preservation **images should contain a small border around the entire image (roughly 3 to 8 mm)**.

Access Derivatives

File formats and specifications for *access derivatives* for use online is more flexible than for preservation files, and are selected here for consistency throughout the CSUJAD project.

Format

Access derivatives of preservation still images will be saved as compressed JPEG (.jpg) or JPEG 2000 (.jp2) format.

Resolution

Still image (non-aerial) resolutions will be 200 ppi. **Aerial images will be 300 to 400 ppi** as necessary to deliver the required detail. For the best full view display of regular images on most monitors, **access images will be scaled to 1024 pixels** along the longest edge. **Aerial images will be scaled to 2048 pixels or greater** as needed to deliver the required detail.

Embedded Metadata

Recommended. Much of the embedded technical metadata for images is automated, however, descriptive metadata pertaining to structure (if part of a compound object), administration (management), and rights are especially useful. Use ITPC for XMP standard.

RELATED RESOURCES: [FADGI Guide Still Image Working Group, IPTC](#), [IPTC Guide](#), [IPTC4XMP](#)

Scanning Specifications - Text

Specifications presented here are for the most common textual documents for the CSUJAD project, which are bound volumes (general collections); unbound manuscripts and other rare or special materials; unbound documents (general collections); and newspapers. See the [FADGI Technical Guidelines](#) for bound volumes of rare and special materials (page 18), and oversize items (page 28).

The PDF Format

The PDF format is the preferred format for documents containing significant amounts of text that can be utilized by databases and assistive technologies (screen readers) during machine processing or tagging, which conforms to the *Americans with Disabilities Act*.

Optical Character Recognition (OCR)

Optical character recognition is often applied automatically at the time a document is photographed and through the scanning software, if such software bundles use third party OCR software such as ABBYY FineReader. In this case, a minimum resolution of 300 ppi is usually required to perform OCR directly from within scanning software.

If you do not have OCR capabilities in your scanning software, you can apply text recognition directly to PDF files. Adobe Acrobat Pro enables text recognition via the tools menu. Click "Tools/ Text
Recognition/In this File" to begin text analysis. There will likely be many recognition errors. If you wish to correct these errors, they can be done manually by clicking "Tools/Text Recognition/Find All Suspects." Follow the prompts to fill in unrecognized characters manually. More information can be found on [Adobe Acrobat OCR](#), and on [Adobe Acrobat OCR](#) [on YouTube](#). **For all typescripts, PDF files should contain OCR of full text.**

Accessibility Tagging

The following instructions apply to Adobe Acrobat for *accessibility* processing. In some cases, structures/reading order is not machine readable, such as hand-written documents, or poor text renderings. In such cases, these documents are not fully accessible unless providing a transcript, and therefore must be saved as PDF/A – 1b, which is the minimum compliance for the PDF/A standard.

Under File/Preferences/Description enter a title for the document and the author (if known). Under File/Preferences/Advanced enter the language of the document. **This is required of all PDF files**.

Using Adobe Acrobat, tag the document structure via "Tabbed Reading Order," set "Alternate Text" for hand-drawn text and/or images using the Accessibility Tool, OR the Action Wizard/ Make Accessible. Save file as **PDF/A – 1a**. If the document could not be successfully tagged as described above, save as **PDF/A – 1b**.

RELATED RESOURCES: Adobe Accessibility Training, PDF/A

Bound Volumes (General Collections)

The following scanning and processing specifications are for materials such as brochures, booklets or other multipage documents that are bound together and NOT considered unique, rare, or otherwise deemed
'special' due to inscriptions, physical materials, or other processes that would require higher scanning standards.

	Preservation Master		Access Derivative
	4 Star (recommended)	3 Star (minimum)	
Master File Format	TIFF, JPEG 2000, PDF/A	TIFF, JPEG 2000, PDF/A	-
Access File Format	-	-	PDF/A
Resolution	400 ppi	300 ppi	150 ppi or higher
Bit Depth	8 or 16	8 or 16	8 or 16
Color Space Options	SRGB Adobe 1998 ProPhoto ECIRGBv2	Gray Gamma 2.2 SRGB Adobe 1998 ProPhoto ECIRGBv2	Grey Gamma 2.2 SRGB Adobe 1998 ProPhoto ECIRGBv2
Color	Color	Grayscale or Color	Grayscale or Color

See FADGI Technical Guidelines (page 21) if assessing tolerance of measurement

parameters. **PDF/A levels**
- **PDF/A-1a** (OCR, tagged structure, tabbed reading order, alternative text), **OR**
- **PDF/A-1b** (minimal compliance to ensure accurate image preservation and rendering)

PDF/A documents over 17 Mb, should be split into 10 Mb parts and submitted as compound objects.

Recommended Technologies
- Planetary book scanners (overhead) with or without glass platens
- Digital cameras

Not Recommended Imaging Technologies
- Flatbed scanners
- Lighting systems that raise the surface temperature of the original more than 6 degrees F (3 degrees C) in the total imaging process. This can cause bowing or cupping of the document.
- Linear scanning processes without glass or plastic platens (scanners and digital scanning back

cameras) are not appropriate because of the potential for the original to flex during the scanning process, producing artifacts that cannot be corrected.

Item reference targets are optional. Targets used for DICE device calibration are recommended.

Cropping

Scanned preservation **images should contain a small border around the entire image**

(roughly 3 to 8 mm).

Unbound Manuscripts and Other Rare/Special Materials

The following scanning and processing specifications are for manuscripts such as correspondence, and other unbound documents deemed unique or rare.

	Preservation Master		Access Derivative
	4 Star (recommended)	3 Star (minimum)	
Master File Format	TIFF, JPEG 2000, PDF/A	TIFF, JPEG 2000, PDF/A	-
Access File Format	-	-	PDF/A
Resolution	400 ppi	300 ppi	300 ppi
Bit Depth	16	8 or 16	8
Color Space Options	Adobe 1998, ProPhoto, ECIRGBv2	Adobe 1998, ProPhoto, ECIRGBv2	Adobe 1998, ProPhoto, ECIRGBv2
Color	Color	Color	Color

See FADGI Technical Guidelines (page 21) if assessing tolerance of measurement parameters.

PDF/A levels

- **PDF/A-1a** (OCR, tagged structure, tabbed reading order, alternative text), **OR**
- **PDF/A-1b** (minimal compliance to ensure accurate image preservation and rendering)

For PDF/A documents over 17 Mb, split into 10 Mb parts and submit as compound object.

Recommended Technologies

- Planetary scanners (overhead)– manually operated
- Digital cameras

Not Recommended Technologies

- Lighting systems that raise the surface temperature of the original more than 4 degrees F (2 degrees C) in the total imaging process. This can cause bowing or cupping of the document.
- Sheet fed scanning systems that contact the recto (face) or verso (back) of the original.

Item reference targets are optional. Targets used for DICE device calibration are recommended.

NOTE: FADGI compliance of special collections materials requires advanced training of personnel in addition to adherence to 4 star conformance levels.

Cropping

Scanned preservation **images should contain a small border around the entire image (roughly 3 to 8 mm).**

Unbound Documents (General Collections)

The following scanning and processing specifications are for documents, usually typescript, NOT deemed unique or rare.

	Preservation Master		Access Derivative
	4 Star (recommended)	3 Star (minimum)	
Master File Format	TIFF, JPEG 2000, PDF/A	TIFF, JPEG 2000, PDF/A	-
Access File Format	-	-	PDF/A
Resolution	400 ppi	300 ppi	150 ppi or higher
Bit Depth	16	8 or 16	8
Color Space Options	Adobe 1998, SRBG ProPhoto, ECIRGBv2	Adobe 1998, SRBG ProPhoto, ECIRGBv2	Grey Gamma 2.2, SRGB Adobe 1998, ProPhoto, ECIRGBv2
Color	Color	Color	Grayscale or Color

See [FADGI Technical Guidelines](#) (page 26) if assessing tolerance of measurement

parameters. **PDF/A levels**

- **PDF/A-1a** (OCR, tagged structure, tabbed reading order, alternative text), **OR**
- **PDF/A-1b** (minimal compliance to ensure accurate image preservation and rendering)

For PDF/A documents over 17 Mb, split into 10 Mb parts and submit as compound

object. **Recommended Technologies**

- Planetary scanners (overhead) with or without glass platens
- Digital cameras
- Pass through manual or automatically fed document scanners
- Flatbed scanners

Not Recommended Technologies

- Lighting systems that raise the surface temperature of the original more than 6 degrees F (3 degrees C) in the total imaging process. This can cause bowing and cupping of the document.

Item reference targets are optional. Targets used for DICE device calibration are recommended.

Cropping

Scanned preservation **images should contain a small border around the entire image (roughly 3 to 8 mm)**.

Oversize Materials: Maps, Posters, or Others

The following scanning and processing specifications are for documents beyond the size facilitated by scanners of legal size or less.

	Preservation Master		Access Derivative
	4 Star (recommended)	3 Star (minimum)	
Master File Format	TIFF, JPEG 2000, PDF/A	TIFF, JPEG 2000, PDF/A	-
Access File Format	-	-	PDF/A
Resolution	400 ppi	300 ppi	300 ppi
Bit Depth	16	8 or 16	8
Color Space Options	Grey Gamma 2.2, Adobe 1998, SRBG ProPhoto, ECIRGBv2	Grey Gamma 2.2, Adobe 1998, SRBG ProPhoto, ECIRGBv2	Grey Gamma 2.2, SRGB Adobe 1998, ProPhoto, ECIRGBv2
Color	Grayscale or Color	Grayscale or Color	Grayscale or Color

See FADGI Technical Guidelines (page 28) if assessing tolerance of measurement

parameters. **Recommended Technologies**

- Planetary scanners (overhead)
- Digital cameras
- Flatbed scanners
- Sheet fed scanners

Not Recommended Technologies

- Pass through style scanning systems that contact the recto or verso of the original in any way
- Lighting systems that raise the surface temperature of the original more than 4 degrees F (2 degrees C) in the total imaging process

NOTE: if item is larger than can be captured in one image, see 'stitching' (page 67) in FADGI Technical Guidelines.

Item reference targets are optional. Targets used for DICE device calibration are recommended.

Cropping

Scanned preservation **images should contain a small border around the entire image (roughly 3 to 8 mm)**.

Newspapers

The following scanning and processing specifications are for common newspaper formats.
For camp newspapers, see 'Unbound Documents (General Collections)'.

	Preservation Master		Access Derivative
	4 Star (recommended)	3 Star (minimum)	
Master File Format	TIFF, JPEG 2000, PDF/A	TIFF, JPEG 2000, PDF/A	-
Access File Format	-	-	PDF/A
Resolution	400 ppi	300 ppi	150 ppi or higher
Bit Depth	8	8	8
Color Space Options	Grey Gamma 2.2 SRGB	Grey Gamma 2.2 SRGB	Grey Gamma 2.2 SRGB
Color	Color	Grayscale or Color	Grayscale or Color

See [FADGI Technical Guidelines](#) (page 31) if assessing tolerance of measurement

parameters. **PDF/A levels**
- **PDF/A-1a** (OCR, tagged structure, tabbed reading order, alternative text), **OR**
- **PDF/A-1b** (minimal compliance to ensure accurate image preservation and rendering)

For PDF/A documents over 17 Mb, split into 10 Mb parts and submit as compound

object. **Recommended Technologies**
- Planetary scanners (overhead)
- Digital cameras
- Flatbed scanners
- Sheet fed scanners

Not Recommended Technologies
- None

Cropping

Scanned preservation **images should contain a small border around the entire image (roughly 3 to**

8 mm).

Capture Specifications - Audio

Preservation Masters

Similar to the revised edition of the Technical Guidelines for Digitizing Cultural Heritage Materials (2016), the FADGI Audio-Visual Working Group produced [Guidelines: Audio Digitization System Performance](#). These guidelines provide a framework with which to apply metrics regarding the performance of equipment used in the digitization of analog audio sources. The recommendations do require advanced equipment and training to apply. If you are seeking to upgrade your capacity for digitizing and preserving recorded audio, please consider consulting these guidelines as part of any planning stage.

Format

Uncompressed Audio Interchange File Format (AIFF) and Broadcast Wave (WAV) are both acceptable formats for audio preservation, however, **Broadcast Wav is preferred** for its ability to embed metadata.

Resolution

Born-digital audio residing on mobile digital storage and optical discs should be migrated and saved as Broadcast Wave (.wav) keeping their native resolution intact without resampling (i.e., digitally re-recording). Increasing the resolution of born-digital audio will not result in improved clarity. For example, audio compact discs store audio as CD-DA files with the resolution of 16-bit, 44.1 kHz. Your preservation copies will keep the native resolution of the original (16/44.1). Similarly, other born-digital audio files will retain their native resolution as preservation copies.

The following audio recording resolutions apply to capturing audio files from analog sources or digital sources for which the bitstream (digital to digital or byte for byte) copy cannot be made.

	Bit depth	Sample rate	Mode
Recommended	24-bit	96 kHz	Stereo or Mono (as in original)
Minimum	24-bit	48 kHz	Stereo or Mono (as in original)

Access Derivatives

Format/Resolution

Compress WAV file as MPEG-2 Audio Layer III (**MP3**), with the following

resolution:

	Bit depth	Sample rate	Mode
Recommended	16-bit	44.1 kHz	Stereo or Mono (as in original)

Embedded Metadata

For more information regarding embedding metadata in WAV files see FADGI [Guidelines: Embed](#) [Metadata in Broadcast WAV Files](#). Audio software such as Adobe Audition enable embedded metadata creation. Limited metadata can be added directly to files via right-clicking the file name and selecting 'Details' (Windows), or 'Get Info' (Mac).

Capture Specifications - Video

The following section describes specifications required for providing medium term preservation of video
(non-film) in analog and digital formats. The conservation and preservation of moving image film is beyond the scope of this document. Should an institution have moving image film sources in need of digital preservation and access use via a digital application, current practice indicates the film should be transferred to video to provide the source for digitization (Library of Congress, 2009).

Preservation Masters

(Analog capture and digital transcoding)

Digitized analog video is *captured* using analog-to-digital converters (IO hardware) and video editing software. Video transferred from one digital source to another (e.g., DVD to hard drive) is *transcoded*
(via software) into a preservation format (see below) to protect against the technological obsolescence
of video codecs that could complicate or prevent future playback. The following specifications apply to both capture and transcoding processes.

Recommended Specifications

The MXF Standard

The Materials Exchange Format (MXF) is an evolving standard for the long term preservation of audiovisual materials. The MXF file format provides a wrapper for all visual, audio, metadata, captions, timecode and other technical settings contained within a digital (or digitized) video. For more information on implementing MXF locally, see Guidelines: MXF Application Specification of the FADGI Audio-Visual Working Group.

Long-term preservation

Container:	**Media Exchange Format (.mxf)**
Codec:	**JPEG2000 lossless**
Resolution:	**640 x 480 px** (for original 4:3 aspect ratio),
	1280 x 720 px (for original HD/16:9 aspect
Sampling	ratio) 4:4:4
Scheme:	30-bit 30 fps
Sampling:	
Framerate:	Progressive 30 MBps
Scanning:	
Rate:	

NOTE: Some compatible software systems for use with the MXF format include: Avid Media Composer (Windows, OSX), Adobe Premiere Pro (Windows, OSX), Sony Vegas, Grass Valley Edius.

Alternative Minimum

Alternative minimum specifications for preserving audio-visual content are derived from [CDL Digital File](#) [Format Recommendations: Master Production Files](#) (pages 4 - 5).

Medium-term preservation

Container:	**Audio Video Interleave* (.avi) OR QuickTime File Format**
Codec:	**(.mov) H.264/MPEG-4 AVC** (compressed)
Resolution:	**640 x 480 px** (for original 4:3 aspect ratio),
	1280 x 720 px (for original HD/16:9 aspect ratio)
Sampling Scheme:	4:2:2
	30-bit
Sampling:	30 fps
Framerate:	Progressive
Scanning:	30 MBps
Rate:	

*Supports XMP

Access Derivatives

Container:	**MPEG – 4 Part 14 (.mp4)**
Codec:	**H.264/MPEG-4**
Resolution:	**640 x 480 px** (for original 4:3 aspect ratio), **720 x 405 px**
	(for original HD/16:9 aspect ratio) 30 fps
Framerate:	Progressive
Scanning:	30 MBps
Rate:	

Embedded metadata

RELATED RESOURCES: CDL-DFFR, Library of Congress Sustainability (H.264/MPEG-4 AVC), FADGI Audio Visual Working Group, ALCTS Minimum Digitization Capture

Video capture, editing and encoding software

MXF native systems:	Avid Media Composer (Windows, OSX) Adobe Premiere Pro (Windows, OSX) Sony Vegas (Windows) Edius (Windows)
Other proprietary:	Adobe Media Encoder

Roxio Media Capture/My DVD
Windows Movie Maker

Open Source:
Pitivi (Linux)
Ingex Studio (Linux)
libMXF (Windows, OSX, Linux)
Handbrake (Windows, OSX, Linux) – transcoding
FFMpeg (Windows, OSX, Linux) – transcoding via command line

Compound Objects

Compound objects are items with multiple parts that make a larger whole. For example, a postcard is a compound object of two digital image files: 1) a front side (recto) displaying an image, and 2) a back side (verso) displaying text. In this case, the two jpeg images (front and back) are described in one metadata record collectively known as a *compound object*. For adding metadata to compound objects see "MetadataGuidelines_CSUJAD.pdf".

For the CSUJAD project, you are most likely to encounter one or more of the following situations in which you will need to add an item as a compound object:

1) Postcards added as JPEGs (see above).
2) Print photographs added as JPEGs that contain markings or annotations on the back side of the photograph that merit their inclusion as a compound object.
3) A physical object containing multiple parts. For example, a sword might have multiple photographs of the sword and sheath.
4) An audio or video file with multiple parts, such as an oral history recording.
5) A PDF document larger than 17MB. If the PDF file was saved as "Adobe PDF Files, Optimized (.pdf)" and the resulting file size is still over 17MB, the file must be split into a multipart PDF and treated as a compound object (See "Splitting a PDF" below).

Splitting a PDF in Adobe Acrobat 11

1. Open PDF file
2. Click Tools
3. Click Pages/Split Document
4. In the prompt, select "File Size" button, and set Max MB to "10".
5. Click OK.
6. Multiple PDFs are automatically saved, appending the filenames with "Part_1", "Part_2", etc. in the directory containing the master PDF from which the multipart PDFs are derived.

Figure 1: Split document page in Adobe Acrobat XI
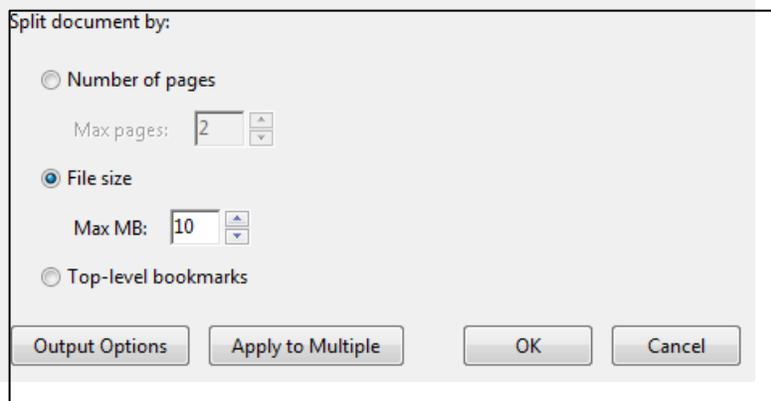
For additional information on Compound Objects, including items for which you want to preserve a document's hierarchy, such as a book, please see the downloadable PDF "Creating Compound Objects" from

OCLC.

Filenames

As referred by the FADGI Still Image Working Group, filenames should be:

- Unique – Not duplicated
- Well-defined – Logically structured, consistent and serviceable to meet existing and future needs
- Persistent – Will not change over time
- Self-referencing – Expresses location of (or relationship to) source material
- Technically observant – Uses legal characters and unbroken strings, etc.
- Sortable – Can be sorted against other files for review or indexing

    Local (i.e. institutional) filenames for objects contributed to the CSUJAD Project will likely follow logical structures that serve local needs. Therefore, a *Local ID* may be used to populate the Local ID field (and subsequent parts of a compound object) to serve the needs of the contributing library. The *Project ID* will be imposed as the filename on all items and serve as the primary unique identifier to provide consistency for users and staff alike.

    The objective of the Project Identifier is to identify:

    1.   Contributing Institution* (institutional acronym/abbreviation)

    2.   Archival Collection~   (Local collection code)

    3.   Number#   (Four digit sequence to identify object from same collection)

Example	Value/Code	Meaning
➢ Institution:	csun	California State University Northridge
➢ Collection:	esm	Eddie S. Muraoka Collection, 1942-1980
➢ Number (####):	0003	Third object added to the project for the ESM collection
➢ **Project ID=**	**csun_ems_0003**	

NOTES:

PLEASE USE UNDERCORES TO LINK CODES for ease of readability and to ensure URLs pointing to files can be used for future contexts.

*Institutional Codes

csub	CSU Bakersfield
csuci	CSU Channel Islands
csudh	CSU Dominguez
csueb	Hills CSU East Bay
csufu	CSU Fullerton
csufr	CSU Fresno
csulb	CSU Long Beach
csun	CSU Northridge

sac	Sacramento State
csusb	CSU San Bernardino
sdsu	San Diego State
sfsu	San Francisco State
slo	Cal Poly San Luis
ssu	Obispo Sonoma State
sjs	San Jose State
	University

Use the collection code assigned to your collections OR use an acronym of the collection (see example above). Whichever you choose, please be consistent.

Begin numbering for each item in a single archival collection with 101 (due to the retroactive numbering needed for the first NEH grant items) or if you anticipate there will be greater than 999 items added from any single collection, begin numbering with 1001.

Creative Commons Licensing

As a project dedicated to the education of the subject materials presented in the CSUJAD database, it is strongly encouraged that archives consider applying Creative Commons licenses to materials for which rights are fully possessed by the institution. By designating a license to use these materials under conditions chosen by the institution, the materials are more likely to be used or remixed by the public, free of charge, and circulate within other learning environments.

Creative Commons licenses are easily assigned. Visit the Creative Commons license chooser online to answer questions and learn about the rights you can give to users of these works fully owned by your institution. Once a license is chosen, text is supplied to copy and paste into your metadata under the 'Rights' field.