



DIGITAL PRESERVATION RESOURCE GUIDE

This document provides some useful online resources for the management and preservation of digital cultural heritage. The document was created as part of the digital preservation session held at the 2015 ATALM Post-conference Sustainable Heritage Workshop in Washington DC, on September 13.

* urls as of October, 2015.

FEDERAL AGENCIES DIGITIZATION GUIDELINES INITIATIVE

This website is a collaborative effort by federal agencies to define common guidelines, methods, and practices for digitizing historical content, specifically still image and audio visual materials.

<http://www.digitizationguidelines.gov/>

[Glossary](#) | [Participants](#) | [Contact Us](#)

Search this site [GO](#)

FADGI Federal Agencies Digitization Guidelines Initiative

[Your Comments](#)
Submit comments on guidelines or on our digitization efforts in general.

[Home](#) | [About](#) | [Guidelines](#) | [Resources](#) | [Review Process](#) | [News & Events](#)

About This Initiative

Started in 2007, this is a collaborative effort by federal agencies to define common guidelines, methods, and practices for digitizing historical content. As part of this, two working groups are studying issues specific to two major areas, Still Image and Audio Visual.

[Learn more about the initiative >](#)

Still Image Working Group

This group is involved in a cooperative effort to develop common digitization guidelines for still image materials.

Audio-Visual Working Group

The goal for this working group is to identify, establish, and disseminate information about standards and practices for the digital reformatting of audio-visual materials.

Digitization Guidelines

[Technical Guidelines for the Still Image Digitization of Cultural Heritage Materials](#)
Review draft for comment | Comments requested by December 1, 2015
Creation of Raster Image Master Files represents shared best practices followed by agencies participating in the Federal Agencies Digitization Guidelines Initiative (FADGI) Still Image Working Group for digitizing cultural heritage material.

[Digitizing Motion Picture Film: Exploration of the Issues and Sample SOW](#)
Review draft for comment | Comments requested by December 31, 2015
The FADGI Audio-Visual Working Group's report on motion picture film scanning combines an introductory essay, a set of tables that describe a range of film "inputs" and digital "outputs," and concludes with a model statement of work for outsourced conversion of film to video.

[MXF Application Specification for Moving Image and other Audio-Visual Content](#)
Review draft for comment | Comments requested by December 31, 2015
A revised version of the the MXF Application Specification for Archiving and Preservation, called AS-07 under the auspices of the Advanced Media Workflow Association (AMWA), is available for public comment. The detailed AS-07 file "wrapper" specification intended to serve reformatting programs includes a list of permitted encoded essences (the underlying content bitstreams), defines a means for the carriage of multiple timecodes; the handling of captions, subtitles, and Timed Text; a minimal core metadata set; program segmentation metadata; and embedded content integrity data.

[File Format Comparison Projects](#)
Posted by both Working Groups | December 2, 2014
Comparisons of file formats for *still-image reformatting, video reformatting, and preservation-oriented treatment of born digital video*. For the reformatting examples, matrices are used to compare formats in terms of about forty factors. For born digital video, current practices in six federal agencies are described in case histories. The offerings include narrative explanatory documents and supplementary reports.

[Print](#) [Subscribe](#) [Share/Save](#)

News & Events

- [Review draft](#) for Digitizing Motion Picture Film: Exploration of the Issues and Sample SOW; comments requested by December 31, 2015
- [Revised version](#) of MXF Application Specification AS-07. Document posted September 8, 2015

Recent blogs:

- [Let's Start at the Very Beginning: Guiding Principles for Creating Born Digital Video](#) (LC, February 2014)
- [AV Artifact Atlas: By the People, For the People](#) (LC, January 2014)
- [Can I Get a Sample of That? Digital File Format Samples and Test Sets](#) (LC, December 2013)
- [On the Road with FADGI: Recent Conference Presentations Highlight Current Audio and Video Projects](#) (LC, November 2013)
- [Connecting Communities: FADGI Still Image Working Group's Impact on the Library of Congress and Beyond](#) (LC, November 2013)
- [One Format Does Not Fit All: FADGI Audio-Visual Working Group's Diverse Approaches to Format Guidance](#) (LC, October 2013)

[See all News & Events >](#)

Resources

The following are provided as resources for the digitization of both audio-visual and still image digital content.

- [JPEG 2000 Lossy Compression](#) (PDF, 1.6 MB)
- [Digital Imaging Standards](#)
- [Evaluating Still Image Digitization and Digitization Equipment](#) (PDF, 2.5 MB)
- [Format Considerations in Audio Visual Preservation Reformatting](#) (PDF, 528 KB)

[View all Resources >](#)

LONG-TERM DIGITAL PRESERVATION REFERENCE MODEL

WHAT IS A DIGITAL ARCHIVE?

The purpose of this section of the LTDP site is to define the entry point for what constitutes the minimum requirements for an archive to be legitimately called a "digital archival system or service."

<http://www.ltdprm.org/reference-model/what-is-an-archive>

The screenshot shows the LTDP website interface. At the top left is the LTDP logo with the text 'Long-Term Digital Preservation Reference Model'. A search bar is at the top right. A navigation menu on the left lists various sections like 'Home', 'News Feeds & Announcements', 'Solutions Guide', 'Reference Model', and 'Reference Publications'. The main content area is titled 'Reference Model > What is a Digital Archive?'. It contains a sub-section 'Defining a "Level-1" Digital Archive System' with a paragraph explaining the purpose of the discussion. Below this are sections for 'What is not a Digital Archive: (Level '0')', 'Required Attributes for a "Level 1" Digital Archive', and 'Proposed "definition"'. A 'Reference Definitions:' section includes a definition for 'Capability Maturity Model' and a diagram showing five maturity levels: 1 Initial, 2 Managed, 3 Defined, 4 Quantitatively Managed, and 5 Optimizing. The diagram shows a progression from level 1 to level 5, with level 1 being 'Process unpredictable, poorly controlled and REACTIVE' and level 5 being 'Focus on CONTINUOUS process improvement'. A 'Digital Archive System' definition is also provided at the bottom right.

What is a Digital Archive?

Defining a "Level-1" Digital Archive System

Practices and processes can be defined in a few paragraphs, unfortunately systems cannot. Systems have a continuum of degrees of maturity, so it is often better to define them in the context of a maturity model. The purpose of this discussion is to define the entry point for what constitutes the minimum requirements for an archive to be legitimately called an "archival system or service." Since "Level-1" refers to the initial maturity state of the entire system, let's use "Level-1" to articulate the minimum system requirements to support the service.

What is not a Digital Archive: (Level '0')

- A specific storage system whether a tier of storage, a cloud-based storage repository, or a specialized storage device.
- A service that provides extensions to an organization's storage for the purposes of offloading (as in tiering) whether onsite, off-site, or at end-of-life as storage expansion
- Any device or practice vendors or IT calls "archiving", "database archive", or "archival storage" - from a "digital archive systems" perspective, these are not "Level 1" systems or services by themselves.

Reference Definitions:

- **Capability Maturity Model:** A maturity model can be viewed as a set of structured levels that describe how well the behaviors, practices and processes of an organization can reliably and sustainably produce required outcomes. - Source: Wikipedia
 - Typically described as a 5-level rating of "maturity" ranging from low maturity to high. However, a CMM is usually applied to the maturity of the organization, not the system.
 - Level "1" in this context is defined as the minimum capabilities required to classify a system, including the organization and its supporting practices as a digital archive system. Level-1 is analogous to an starting point and unlike a rating of '0' which implies no awareness or no capabilities.

Required Attributes for a "Level 1" Digital Archive

- A digital archive systems has three essential components: the digital records being retained, the organization that administers those records, and the infrastructure that supports the retention, protection, and accessibility of those records over their designated retention periods.
- The infrastructure that supports a digital archive needs to have these preservation-specific functions (the "service-catalog") at minimum:
 - Digital Auditing (using digital signatures) and verification on a periodic basis to assure record authenticity, digital integrity, and data protection
 - Physical and logical Migration capabilities
 - Infrastructure supporting Access to records, including Search/Discovery and re-use
 - Minimal data loss and means to audit, detect, and repair damage, corruption, or loss over extended periods of time.

Proposed "definition"

- An IT system and its supporting organization established for the digital preservation of information and data assets over extended periods of time, generally greater than 10 years. A digital archive is not a tier of storage or a special class of storage. Rather, what distinguishes a digital archive is that it is a policy-based system with content management, policy management, and preservation services that define and control the supporting ingestion database infrastructure storage

Digital Archive System

- Archives are long term repositories for the storage of records. Electronic archives preserve the content, prevent or track alterations and control access to

The diagram illustrates the Capability Maturity Model (CMM) with five levels of maturity, each represented by a blue box with a number and a description. The levels are arranged vertically, with level 1 at the bottom and level 5 at the top. A dashed blue arrow curves from level 1 up to level 5, indicating the progression. Level 1 is labeled 'Initial' and described as 'Process unpredictable, poorly controlled and REACTIVE'. Level 2 is 'Managed' and 'Process characterized for PROJECTS and is MANAGEABLE'. Level 3 is 'Defined' and 'Process characterized for the ORGANIZATION and is PROACTIVE'. Level 4 is 'Quantitatively Managed' and 'Process QUANTITATIVELY measured and controlled'. Level 5 is 'Optimizing' and 'Focus on CONTINUOUS process improvement'.

UNIVERSITY OF KENTUCKY LIBRARIES

LOUIE B. NUNN CENTER FOR ORAL HISTORY: DIGITAL AUDIO CALCULATOR

Digital audio can occupy an enormous data footprint. The Nunn Center has developed the Digital Audio Calculator for estimating the file size for your project or your digitization initiative. At this time, the calculator computes sizes for uncompressed audio.

https://libraries.uky.edu/libpage.php?lweb_id=856&llib_id=13

The screenshot shows the University of Kentucky Libraries website. At the top left is the UK logo. A search bar is at the top right. The main navigation bar includes 'UNIVERSITY OF KENTUCKY Libraries' and links for 'Services', 'Campus Libraries', 'Interlibrary Loan', 'Research Guides', 'Hours', 'My Library Account', 'Off-Campus Access', and 'Ask Us'. Below this is a 'Selected Databases' dropdown menu. The main content area features a 'Special Collections Library' sidebar with a list of links including 'Research Services', 'Digital Library Services', 'Learning Lab', 'Exhibits', 'Teaching with Primary Sources', 'Donations', 'News & Information', 'History', 'Projects and Grants', 'Directions', 'Contact Us', and 'SCRC Jobs'. The central focus is the 'Louie B. Nunn Center for Oral History: Digital Audio Calculator'. It includes a description of the tool, a 'Return Home' link, and a form with the following fields: 'Format' (radio buttons for Uncompressed Audio (.wav or .aiff)), 'Time' (input field with '1' and 'hour(s)'), 'Bit Depth' (radio buttons for 8, 16, 24), 'Sample Rate' (radio buttons for 44.1 kHz, 48 kHz, 96 kHz, 192 kHz), and 'Channels' (radio buttons for Mono, Stereo). The 'Data Size' field shows '635.04' with unit options (KB, MB, GB, TB). 'Compute' and 'Reset' buttons are at the bottom of the form. A footer contains four columns: 'RESEARCH HELP' (Ask Us for Research Help), 'COURSE RELATED' (Course Guides), 'LIBRARY INFORMATION' (About UK Libraries), and 'LIBRARY RESOURCES' (Book Purchase Request).

VIDEO SPACE CALCULATOR

This tool is intended to give an indication of the amount of space a given video format will take up on disk. The actual space taken up may differ slightly due to embedded audio, differing frame sizes and aspect ratios, and inter-frame compression / pulldown.

<https://www.digitalrebellion.com/webapps/vidoealc>

Video Space Calculator

This tool is intended to give an indication of the amount of space a given video format will take up on disk. The actual space taken up may differ slightly due to embedded audio, differing frame sizes and aspect ratios, and inter-frame compression / pulldown.

Format ProRes 422 HQ 1080

Resolution 1920x1080

Frame rate 30

Video length 1 hours

Total space: 96.78 GB

Video Space Calculator

This tool is intended to give an indication of the amount of space a given video format will take up on disk. The actual space taken up may differ slightly due to embedded audio, differing frame sizes and aspect ratios, and inter-frame compression / pulldown.

Format ProRes 422 HQ 1080

Resolution 1920x1080

Frame rate 30

Video length 40 hours

Total space: 3.78 TB

AUDIO FILE SIZE CALCULATIONS

These calculations will help you to estimate the size of audio files.

<http://www.audiomountain.com/tech/audio-file-size.html>

AudioMountain.com
Audio, Music, Tools & Information

[Home](#) | [Alma Fest](#) | [Music](#) | [News](#) | [Tech](#) | [Tools](#) | [Resources](#)







Audio File Size Calculations

These calculations will help you to estimate the size of audio files.

NOTES:
 1) "x" means "multiplied by". Forward slash (/) means "divided by".
 2) Kbps means "Kilobits per second" (1,000 bits per second). KB means KiloBytes (1,000 Bytes). There are 8 bits in a byte. Note the uppercase "B" for bytes in "KB". A lowercase "b" (Kb) would indicate bits.
 3) Calculations for MP3 files generally include both left and right stereo channels. Calculations for PCM files must be multiplied by 2 to allow for both left and right stereo channels.
 4) We took reasonable care in computing these figures, but you may want to doublecheck them if you're working on an important project. We did a casual experiment with a one-minute PCM file, and Windows Explorer reported it to be slightly smaller (10,337 KB) than what was predicted by the calculations below (10,584 KB). This may be due to differences between binary and decimal counting systems (multiples of 2 versus multiples of 10). In addition, file header information and ID3 tags (for MP3s) will have an effect on the numbers. In general, however, we have found the figures below to be useful for approximating audio file sizes.

MP3 File Size Calculations

Formula:
 Kbps = bits per second / 8 = Bytes per second x 60 seconds = Bytes per minute x 60 minutes = Bytes per hour

Bitrate	File size per second	File size per minute	File size per hour
8 Kbps	1 KB	60 KB	3.6 MB
16 Kbps	2 KB	120 KB	7.2 MB
32 Kbps	4 KB	240 KB	14.4 MB
40 Kbps	5 KB	300 KB	18.0 MB
48 Kbps	6 KB	360 KB	21.6 MB
56 Kbps	7 KB	420 KB	25.2 MB
64 Kbps	8 KB	480 KB	28.8 MB
80 Kbps	10 KB	600 KB	36.0 MB
96 Kbps	12 KB	720 KB	43.2 MB
112 Kbps	14 KB	840 KB	50.4 MB
128 Kbps	16 KB	960 KB	57.6 MB
160 Kbps	20 KB	1.20 MB	72.0 MB
192 Kbps	24 KB	1.44 MB	86.4 MB
224 Kbps	28 KB	1.68 MB	100.8 MB
256 Kbps	32 KB	1.92 MB	115.2 MB
320 Kbps	40 KB	2.40 MB	144.0 MB

PCM File Size Calculations

Here are some file size calculations for common PCM audio settings. PCM stands for Pulse Code Modulation and commonly uses the file extensions .wav or .cda. There are quite a few other combinations of bits per sample and samples per second which may be used. We have included calculations for the most common mono (one channel) and stereo (two channel) settings.

Mono

Formula:
 Bits per sample x samples per second = bits per second / 8 = Bytes per second x 60 seconds = Bytes per minute x 60 minutes = Bytes per hour of mono.

Settings	Bitrate	File size per second	File size per minute	File size per hour
16 bit, 44.1 KHz	705.6 Kbps	88.2 KB	5.292 MB	317.52 MB
16 bit, 48 KHz	768 Kbps	96 KB	5.750 MB	345.60 MB
24 bit, 48KHz	1,152 Kbps	144 KB	8.640 MB	518.40 MB
24 bit, 96KHz	2,304 Kbps	288 KB	17.280 MB	1.0368 GB

Stereo

Formula:
 Bits per sample x samples per second = bits per second x 2 channels = bits per second of stereo / 8 = Bytes per second of stereo x 60 seconds = Bytes per minute of stereo x 60 minutes = Bytes per hour of stereo.

Settings	Bitrate	File size per second	File size per minute	File size per hour
16 bit, 44.1 KHz	1,411.2 Kbps	176.4 KB	10.584 MB	635.04 MB
16 bit, 48 KHz	1,536 Kbps	192 KB	11.520 MB	691.2 MB
24 bit, 48KHz	2,304 Kbps	288 KB	17.28 MB	1.036 GB
24 bit, 96KHz	4,608 Kbps	576 KB	34.56 MB	2.0736 GB

[Calculations](#)