



Community Archiving Workshop

Webinar 4: Digitization in-house vs out-of-house

Southwest Region IMLS + NEH Cohorts

Digitization Planning

May 26, 2021

Presenters: Kelli Hix & Marie Lascu

In this session, participants learn about planning for digitization, with an emphasis on the differences between in and out of house digitization.

#cawesome

What has been covered so far:

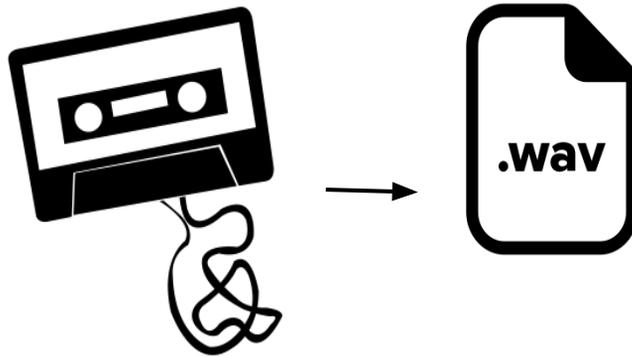
- Identifying, inspecting, and assessing film & magnetic media
- Inventorying film & magnetic media
- Metadata for creating inventory templates

This webinar will cover:

- Planning for digitization
- Determining whether to pursue
 - in-house digitization
 - out-of-house digitization with a vendor
 - a combination of both
- Utilizing your inventory for
 - preservation planning
 - tracking preservation work

What is Digitization?

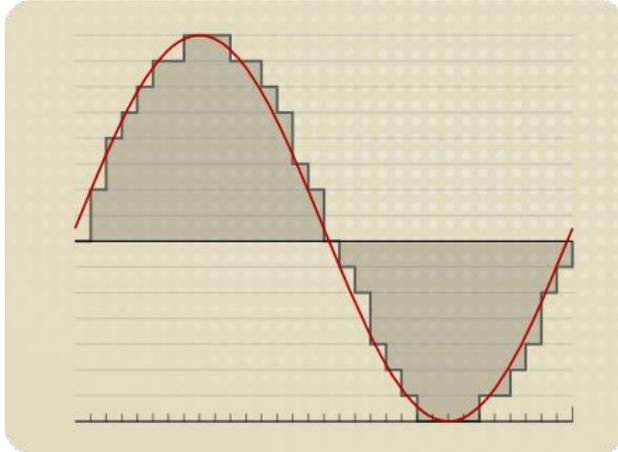
Analog signal (e.g. light, sound) → Digitally encoded format.



The conversion of text, pictures, or sound into a digital form that can be processed by a computer. (Oxford English Dictionary)

What is Digitization?

Analog waveform is approximated digitally.



Converting analog media to digital form is an approximation of the original visually or aurally.

The diagram above is in reference to magnetic media signals. Notice the step-like structure of the digital waveform vs the analog. Computers are essentially machines where switches are either on or off while capturing data so nothing will ever be perfectly smooth.

We seek to reproduce the video or audio signal recorded on tape, or the film scanned as nearly as possible.

Why Digitize?

A step toward preservation:

- Many analog formats are obsolete, and need to be reformatted to non-obsolete formats to be usable.
- Many analog formats are unstable and degrade over time.
- Digital content can be copied exactly over and over again without “generation loss.”

For easier access:

- It is easier to make copies of digitized content.
- Digitized content can be shared electronically / online.

Digitization vs. Digital Preservation

Digitization:

- Reformatting an analog object into a digital representation.
- Migrating a media carrier-based digital signal to a digital file form.

Digital Preservation:

- Digitization is ONE step in the preservation process.
- Multiple steps required to ensure preservation.
- Digital files cannot survive through passive preservation.
- Preservation is long-term.

Digitization alone is NOT preservation.

Plan Before You Digitize!

Avoid common pitfalls:

- Digitizing to a file format that you cannot play or open
- Digitizing with a vendor that does not meet standards
- Digitizing at low quality
- Digitizing to optical media (not a preservation format)



Plan before you digitize! That is the key message of this presentation. Planning helps to avoid common pitfalls, and it helps to avoid re-digitizing in the future. Common pitfalls include:

- Digitizing to a file format that you cannot play or open
- Digitizing with a vendor or set-up that does not perform to good /archival standards. Vendors like Walmart, Costco, or local home movie transfer companies, or an in-house set-up like shooting footage with a camera off of a screen offer low cost way to digitize, and might create a decent reference file for you. But these vendors and methods will not give you the quality that you are likely to want down the road.
- Digitizing at low quality is another one of the common pitfalls of digitization projects. Why is digitizing at high quality important? A low quality image might tell you what is on the film or tape, but the magic happens when you have higher quality. This is when we can read writing such as street signs, license plates, books, and newspapers. We can see facial expressions and gestures that are at the heart of why you may want the footage. Remember that with some of this material you may only have one chance to digitize it if it is fragile. So digitizing at the highest quality that you can is

- important.
- Another common pitfall is digitizing to optical media-- CDs and DVDs. Optical media is not a preservation format. It is fragile, it scratches easily, and often compresses the files, causing low quality.

Planning Foundations: Create an Inventory

D	E	F	G
Capture any information on the label or item that identifies the date of the recording. Be consistent with date format. Example: YYYY-MM-DD or YYYY-MM	Capture run time, if known. If you do not know, capture tape length (written on the box or tape). Be consistent with format.	Capture clearly indicated title on label or asset, or if specified by collection stewards.	Capture information regarding content: on audio or video info may be recorded on labels, on glass slides or film elements - visible at a glance, recorded on labels. If time code is used be consistent. For example: (hours: minutes: seconds) 00:00:00
Date on Item	Duration of Content	Title on Item	Content

Let's look at the foundations for digitization planning that can help your project to be successful. First, an inventory is key. Inventories are the foundation of all of the work that you do in the future, and that is why creating an inventory is central to this project. Your inventory tells you how many hours of footage you have, what condition your assets are in, what formats you have, and what content you have. It's the central place where you track and organize your collection. All of this information is vital for developing a budget and timeline for your project.

Planning Foundations: Develop Preservation Priorities

- Which of our collections have the most important content to our organization?
- What do we have that is unique to our collection?
- What is the most unstable / endangered part of the collection?

Understanding your priorities is also important for preservation planning. A Preservation Plan is a working list of what is most important for you to digitize. This helps you to know what you want to digitize first and where to start. During this project, in the Fall, you will be working with your mentor to create a preservation priority plan. It answers questions like:

Planning Foundations: Determine Digitization Strategy

In- House

Out of House

Combination



Another foundation of digitization planning is determining whether you will digitize in-house, out of house, or using a combination.

- 1) Financial Considerations**
- 2) Collection Considerations**
- 3) Human / Quality of Life Considerations**

Let's go into some detail about what to consider when you are making this decision. I divide the process into three main considerations: Financial, Collection, and Human considerations.

What method is more cost effective?

1. Labor costs
2. Equipment costs
3. Vendor costs

We'll start with financial considerations. Which is more cost effective -- digitizing in house or out of house? In order to determine this, you will need to estimate the cost of labor, the cost of equipment, and the costs to a vendor and compare the estimates. This is a rough number meant to guide your decision, so try not to get too caught in the weeds. It's important to start somewhere.

For example, If you have a large collection of a common media format in good condition, it may be more cost effective to digitize in-house, even after purchasing equipment and training staff. If you have few items in a rare or older format in poor condition, it will probably cost less to send it to a vendor who already is equipped and trained to handle the material. For example, if you have 10 moldy cylinder recordings, it's probably cost-effective to send them to a vendor. If you have 5,000 VHS tapes in fair condition, it can be more cost-effective to digitize them in-house. If you have both the VHS and the cylinder collections, you may want to choose a combination approach.

Factors to include in cost estimate:

In-house:

- 3 hours labor / 1 hour of content
- Training cost
- Equipment cost

Out of House:

- Vendor estimate
- 2-3 hours of admin and quality control labor per hour of content
- Equipment / computers for performing quality control

Here is a handy rule of thumb for creating cost estimates. For in-house estimates, most of your cost is in labor and equipment. To estimate labor, a good rule of thumb is the 3:1 rule: three hours of labor for every hour of content. This allows time to digitize -- tapes must be digitized in real time -- plus time to perform troubleshooting and follow-up tasks such as quality control. Also factor in the cost for a period of training for your project. Finally, add your equipment costs. We aren't going into detail on equipment today, but you will need playback decks, and analog to digital converter, a computer, and some other items to control the video and audio signals.

When estimating the cost to digitize using a vendor, it's important to remember that the estimate your vendor gives you is only the tip of the iceberg. You also need to factor in about 2-3 hours of labor to prepare and ship assets, administer the project, and perform quality control to check the vendor's work.

Advantages of Digitizing In-House:

- No need to ship assets
- You have control over quality
- Sensitive materials / materials with access restrictions stay with your organization.
- Improved knowledge about collection content for your organization

Advantages of Digitizing with a Vendor:

- Specialized training and experience
- Specialized equipment
- May be able to handle mold, damage, and higher levels of decay, etc.
- The vendor has built-in technical support

Though I am a fan of performing in-house digitization when you can, both have advantages to consider. Advantages of digitizing in-house include: you do not need to ship fragile assets; you have control over the quality of the digitization process; sensitive materials or materials with access restrictions stay in your care. Perhaps the biggest advantage is that, during digitization, your staff will see and hear the footage being digitized. This leads to increased understanding of the collection content that is invaluable. Digitizing with a vendor also has advantages: the vendor can handle broken or moldy or damaged items; the vendor performs troubleshooting and deals with technical issues. The vendor has technical support.

What is best for your staff and community?

In-house benefits:

- Skill building
- Job creation
- Quality of life
- Building Community
- Building knowledge about the collection

Out of house benefits:

- No need to purchase software, hardware, equipment
- Potential for less technical skills hand-on

The final group of considerations are less tangible. I call them human or quality of life considerations. These can also be described as how well does the approach meet your mission statement. Here are some advantages of both methods in this category.

Digitizing in-house can provide learning opportunities for staff; provides local jobs; can improve quality of life for those wishing to take on the task. I want to make the point here that digitization can be challenging, but it also can be learned, especially if there is a network and a community. This is also a great opportunity to share local resources, like equipment, skills, and experience, and to pass that on to others in the community.

Out of house projects can be beneficial if you do not want to purchase software and hardware, learn the full process, and do not

want to or cannot perform digitization on-site. I do want to emphasize that you will need to learn technical skills no matter which method you choose, because it is vital to the success of your project that you can communicate what you want with your vendor, understand your options, and check on the quality of the service.

Consider the space requirements

- + Soundproofing
- + Internet access
- + Power
- + Permanent set-up

Have backup equipment

Have an extra computer for file processing

Have a dedicated staff member

Have a network or advisor that can help with troubleshooting.

Consider sharing a digitization station or equipment - use existing community resources

Kelli - Once you have made a decision, consider these factors as part of your planning. Space requirements

- + Soundproofing
- + Internet access
- + Power
- + Permanent set-up

Have backup equipment

Have an extra computer for file processing

Have a dedicated staff member

Have a network or advisor that can help with troubleshooting.

Consider sharing a digitization station or equipment - use existing community resources

- Create a vendor agreement
- Get an estimate
- Plan staff roles
- Test the workflow with a pilot project
- Do not have assets shipped back until you have approved them.
- And so much more...

If you have decided to digitize in-house, consider these aspects of the project:

Create a vendor agreement, Get an estimate, plan staff roles (quality control, etc.) Test the workflow with a pilot project

- Do not have assets shipped back until you have approved them.

Digital Files

Project Planning includes:

- Target formats for digitization
- File naming conventions
- File types and version control
- Quality control considerations
- Use your inventory for:
 - Estimating digital storage
 - Tracking digital preservation information

Questions to ask before choosing target file formats:

- Can you store large digital files?
- Can you back up digital media?
- What are your long term goals for saving this media – both physical and digital?
- What type of access do you want for your media?
- How do you want to use your digital media?

Paraphrasing about Evaluation Criteria for selecting target file formats - sourced from the *Recommendations on Preservation Files for Use in the Digitization of Analog Audio and Video Recordings and Motion Picture Films* document created by the National, Provincial and Territorial Archivists Conference Audiovisual Preservation Working Group in January 2018: <https://www.bac-lac.gc.ca/eng/about-us/publications/Documents/preservation-file-formats.pdf>

An organization should adopt formats that can be sustained. Most organizations will only have a single opportunity to digitize analog materials and should try to future-proof their work by making informed decisions around file format and codec choices.

Losslessly compressed or uncompressed preservation files will capture the maximum number of attributes of the original - but the

declaration that lossless uncompressed preservation files are the ONLY option to be considered ignores the realities facing most organizations tasked with preserving a/v collections. The resources you have available to you may render the ideal preservation level file unsustainable. There has to be a balance struck between quality and sustainability.

The ideal: Preservation-level file

a high quality digital transfer that would capture as much information as possible from the original source.

This is not the file used for any editing, screening, or other access-related projects.

In addition to the preservation file, you will want to create access copies - *compressed or smaller formats* - generated using your preservation file.

There are multiple options for your preservation file format, and your choice will depend on where you are storing the file long-term, the internal standards of your organization or the standards you've set for your own project. We are also encouraging you to consider the standards/practices within the media preservation community.

The following target file formats are either Recommended or Acceptable. These options do not cover everything, and do not cover more compressed common access formats such as mp3 or mp4

Target File Formats: Film

35mm

File Format (image):

Digital Picture Exchange (DPX)

Codec: Uncompressed

Resolution: 4k
(4096 pixels)

Bit depth / Sample rate:
10 bit RGB

File Format (audio): BWA

Codec: PCM

Bit depth / Sample rate: 24-bit,
48kHz

16mm, 8mm, Super 8

File Format (image): Digital
Picture Exchange (DPX)

Codec: Uncompressed

Resolution: 2k
(2048 pixels)

Bit depth / Sample rate:
10 bit RGB

File Format (audio): BWA

Codec: PCM

Bit depth / Sample rate: 24-bit,
48kHz

We are not covering the complexities of digital file formats, but suggested formats are depicted here and in the following examples with broken out information that will typically need to be selected in most digitization capture software, or will need to be discussed when working with a vendor.

Target File Formats: Video

½" open reel, U-Matic, Betacam, Hi-8, VHS...

File Format:

QuickTime (MOV)

Codec (Image): Uncompressed
4:2:2

Bit depth / Sample rate:
(recommended) 10-bit
(minimum) 8-bit

Codec (audio): PCM

Bit depth / Sample rate:
(recommended) 24-bit, 48kHz
(minimum) 16-bit, 48kHz

File Format:

Matroska (MKV)

Codec (Image): FFV1 version 3

Bit depth / Sample rate:
(recommended) 10-bit
(minimum) 8-bit

Codec (audio): PCM

Bit depth / Sample rate:
(recommended) 24-bit, 48kHz
(minimum) 16-bit, 48kHz

...DV, MiniDV, DVCam...

Target File Formats: Video

1/2" open reel, U-Matic, Betacam, Hi-8, VHS...

File Format:

QuickTime (MOV)

Codec (Image): ProRes 422

Bit depth / Sample rate:

(recommended) 10-bit HQ

(minimum) 8-bit

Codec (audio): PCM

Bit depth / Sample rate:

(recommended) 24-bit, 48kHz

(minimum) 16-bit, 48kHz

DV, MiniDV, DVCam.

File Format:

Quicktime (MOV)

Codec (Image): DV-NTSC (DV-25)

Bit depth / Sample rate: 8-bit

Codec (audio): PCM

Bit depth / Sample rate:

16-bit, 48kHz

Target File Formats: Audio

¼" open reel, Compact audio cassette

File Format:

Broadcast Wave (BWF)

Codec: LPCM

Bit depth / Sample rate:

(recommended) 24-bit, 96kHz

(minimum) 24-bit, 48kHz

File Format:

Waveform (WAV)

Codec: LPCM

Bit depth / Sample rate:

(recommended) 24-bit, 96kHz

(minimum) 24-bit, 48kHz

(This should sound familiar!)

Develop File Naming Conventions for digital film/video/audio files:

- Name files/folders in a consistent way
- Keep file/folder names relatively short
- Use unique IDs
- Name in a way that makes files/folders easily searchable
- Avoid characters like #*\$&%/?!:@

This should sound very familiar after the previous webinar on Metadata.

1. Name in a consistent way
(established consistencies, or “naming conventions” are important for maintaining organizational control over collections throughout the lifespan of an Organization/Collective/Institution, as well as making them more discoverable both internally and externally)
2. Keep names relatively short
(your naming conventions are intended to clarify content and maintain organization, not do the work of a catalog record)
3. Use unique IDs - the ID associated with the physical media asset, it might also be appropriate to add an additional unique ID for individual digital files.
4. Name in a way that makes files/folders easily searchable

(this speaks to the point of “discoverability” and maintaining organization, but not intending to provide contextualization the way a full catalog record would. Also version control!)

1. Do not use characters like #*\$%/?:!@

(This refers to characters that a computer will not process correctly, which would result in “missing” files or simply information being processed incorrectly)

EXAMPLE 1: File Naming Convention example

Established convention: UniqueID_DateYYYYMMDD_CityName

Name
9GmrgPuC4ce_20130420_Aleppo.mov
IHfdhskdcjcdio_20110112_PhnomPenh.mov
OBNfdfsifwnRb_20130206_Oakland.mov
XedDOSnmsn_20120605_Madrid.mov

Unique ID Date City Name

EXAMPLE 2: Folder structure for digital production elements

Name	Size
P-OTF	4 GB
Elements	316.2 MB
Outputs	25.23 GB
Projects	143.9 MB
Raw	177.46 GB
E002277_3864	12.74 GB
E002515_3282	12.88 GB
E002517_3284	12.7 GB
E002519_3285	12.84 GB
E002526_3283	2.01 GB
E002556_3869	

*Examples sourced from archiving.witness.org/archive-guide

Example 1: File Naming Convention example

Establish and define file naming conventions in a similar manner to establishing fields for the inventory template - select intentionally for your purposes, define clear and easy to follow standards. Document these decisions so anyone working on the collection can reference and maintaining the practice.

In this particular example, the unique IDs are YouTube video IDs that were auto generated by the app, however they are unique to the individual file. It was then decided to include the date the footage was recorded and the city it was recorded in - this information could also be embedded in the file. As was discussed in the metadata webinar, it is not necessary to include descriptive metadata in the file name but if it's decided that having such information in the filename is necessary, keep it as simple and structured as possible to support the organization/navigation of the

files.

Example 2: Folder structure for digital production elements

Not to create confusion, but this second example is primarily using raw digital camera footage. It gives a good sense of how to organize multiple elements that may be specific to one production or collection.

The sub folders under the Raw footage sub folder are utilizing the assigned camera file names - digital cameras auto create different metadata files and derivative files and its often important to keep all of that together and to maintain the original file name. This is also a good example, raw digital camera footage aside, of unique IDs being applied consistently to distinguish individual files - you will notice there is no descriptive information whatsoever.

	Unique							
	ID Title DateYYYY-MM-DD FootageType FileType/Version							
EXAMPLE 1: Assigning file types/versions	000001_Demo_2015-01-25_Raw_Preservation.mov 000001_Demo_2015-01-25_Raw_Mezzanine.mov 000001_Demo_2015-01-25_Raw_Access.mp4 000001_Demo_2015-01-25_Edit001.mov 000001_Demo_2015-01-25_Edit002.mp4 000001_Demo_2015-01-25_Raw_Transcript.doc							
EXAMPLE 2: Assigning file types/versions	<table border="1"> <thead> <tr> <th>Unique ID(item)_UniqueID(file)_Title_DateYYYY-MM-DD_FootageType_FileType/Version</th> </tr> </thead> <tbody> <tr> <td>000001_001_Demo_2015-01-25_Raw_Preservation.mov</td> </tr> <tr> <td>000001_003_Demo_2015-01-25_Raw_Mezzanine.mov</td> </tr> <tr> <td>000001_004_Demo_2015-01-25_Raw_Access.mp4</td> </tr> <tr> <td>000001_005_Demo_2015-01-25_Edit001.mov</td> </tr> <tr> <td>000001_006_Demo_2015-01-25_Edit002.mp4</td> </tr> <tr> <td>000001_009_Demo_2015-01-25_Raw_Transcript.doc</td> </tr> </tbody> </table>	Unique ID(item)_UniqueID(file)_Title_DateYYYY-MM-DD_FootageType_FileType/Version	000001_001_Demo_2015-01-25_Raw_Preservation.mov	000001_003_Demo_2015-01-25_Raw_Mezzanine.mov	000001_004_Demo_2015-01-25_Raw_Access.mp4	000001_005_Demo_2015-01-25_Edit001.mov	000001_006_Demo_2015-01-25_Edit002.mp4	000001_009_Demo_2015-01-25_Raw_Transcript.doc
Unique ID(item)_UniqueID(file)_Title_DateYYYY-MM-DD_FootageType_FileType/Version								
000001_001_Demo_2015-01-25_Raw_Preservation.mov								
000001_003_Demo_2015-01-25_Raw_Mezzanine.mov								
000001_004_Demo_2015-01-25_Raw_Access.mp4								
000001_005_Demo_2015-01-25_Edit001.mov								
000001_006_Demo_2015-01-25_Edit002.mp4								
000001_009_Demo_2015-01-25_Raw_Transcript.doc								
EXAMPLE 3: Assigning file types/versions	<table border="1"> <thead> <tr> <th>Unique ID(item)_UniqueID(file)_FileType/Version</th> </tr> </thead> <tbody> <tr> <td>000001_001_Preservation.mov</td> </tr> <tr> <td>000001_003_Mezzanine.mov</td> </tr> <tr> <td>000001_004_Access.mp4</td> </tr> <tr> <td>000001_005_Edit001.mov</td> </tr> <tr> <td>000001_006_Edit002.mp4</td> </tr> <tr> <td>000001_009_Transcript.doc</td> </tr> </tbody> </table>	Unique ID(item)_UniqueID(file)_FileType/Version	000001_001_Preservation.mov	000001_003_Mezzanine.mov	000001_004_Access.mp4	000001_005_Edit001.mov	000001_006_Edit002.mp4	000001_009_Transcript.doc
Unique ID(item)_UniqueID(file)_FileType/Version								
000001_001_Preservation.mov								
000001_003_Mezzanine.mov								
000001_004_Access.mp4								
000001_005_Edit001.mov								
000001_006_Edit002.mp4								
000001_009_Transcript.doc								

These examples continue with establishing file naming conventions, while explicitly adding on file types and or file versions in order to more easily distinguish the preservational level file from the file intended for a web upload, or the file that has been edited for a promotional effort. This example also includes a word document, in this case a transcript, to highlight that supplementary material can be associated with audiovisual material in the digital realm, just as it can be in the analog realm.

Example 1:

The file naming convention is established as **Unique ID_Title_DateYYYY-MM-DD_FootageType_FileType/Version**. In this instance, the unique ID assigned is the same ID that would appear on the physical item that has been digitized. It continues to be applied to all versions of the listed files in order to associate all related digital files to that single original item.

This example also includes minimal descriptive metadata, such as a brief Title and Date. The footage type is also indicated, using the term 'raw' for unedited recorded footage. File types are indicated as 'Preservation' (the primary archival file), Mezzanine (a term often used for a compressed but still high quality version of the primary file - commonly used for editing projects or projection), and Access (the most compressed file, intended for public access - whether for researchers or to upload to a website). File Versions are also indicated with two edited files, a very simplified way of tracking a newly created version of the original material.

Example 2:

This example uses the same structure as above, with the addition of a second unique ID intended to distinguish the individual files while still keeping them associated with the physical item unique ID.

The file naming convention is established as **Unique ID(item)_UniqueID(file)_Title_DateYYYY-MM-DD_FootageType_FileType/Version**

Example 3:

This example is even simpler, utilizing only the unique IDs, removing any descriptive information (including footage type), and utilizing only the File Type/Version designations. This assumes files can be easily associated to recorded information in an inventory to database, and allows for quick navigation assuming the user knows which file type or version they are looking for.

Making Sure Your New Digital Files Meet Your Specifications

Quality Control -

Procedures need to be established so that methods are consistent. QC must occur within a reasonable timeframe after files have been created (especially if created by outside vendor).

Responding to Vendor Errors -

Files from vendors must be reviewed immediately in case there are problems with the file. Errors should be clearly documented before communicating with the vendor. Make sure to know what the vendor's policy is in dealing with errors (will they re-do for free, etc.).

Quality control procedures should be adopted and consistently applied across all preservation projects. Quality control actions should take place within a reasonable window of time, following preservation transfer, to allow you to request mediation- ideally no less than 30 days. Digitized files that remain unreviewed for a long period of time introduce the possibility of file corruption and reduce the possibility of mitigating errors and negotiating with vendors.

- **Is the filename correct?**
- **Open/play the file** - confirm the file works!
- **Spot check content** -
 - Content correct in relation to inventory? (possible content was unknown!)
 - View/listen to file in at least 3 spots - (beginning, middle, and end)
 - Audio in sync? other errors?
 - Video error from analog or transfer?
- **Confirm Technical Specifications**
 - Correct file format?
 - Correct aspect ratio?
 - Audio and color information correct?
 - Correct additional embedded metadata?
- **How to deal with errors found...**

QC Procedures to implement should include:

- Confirming the assigned file name is correct.
- Confirming the file is playable - can actually open with needed video players on operating systems currently in use.
- Scanning through file to confirm -
 - Does the content match what was inventoried (though it's possible content was previously unknown)
 - Confirming that audio is in sync with the image, or generally that audio sounds as expected
 - Because it is often unlikely someone can sit and watch or listen through every single new file from beginning to end, it is considered good practice to choose at least 3 points in a file (depending on the duration) to check content for up to 2-5 minutes at a time
 - Yes, you may miss something, but you have to weigh resources and time.
 - Any potential errors logged need to be identified either as originating on the analog item or originating during

- the digitization process - it is not always possible to determine this, especially if you cannot compare the digitized content to the analog content.
- Confirming other technical specifications
 - Correct file format?
 - Correct aspect ratio? Correct frame rate? Correct audio and color information?
 - in the case of film, you should decide if a preservation file will contain any level of color correction or scratch reduction. You may decide to only include such alterations in reference or production files.
- Procedures for following up on problems found:
 - If the file has an unstable signal, the original asset can be checked to see if the problem is in the asset or in the transfer. Other issues may require re-digitizing the asset. If you are working with a vendor, the issue should be documented and reported to the vendor, who may need to re-transfer the asset.

Responding to Vendor Errors:

Review files from a vendor immediately
(recommended window is 30 days)

- **Receiving staff should check for:**
 - Correct number of files
 - Confirm requested technical specifications
 - Confirm requested additional metadata

Vendor errors to note are same as in-house work:

- sync issues not present in the original asset
- discrete audio channels that have been mixed
- misidentified or improperly attributed file names
- Etc. etc.

QC Procedures to implement should include: Responding to Vendor Errors

- **General Recommendations:**
 - Digitized assets received from a vendor should be reviewed immediately (recommended window is 30 days) in the event that any errors are discovered. Receiving staff members should check for the correct number of files, correct codecs and wrappers, and perform a quality control review of each file. This can consist of viewing the file or running it through a program such as QCTools.
 - Vendor errors to note include sync issues not present in the original asset, maladjusted chrominance/luminance, discrete audio channels that have been mixed, misidentified or improperly attributed file names, missing files, etc.
 - Errors that can be mediated—but have not—

- include head clogs, video dropout, flagging, vertical roll, and tracking errors.
- Errors need to be clearly documented in order to communicate them to the vendor. If the error in question is related to specifications established with the vendor prior to the project's initiation (file capture preferences, requested derivatives), this documentation should be referenced when communicating discovered errors with the vendor.
- Generally speaking, vendors have a window of time in which they need to be made aware of errors which will allow for re-transfers free of charge. It is recommended that the unit discuss error procedures with a vendor before the project begins.

Ideal: Enough digital storage available for files in advance, including for back-up

- Plan in advance to avoid maxing out drives or having to halt digitization processes until new storage can be added
- Digitizing in-house can allow for increasing storage incrementally
- Material is sent in bulk to a vendor requires storage to be available as soon as vendor completes work

Estimating Digital Storage: Film

35mm

File Format (image): DPX

Codec: Uncompressed

Resolution: 4k
(4096 pixels)

Bit depth / Sample rate:
10 bit RGB

Storage estimate: ~4.3 TB/hr

File Format (audio): BWA

Codec: PCM

Bit depth / Sample rate: 24-bit, 48kHz

Storage estimate: ~1 GB/hr

16mm, 8mm, Super 8

File Format (image): DPX

Codec: Uncompressed

Resolution: 2k
(2048 pixels)

Bit depth / Sample rate:
10 bit RGB

Storage estimate: ~765 GB/hr

File Format (audio): BWA

Codec: PCM

Bit depth / Sample rate: 24-bit, 48kHz

Storage estimate: ~1 GB/hr

Estimating Digital Storage: Video

½" open reel, U-Matic, Betacam, Hi-8, VHS...

File Format:

QuickTime (MOV)

Codec (Image): Uncompressed 4:2:2

Bit depth / Sample rate:

(recommended) 10-bit

(minimum) 8-bit

Codec (audio): PCM

Bit depth / Sample rate:

(recommended) 24-bit, 48kHz

(minimum) 16-bit, 48kHz

Storage estimates:

10-bit ~100 GB/hr

8-bit ~76 GB/hr

File Format:

Matroska (MKV)

Codec (Image): FFV1 version 3

Bit depth / Sample rate:

(recommended) 10-bit

(minimum) 8-bit

Codec (audio): PCM

Bit depth / Sample rate:

(recommended) 24-bit, 48kHz

(minimum) 16-bit, 48kHz

Storage estimates:

10-bit ~40 GB/hr

8-bit ~40 GB/hr

Estimating Digital Storage: Video

1/2" open reel, U-Matic, Betacam, Hi-8, VHS...

File Format:

QuickTime (MOV)

Codec (Image): ProRes 422

Bit depth / Sample rate:

(recommended) 10-bit HQ

(minimum) 8-bit

Codec (audio): PCM

Bit depth / Sample rate:

(recommended) 24-bit, 48kHz

(minimum) 16-bit, 48kHz

Storage estimates:

10-bit ~27 GB/hr

8-bit ~18 GB/hr

DV, MiniDV, DVCam

File Format:

Quicktime (MOV)

Codec (Image): DV-NTSC (DV-25)

Bit depth / Sample rate: 8-bit

Codec (audio): PCM

Bit depth / Sample rate:

16-bit, 48kHz

Storage estimate:

8-bit ~13 GB/hr

Estimating Digital Storage: Audio

1/4" open reel, Compact audio cassette

File Format:

Broadcast Wave (BWF)

Codec: LPCM

Bit depth / Sample rate:

(recommended) 24-bit, 96kHz

(minimum) 24-bit, 48kHz

Storage estimate: ~2 GB/hr

File Format:

Waveform (WAV)

Codec: LPCM

Bit depth / Sample rate:

(recommended) 24-bit, 96kHz

(minimum) 24-bit, 48kHz

Storage estimate: ~1 GB/hr

- Did you begin the project with enough storage?
- Is storage filling up faster than anticipated?
- Can you make adjustments to the established workflow for future projects?
- Do you have the ability to deal with sudden hard drive failures (i.e. purchase immediate replacements?)
- It is impossible to anticipate every mistake.
- It is important to be flexible, especially when dealing with ever changing technology.

Track this all in your inventory!

The screenshot shows a spreadsheet application with a toolbar at the top. The active cell is A1, containing the text 'object ID'. The table below has the following structure:

	A	BA	BB	BC	
	object ID	Digital Preservation / Master File name	Master File size	Digital reference file name	Referer
1					
27	2015.0015	2015_0015FootageRegistration Desks_master.mp4		2015_0015FootageRegistration Desks_ref.mp4	
28	2015.0007				
29	2015.0040				
30	2015.0041				
31	2015.0035				
	2015.0014	2015_0014_SelmaRaceCarsEt c_master.mov		2015_0014_SelmaRaceCarsEt c_ref.mp4	

Last but not least, keep track of your files. The best way to do this is have consistent file naming conventions, and to track them in the inventory. You can simply add a column that notes the preservation file names, or you can add further data, like file size.

We Could Go On!!

Questions?

Thank you!



Community Archiving Workshop

