



**Federal Agencies
Digitization Guidelines Initiative**

Creating and Archiving Born Digital Video

Part III. High Level Recommended Practices

December 2, 2014

The FADGI Audio-Visual Working Group
<http://www.digitizationguidelines.gov/audio-visual/>

Creating and Archiving Born Digital Video III: High Level Recommended Practices

By the Federal Agencies Digitization Guidelines Initiative Audio-Visual Working Group
<http://www.digitizationguidelines.gov/audio-visual/>

Version 1.1, December 2, 2014

TABLE OF CONTENTS

Introduction	4
<i>What is this Document?</i>	4
<i>Key to Case History References</i>	4
<i>Understanding the Recommended Practices</i>	4
Part 1. Advice for File Creators	6
<i>Plan for High Quality Video Files and Metadata</i>	6
RP 1.1 Select a camera and other recording equipment with the capability to capture at high quality levels...6	
RP 1.2 Provide the means to collect and submit metadata starting at the video shoot.....6	
RP 1.3 Capture video data to stable storage devices that allow for streamlined file transfer into managed storage.....7	
<i>Create the Highest Quality Video Files You Can Afford to Make and Maintain</i>	7
RP 1.4 Select High Definition (HD) video encoding over Standard Definition (SD).....7	
RP 1.5 Select larger picture sizes over smaller picture sizes.....8	
RP 1.6 Select higher bit rates over lower bit rates.....8	
RP 1.7 Select higher bit depths over lower bit depths.....9	
RP 1.8 Use higher chroma subsampling ratios rather than lower.....10	
RP 1.9 Generate a high integrity and continuous master timecode.....10	
RP 1.10 Stay within the range of common frame rates of 24-30 frames per second (fps).....11	
Part 2. Advice for File Archivists	12
<i>Document Provenance and Relationships</i>	12
RP 2.1 Document the original order, especially camera-created file structures.....12	
RP 2.2 Document relationships between the video object and other files, such as closed captions, scripts, location notes and other supplemental material.....12	
RP 2.3 Identify the file characteristics at the most granular level possible, including the wrapper and video stream encoding.....13	
<i>Understand the Impacts of Changing the Technical Characteristics</i>	13
RP 2.4 Develop selection criteria based on business needs to inform decisions on what files and/or formats to keep, especially if the same content is submitted in multiple video files.....13	

RP 2.5 Determine and document criteria for when (if ever) it is appropriate to change the video file's technical properties (including normalization)	14
RP 2.6 Retain the original video file as submitted if transcoding, normalizing or otherwise changing the video stream to meet business needs	14
RP 2.7 Select appropriate technical characteristics for the video encoding if transcoding, normalizing or otherwise changing the video stream to meet business needs.....	15
RP 2.8 Generate a new high integrity and continuous master timecode, especially if there is no timecode in the original material	15
RP 2.9 Retain original timecode(s) if provided, even if you generate a new high integrity continuous master timecode.....	15
RP 2.10 Retain all the data from the original file if the video file structure is changed	15
RP 2.11 Retain the original chroma subsampling if the video data is transcoded	16
RP 2.12 Retain original frame rates if the video data is transcoded, even when they are beyond the standard 24 - 30 fps.....	16
<i>Use Stable and Managed Digital Storage.....</i>	<i>17</i>
RP 2.13 Move video files off removable data carriers to stable and managed storage as soon as possible	17
Part 3. Advice For File Creators and File Archivists	18
<i>Create and Use Metadata to Facilitate Life Cycle Management.....</i>	<i>18</i>
RP 3.1 Use metadata data models with strong support for digital video	18
RP 3.2 Document and use technical metadata	18
<i>Selecting File Formats (Wrappers, Containers and/or Encodings)</i>	<i>19</i>
RP 3.3 Select uncompressed video encoding over compressed encoding	20
RP 3.4 If compression is used, select mathematically lossless compression over visually lossless or lossy compression	21
RP 3.5 Avoid multiple compressions and decompressions steps	22
RP 3.6 Stay within the same codec family if the video data is transcoded.....	23
RP 3.7 Select video encoding and wrapper formats that are well-supported now and future focused	24
RP 3.8 Select video encoding and wrapper formats that are non-proprietary	24
RP 3.9 Select video encoding and wrapper formats that are supported by downstream applications	24
RP 3.10 Select video formats that are standardized and well-documented	25
RP 3.11 Select video formats with capacity for robust and detailed technical metadata.....	26
RP 3.12 Select video formats with greater capacity for embedded metadata over less metadata capacity	26
RP 3.13 Select formats that can contain and label complex audio configurations including multiple channels and sound fields beyond mono and stereo	26
RP 3.14 Select formats that can support robust timecode data.....	27
<i>Plan for Access</i>	<i>28</i>
RP 3.15 Create access, viewing or proxy copies with appropriate technical characteristics to meet expected use cases	28

INTRODUCTION

WHAT IS THIS DOCUMENT?

This is one of four documents examining aspects of the current practice for creating and archiving born digital video at selected institutional members of the Federal Agencies Digitization Guidelines Initiative Audio-Visual Working Group. The three companion documents are:

- *Creating and Archiving Born Digital Video I: Introduction (Version 1.1, 12/2/14)*
- *Creating and Archiving Born Digital Video II: Eight Federal Case Histories (Version 1.1, 12/2/14)*
- *Creating and Archiving Born Digital Video IV: Resource Guide (Version 1.1, 12/2/14)*¹

This document outlines a set of high level Recommended Practices (RP) for creating and archiving born digital video. Each RP includes the rationale which explains why the FADGI members endorse this practice as well as how the RP is reflected in the accompanying case histories. Some RPs also include examples or other notes.

KEY TO CASE HISTORY REFERENCES

- **LC-AFC-CRHP:** Library of Congress American Folklife Center Civil Rights History Project
- **LC-NAVCC-VEF:** Library of Congress Packard Campus of the National Audio-Visual Conservation Center Video Evergreen Format
- **LC-WebArch-YouTube:** Library of Congress Web Archiving YouTube Harvesting
- **NARA-BRCC:** National Archives and Records Administration Base Realignment and Closure Commissions project
- **NOAA-OkEx:** National Oceanic and Atmospheric Administration *Okeanus Explorer*
- **SIA-DVD:** Smithsonian Institution Archives Authored DVD project
- **SI-DAMS:** Smithsonian Institution Digital Asset Management System
- **VOA-MMAM:** Voice of America Metadata for Media Asset Management

UNDERSTANDING THE RECOMMENDED PRACTICES

As discussed in greater detail in *Creating and Archiving Born Digital Video I: Introduction*, the Recommended Practices are intentionally high level and not intended to be comprehensive. They reflect the range of choices encountered by the eight case history projects but do not cover every issue that other projects might encounter when creating or archiving born digital video. The Recommended Practices aim to highlight the advantages of selecting one option over another when choices are available. They are also tightly scoped to issues pertinent to creating and archiving born digital video. Concerns common to digital preservation as a whole, such as consistent file naming protocols or repository actions, are not addressed.

The Recommended Practices are organized into three groups:

- Advice for File Creators,
- Advice for File Archivists, and
- Advice for File Creators and File Archivists

¹ The URLs for the three documents are:

(II) http://www.digitizationguidelines.gov/guidelines/FADGI_BDV_p1_20141202.pdf

(III) http://www.digitizationguidelines.gov/guidelines/FADGI_BDV_p2_20141202.pdf

(IV) http://www.digitizationguidelines.gov/guidelines/FADGI_BDV_p4_20141202.pdf

Advice for File Creators, also known as “advice for shooters,” focuses on providing video content producers, including videographers and, by extension, the project managers within cultural heritage institutions who are responsible for the creation new born digital video files, with a set of practices that emphasize the benefits of aiming for high quality and planning for archival repository ingest from the point of file creation. Advice for File Archivists seeks to provide guidance about video-specific issues which come into play when ingesting the files into a managed storage repository. Advice for File Creators and File Archivists are grouped together because they transcend specific life cycle points. This guidance focuses on selecting sustainable encodings and wrappers whether at initial file creation or during normalization upon ingest.

PART 1. ADVICE FOR FILE CREATORS

The goal of the Recommended Practices for File Creators is to emphasize the benefits of planning for archival repository ingest from the point of file creation. These Recommended Practices are aimed not just at videographers and camera operators, but also at the project managers, archivists, metadata specialists and technologists who develop Statements of Work and oversee project plans.

Born digital video files should be the highest quality that the institution can afford to make and maintain over the long term. While current limitations are important to understand and acknowledge, file creators should look beyond present constraints towards future, and most likely more advanced, capabilities. High quality, data-rich files stored in standardized and well-used formats alongside robust metadata will allow for flexibility over the long term, including after ingest into a managed storage repository.

PLAN FOR HIGH QUALITY VIDEO FILES AND METADATA

Project planning should include capabilities to create high quality digital video files and metadata from the outset. Select capture equipment such as cameras and microphones that not only meet the immediate needs of the project at hand but also could be reused for other projects later on. Since current technology constraints are likely to be relatively short-lived in many cases, look for equipment that has expanded capabilities for improved image and data capture. Set up the workflow processes so that the outcomes (often large video files and XML-based metadata) are easily integrated into appropriate systems and structures without excessive additional interaction. These RPs focus on setting up the framework to facilitate the creation of high quality digital video files and metadata.

RP 1.1 Select a camera and other recording equipment with the capability to capture at high quality levels	
Rationale	While the cost of high quality cameras continues to come down, they are still a significant financial investment so it pays to get a camera that can be useful beyond the project at hand. . Avoid cameras or other recording equipment that only satisfy the narrow current need. Instead, think about possible future requirements and leave your technical capture options more open and flexible. Select a camera that is "future-looking" with options for capturing higher quality image data at higher resolutions and bit depths, with more encoding and wrapper format options, more internal storage, etc. You may not use all the functionality now but you'll be ready for more complex projects in the future.
Case History Details	<ul style="list-style-type: none"> • LC-AFC-CRHP: All capture decisions are a balance of technical capacity in the digital capture equipment and the project goals for quality levels. It is best to start with a camera or capture device such as the Sony XDCam EX-1 used for this project that errs on the side of technical complexity, rather than be limited in capture specifications due to a device that is too basic for the job. • NOAA-OkEx uses a custom built camera for submersible filming, capable of supporting 1080i.

RP 1.2 Provide the means to collect and submit metadata starting at the video shoot	
Rationale	Provide a sustainable way to collect and submit the metadata from the very beginning of the life cycle of the file - the video shoot. In some cases, this can be as simple as a spreadsheet but in other cases, especially when the project is more complex and multiple people and systems are interacting with the data, this mechanism should be scaled appropriately. It can be difficult, time-consuming and error-prone to switch data collection systems once the project is underway. Take the time to map out a metadata collection and submission plan.

Case History Details	<ul style="list-style-type: none"> • LC-AFC-CRHP uses a collaborative cataloging application to collect and submit metadata records in the field in real time. • NOAA-OkEx: The OER data management team generates metadata at the end of each cruise. • VOA-MMAM uses its custom built AMF (Asset Manager Form) with fields tailored for each asset type.
-----------------------------	---

RP 1.3 Capture video data to stable storage devices that allow for streamlined file transfer into managed storage	
Rationale	<p>Data kept on removable storage devices such as videotape (including BetaCam, VHS, etc.) and optical discs (DVDs) is highly vulnerable because 1) these formats are at risk of technical obsolescence due to the specific combinations of hardware and software required to access the data, 2) the physical data carriers themselves are often unstable and prone to condition issues that negatively impact data access, and 3) these data carriers may not have strong data protection mechanisms that prohibit accidental loss of data by over-writing. Moreover, it can be time and labor intensive to migrate large complex video files from one data carrier to another.</p> <p>Evaluate the longevity, capacity, viability, obsolescence and susceptibility of the camera's native data storage and design the video capture workflow so that the video data is as secure as possible until it is ingested into a managed storage system that provides stability through redundancy, scheduled audits, and hardware replacement.</p>
Case History Details	<ul style="list-style-type: none"> • LC-AFC-CRHP interviews are captured via the camera directly to an external hard disk (the AJA KiPro) at the point of recording. • NOAA-OkEx video is captured directly to spinning disk on an EVS instant replay system. Video is then clipped out and saved to a large Nexsan SAN array.

CREATE THE HIGHEST QUALITY VIDEO FILES YOU CAN AFFORD TO MAKE AND MAINTAIN

High quality video typically means video with robust amounts of visual information such as a large picture size, a wide range for storing color information, high bit rate and continuous timecode stored in a well-defined and flexible format. Higher quality files often equates to big files. These large data rich files can also be costly to create, store and maintain so compromises are often needed on some parameters. Strive to capture and maintain the maximum amount of picture and sound data which your systems can support. The RPs in this section define some of the technical parameters that improve the quality of digital video files.

RP 1.4 Select High Definition (HD) video encoding over Standard Definition (SD)			
Rationale	<p>HD formats provide options to carry more visual information than SD formats including larger picture or frame size, more flexible and higher frame rates as well as options for scanning methods (interlaced and progressive).</p>		
Case History Details	<table border="1" style="width: 100%;"> <tr> <td data-bbox="381 1707 727 1841"> Project Goals Permit Following Recommended Practice </td> <td data-bbox="727 1707 1443 1841"> LC-AFC-CRHP captures in HD (1080/60i). </td> </tr> </table>	Project Goals Permit Following Recommended Practice	LC-AFC-CRHP captures in HD (1080/60i).
Project Goals Permit Following Recommended Practice	LC-AFC-CRHP captures in HD (1080/60i).		

	Project Goals Preclude Following Recommended Practice	VOA-MMAM, following the house style, captures in SD.
	Note	LC-WebArch-YouTube has an interesting perspective on this Recommended Practice which is worth mentioning even though the LC-WebArch-YouTube case study does not address file creation: Currently the LC Web Archiving Team has opted to collect the high definition version of nominated YouTube videos. However, given the volume of video content that the LC Web Archiving Team collects and the storage/processing limitations that they face, there is an argument to be made for why a sizeable Web Archive would specifically choose not to pursue HD quality when an SD option is available.

RP 1.5 Select larger picture sizes over smaller picture sizes		
Rationale	Larger picture sizes or frame sizes have more lines in the vertical and more pixels in the horizontal display resolution to carry detailed visual information. The more lines per frame, the higher the vertical image resolution. The more pixels per line, the higher the horizontal image resolution.	
Case History Details	Project Goals Permit Following Recommended Practice	LC-AFC-CRHP and NOAA-OkEx use 1920 x 1080.
	Project Goals Preclude Following Recommended Practice	VOA-MMAM, following the house style, uses 720 x 480.

RP 1.6 Select higher bit rates over lower bit rates	
Rationale	Higher bit rates, the number of bits that are conveyed or processed per unit of time and often expressed in kilobits per second (kbit/s or kbps, 10 to the third power), megabits per second (Mbit/s or Mbps, 10 to the sixth power), or gigabits per second, allow for the capture of more information and higher spatial resolution. Higher bit rates result in larger file sizes.
Example	50 Mbps is preferred over 8 Mbps.

Case History Details	Project Goals Permit Following Recommended Practice	<ul style="list-style-type: none"> • LC-AFC-CRHP data rates almost uniformly achieve 220 Mbps for each segment or video file. • NOAA-OkEx video is captured at 145 Mbps in ProRes 422. The ProRes family² includes a higher bit rate codec, ProRes 422 HQ, as well as lower bit rate codecs in ProRes LT and ProRes Proxy. ProRes 422 and its 145 Mbps target bit rate meets the need for gathering the highest quality video in limited storage.
	Project Goals Preclude Following Recommended Practice	VOA-MMAM's internal house standard is 25 Mbps, which meets VOA-MMAM's business needs and is supported by essential internal VOA-MMAM systems.

RP 1.7 Select higher bit depths over lower bit depths	
Rationale	<p>The number of bits used per sample determines how much intensity variation is possible within the signal. In 8-bit video, shades of gray or a color is represented on through an active range of 16 (darkest value) to 235 (white/lightest value). This active range extends from 64 to 940 for 10-bit video and 264 to 3760 for 12-bit video. Values outside the active range (1–15 and 236–254 for 8-bit video, 1-63 and 941-1023 for 10-bit video and 0-263 and 3761- 4096 for 12-bit video) are retained for special functions. The key element is that higher bit depths permit a wider range of possible gradations which means much more subtle variations in intensity can be recorded.</p> <p>Higher bit depths allow for the capture of more color information per pixel, but result in larger file sizes. Lower bit depths can contribute to exhibit banding in saturated colors but they may be an acceptable choice if there's limited color saturation or effects.</p> <p>Many digitizing systems offer the option of recording either 8 or 10 bits per sample. FADGI generally encourages the use of 10-bit sampling for the sake of higher image quality. Some archives use 8-bit sampling for certain classes of material in order to keep file sizes low. However, with 8-bit sampling, there is greater risk that imagery will show abrupt changes between shades of the same color. Image elements that feature natural gradients like blue skies or areas of (seemingly) solid tonality can show what is called banding or contouring. In these cases, not every change in the continuous gradient can be shown because there are insufficient bits to represent all of the shades. The risk of banding is reduced by increasing the number of bits per sample.</p>
Example	10 bits per sample is preferred over 8 bits per sample.
Case History Details	<ul style="list-style-type: none"> • LC-AFC-CRHP uses 10-bit. • LC-NAVCC-VEF currently uses 10-bit for video because this matches the specifications laid out for serial digital interface as defined by SMPTE starting with ST 259M (although NAVCC is preparing to accommodate native bit depths beyond 10-bit). Additionally, 10-bit encoding is preferred over 8-bit as a harmonization encoding so that decoder software writers do not have to accommodate both.

² <http://www.digitalpreservation.gov/formats/fdd/fdd000389.shtml>

RP 1.8 Use higher chroma subsampling ratios rather than lower		
Rationale	<p>The signal representation of color in video is called chrominance, or chroma for short. Chroma subsampling describes the use of lower spatial and/or temporal sampling frequencies for color information than for brightness (luma) information. Chroma subsampling (or color compression) is rather common due to the fact that the eye is less sensitive to color detail than to brightness detail. However, while chroma subsampling results in reduced file sizes, color detail and reproduction are also permanently reduced.</p> <p>The widely-used 4:2:2 subsampling provides higher image quality than 4:2:0 and 4:1:1 options. In addition to inherently better initial image quality, 4:2:2 also provides benefits if material is reformatted over time, in what professional broadcasters sometimes call a <i>cascading</i> scenario. A <i>cascade</i> may be encountered in a chain of connected broadcast elements with the same risks of quality loss as in a cascade over time. The 4:4:4 ratio offers even higher quality but practical considerations in terms of available equipment and interfaces generally preclude its use in video reformatting. The use of 4:4:4 also produces significantly larger files.</p>	
Example	4:2:2 is preferred over 4:2:0 and 4:2:0 is preferred over 4:1:1.	
Case History Details	Project Goals Permit Following Recommended Practice	LC-AFC-CRHP and NOAA-OkEx projects capture in 4:2:2 which is supported by ProRes.
	Project Goals Preclude Following Recommended Practice	VOA-MMAM's internal house standard is 4:1:1 which meets VOA's business needs and is supported by essential internal systems.

RP 1.9 Generate a high integrity and continuous master timecode	
Rationale	<p>Timecode values are important to the provenance and discovery of the video object. Some files may contain more than one timecode stream. At a minimum, capture a continuous master timecode stream to serve as the canonical point of reference for all timecode-dependent activities.</p> <p>If a time-of-day timecode is required, it is strongly advised that you capture a continuous timecode stream as well.</p>
Case History Details	<ul style="list-style-type: none"> • LC-AFC-CRHP creates SMPTE timecode. • NOAA-OkEx: Captured video uses SMPTE timecode (Control Clock set to UTC).

RP 1.10 Stay within the range of common frame rates of 24-30 frames per second (fps)	
Rationale	Frame rates within the usual range of 24-30 frames per second are appropriate for non-theatrical productions. Higher frame rates, such as 48, 60 fps and more, may result in smoother transitions for high movement shots but will also significantly increase the file size. It is unlikely at this point that these higher frame rates would be helpful or necessary for non-theatrical, non-broadcast files.
Case History Details	LC-AFC-CRHP, NOAA-OkEx and VOA-MMAM projects all use 29.97 fps.

PART 2. ADVICE FOR FILE ARCHIVISTS

Advice for File Archivists seeks to provide guidance about video-specific issues which come into play when ingesting the files into a managed storage repository. These Recommended Practices are aimed at digital preservationists including archivists, librarians, digital asset managers, and other staff within cultural heritage institutions that receive born digital video from creators and inherit the responsibility of describing, preserving, and providing access to those files. The goal of these Recommended Practices is to provide guidance about video-specific issues that come into play when ingesting the files into a managed storage repository. These include retaining essential camera-created data and making changes to acquired files (e.g., migrate or normalize file formats). The Recommended Practices offer recommendations on safe storage practices, choosing sustainable formats, embedding metadata, and configuring workflows for systems interoperability to allow File Archivists to plan for future migration and support.

DOCUMENT PROVENANCE AND RELATIONSHIPS

One of the most important functions of archival repositories is to document their holdings. The RPs in this section highlight the importance of documenting the context and provenance of the digital video file, including its technical structure, history of change actions and relationships to other objects.

RP 2.1 Document the original order, especially camera-created file structures	
Rationale	In some collections, the original order provides essential information about the individual files and the relationships among files. This especially is relevant for nested folder structures created by the camera/capture device which contain the video files. While it may not be possible to retain the original order in a processing or repository environment, it's important to harvest and retain metadata about these relationships.
Case History Details	<ul style="list-style-type: none">• SIA-DVD follows the VOB order when creating a MPEG-2 access file.• SI-DAMS: NAA archivists using the DAMS have created a folder structure to replace (not entirely replicate) the original file directory structure from within the SI-DAMS interface and files will be placed in this structure for easy retrieval and to retain original order.

RP 2.2 Document relationships between the video object and other files, such as closed captions, scripts, location notes and other supplemental material	
Rationale	Born digital video files are often accompanied by supplemental material such as closed captions, scripts, and location notes. Some, including the closed captions, are closely tied to the timeline of the video object and need to be closely aligned.
Case History Details	The SI-DAMS allows for relating files to each other (effectively linking records together in the database). The files are linked to each other with parent-child links (derivatives, text files, etc.), and relate-to links (in the event there are multiple video files for one shoot).

RP 2.3 Identify the file characteristics at the most granular level possible, including the wrapper and video stream encoding	
Rationale	<p>Just knowing the file extension (e.g. *.mov) is not enough, especially for digital video objects because they contain a mix of audio and video essence encodings, technical and embedded metadata and other data structures all contained within a wrapper. File identification tools can also aid in discoverability when evaluating stability of codecs and wrappers for migration and normalization planning.</p> <p>FOURCC (four character) codes may be helpful in accurately identifying video codecs. A FOURCC code is a sequence of four bytes, usually four concatenated ASCII characters, used to uniquely identify data formats. For example, the FOURCC code for YUY2 video is 'YUY2'. The FOURCC codes can be used efficiently in program code as integers as well as providing human reading cues in binary data streams when inspected.</p>
Example	<p>Example tools: FITS, JHOVE, MediaInfo. See Resource Guide.</p> <p>Example characteristics: Wrapper, encoding, audio bit rate, video bit rate, frame rate, chroma subsampling, etc.</p>
Case History Details	LC-AFC-CRHP, LC-NAVCC-VEF, NARA-BRCC, SIA-DVD and SI-DAMS all use various format identification tools.

UNDERSTAND THE IMPACTS OF CHANGING THE TECHNICAL CHARACTERISTICS

Business needs, “house rules,” technical infrastructure or staff can require that files be transcoded (change the encoding) and/or transwrapped (change the wrapper or container) to facilitate normalization or format migration for long term archiving or access. It's essential in an archival environment to understand why changes to the technical characteristics of the file are needed and the impacts of these changes on the data. Equally as important is to document all the changes to order to document provenance.

RP 2.4 Develop selection criteria based on business needs to inform decisions on what files and/or formats to keep, especially if the same content is submitted in multiple video files		
Rationale	Archival repositories may get the same content in different video formats. A common set for one title might be a high quality archival master, a digital intermediate or proxy and multiple system-specific access copies. Some archives may choose to keep all variations but others may need to make informed curatorial decisions to save on storage space or other restrictions. A good rule of thumb is to keep at least the highest quality, most data rich, file(s). It may not be prudent to keep lower quality viewing copies if they can be regenerated from the higher quality files.	
Case History Details	Project Goals Permit Following Recommended Practice	SI-DAMS retains the original, master files MXF video files as well as the derivative MOV files (the latter for easy retrieval by the producers). Other files including F4V, XML, CIF and SIF are not retained.
	Project Goals Preclude Following Recommended	As demonstrated in the LC-WebArch-YouTube Case History, the practice of following strict and comprehensive format and

	Practice	file quality selection is generally counterintuitive given the context of web archiving. Selection criteria for web archiving activities focuses on content, and the nature of the internet implies that we will encounter, and must embrace, a broad diversity of formats.
--	----------	---

RP 2.5 Determine and document criteria for when (if ever) it is appropriate to change the video file's technical properties (including normalization)	
Rationale	Consider whether or not a need for transcoding the original format exists, including: is the file proprietary, unsupported format? Is the strategy at your archive to normalize media files to one or more specific sets of characteristics? Is the file "at risk" in its current form? Can you isolate the risk to the encoding or wrapper so you know that you need to transcode, transwrap or both? While it may not be possible to create hard and fast rules, documented the criteria will be helpful in the decision-making process to help assure consistency and continuity.
Example	The NDSA "Levels of Preservation" may be useful in developing these criteria. See Resource Guide.
Case History Details	<ul style="list-style-type: none"> • NARA-BRCC and LC-NAVCC-VEF case histories detail the rationale for normalizing file types. • SI DAMS: Upon ingest, a preview copy of each file is made with the DAMS transcoders. Only major formats are supported, allowing us to target and troubleshoot proprietary and/or obsolete formats, and then make decisions about moving forward with them. Files in this case study are supported files and can be ingested to the DAMS with preview files made. We do not normalize and we do not currently migrate at the point of acquisition/ingest. • SIA-DVD case history documents the rationale for extracting video files off DVDs because it's not a long-term storage option for preservation or access.

RP 2.6 Retain the original video file as submitted if transcoding, normalizing or otherwise changing the video stream to meet business needs	
Rationale	It is sound archival practice to retain digital objects as they were submitted to the archive, even if the archive cannot provide functional access to the file at the current time. This is especially important if the content is transcoded or transwrapped or otherwise altered by necessary processing. Future uses of the file can't be known so it's best to retain the original object for the highest degree of flexibility.
Case History Details	<ul style="list-style-type: none"> • NARA-BRCC and LC-NAVCC-VEF retain the originals as submitted. • SIA-DVD retains the submitted original DVDs and creates ISO preservation masters as part of its best practices

RP 2.7 Select appropriate technical characteristics for the video encoding if transcoding, normalizing or otherwise changing the video stream to meet business needs	
Rationale	Some technical characteristics should migrate unchanged from the original material such as bit depth, chroma sampling, frame rate, and color encoding. For example, creating an HD version of an SD original file might not be appropriate since there's only so much data in the original file.
Case History Details	NARA-BRCC, LC-NAVCC-VEF and SIA-DVD case histories detail the rationale for selecting file characteristics for normalizing ingested files.

RP 2.8 Generate a new high integrity and continuous master timecode, especially if there is no timecode in the original material	
Rationale	Timecode facilitates file-based workflows and search and discovery. Some files may contain more than one timecode stream. At a minimum, create a continuous master timecode stream to serve as the canonical representation of references into the essence for all timeline-dependent activities such as closed captions, scene changes, audio-sync, etc.
Case History Details	LC-NAVCC-VEF plans to create a continuous master timecode as part of the upcoming MXF AS-07 implementation.

RP 2.9 Retain original timecode(s) if provided, even if you generate a new high integrity continuous master timecode	
Rationale	Timecode values are important to the provenance and discovery of the video object. If timecode values exist in the source file, it's important to retain them for historical and provenance purposes. The selected target format should be able to contain multiple internal timecodes streams or be able to track and refer to timecode data through metadata.
Case History Details	SIA-DVD and LC-NAVCC-VEF case history projects retain timecode if present when normalizing files.

RP 2.10 Retain all the data from the original file if the video file structure is changed	
Rationale	<p>When migrating archival data from one structure to another, choose workflows and processes that retain all the data from the original file. This might mean keeping uncompressed data streams as is, or if business needs require compression, selecting lossless compression over lossy compression. Other important data to maintain might include embedded metadata, timecode(s), and captions.</p> <p>While it's a generally accepted archival practice to retain the original(s) digital objects as submitted, especially if normalizing or performing other actions that change the structure or</p>

	composition of the archived version of the file, this RP focuses on the data contained within the original digital object, not necessarily the object itself as delivered.
Case History Details	VOA-MMAM case history touches upon associated and embedded data.

RP 2.11 Retain the original chroma subsampling if the video data is transcoded	
Rationale	<p>Chroma subsampling, in which the chroma color components (Cb and Cr) are compressed by sampling them at a lower rate than the luma or brightness (Y), reduces the color resolution in digital component video signals. Only the colors are compressed, not the luma, because the human eye is more sensitive to brightness than to the color components. This lossy compression, in which the color information is permanently discarded, results in reduces file sizes and is often implemented to accommodate storage and bandwidth limitations.</p> <p>Retain the original chroma subsampling if transcoding is required to meet business needs to avoid permanent data loss via additional compression which can result in visual artifacts.</p>
Case History Details	LC-NAVCC-VEF retains the native chroma subsampling when it's declared or knowable through metadata extraction.

RP 2.12 Retain original frame rates if the video data is transcoded, even when they are beyond the standard 24 - 30 fps					
Rationale	Higher than typical frame rates (e.g., 46 fps, 60 fps or higher) or even more complex frame rates such as 24 fps per eye for 3D are especially challenging and may be beyond the current capabilities of many archival repositories. Best advice is to retain the original frame rate and not try to normalize to standard frame rates. This may mean, in some cases, less than optimal functionality until other workflow systems catch up with their functionality but it's more important to retain the original "as is" and expect technology to improve than to force a reduction in the information to fit into today's technology limitations.				
Case History Details	<table border="1"> <tr> <td>Project Goals Permit Following Recommended Practice</td> <td>LC-NAVCC-VEF and SIA-DVD case histories, which deal with format normalization in different environments, retain the native frame rates.</td> </tr> <tr> <td>Project Goals Preclude Following Recommended Practice</td> <td>NARA-BRCC project adjusted the original frame rate on one item from 15 fps to a more standard 29.97 and ensured that no unwanted artifacts or changes were introduced.</td> </tr> </table>	Project Goals Permit Following Recommended Practice	LC-NAVCC-VEF and SIA-DVD case histories, which deal with format normalization in different environments, retain the native frame rates.	Project Goals Preclude Following Recommended Practice	NARA-BRCC project adjusted the original frame rate on one item from 15 fps to a more standard 29.97 and ensured that no unwanted artifacts or changes were introduced.
	Project Goals Permit Following Recommended Practice	LC-NAVCC-VEF and SIA-DVD case histories, which deal with format normalization in different environments, retain the native frame rates.			
Project Goals Preclude Following Recommended Practice	NARA-BRCC project adjusted the original frame rate on one item from 15 fps to a more standard 29.97 and ensured that no unwanted artifacts or changes were introduced.				

USE STABLE AND MANAGED DIGITAL STORAGE

The RPs in this section outline the advantages of stable and managed storage systems in archival repositories.

RP 2.13 Move video files off removable data carriers to stable and managed storage as soon as possible	
Rationale	<p>Archivists are likely to come across video files stored on many different types of media including removable storage devices such as memory cards, flash drives, videotape, and optical media. It's essential to migrate the video data off this removable media in part because of the removable media issues outlined in <i>RP 1.3 Capture video data to stable storage devices that allow for streamlined file transfer into managed storage</i> including risk of technical obsolescence, condition issues with the physical data carriers and the lack of strong data protection mechanisms. Other issues with data only residing on the original removable media devices is the lack of redundancy in case of media failure and the limited opportunities for access.</p> <p>Stable and managed storage, which provides stability not only through redundancy but also scheduled audits, and hardware replacement, is essential to the act of archiving. Digital objects on more vulnerable media need to be moved to more stable media as soon as possible. While this process can be as simple as a file transfer, it can be more challenging for complex objects like authored DVDs which may require the creation of an ISO disc image.</p>
Case History Details	<ul style="list-style-type: none"> • LC-AFC-CRHP: Interviews recordings were saved to an external hard disk recorder in the field. At the end of each day, these files were then copied to two separately stored hard drives to ensure redundant copies. • LC-WebArch-YouTube: After each crawl is completed by the Internet Archive, we must use in-house tools to migrate the content to our local network, copy it to long term storage, process it for public access, and then copy it to public access storage. • NARA-BRCC, LC-NAVCC-VEF and SIA-DVD transfer video off optical discs and external hard drives into managed storage as soon as it can after accession of the collection. • NOAA-OkEx: Video is duplicated across two shipboard SAN arrays that are each configured as RAID5 to guard against data loss. In addition at the end of a season video is offloaded from the ship and stored shore side on an additional SAN. As of FY2013 OER became involved in a NOAA pilot program to address the issue of archiving large volumes of these datasets in a near line access model. • SI-DAMS: As archivists processing collections, working with the content creators is not always an easy process, and often decisions that are best for the collection are made without in-depth consultation with content creators, due to the size of the collection and the expedient need to process it because of technological risk. Having a DAMS application allows for at risk media to be secured and moved off of local curators' local storage. • VOA-MMAM uses the Dalet Plus News Management System which is connected to two Omneon Spectrum playout servers (one for studios and one for TV Master Control) for asset archiving. Dalet operates with the Front Porch Archiving system and Spectra Logic LTO storage.

PART 3. ADVICE FOR FILE CREATORS AND FILE ARCHIVISTS

Many Recommended Practices are valid for both File Creators and File Archivists because they transcend specific life cycle points. These include RPs related to selecting encodings or wrappers for digital video, either at the point of initial capture on the camera or for normalizing to a common file format to meet business needs. Other RPs in this category include advice for creating and harvesting metadata in sustainable and structured ways to facilitate downstream workflows. Metadata creation starts at the point of file creation and continues through repository ingest.

CREATE AND USE METADATA TO FACILITATE LIFE CYCLE MANAGEMENT

The RPs in this section emphasize the value in creating structured metadata to that it can be used, especially by automated systems, to facilitate the discovery of relevant information.

RP 3.1 Use metadata data models with strong support for digital video	
Rationale	<p>Born digital video objects have unique metadata requirements. These include documenting: 1) relationships of objects within a file such as video and audio streams, 2) timeline-based activities like closed captioning or timecode, and 3) relationships between separate but associated files such as episodes in an ongoing series.</p> <p>Prepare your metadata for reuse, especially machine-oriented reuse, either from the initial capture or through mapping, by using data models that have tailored support for digital video. Typical data models include XML and RDF. Although these models for digital video are still emerging, examples include the XML-based schemas for reVTMD, PB Core2.0 and videoMD.</p>
Example	<p>reVTMD, PB Core and videoMD are examples of XML-based schemas for video. See Resource Guide.</p>
Case History Details	<ul style="list-style-type: none">• SI-DAMS has a standard XML format for each asset.• VOA-MMAM uses standard XML-compliant metadata templates which evolved over time and were adjusted as needed throughout the project.

RP 3.2 Document and use technical metadata	
Rationale	<p>The FADGI glossary defines technical metadata as a “generic term for technical information about the digital files and multi-file objects, as further defined by three terms for important aspects of technical information: (1) file-characteristics metadata for technical information about the formatted digital file in hand; (2) source metadata for technical information about the source item, whether analog or digital; and (3) process metadata for information about the technical processes used to convert the source item into the digital file that is described in (1).”³</p> <p>A wider understanding of the term also includes preservation and administrative actions performed on the file. For born digital video, technical metadata is primarily concerned with how the file is structured, how it was created and what has happened to it during its life cycle.</p>

³ <http://www.digitizationguidelines.gov/term.php?term=metadatatechnical>

	<p>This information is then used by technologies and collection managers in a variety of ways including reporting (“how many *.mov files are in this collection?”), quality control (“was the format normalization successful?”) and process monitoring (“what was the last action performed on these files?”).</p> <p>Technical metadata in digital video files can be viewed and/or harvested by tools such as MediaInfo and FFprobe. In addition, these tools also allow for the generation of reports on technical specifications useful for short and long term obsolescence monitoring.</p>
Example	Sample tools that read technical metadata: MediaInfo and FFprobe. See Resource Guide.
Case History Details	<ul style="list-style-type: none"> • LC-NAVCC-VEF reports that metadata is perhaps the area in need of greatest focus: metadata pathways through the entire file-based workflow systems must be enabled and made reliable; metadata embedding standards need to be developed. (The FADGI AS-07 project is a big step in the right direction.); metadata schema standards need to be developed in coordination with industry. • LC-WebArch-YouTube uses BagIt metadata fields to store information on the storage containers (bags) that house the crawl content as well as separately generating a range of data reports for each crawl that provide more content-specific statistics. • NARA-BRCC technical metadata was extracted from the files and collected in a spreadsheet to inform normalization decisions. • SIA-DVD metadata is noted on spreadsheets and other documentation is created. • SI-DAMS uses MediaInfo and ExifTool to read the technical information from the files at the point of ingest, including limited supported descriptive fields (XMP). The DAMS ingest transcoders create proxy video files for viewing in the DAMS, at which point technical information (video and audio codecs, bit rate, frame rate, frame height and width, number of audio tracks) is generated. This makes them reportable for obsolescence monitoring. More fields are recommended for manual data entry: originating format, coding history, color space, capture device, caption format. Mapping to PREMIS can easily be done with the SI-DAMS standard XML format for each asset. • VOA-MMAM utilizes a sophisticated workflow to using an XSLT translation script to populate to mapped fields in their customer Asset Manager Form with the appropriate data.

SELECTING FILE FORMATS (WRAPPERS, CONTAINERS AND/OR ENCODINGS)

While the Recommended Practices for selecting digital formats are relevant throughout the life cycle (creation, normalization prior to or after submission to repository, access delivery, etc.), it's always important to consider the intended end use of the file. Files intended for long term retention require open, flexible and transparent structures; files intended for specific distribution systems will need to conform to constraints of downstream applications. The Library of Congress' Sustainability of Digital Formats⁴ website and the FADGI resources comparing formats and wrappers for reformatted video⁵ will be helpful in the decision-making process.

⁴ <http://www.digitalpreservation.gov/formats/index.shtml>

⁵ http://www.digitizationguidelines.gov/guidelines/File_format_compare.html

RP 3.3 Select uncompressed video encoding over compressed encoding		
Rationale	<p>Uncompressed video retains all the visual information captured at the selected resolution whereas compressed video reduces the amount of visual information stored in a file or stream (see RP 3.4 for more information on compression). Uncompressed video has the advantage of less processing complexity but results in larger files, which can be more expensive to store and difficult to transfer. Compressed video files are generally smaller and easier to move around but might be more technically complex and more challenging to process.</p> <p>For archival workflows, the goal is to retain all the visual information present in the initial capture process so uncompressed video generally is preferred. Compressed video however is often a reasonable and responsible decision depending on the scale, scope and goals of the project.</p>	
Case History Details	Project Goals Permit Following Recommended Practice	<p>Interestingly, none of the case history projects use uncompressed video streams - for good reasons. In each case, the need for smaller files and/or systems-specific compressed formats outweighed the need for uncompressed video. The Recommended Practice for selecting uncompressed video when there's the option to do so is still a valid one since archival repositories want to receive and retain data-rich files which will allow the most flexibility in the future. For these particular projects, the compromises were made in order to meet the project goals.</p>
	Project Goals Preclude Following Recommended Practice	<ul style="list-style-type: none"> • LC-AFC-CRHP implements compression because the relatively simple content, oral history interviews, did not warrant the larger file sizes from uncompressed video. • LC-NAVCC-VEF: With collections at the scale of those at the Packard Campus, implementing compression translates into considerable savings in digital storage space and costs compared to storing uncompressed files. • NARA-BRCC: This case history project did not select an uncompressed target format because the source material already was highly compressed. In addition, data storage was limited so the increased file size would be problematic. • NOAA-OkEx implements compression because the finite physical space aboard ship limits the number and configuration digital storage devices and there's an institutional directive to capture as much subsurface video as possible in as high a bit rate as is operationally feasible. • VOA-MMAM implements compression because the large scale and scope of their broadcast workflows make uncompressed video infeasible from a practical perspective.

RP 3.4 If compression is used, select mathematically lossless compression over visually lossless or lossy compression

<p>Rationale</p>	<p>The FADGI Glossary defines lossless compression in this way: “Data compressed using a lossless compression technique will allow the decompressed data to be exactly the same as the original data before compression, bit for bit. The compression of data is achieved by coding redundant data in a more efficient manner than in the uncompressed format. The compression ratios that can be achieved with lossless compression are generally much lower than those that can be achieved using lossy compression techniques. Data compressed using a lossy compression technique results in the loss of information. The decompressed data will not be identical to the original uncompressed data. Conservative lossless compression can result in a form of lossy compression referred to as visually lossless compression.”⁶</p> <p>In a nutshell, lossless compression works by removing redundant information that can be recreated from the remaining data; lossy compression works by permanently removes non-essential data. For digital video files destined for archival repositories, lossless compression is preferred over lossy because lossless compression retains all the original data while at the same time achieving practical and financial goals.</p> <p>One key distinction in lossless compression is mathematically lossless verses visually lossless. In mathematically lossless compression (such as that in use at the Packard Campus in their evergreen format JPEG2000 lossless encoding reversible 5/3 in MXF OP1a), the video is identical to uncompressed HD-SDI video. Visually lossless compression however is really a term of art because it is just another form of lossy compression. The FADGI glossary defines visually lossless compression as “a form or manner of lossy compression where the data that is lost after the file is compressed and decompressed is not detectable to the eye; the compressed data appearing identical to the uncompressed data.”</p> <p>Within compression algorithms, there’s a further distinction between intraframe and interframe compression. An intraframe codec applies compression to each individual frame and does not take data from other frames into account. Interframe compression, on the other hand, is based on the idea that although action is happening, the background remains mostly stable across a scene in adjacent frames so a great deal of the data is redundant. Compression is started by creating a reference frame. Each subsequent frame of the video is compared to the previous frame and the next frame, and only the difference between the frames is stored. In interframe compression, the redundant data is permanently lost. For these reasons, intraframe compression is preferred because it results in less data loss.</p>	
<p>Case History Details</p>	<p>Project Goals Permit Following Recommended Practice</p>	<p>LC-NAVCC-VEF: The profile of JPEG2000 used in the Packard Campus Evergreen format is reversible and mathematically losslessly compressed.</p>
	<p>Project Goals Preclude Following Recommended Practice</p>	<ul style="list-style-type: none"> • LC-AFC-CRHP uses ProRes HQ (422) lossy compressed because the relatively simple content, oral history interviews, did not warrant the larger file sizes from uncompressed video. In addition, Pro-Res compresses each frame individually (intraframe) as compared to the GOP (interframe) approach. • NARA-BRCC: The case history uses MPEG-2 at

⁶ <http://www.digitizationguidelines.gov/term.php?term=compressionlossless>

		<p>50Mbps with I-frames only. This is a moderately high bitrate encoding scheme that relies only on intraframe compression (less lossy than interframe). It is a visually lossy compression, because lossless compression would have resulted in larger file sizes that weren't justified by the source material.</p> <ul style="list-style-type: none"> • NOAA-OkEx uses ProRes, a lossy format, because the ship's capacity limits the available storage space. Moreover, ProRes was preferred because it supports I-frames. Thus, although a lossy format, each individual frame of the video is independent of proceeding or subsequent frames to determine its individual pixel composition. • VOA-MMAM's house standard is DV25. While DV25 is a lossy codec, it meets VOA-MMAM's business needs and is supported by essential internal VOA-MMAM systems.
	Notes	<p>It is interesting to note that, with the exception of LC-NAVCC-VEF, none of the case history projects support the suggested Recommended Practice for valid reasons. In each case history, the use of visually lossy compression was evaluated against lossless and uncompressed and for both business and workflow reasons, visually lossy compression best met the needs of each project.</p> <p>The Recommended Practice for selecting lossless compression over lossy compression when it makes business sense to do so is still a valid one because archival repositories may need to save storage space by using compression but also strive to not lose data. For these particular projects, the compromises were made in order to meet the project goals.</p>

RP 3.5 Avoid multiple compressions and decompressions steps		
Rationale	When migrating to a different codec, move from the source codec to the target codec in one transformation. Compressing and decompressing the data multiple times increases the likelihood of visual artifacts and other quality and technical issues. This is especially true for lossy codecs where repeatedly compressing and decompressing the file will cause it to progressively lose quality. Generation loss is not generally an issue for lossless codecs but the repeated manipulation of the bitstream introduces more opportunity for error.	
Case History Details	Project Goals Permit Following Recommended Practice	<ul style="list-style-type: none"> • NOAA-OkEx: The source video is captured as ProRes at 145Mbps and stored as the same. • NARA-BRCC uses single-step transformation to move from the source files directly to the normalized intermediate formats.
	Project Goals Preclude Following Recommended	<ul style="list-style-type: none"> • LC-WebArch-YouTube: Given the dependence on the Wayback Machine to replay the archived content, it is

	Practice	<p>necessary that to store that content in compressed WARC files. This provides the added benefit of reducing the total volume of our bulk storage, while supporting stewardship and the standardization methods of the greater Web Archiving community.</p> <ul style="list-style-type: none"> • SIA-DVD: The video content on the authored DVDs is lossy and compressed and is normalized to MPEG-2 to meet business needs.
--	----------	--

RP 3.6 Stay within the same codec family if the video data is transcoded		
Rationale	<p>Different compression schemes make different coding mode decisions regarding <i>what</i> and <i>how much</i> information is discarded to reduce the file's size. Lossy coding, with its complex array of decision parameters, varies greatly depending on the implementation and dynamic characteristics of the data. Even lossless coding, where discarded redundant data can be reinterpreted upon decoding, has options in how the data is compressed. Compressed bitstreams can result in many changes to the characteristics of the picture including image data lost from the picture and artifacts like blocking and noise introduced into the picture.</p> <p>Staying within the same codec family when transcoding assures that at least the same compression coding technique is used to achieve the space savings. Moving from one compressed codec family to another (e.g., from MPEG to JPEG) might result in increased risk of data loss and introduction of visual artifacts because the data reduction criteria is not the same for both codec families. Data would be lost on the initial MPEG compression and different data might be lost on the transcode to JPEG.</p>	
Examples	<p>Examples of codec families include the MPEG-2 family,⁷ the DV (DV25, DC50, DVCAM, DVCPRO) family⁸ and the JPEG 2000 family.⁹ See <i>Sustainability of Digital Formats</i> on the Resource List for more information.</p>	
Case History Details	Project Goals Permit Following Recommended Practice	<ul style="list-style-type: none"> • NARA-BRCC: Higher quality MPEG-2 source files remain in MPEG-2 for the target format, MPEG-2 at 50Mbps.
	Project Goals Preclude Following Recommended Practice	<ul style="list-style-type: none"> • LC-NAVCC-VEH project makes use of the house standard JPEG2000 in MXF OP1a target normalization format for all digital video files, regardless of the source format. • LC-WebArch-YouTube project stores all content in compressed WARC files regardless of its source format.

⁷ <http://www.digitalpreservation.gov/formats/fdd/fdd000335.shtml>

⁸ <http://www.digitalpreservation.gov/formats/fdd/fdd000183.shtml>

⁹ <http://www.digitalpreservation.gov/formats/fdd/fdd000138.shtml>

RP 3.7 Select video encoding and wrapper formats that are well-supported now and future focused	
Rationale	When selecting a target video format (encoding and/or wrapper) for use in an archive, either as a direct deposit format or a target normalization format, choose formats that have a robust future ahead of them. Practically, this means formats that are strongly adopted with active user communities and large and diverse options for both commercial and open source tool sets.
Case History Details	<ul style="list-style-type: none"> • LC-AFC-CRHP: ProRes is established, well documented and widely used, particularly in the postproduction environment. • LC-NAVCC-VEH: JPEG 2000 lossless and MXF are both well supported in current toolsets and are likely to continue to be so. • NARA-BRCC: MPEG-2 @ 50Mbps is a common implementation of the MPEG-2 standard and should be supported most applications. • NOAA OkEx: Both Quicktime and ProRes are established, well documented and widely used, particularly in the postproduction environment. • SIA-DVD uses MPEG-2, which should work with many video players including WMP, VLC, and QuickTime. ISO disc images can be mounted on computers for playback. With an ISO copy, SIA can revisit the file for additional processing as software is developed to possibly create a “more accurate” access file, such as better audio conversion.

RP 3.8 Select video encoding and wrapper formats that are non-proprietary	
Rationale	When selecting a target video format (encoding and/or wrapper) for use in an archive, either as a direct deposit format or a target normalization format, choose formats that are open and non-proprietary. Non-proprietary formats are less likely to change dramatically without user input, be pulled from the marketplace or have patent or licensing restrictions.
Case History Details	Project Goals Permit Following Recommended Practice <ul style="list-style-type: none"> • SIA-DVD and NARA-BRCC case histories use MPEG-2, which is standardized through the ISO 13818 document set and is not proprietary. • LC-NAVCC-VEF: Both JPEG 2000 and MXF are international standards and not proprietary.
	Project Goals Preclude Following Recommended Practice <ul style="list-style-type: none"> • LC-AFC-CRHP uses Pro-Res, a proprietary format, because it retains the 4:2:2 information and compresses each frame individually (intraframe) as compared to the GOP (interframe) approach. • NOAA-OkEx: QuickTime and ProRes are proprietary formats, but there are active open source projects actively supporting both.

RP 3.9 Select video encoding and wrapper formats that are supported by downstream applications	
Rationale	Digital video files are often very large, requiring significant time and resources to move across

	data storage systems, and structurally complex. Choosing file formats (wrappers and encodings) that are supported by downstream applications will mean less processing time and reduce the complexity of the workflow.
Case History Details	<ul style="list-style-type: none"> • LC-AFC-CRHP: the HQ codec is fully compatible with NLEs using Final Cut Pro software presently installed in the suite and eliminates the need to render the materials into a lower bit rate/ resolution for editing and production, as has been the case previously. • LC-NAVCC-VEF: Because NAVCC normalizes all video data to the same encoding and wrapping formats, all toolsets are geared to work with these format selections. • LC-WebArch-YouTube: This principle applies to the Library’s Web Archive in that our diversity of content must be store in compressed WARC formatted files so that the Wayback Machine application can replay it. • NARA-BRCC: MPEG-2 @ 50Mbps is a common implementation of the MPEG-2 standard and should be supported most applications. • NOAA OkEx: ProRes video in a Quicktime wrapper is readily ingestible by nonlinear video editors such as Adobe Premiere and Apple Final Cut Pro in post-production editing. • SIA-DVD File remains MPEG-2. VOB is based on the MPEG program stream. • VOA-MMAM has a controlled internal system which supports the specified house formats.

RP 3.10 Select video formats that are standardized and well-documented	
Rationale	Standardized and well-documented formats are generally more stable. They have fewer variants and are more widely supported through diverse toolsets and will have a much greater longevity than those formats subject to frequent change. Ideally, standards are published through an international standards body such as ISO, SMPTE or AES which includes a wide review and voting prior to finalization. Documentation through vendors and proprietary companies is valuable but is less stable because it can be changed or removed with little to no notice or process.
Example	Documentation for digital video formats is available on the Sustainability of Digital Formats website. See Resource Guide.
Case History Details	<ul style="list-style-type: none"> • LC-AFC-CRHP uses ProRes which is standardized and well documented. • LC-NAVCC-VEF uses JPEG2000 encoding in the MXF OP1a wrapper for a variety of reasons including the fact that both are international and well-documented standards and preservation of content in a given digital format over the long term is not feasible without an understanding of how the information is represented (encoded) as bits and bytes in digital files. JPEG2000 is standardized in ISO 15444. MXF OP1a is standardized through SMPTE 377-1 and SMPTE 378M-2004. A compelling benefit of adopting internationally standardized wrappers and codecs is that vendors build tools and applications to meet these specifications because that will help wide market adoption. Another reason is that preservation planning is simplified because institutions only need to maintain the documentation on one set of standards instead of many. • NOAA-OkEx: Both Quicktime and ProRes are established and well documented. • SIA-DVD and NARA-BRCC: Both MPEG-1 and MPEG-2 are well established and well documented through ISO.

RP 3.11 Select video formats with capacity for robust and detailed technical metadata	
Rationale	Technical metadata includes details on the file characteristics. Understanding the technical layout and components of the file is essential to its long term preservation because you need to know what you have before you know what you can do with it.
Case History Details	LC-NAVCC-VEF: The MXF wrapper was specifically designed to aid interoperability and interchange between different vendor systems, especially within the media and entertainment production communities which are the primary content providers to Packard Campus collections. The file specification was standardized by the SMPTE (Society of Motion Picture & Television Engineers) & AMWA (Advanced Media Workflow Association) and allows different variations of files to be created for specific production environments and can act as a wrapper for metadata & other types of associated data.

RP 3.12 Select video formats with greater capacity for embedded metadata over less metadata capacity	
Rationale	<p>Embedded metadata can provide information to and support functionality for various persons and systems at a variety of points in the content life cycle. For example, it can help the digitizing unit or organization as it produces and preserves content. It can serve persons or systems who receive content that is disseminated by the digitizing unit or organization. Some metadata elements are especially valuable to internal actors, some to external, and some to both. Embedded metadata, of course, is rarely an agency's only metadata. In most archiving and preservation programs, workflow and archiving are supported by one or more databases, cataloging systems, finding aids, and the like, each of which contains metadata. Many if not all metadata elements turn up in more than one place, a good thing since redundancy supports long-term preservation.</p> <p>It should be noted that embedding metadata that duplicates metadata held in systems outside the file (like databases, finding aids or catalogs) can make it difficult for the metadata to stay in sync unless this is supported in an automated way by an organization's technical infrastructure.</p>
Case History Details	<ul style="list-style-type: none"> • LC-NAVCC-VEF: The MXF wrapper was specifically designed to aid interoperability and interchange between different vendor systems, especially within the media and entertainment production communities which are the primary content providers to Packard Campus collections. The file specification was standardized by the SMPTE (Society of Motion Picture & Television Engineers) & AMWA (Advanced Media Workflow Association) and allows different variations of files to be created for specific production environments and can act as a wrapper for metadata & other types of associated data. • SIA-DVD: The MPEG-2 allows for a sidecar XMP file, which lives with the corresponding file and is also imported into the DAMS.

RP 3.13 Select formats that can contain and label complex audio configurations including multiple channels and sound fields beyond mono and stereo	
Rationale	Some of the most common audio channel configurations are mono (single channel) and stereo

	<p>(left and right channels) both of which are widely accommodated in many video formats. More complex configurations such as 5.1 surround (left, center, right, left surround, right surround, and low-frequency effects channels) and 7.1 surround (left, center, right, left surround, right surround, left back, right back, and low-frequency effects channels) and others may not be supported in less structured formats. Multiple channel sound may also feature two or more complementary signal streams that provide alternate or supplemental content, e.g., narration in French and German, commentary from film directors or actors, sound effects separate from music, karaoke content, or the like.</p> <p>It's essential to capture and retain the original audio channel layout so choose a file structure that can duplicate the structure of the source material.</p>
Example	FFmpeg has a good standard audio channel layout listing. See Resource Guide.
Case History Details	<ul style="list-style-type: none"> • LC-AFC-CRHP uses ProRes which supports 1 video stream and up to 4 audio channels. • LC-NAVCC-VEF uses MXF which has strong support for multiple and complex audio configurations. • NOAA-OkEx: The ProRes format used by the <i>Okeanos</i> program supports 1 video stream and up to 4 audio channels (all utilized in the EVS clipping process).

RP 3.14 Select formats that can support robust timecode data	
Rationale	Timecode is an electronic signal used to identify a precise location in a time-based media file or tape. Its primary use is synchronization of various data streams but it can also have important uses in search and discovery. Select formats that support timecode data, preferably multiple timecode streams, for both new video files and normalized or migrated video files.
Case History Details	SIA-DVD and LC-NAVCC-VEF case history projects discuss timecode issues.

PLAN FOR ACCESS

Because of their often large size and structural complexity, high quality born digital video files may need additional processing to be made widely available especially over the Internet. The RPs in this section explore the some of the issues involved with improving access to high quality born digital video files.

RP 3.15 Create access, viewing or proxy copies with appropriate technical characteristics to meet expected use cases	
Rationale	<p>Because high quality video files are often large, complicated and difficult to display through many common access points including the Internet, viewing copies are often created to facilitate access. Different end uses will require different technical characteristics and file properties. In some cases, like broadcast over commercial television where fine detail and high resolution is expected, a robust and data rich file would be required. In other cases, a lower quality, smaller file suitable for streaming over the web meets access expectations. The end use will inform the selection of the access derivative's file characteristics. Research the technical specifications for your common delivery systems and access use cases, such as supported encoding and wrapper formats, and bit rate, to be sure the access file will meet the expected needs of the delivery systems. Typical end uses might include posting on video sharing website such as YouTube and access through your institution's online catalog.</p>
Case History Details	<ul style="list-style-type: none">• LC-NAVCC-VEF and SIA-DVD case histories detail the technical characteristic of access copies.