

Fundamentals of AV Preservation

Northeast Document Conservation Center | NEDCC
Andover, Massachusetts | 2022

PREFACE

NEDCC is pleased to offer the Fundamentals of AV Preservation textbook for self-study to anyone with internet access. An HTML version is available online at www.nedcc.org/av-textbook, and downloadable PDFs can be found at www.nedcc.org/publications. Covering the core topics in caring for and reformatting audiovisual collections, this resource supports cultural heritage professionals in their efforts to steward audiovisual materials.

- Chapter 1: Care and Handling of Audiovisual Collections
- Chapter 2: Inventory and Assessment
- Chapter 3: Planning, Preparing, and Implementing Reformatting Projects
- Chapter 4: Managing Digital Audiovisual Collections
- Chapter 5: Disaster Preparedness and Response
- Glossary

Credits

The content for each chapter of the Fundamentals of AV Preservation textbook was created in 2017 by staff members of AVPreserve (now AVP) and edited by NEDCC staff members. The textbook served as the foundation of a multi-session, instructor-facilitated online course launched by NEDCC in 2017. The development of the course and textbook were subsidized by the National Endowment for the Humanities, and in 2022, additional NEH funding supported the transformation of the HTML textbook into downloadable PDFs.

AVPreserve — Authors

Chris Lacinak, *President (Ch. 1 & 3)*
 Rebecca Chandler, *Consultant (Ch. 1 & 2)*
 Amy Rudersdorf, *Senior Consultant (Ch. 4)*
 Kara Van Mallsen, *Partner and Senior Consultant (Ch. 5)*

NEDCC — Editors

Ann Marie Willer, *Director of Preservation Services*
 Sean Ferguson, *Preservation Specialist*
 Becky Geller, *Preservation Specialist*
 Frances Harrell, *Senior Preservation Specialist*
 Kim O’Leary, *Technology and Events Coordinator*
 Danielle Spalenka, *Preservation Specialist*



NATIONAL
 ENDOWMENT
 FOR THE
 HUMANITIES

CONTENTS

i	Preface
ii	Contents
1	Chapter 1: Care and Handling of Audiovisual Collections
1	Care and Handling
2	Grooved Media
7	Magnetic Media
12	Film
15	Optical Media
17	Shipping of Media Carriers
18	Additional Resources
21	Chapter 2: Inventory and Assessment
21	Inventory
28	Selection for Digitization
29	Prioritization for Digitization
33	Chapter 3: Planning, Preparing, and Implementing Reformatting Projects
34	The Digitization Signal Path
35	The Request for Proposal (RFP) Process
36	Internal Project Planning
39	Drafting a Statement of Work
58	Post-Digitization
59	Chapter 4: Managing Digital Audiovisual Collections
60	Digital Preservation Concepts
63	Organizational Infrastructure
65	Storage Infrastructure
70	Active Management
74	Metadata
78	Planning for Obsolescence
80	Prioritization and Phasing
83	Chapter 5: Disaster Preparedness and Response
83	Disaster Prevention and Mitigation
87	Disaster Planning
90	First Response Steps
96	Glossary

CHAPTER 1

CARE AND HANDLING OF AUDIOVISUAL COLLECTIONS

by Chris Lacinak, President, AVPreserve and Rebecca Chandler, Consultant, AVPreserve

THIS CHAPTER DISCUSSES THE CARE AND handling of physical audiovisual collections, which is an essential component of any preservation strategy. While proper care and handling serve as great tools for combatting degradation and damage, they do not address obsolescence in any way. It is clear that preservation and access of information recorded on physical audiovisual media requires digitization. It is also clear that the need is great and that prioritization for digitization over a longer period of time is the reality for many organizations. Until digitization can occur, proper care and handling of the original recording is the primary mechanism of preservation and helps ensure that the information recorded on the physical media will be able to be digitized with the highest degree of integrity possible.

However, even following digitization the need to care for the physical media containing the original recording persists, as it plays an integral role in the overarching preservation strategy of the recording. There are multiple scenarios that may lead to utilizing that physical media post-digitization. These could and have included disasters resulting in loss of the digital copies, critical errors in digitization or digital preservation implementation, or technological improvements that yield significantly better reproduction and digitization. Central to preservation is the concept of risk management and proper care and handling of the original recordings, before and after digitization, which greatly mitigates the risk of loss.

SECTION 1: CARE AND HANDLING

Physical care

There are two primary components that make up proper care and handling: storage of the objects and handling of the objects by staff who interact with the media. Storage factors are explored in depth in our Preservation 101 guide¹ and will be related to specific AV materials throughout the course of this chapter.

Many materials may end up in multiple storage circumstances throughout their life. The more time they spend in advantageous conditions, the greater the chance of significantly slowing down degradation mechanisms.

There are a number of general practices that improve the longevity of media collections. These include limiting exposure to:

- Heat
- Humidity
- Liquids
- Dirt, dust, and particulates
- Light
- Mold and fungus
- Pollutants and contaminants

1 <https://www.nedcc.org/preservation101/session-2/2storage-environment>

You should also be careful to avoid:

- Handling portions of the media where the signal is recorded without appropriate gloves
- Putting weight on the media
- Leaving media unhoused
- Using machine transports that are not functioning properly
- Using machine transports that have not been properly calibrated and aligned

Storage approaches specific to media types will be addressed later in this chapter.

Handling of materials

Throughout an object's time held within your archive, it is likely to be handled by multiple people for different reasons. At each of these points, how the materials are handled and treated has implications for the well-being of the recording stored on the original media. Salient activities performed in the life of an object that involve handling include objects being:

1. Accessioned
2. Moved
3. Described and assessed
4. Inspected, surveyed, and inventoried
5. Retrieved and returned to storage for collection management or patron reasons
6. Treated and repaired
7. Reproduced and digitized

Determining how an object should be stored and handled is dependent on the physical and chemical composition of the media, as well as the way information is recorded onto and played back from the media. It is not critical to understand the finer details of every format and media type in your collections in order to properly care for them, but having a basic understanding will help bolster your knowledge of how to care for an object once you have properly identified the format. Understanding these aspects and knowing how to act on them is the subject matter of the remaining sections of this chapter. Please note that for

the sake of simplicity and understanding, this text will focus on the most predominant material types and formats only, including magnetic media, grooved media, optical media, and film. Fringe media types and esoteric formats are not within the purview of this chapter.

Knowing your collection

When looking at particular physical audiovisual media, there are three primary characteristics that affect proper storage, care, and handling practices. These are carrier types, media types, and the method of reproduction.

Although this text presents a number of reasons to be cautious when handling media objects, it is also important to become comfortable with these items. Reading about how handling can cause damage is much different than experiencing it. If you have access to or can purchase disposable, non-unique material, try bending, stretching, scratching, tearing, and breaking it to become more familiar with the look and feel of these objects and the damage that can occur to them.

The following section is an overview of media objects likely to be held in your archive. Each media type's materials and associated preservation issues will be discussed, but specific format identification will be discussed in Chapter 2.

SECTION 2: GROOVED MEDIA

History

DISCS

Discs, also commonly referred to as records, have a long history spanning from the late 1800s with the introduction of Berliner discs to the present with the ongoing manufacturing of vinyl records. While the form has essentially stayed the same over time, there have been several variations in material composition and construction, method of recording, and recording specifications. For simplification, this text will focus on the three types of discs that you are most likely to encounter in collections: shellac, instantaneous, and vinyl.

Shellac discs were used in the earliest days of recording. Discs were originally recorded using an acoustic-mechanical method. Think here of the pictures you have undoubtedly seen of a group of musicians or a speaker in front of a horn.

The acoustic energy was captured by the horn which would move a diaphragm at the small end of the horn, attached to a cutter that would cut into the disc. This disc would then be replicated for commercial distribution; these discs were referred to as shellac. In the mid-1920s, the creation of the microphone led to the recording of discs using an electro-mechanical recording method, converting electrons to mechanical energy used to cut into the disc. It was also about this time that the industry involved in the creation of these discs standardized playback speed at around **78 RPM**. Prior to this, speed varied quite a bit. Shellac discs are identifiable by their uniform thickness throughout the disc, their rigidity, and a matte finish when compared to vinyl and lacquer. The most common diameters are 10" and 12". Although the distinction is made between acoustic-mechanically recorded and electro-mechanically recorded discs, this has little bearing on how these discs are stored and handled and despite the variance in recording method, they will share a common reproduction method.

In the 1930s, instantaneous discs were introduced, fulfilling the need to create a recording that could immediately be played back. They did not require a replication process and they were not intended for distribution. This spawned a plethora of documentary recordings, capturing everything from ethnomusicological field recordings to home recordings, interviews, oral histories, off-air radio recordings, and more.

Vinyl discs were introduced in the late 1940s. Similar to shellac recordings, an original is recorded and then used to start an intricate process that results in the creation of copies for distribution. Compared to shellac and instantaneous discs, vinyl is flexible, thin, and lightweight. A vinyl disc's surface is softer, making it more prone to scratches.

CYLINDERS

Cylinders date from the late 1870s with the invention of the phonograph and were common through 1929 when production ended in favor of discs. The composition of

FIGURE 1.1

President Harding speaking into recording apparatus.

Credit: Library of Congress <https://www.loc.gov/item/96522821>



cylinders varies greatly within this time period—from tinfoil covered metal tubes to molded wax to celluloid (a type of plastic). The method of recording remained constant; however, cylinders, like early discs, were recorded using an acoustic-mechanical process. A standard cylinder is typically 2.25" in diameter and between 4" and 4.25" in length, but specialty sizes can range from 1.33" to 5" in diameter and 0.5" to 8" in length. Cylinders may or may not have a cardboard core.

Playback speed varies greatly depending on the manufacturer and technology for commercial recordings and the equipment and operator for home recordings. Rotation speeds for commercial recordings include 120 RPM, 144 RPM, and 160 RPM. Cylinders with differing diameters must be played back on a matching **mandrel**.

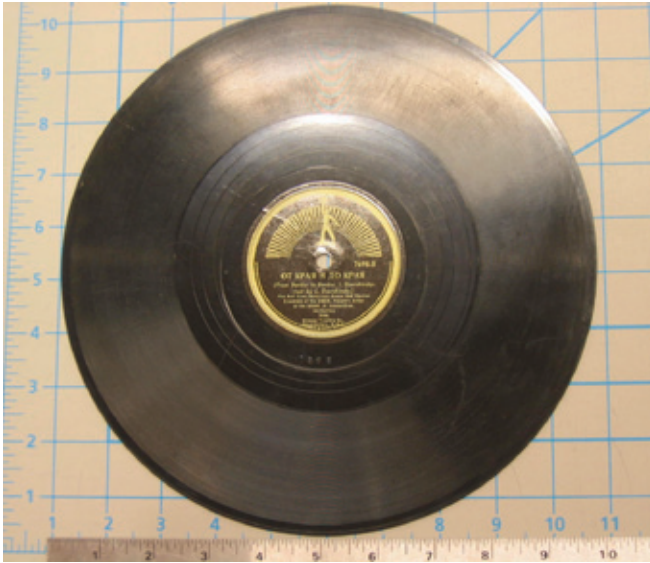
Materials and preservation risks

DISCS

Shellac discs have no significant degradation mechanisms under normal conditions and are a robust format in all regards except one. Their rigidity and composition make them prone to breakage through mishandling. Excessive pressure or dropping a shellac disc will almost certainly result in breakage.

FIGURE 1.2
Shellac disc.

Credit: AVP



There are three main types of instantaneous discs that are found in collections. These are aluminum-based lacquer discs, glass-based lacquer discs, and aluminum discs. Glass-based lacquer discs replaced aluminum during World War II when the demand for metal became great.

FIGURE 1.3
Laquer disc.

Credit: AVP

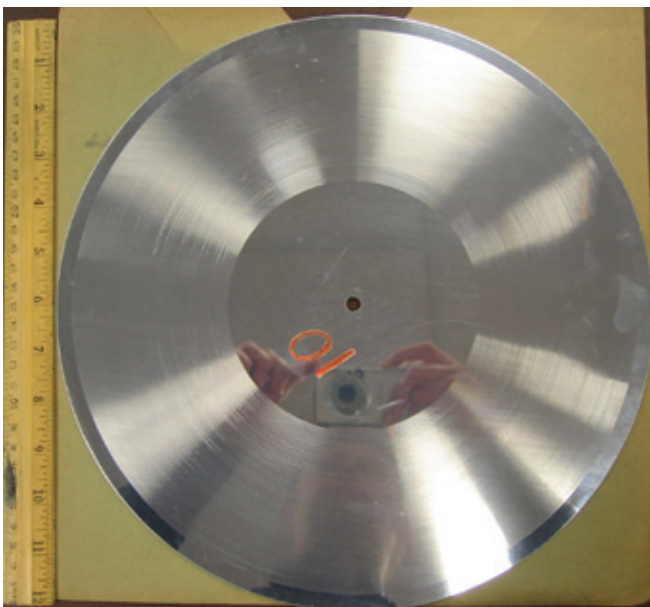


FIGURE 1.4
Laquer disc.

Credit: AVP



Lacquer discs have multiple degradation mechanisms. Over time, the lacquer layer can undergo chemical and dimensional shifts that cause it to separate from the base glass or aluminum layer, leading to what is referred to as **delamination**.

FIGURE 1.5
Delaminated disc.

Credit: AVP



Once delamination has occurred there is no way repair the damage; however, there are warning signs that can alert you in time to successfully transfer the signal on those discs about to suffer delamination. Early signs consist of a whitish film covering the disc, consisting primarily of palmitic acid. Palmitic acid is a chemical used in the creation of lacquer discs; this leaching out of the palmitic acid is a major contributor to delamination. At this stage, where the palmitic acid is leaching out, but the lacquer has not yet begun to crack, it may still be possible to have the disc cleaned by a professional and the signal transferred from the disc.

FIGURE 1.6
Palmitic acid, an early sign of delamination.

Credit: AVP



Lacquer discs are also fragile and rigid, making them prone to breakage. This is particularly true for glass-based lacquers.

Aluminum discs do not tend to have degradation mechanisms, but they do have one issue that can manifest through mishandling and improper storage. The grooves cut into aluminum discs can be shallow due to the way they were recorded. Combined with the softness of aluminum, significant pressure against the side of an aluminum disc can make the grooves shallower than when originally recorded. This diminishes the integrity of the original recording and sometimes may render it unplayable.

Improper storage, care, and handling practices may result in the following disc issues, impairing reproduction:

- Matter such as dirt, dust, hair, and other particulates can scratch the surface of the disc, altering the groove or creating separation between the stylus and the groove.
- Fingerprints, where they exist on the surface of a disc, can speed up or cause degradation.
- A layer of white film (palmitic acid) covering the disc, indicating the beginning stages of delamination, which must be cleaned off before reproduction.
- Delamination, resulting in partial or total loss of the signal.
- Warping, causing audio artifacts during impact due to speed variation.
- Flattened groove walls from too much pressure on the side of the disc, resulting in distorted sound and potentially an inability to track the groove with a stylus.
- Breakage, resulting in partial or total loss of the signal.
- Scratches from improper handling and/or transport.

CYLINDERS

In general, cylinders are fragile, and care should be taken not to drop or bump them. Prior to 1902, when more rigid wax formulas were introduced, cylinders were typically made from brown wax. Brown wax cylinders are extremely soft and fragile, are prone to surface scratching and groove wear, and are susceptible to fungal growth. Home recordings made on brown wax are especially high risk; users often shaved off layers of wax in order to reuse the cylinders, making these particular cylinders extremely thin and prone to breakage. Improper storage and handling of these softer wax cylinders may also result in a warping of the grooved surface, resulting in sound artifacts due to speed variations.

Improper storage, care, and handling practices may result in the following cylinder issues, impairing reproduction:

- Matter such as dirt, dust, hair, and other particulates can scratch the surface of the cylinder, altering the groove, or creating separation between the stylus and the groove.
- Fingerprints, where they exist on the surface of a cylinder, can speed up or cause degradation.

- Warping, causing audio **artifacts** during impact due to speed variation.
- Breakage, resulting in partial or total loss of the signal.
- Scratches from improper handling and/or transport.

Reproduction methods

DISCS

Presently, there are two different ways that the sound recorded on discs is reproduced. The traditional method consists of using a turntable, otherwise known as a record player, which utilizes an electro-mechanical method of reproduction. A stylus of the appropriate size and shape is selected according to the groove dimensions. Any movement that the stylus makes which is not part of the actual recorded signal adds noise and distortion to the audio reproduction. Given the miniscule movements involved in this process, even a small amount of particulate matter or a scratch can significantly impact the resulting sound. It is easy to see how delamination, breakage, or significant scratches may make it impossible to play a disc back.

The other method that is used for reproduction involves scanning or imaging the groove. Once the image of the groove is captured, the mechanical variation is able to be identified and calculated. These calculated variations are then plugged into an algorithm that generates a digital audio file representing the originally recorded signal. This method is particularly useful when discs are damaged or degraded to the extent that a stylus would not be able to stay in the groove of a disc.

Reproduction via turntable and via imaging are significantly different, but the recorded signal in the original disc results in similar audio results when particulate matter, scratches and other things obscure the groove. Ultimately, careful storage and handling practices are important to the integrity of the content regardless of which reproduction method you choose.

CYLINDERS

Like discs, there are currently two playback mechanisms for cylinders: the traditional electro-mechanical method and the imaging method. First, a correctly sized mandrel must be chosen for playback; the diameter of the mandrel will correspond to the interior diameter of the cylinder. Then an appropriately sized and shaped stylus must be chosen based on the groove dimensions. As with discs, the movement of the stylus generates electrons and any additional movement caused by scratches or particulate matter will affect the sound.

Cylinders differ from most discs, however, in that their grooves are vertical as opposed to discs' lateral grooves. The vertically oriented grooves create hill and dales which must be imaged using 3D technology in order to obtain the information necessary to generate a digital audio file representing the originally recorded signal.

Best practices for storage and handling²

ENVIRONMENT

- Discs and cylinders should be stored at temperatures between 33 and 54 degrees Fahrenheit, with RH between 30%-50%. Temperatures should not fluctuate more than ± 2 degrees within a 24-hour period. RH should not fluctuate more than $\pm 5\%$ within a 24-hour period.
- Discs and cylinders should be kept away from sources of heat, including sunlight or room lighting.

HOUSING

- Discs and cylinders should be stored on their edge or end so that they stand vertically. This will avoid pressure that may cause breakage, warping, groove wall flattening, or otherwise place undue stress on the surface of the disc.
- Discs should be in sleeves and housing that protect them from external elements and do not generate particulates or react chemically with the disc.

² The storage and handling best practices used in this chapter are based on various ISO standards. The reality is that any improvement in environmental conditions will be beneficial to your collections.

- Cylinders should be removed from any original cardboard housings. These cardboard housings are known to foster mold growth in humid storage conditions.
- Cylinders should be stored in appropriate housings that protect them from external elements and do not generate particulate or react chemically with the disc. Cylinders should be housed in specially designed archival boxes containing a foam tube that fits within the cylinder, keeping the cylinder steady without allowing the grooves to touch the interior of the box.

HANDLING

- Handle the disc from the edges only. Do not touch the surface of the disc without using non-abrasive, non-shedding gloves.
- Handle the cylinder from the ends or interior only. Do not touch the surface of the cylinder without using non-abrasive, non-shedding gloves. To pick up a cylinder, insert two fingers into the cylinder and spread your fingers.

SECTION 3: MAGNETIC MEDIA

History

TAPE

The ability to record audio on magnetic tape in a consistent way first came on the scene in Germany in the late 1920s. It wasn't until the late-1930s that the tapes could be mass produced and the playback devices were stable enough for general use. The quality of recordings began quite poor but had improved dramatically by the early 1940s. However, it wasn't until after WWII that magnetic tape recorders made it to the United States. Singer and actor Bing Crosby heavily invested in magnetic tape technologies and helped to bring the medium to broadcast radio in the United States. Magnetic tape recordings were a huge improvement over the transcription discs that had previously been used in broadcast radio: they were of better quality and easier to edit.

Magnetic tape recording is built on the same basic concepts on which the telephone was created: mechanical

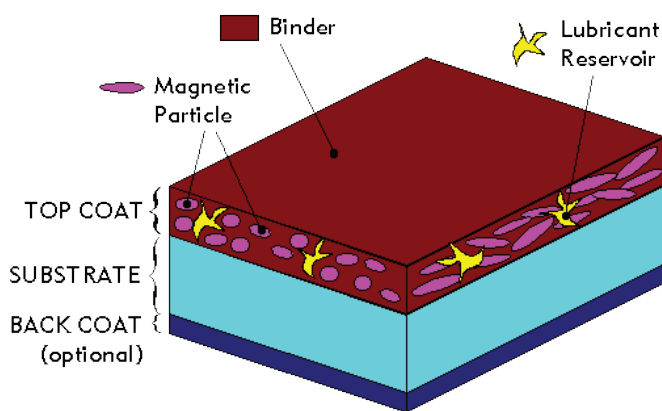
sound waves are converted into electrical current, transported through cables, then converted back to sound waves. Magnetic tape recordings add an additional step—a magnetizing coil is used to record the signal. In the early days of magnetic tape recording, the signal was recorded onto paper coated with iron oxide; the paper was later replaced with the acetate and polyester base with which we are all more familiar.

For all tape there is one side that contains the magnetic particles on which information is recorded. These magnetic particles are adhered either through use of a binder system or the **metal evaporated tape process**. The opposite side of the tape will either be the bare side of the **substrate** or a **back coat** that is added on.

Soon after the introduction of magnetic tape audio recordings to the United States, the same principle was applied to video. Video signals are similar to audio signals, but have a much higher bandwidth, meaning a lot more information must be recorded in order to properly reproduce the signal. This hurdle was overcome with the invention of the spinning recording head that allowed for much more information to be recorded and read. It wasn't until

FIGURE 1.7
Cross Section of Magnetic Tape.

Credit: The Commission on Preservation and Access and the National Media Laboratory, *Magnetic Tape Storage and Handling: A Guide for Libraries and Archives*, https://www.clir.org/pubs/reports/pub54/2what_wrong.



much later that further advancements and improvements were made so that the medium became compact and easy enough to handle that video was able to enter the home market.

WIRE

Magnetic wire recordings actually pre-date magnetic tape recordings. They were first invented in 1898 and were meant to record dictation. Wire recording did not really catch on as a major format until the mid-1940s when technical improvements made it both better quality and more affordable. It began to be marketed not only for office use, but also for home recording. By the mid-1950s, however, magnetic tape recorders had caught up in quality and price and quickly surpassed the wire recorder in use.

Again, the recording method recalls the principles of the telephone. The electrical audio signal magnetizes the wire as it passes quickly over the recording head. The magnetized signal corresponds to the intensity and polarity of the original signal. During playback, the wire is run over the same head, but since there is no signal being supplied during playback, the changing magnetic field present on the wire reproduces the original signal.

MATERIALS AND PRESERVATION RISKS

The base materials for tape fall into one of four basic categories. These are acetate, polyester, PVC, and paper. Acetate, PVC, and paper will only be found on open reel audio. All audiotape housed in cassettes and all videotape is polyester based.

Paper-based tape is notable in look and feel and is somewhat rare to come across. The associated handling follows basic common sense regarding the sensitivities of paper-based materials and the possibility of tearing. Acetate can be identified by holding a reel of tape sideways between your eye and a light. If it is translucent then you are holding an acetate tape. If it is opaque, then it is either polyester or PVC.³

When put under excessive tension, acetate tape will simply break. Acetate that has degraded and become brittle will be prone to breaking more easily. Breaks in paper and acetate tape can be easily repaired using **splices**, but a significant number of breaks is a sign of an issue with the tape or with the playback machine and will cause notable audio artifacts. PVC- and polyester-based tapes, when put under excessive tension, will stretch and become

FIGURE 1.8

A reel of acetate tape will appear translucent when held sideways between your eye and a light.

Credit: AVP



deformed resulting in audio artifacts, as well. For analog audio, stretching will cause audio artifacts. For analog videotapes, digital audiotapes, and digital videotapes, stretching can lead to artifacts or a total inability to retrieve a signal in the stretched portions. If the tape is visibly stretched, it can also damage the transport of the playback machine. An issue that primarily impacts polyester and PVC tapes is failure of the **binder system**. The most common problem that falls under this category of failure is known as **sticky shed syndrome**, which results in low cohesion and high friction during transport. When the tape is put onto a machine and wound or played back, sticky residue comes off onto the transport of the machine and the increased friction in the transport causes audio distortion upon playback. Excessive winding of a tape that experiences binder system failure can lead to stretching of the tape and edge damage.

Carrier types and preservation problems

Tape, whether audio or video, analog or digital, comes in two basic forms: open reel and cassette. Open reel tape has tape wound around a hub with **flanges** on the sides, with tape ranging in width from ¼" to 2". Audiotape is sometimes stored on a hub with no flanges attached. This is referred to as a “pancake” and requires special

³ It is difficult to differentiate between polyester and PVC tapes. For the purposes of care and handling, their needs are similar, meaning you do not need to identify these to properly care for them.

attention because it is particularly prone to coming unwound.

Cassettes come in a range of sizes, typically housing tape widths from 1/8" to 3/4". There are also cartridge formats that are more blockish in appearance, but the construction is more or less the same.

An issue pertaining primarily to audio open reel is that of degrading splices. This pertains primarily to audio open reel because it is rare for cassette-based formats to have splices and uncommon for video open reel to have splices.

FIGURE 1.9
A tape with sticky shed syndrome has left this sticky residue on the playback machine.

Credit AVP

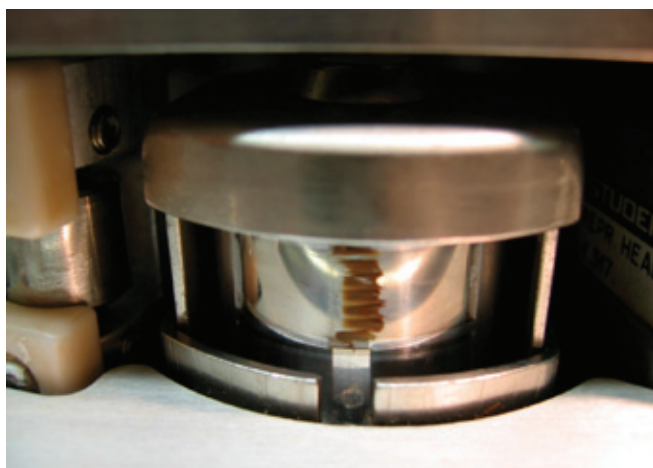


FIGURE 1.10
Diagram of a Tape Reel.

Credit: The Commission on Preservation and Access and the National Media Laboratory, *Magnetic Tape Storage and Handling: A Guide for Libraries and Archives*, https://www.clir.org/pubs/reports/pub54/2what_wrong/.

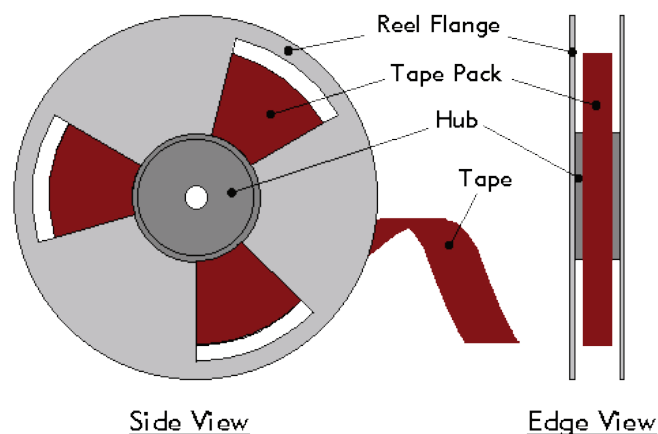


FIGURE 1.11
Unwound magnetic tape.

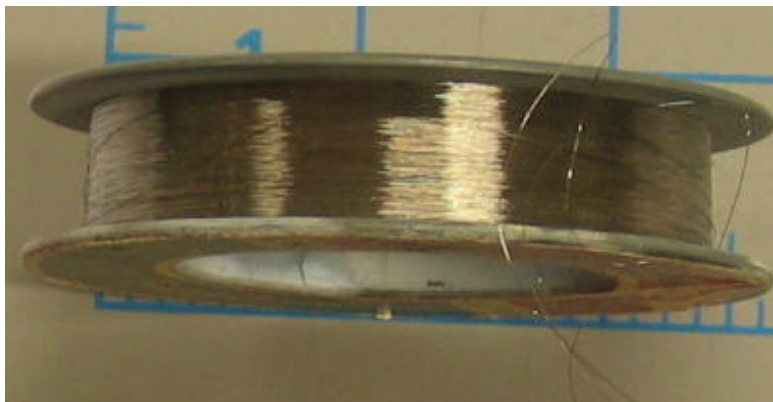
Credit: AVP



If there are splices in video open reel they tend to be few in number. Audio open reel tapes on the other hand have been known to have anywhere from none to hundreds of splices. Issues with degrading splices can range from splicing tape that simply dries up and breaks to extruding adhesive, causing multiple adjacent layers to stick together.

Wire recordings are specific to audio and typically consist of stainless-steel wire that is between .004 and .006 inches in diameter wrapped around a spool. Spools of wire are commonly 3" in diameter and .75" wide, although these dimensions can vary depending on when and for which application the wire was made.

The wire itself does not have degradation mechanisms, as it is made of steel. The biggest risk factor to the physical state of wire recordings is becoming tangled in transport, which can lead to breakage or the need to cut out sections in order to resolve the tangle. For this reason, it is important to make sure wire recordings are played back and wound on a transport that is known to be in good working order. Adhesives from splicing are not a concern, because the process for wire recordings involves tying a square knot.



FIGURES 1.12 & 1.13

Wire recording.

Credit: AVP

Improper storage, care, and handling practices may result in the following tape and/or wire issues, impairing reproduction:

- Deposits of matter such as dirt, dust, hair, and even fingerprints that create separation between the tape or wire and the head.
- Edge damage and tape deformation caused by mis-handling, degradation, poor storage, or a bad transport. Even when slight, this type of damage can cause variation in the tape's contact with the head, resulting in the creation of artifacts.
- Broken or poorly repaired splices on both tape and wire.
- Adhesive extruding from splices and adhering to adjacent layers of tape.
- Binder system failure, including sticky shed syndrome, causing poor cohesion and high friction during transport of tape.
- A broken cassette or reel.
- Tangled wire.
- Magnetic particles that were originally recorded to tape, but scratched off or otherwise removed .
- Magnetic particles that were originally recorded but were exposed to a magnetic force that changed their magnetic state on both tape and wire.

Reproduction methods

The reproduction method for magnetic tape and wire is electromagnetic, meaning that the tape or wire reproducer head in the playback machine converts the magnetic energy stored on the tape or wire to electrons or voltage which is ultimately converted back to light and sound. It is critical that there be good and precise contact between the tape/wire and the playback head in order for the magnetic energy to be properly read and converted to electrons. It is also critical that the magnetic particles containing the recorded information remain in place and do not have their magnetic state altered.

Best practices for storage and handling

TAPE

Environment

- Store at temperatures between 32 and 61 degrees Fahrenheit,⁴ with a relative humidity (RH) between 30%-50%. Temperatures should not fluctuate more than ± 3 degrees within a 24-hour period. RH should not fluctuate more than $\pm 5\%$ within a 24-hour period.⁵
- Keep away from magnets, items containing magnets, and items generating magnetic force.

Housing

- Store on edge standing vertically.
- Containers should support reels by the hub.
- Place in housings that protect from external elements.
- Open reel tape should be stored tails out.
- The end of an open reel tape should be taped down, either to itself or the reel flange, using “hold down” tape intended for this purpose.
- Use only splicing tape for splices; use only hold down tape to secure tape ends.
- Wind cassettes all the way to one end.
- Avoid excessive pressure (enough to press a reel flange or cassette housing against the tape) on any side of the container or housing.

Handling

- Wear non-abrasive, non-shedding gloves if touching areas of the tape containing recorded information.

WIRE

Environment

- Store at temperatures between 33 and 54 degrees Fahrenheit, with RH between 30%-50%. Temperatures should not fluctuate more than ± 3 degrees within a 24-hour period. RH should not fluctuate more than $\pm 5\%$ within a 24-hour period.
- Keep away from magnets, items containing magnets, and items generating magnetic force.

Housing

- Place in housings that protect from external elements.

Handling

- Do not touch areas of the wire containing recorded information without using non-abrasive, non-shedding gloves.
- Tie down end to keep it from coming loose and unwinding.
- Use only a square knot for splices.



FIGURES 1.14 & 1.15

Store magnetic tape on edge, standing vertically.

Credit: AVP

4 Acetate-based magnetic tapes benefit from temperatures at the lower end of this range, polyester from the higher end.

5 ISO 18934:2011 http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=55518

SPLICING WIRE

If it is desired to remove or insert a section for editing reasons, or if the wire is accidentally broken, the ends may be spliced by tying them together with a simple square knot as shown. Pull the knot tight and cut off the loose ends close to the splice. The knot will pull across the recording head without catching.

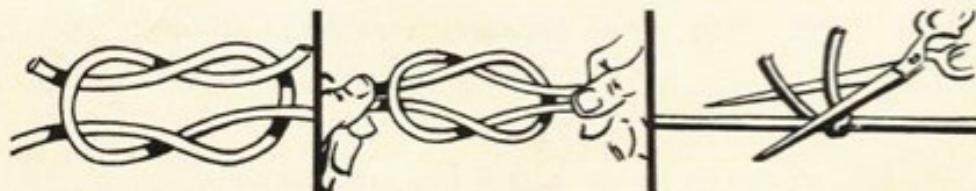


FIGURE 1.16
A splice in a magnetic wire recording.

Credit: Webster-Chicago Electronic Memory Portable Wire Recorder brochure, held by the Museum of Magnetic Sound Recording, Austin, TX.

SECTION 4: FILM

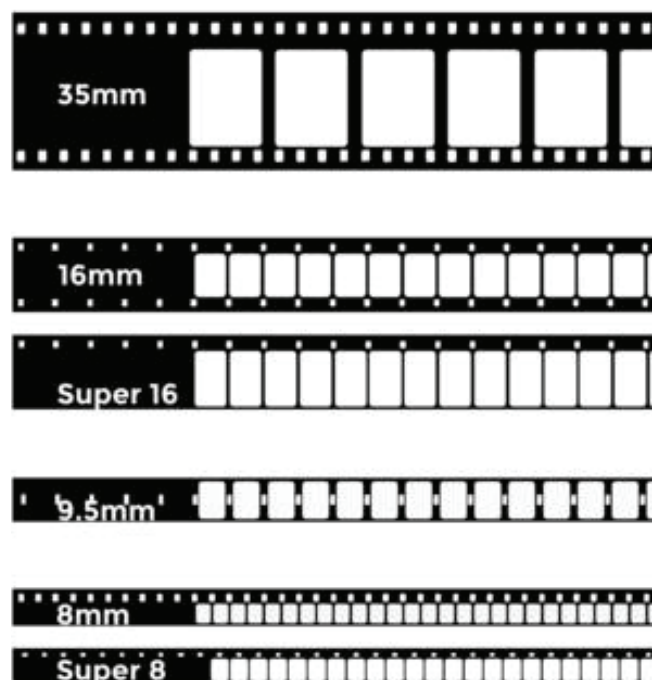
History

Film has a long history dating from the 1890s, lasting through to today. Due to technological limits, films began in the 1890s at less than a minute long and remained silent until 1927. The first commercially exhibited film in the United States was 35mm, which remains the industry standard today. Smaller gauges, such as 16mm and 8mm, are less expensive to use and are more popular in the amateur market. 16mm was introduced in 1923 by Kodak and is the most common gauge present in today's libraries and archives. 8mm was introduced by Kodak in 1932 and marketed to amateur film enthusiasts. Other gauges, such as 28mm and 9.5mm, are much rarer but may be present in your collections.

Film can be broken down into three main types consisting of image, optical sound, and magnetic sound. Each of these types has multiple sub-types and variants. While it is true for all media types, film more than any other demands an understanding of production and post-production processes and workflows in order to properly interpret what a given piece of film media is and how it should be considered, prioritized, and treated. A deep dive into the particulars of film production and post-production workflows and the corresponding media outputs is beyond the scope of this chapter, although the reader is strongly advised to review "Section 8: Additional Resources," which covers these topics.

FIGURE 1.17
Film gauge comparison chart.

Credit: Ryan Edge (CC BY-SA).



Materials and preservation risks

The three most common bases for film consist of nitrate, acetate, and polyester. Contrary to audiotape, polyester film is translucent when held up to the light and acetate is opaque.

The earliest of the three film bases in production was nitrate; notorious for its combustible properties, it became well known for the cause of major fires in projection booths, storage spaces, and archives. When nitrate is subjected to high heat and humidity, it can combust. Nitrate may be identified through edge printing or edge code if the manufacturer took advantage of this convention. Readers should note that it is also possible for the edge printing on nitrate film to be transferred to other film stocks in the film-to-film reformatting process. If nitrate is identified, it should be moved to a cold storage environment and prioritized for preservation reformatting. Decay in nitrate film is identified through discoloration, stickiness, odor, self-adhesion, deformation, and ultimately disintegration.

In 1909, acetate film⁶ followed nitrate film, introduced as “safety film” due to its noncombustible properties. While safer than nitrate, acetate eventually demonstrated its own significant degradation mechanisms, the most common of which is known as “vinegar syndrome” due to the

resulting odor. Promoted by storage conditions, the vinegar smell is a result of acetic acid production from a catalytic degradation process that is logarithmic in nature. At a certain point of degradation, the speed of decline is greatly increased, progressing from brittleness and a vinegar odor to delamination and decomposition. An early form of acetate known as diacetate can give off an odor of camphor or moth balls as it degrades. This process is known as “naphthalene syndrome.”⁷ While not as widely used as acetate, diacetate film may exist in archival collections.

Polyester was the latest film base to be introduced in the 1950s and has not yet exhibited significant degradation mechanisms, though discoloration and fading may occur.

Film image and sound can be together on one film (composite) or separated onto multiple films (separations) depending on at what type of recording it is and at what point in production or post-production process it was created.

The aforementioned film deformation resulting from degradation will negatively impact the ability to reproduce film images and magnetic and optical sound alike. Discoloration and fading can be an issue with polyester and acetate film bases, impacting both film images and optical sound tracks.

Depending on the type of film recording and where in the production or post-production workflow the film was created, there may be significant splices. These may consist of either tape splices or cement splices. Poor quality splices or degraded splices can result in the failure to bind two pieces of film together or cause an artifact in the image or sound.

Film is often stored on film cores without a reel or flanges, just like an open reel audio pancake. (See the polyester reel pictured above.) As with pancakes, film on cores is prone to becoming loose and unspooling, resulting in a mess of film that is difficult to put back in order.

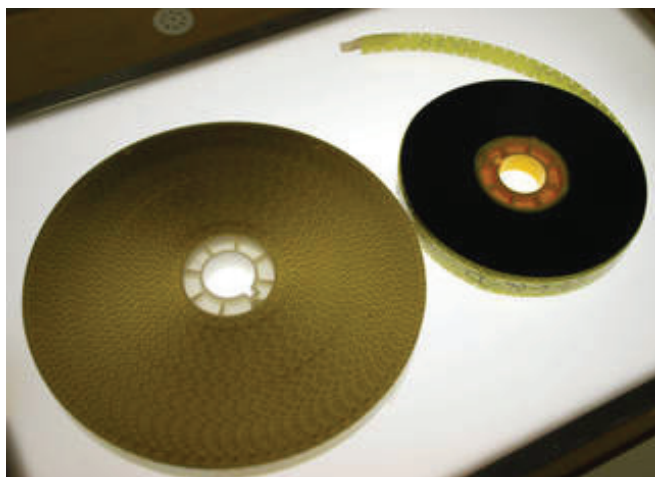
Improper storage, care, and handling practices may result in the following film issues, impairing reproduction:

- Scratches, resulting in loss of image content and artifacts in image and sound.

FIGURE 1.18

Polyester film (left) vs. acetate film (right).

Credit: <https://psap.library.illinois.edu/collection-id-guide>



6 The term acetate is a collective name for Cellulose triacetate and Cellulose diacetate film. National Film and Sound Archive of Australia, “Acetate film,” <https://www.nfsa.gov.au/preservation/preservation-glossary/acetate-film>

7 National Film and Sound Archive of Australia, “Naphthalene Syndrome,” <https://www.nfsa.gov.au/preservation/preservation-glossary/naphthalene-syndrome>

FIGURE 1.19**A film with optical sound.**

Credit: AVP



- Fading or discoloration of the image or optical sound.
- Deformation of the film, causing distortion of the light as it passes through the film or poor contact between magnetic film and the magnetic head.
- Decomposition and delamination of the emulsion layer, resulting in loss.
- Matter such as dirt, dust, hair, and even fingerprints that create separation between the magnetic film and the head.
- Broken or poorly repaired splices.
- Film that is unwound from cores.
- Magnetic particles that were originally recorded but were exposed to a magnetic force that changed their magnetic state.

Reproduction methods

The reproduction method for film images, magnetic sound, and optical sound are all different. Like magnetic tape, magnetic film sound uses electromagnetic means of reproduction. Therefore, the same concerns exist regarding contact with the head and potential factors that may disrupt good contact.

There are two types of optical film sound, consisting of variable density and variable area. Both utilize an opto-electronic method of reproduction, shining light through

the film into a light sensor on other side that converts light values to electrons.

Film images are reproduced through the projection of light through an image; the light that passes through the image may end up on a screen for viewing purposes. When digitizing, the light that passes through the film is read by photoelectric sensors that convert the light to electrons. Whether a film is a positive or a negative will impact the digitization process.

Best practices for storage and handling

Environment

- Film image and optical sound should be stored at temperatures between -4 and 46 degrees Fahrenheit, with RH between 30%-50%. Temperatures should not fluctuate more than ± 3 degrees within a 24-hour period. RH should not fluctuate more than $\pm 5\%$ within a 24-hour period.
- Film magnetic sound should be stored at temperatures between 32 and 61 degrees Fahrenheit,⁸ with RH between 30%-50%. Temperatures should not fluctuate more than ± 3 degrees within a 24-hour period. RH should not fluctuate more than $\pm 5\%$ within a 24-hour period.

8 Acetate-based magnetic tapes benefit from temperatures at the lower end of this range, polyester from the higher end.

FIGURE 1.20**Store film on its side, lying horizontally.**

Credit: AVP



- Film images and optical sound should be kept out of light when not being used.
- Magnetic film sound should be kept away from magnets, items containing magnets, and items generating magnetic force.

Housing

- Film should be stored on its side so that it lays horizontally.
- The end of a film should be taped down, either to itself or the reel flange, using “hold down” tape intended for this purpose.
- Film should be in housing that protects it from external elements.

Handling

- Do not use any other type of tape or solution for splices other than film splicing tape and cement solution.
- Handle film by the edge. Do not touch areas of the film containing images or sound without using non-abrasive, non-shedding gloves.

SECTION 5: OPTICAL MEDIA**HISTORY**

Audio compact discs (CDs) date from 1982 to the present. They represent the first successful consumer digital audio format. Composition varies between pressed CDs and DVDs (digital versatile discs) and those that are writeable and rewriteable (CD-R and CD-RW).

Pressed CDs are comprised of a protective lacquer layer, a metal data layer, and a polycarbonate plastic layer. Digital information is expressed in binary, and the metal data layer is molded to create pits; these pits reflect light differently than unaffected areas called lands. Playback is achieved through optoelectronic means: an optical stylus laser reads the pits and lands to reproduce the encoded information.

CD-Rs have a protective lacquer layer, an organic dye layer, a gold or silver metal reflective layer, and a polycarbonate plastic layer. The organic dye layer functions in much the same way as the metal data layer in a pressed CD; however, instead of being molded, the dye layer allows or blocks light transfer through the data layer, functioning in the same manner as the pits and lands.

CD-RWs have a protective lacquer layer, metal alloy recording layer, aluminum reflective layer, and a polycarbonate plastic layer. The metal alloy recording layer functions similarly to the dye layer in CD-Rs. The phase-changing metal alloy film can be altered in the same way the organic dyes can to reflect light.

DVDs were introduced to the market in 1995 and continue to be used today. They may contain any type of digital file, but commercially produced DVDs containing videos are common. Pressed DVDs are essentially two CDs glued together: a polycarbonate plastic layer, adhesive, a metal data layer, and a polycarbonate plastic layer. (The protective lacquer present on a CD is not needed on a DVD because the second polycarbonate layer acts as a protective layer.) DVDs can be two sided, meaning there may or may not be two recording layers present. Digital information is recorded in the same way as on a pressed CD: binary information is pressed into the metal data layer. Again, playback is through optoelectronic means using an optical stylus laser.

DVD-Rs have a polycarbonate plastic layer, an organic dye layer, a gold or silver metal reflective layer, and a polycarbonate plastic layer. The organic dye layer works the same as in a CD; the dye layer allows or blocks light transfer through the data layer. These, too, can be two-sided, with two dye layers and two metal reflective layers.

DVD-RWs have a polycarbonate plastic layer, metal alloy recording layer, aluminum reflective layer, and a polycarbonate plastic layer. The metal alloy recording layer functions the same as in CD-RWs; the phase-changing metal alloy film can be altered to reflect light. CD-RWs can be two-sided with two metal alloy recording layers and two reflective layers.

Materials and preservation risks

The polycarbonate layers in all optical disc types are easy to scratch or smudge, which may cause read errors or, if the damage is great enough, prevent successful playback. In addition, inks and adhesives used to label optical discs can harm the data stored in the top layer of the disc. The organic dyes used in optical discs will degrade over time, eventually making the disc unreadable. Delamination is possible through large swings in temperature and humidity; oxygen is introduced through the delamination process and can damage the data layer in CD- and DVD-RWs.

Improper storage, care, and handling practices may result in the following issues, impairing reproduction:

- Scratches or other means of damaging the data layer, therefore removing the data itself.
- Scratches on the polycarbonate layer, resulting in the inability of light to reach the data layer without distortion.
- Warping of the disc.
- Breakage of the disc.
- Discoloration of the layers resulting in the distortion of light values and the inability to retrieve the data.

Best practices for storage and handling

Environment

- Optical discs should be stored at temperatures between 41 and 68 degrees Fahrenheit, with a relative humidity (RH) between 30%-50%. Temperatures should not fluctuate more than ± 2 degrees within a 24-hour period. RH should not fluctuate more than $\pm 5\%$ within a 24-hour period.
- Optical discs should be kept away from sources of heat, including sunlight or room lighting. Exposure to light greatly accelerates the fading of dyes.

Housing

- Optical discs should be stored on their edge so that they stand vertically.
- Optical discs should be in individual cases or Tyvek envelopes.
- Paper inserts should be removed and stored alongside the case or envelope. Paper can attract and hold moisture in humid environments.

Handling

- Handle the optical disc from the interior and exterior edges only. Do not touch the surface of the optical disc without using non-abrasive, non-shedding gloves. Fingerprints, smudges, scratches, particles of dirt or dust, solvents, or excessive moisture all may interfere with the ability of the laser to read the data layer.

FIGURE 1.21

Store optical discs on edge, standing vertically.

Credit: AVP



SECTION 6: SHIPPING OF MEDIA CARRIERS

Sometimes collection materials must be moved in-house or to a third-party location. Even when transporting media carriers within the archive, careful handling is essential to protect the media from climatic changes, stray magnetic fields, and shock. Be sure items are secured on rolling carts and avoid bumps and uneven floors when moving collection items.

When shipping collection materials, ensure that the ambient temperature in the transport vehicle does not exceed 75 degrees Fahrenheit. Temperatures between 65 and 75 degrees F are acceptable, and this may require that materials are stored in containers that can maintain this temperature range while en route. In general, it is best to transport materials when outside temperatures are similar to materials' storage conditions, although this is not always possible. Avoid large swings in temperature and humidity.

Audio and Video Tapes

If you are going to ship audio and video cassettes and open reel tapes, be sure that each tape is supported by its container. Ship them the way you store them: on edge. It is most important to avoid shifting and shock. When packing smaller items within bigger boxes, be sure to surround them with enough packing materials so that they do not move or rub together during transport.

Discs

Discs of the same size can be stacked together horizontally in groups of five to ten. Separate each disc from its neighbors by placing each disc in a sleeve or by placing a smooth piece of paper between each pair of discs. Place each stack of five to ten discs between two rigid boards, creating a "sandwich." 1/2-inch honeycomb board or multiple layers of corrugated cardboard will work. Tape the edges together, taking care that no adhesive directly touches a disc. Wrap the entire sandwich in bubble wrap, and tape it closed. Place this package in a larger box and surround it with packing peanuts or bubble wrap. Multiple sandwiches may be placed in one larger box as long as there is room for padding between the sandwiches and the exterior box.

Cylinders

Cylinders may be shipped in their boxes. Wrap each box individually in bubble wrap, then place the wrapped boxes upright in a larger box. Make sure these boxes are tightly packed; if there is room for motion, fill in the spaces with bubble wrap or packing peanuts. Tape the box closed, then place the box in a larger box. Fill the spaces with packing peanuts or bubble wrap.

Film

Film can be damaged during transport if not packed correctly. It is important to protect film rolls against slipping. Check each film can for space and fill any existing space with bubble wrap or acid-free paper so it fits snugly. Place like-sized cans horizontally in a box, filling any gaps with bubble wrap or packing peanuts to ensure the films will not move during transport.

SECTION 7: CONCLUSION

Effective care of audiovisual materials requires an understanding of a format's history, physical components, reproduction methods, and best practices for storage and handling. Several different formats may be stored within a single repository, so this information can become overwhelming. One strategy is to use the information covered in this Chapter as a reference instead of attempting to memorize it all. Focus on learning the points relevant to the materials in your collection, and aim to improve your handling procedures and storage methods over time.

An important takeaway from this Chapter is that most formats cannot be treated in exactly the same way. Wax cylinders and vinyl discs need different types of storage containers to prevent physical damage, and film and video-tape are reproduced through extremely different methods. Treating these materials with their differences in mind and referring back to this Chapter for specific information will greatly improve the stewardship of your collections.

It is also important to know exactly what you have in your collections so that you can improve their storage conditions and effectively select items for digitization. In "Chapter 2: Inventory and Assessment of Audiovisual Collections," you will learn how to develop an inventory of your collection

that contains the data needed to care for and evaluate your collections for reformatting projects. You will be introduced to the required fields for inventory as well

as additional fields that you may find helpful, and you will learn about the considerations that guide prioritization for digitization projects.

SECTION 8: ADDITIONAL RESOURCES

TABLE 1.1
Carrier Type, Media Type, and Reproduction Method Summary

Media Type	Format	Reproduction Method				
		Electro magnetic	Electro mechanical	Imaging	Opto electronic	Photo electronic
Magnetic Media	Tape	X				
	Wire	X				
Grooved Media	Disc		X	X		
	Cylinder		X	X		
Optical Media	CD				X	
	DVD				X	
Film Media	Image					X
	Optical Sound			X	X	
	Magnetic Sound	X				

TABLE 1.2
Storage Temperature and Humidity Summary

Storage Conditions		Magnetic Tape		Disc	Cylinder	CD, DVD	Film				
		Acetate	Polyester				Nitrate	Acetate		Polyester	
								B&W	Color	B&W	Color
Room	60°F-74°F 30%-50% RH	No	No	Fair	Fair	Fair	No	No	No	Fair	No
Cool	46°F-60°F 30%-50% RH	Fair	Good	Very Good	Very Good	Good	No	No	No	Good	No
Cold	32°F-46°F 30%-50% RH	Good	Fair	Good	Good	Good	Good	Good	Good	Very Good	Good
Subzero	-4°F-32°F 30%-50% RH	No	No	No	No	No	Very Good	Very Good	Very Good	Very Good	Very Good

Resources from the community

AUDIO RESOURCES

- NARA, How Do I Identify Audio Formats? <https://www.archives.gov/preservation/formats/audio-identify-formats.html>
- NARA, What are some important characteristics of audio formats? <https://www.archives.gov/preservation/formats/audio-important-characteristics.html>
- FACET: The Field Audio Collection Evaluation Tool, Format Characteristics and Preservation Problems Version 1.0 http://www.dlib.indiana.edu/projects/sounddirections/facet/facet_formats.pdf
- ARSC Guide for Audio Preservation, Chapter 2, Audio Formats: Characteristics and Deterioration <https://www.clir.org/pubs/reports/pub164/pub164.pdf>
- IASA TC-05, Chapter 2: Type of Carriers, Recording Principles, Composition, Physical and Chemical Stability, Deterioration by Replay <http://www.iasa-web.org/tc05/2-type-carriers-recording-principles-composition-physical-and-chemical-stability-deterioration>
- The History of Sound Recording Technology <https://recordinghistory.org>
- Chapter 3 (Disc Structure) of the CLIR report *Care and Handling of CDs and DVDs: A Guide for Librarians and Archivists* (PDF)
- An Introduction to Optical Media Preservation <http://www.avpreserve.com/wp-content/uploads/2014/04/OpticalMediaPreservation.pdf>
- IASA TC-05, Handling and Storage of Audio and Video Carriers, Chapter 2.3: Optical carriers <http://www.iasa-web.org/tc05/23-optical-carriers>
- ARSC Guide for Audio Preservation, Chapter 2.4 Optical Disc Formats <https://www.clir.org/pubs/reports/pub164/pub164.pdf>

VIDEO RESOURCES

- Mona Jimenez and Liss Platt, Videotape Identification and Assessment Guide. Texas Commission on the Arts, 2004 <http://www.arts.texas.gov/wp-content/uploads/2012/04/video.pdf>
- Timothy Vitale and Paul Messier, Video Preservation Website, Video Format Identification Guide on the Video Preservation Website, 2007 http://videopreservation.conservation-us.org/vid_id/index.html
- NARA, How Do I Identify Video Formats? <https://www.archives.gov/preservation/formats/video-identify-formats.html>

FILM RESOURCES

- NEDCC: A Short Guide to Film Base Photographic Materials <https://www.nedcc.org/free-resources/preservation-leaflets/5.-photographs/5.1-a-short-guide-to-film-base-photographic-materials-identification,-care,-and-duplication>
- IPI Storage Guide for Acetate Film https://s3.cad.rit.edu/ipi-assets/publications/acetate_guide.pdf
- Film Forever: The Home Film Preservation Guide <http://www.centerforhomemovies.org/filmforever>
- The Film Preservation Guide: The Basics for Libraries, Archives and Museums, National Film Preservation Board <http://www.filmpreservation.org/preservation-basics/the-film-preservation-guide-download>
- NARA, How Do I Identify Motion Picture Film Formats? <https://www.archives.gov/preservation/formats/motion-picture-film-identify-formats.html>
- The National Film and Sound Archive, Australia's website on Film Gauges <http://www.nfsa.gov.au/preservation/glossary/film-gauges>
- Richard Wright, DPC Technology Watch Report 12-01: Preserving Moving Pictures and Sound, 2012 <http://dx.doi.org/10.7207/twr12-01>

STORAGE AND SHIPPING REFERENCES

- Peter Z. Adelstein, IPI Media Storage Quick Reference, 2nd edition <https://s3.cad.rit.edu/ipi-assets/publications/msqr.pdf>
- Fred Byers, “Conditions that Affect CDs and DVDs” <https://www.clir.org/pubs/reports/pub121/sec5.html>
- IASA TC-05 Handling and Storage of Audio and Video Carriers, Chapter 3: Passive Preservation: Environmental Factors, Handling and Storage <http://www.iasa-web.org/tc05/3-passive-preservation-environmental-factors-handling-and-storage>
- NARA Directive 1571 Archival Storage Standards from the National Archives <https://www.archives.gov/foia/directives/nara1571.pdf>
- IASA TC-05 Handling and Storage of Audio and Video Carriers, Chapter 4.5.2: Storage Facilities and Transport <http://www.iasa-web.org/tc05/46-shelving>
- Film Forever: Chapter 8.2, Freezing your film, Five Easy Steps <http://www.filmforever.org>
- Film Preservation Guide, Chapter 5.8, Shipping your film http://www.filmpreservation.org/userfiles/image/PDFs/fpg_5.pdf
- SMPTE RP 103:1995 - SMPTE Recommended Practice —Care, Storage, Operation, Handling and Shipping of Magnetic Recording Tape for Television <http://ieeexplore.ieee.org/document/7291326/>

CHAPTER 2

INVENTORY & ASSESSMENT

by Rebecca Chandler, Consultant, AVPreserve

ORGANIZATIONS RANGING FROM MUSEUMS to government agencies maintain moving image and sound collections that are used daily for research, education, communication, and the fulfillment of institutional missions. Unfortunately, these collections are often inadequately documented and frequently inaccessible due to various resource constraints, difficulties in capturing relevant data from objects, and the effects of obsolescence and physical degradation.

The lack of intellectual control is a very real threat to the preservation of audiovisual materials. If an organization does not know what content and formats it has, it will not know how to plan or prioritize nor know what materials are most at-risk. Common practice across archives has been to drive preservation by patron request or internally identified priorities that fit under a particular budget allocation. However, such initiatives cannot be carried out without some basic knowledge of collection contents; without such insight, unidentified or unprocessed materials will sit on the shelf, not because the content is unimportant, but because it has been unfindable and inaccessible. As a steward for the assets in your care, you are responsible for making sure they are findable, accessible, and sustainable; this important work requires an accurate inventory.

An accurate inventory enables you to make informed decisions regarding your collections. It allows you to make smart choices around selection and prioritization for digitization. Creating an inventory will allow you to learn more about your collection and reveal a path forward to its preservation.

SECTION 1: INVENTORY

A basic item-level inventory is vital to the successful stewardship of audiovisual assets. At a minimum, a unique identifier and a format-type must be identified and recorded for each asset in order to responsibly track items throughout their lives. However, gathering as much data as possible in one pass will provide a baseline set of information that will support selection, prioritization, and initial discovery, as well as minimize the number of times the assets need to be referenced and handled.

Establishing a minimal set of required fields will help to streamline the process and manage expectations. Minimal sets are meant to be used for basic identification, selection, prioritization, and digitization planning. Information beyond the required fields should only be recorded if it is easily discoverable by simply examining the physical asset or consulting an existing record. Further descriptive and technical information can be recorded post-digitization at your organization's discretion.

As archivists, we want to provide our users with as much information regarding our collection as possible; however, richer descriptions do not play a part in the inventory process. At this point, we want to record as much information as can be quickly and easily gleaned from the asset itself. There will be a time for more detailed description at a later point, should you and your institution decide that it is necessary. At this beginning stage, however, time is of the essence.

What follows is a list of suggested fields to capture in your inventory, along with a rationale for their capture. These fields are divided between ones that are so important as to be considered “required” and those that are optional but helpful. Ultimately, the decision of what fields to record lies with you and what is needed for your organization to be able to adequately plan for preservation.

Required Fields

UNIQUE ID

Description

A free text field for a unique number assigned to each asset. This number may be an existing ID number already associated with the asset or it may be created from scratch.

Rationale

Each asset must be assigned a unique ID in order to allow responsible stewardship throughout its life. Assigning a unique ID will make it easier for you to record an asset’s location for retrieval purposes and to track its process throughout the digitization process, especially when sending the asset to an outside digitization vendor.

Procedure

If the item has already been logged or cataloged in an existing system, a unique ID may already exist. If this is the case, simply capture this information in your new inventory. If a unique ID does not exist, you will have to create one. It could be an alphanumeric string that follows the IDs you use in other systems. Or you could make up a system yourself. In the absence of an existing system, you might work with something like “000001, 000002, 000003...,” with the number of leading zeros dependent on the estimated size of the collection. No matter what you choose, remain consistent and follow a system that will not allow accidental repeats of ID numbers.

If the unique ID is not indicated on the item itself, it is a good idea at this point in the process to add a barcode to the asset capturing this data. It will save time in the future when you move on to digitization projects, and it makes sense to do it now as you are handling the materials. Print

two of each unique ID—one for the asset case and one for the asset itself.

LOCATION

Description

A free text field containing the asset’s location.

Rationale

Recording an asset’s location will allow you to locate it more easily whenever necessary.

Procedure

Once again, consistency is the goal. Creating a system for identifying locations that can be used across your collection will be most helpful to you in the future. If your collection is large and stretched across several buildings, start with a building code and then perhaps a room code. If your items are on shelves, perhaps a row/bay/shelf code would make sense next. If they are in boxes on shelves, add a box number. Here is an example:

Reed:R102:04:B:3:842

This item is located in the Reed Building, in Room 102, in row 4, bay B, on shelf 3, in box 842.

This is simply one approach among many. There is no wrong way to create location codes; think through what will work best for your institution and remain consistent.

MEDIA TYPE

Description

A field that designates whether the record is for an audio, video, or film asset.

Rationale

Indicating the media type will allow you to organize and group like assets when planning for digitization.

Procedure

Media type can be determined by inspecting the physical item and goes hand in hand with identifying the format. Use a controlled vocabulary such as an internal list or

PBCore’s instantiationMedia Type list: http://metadata.registry.org/concept/list/vocabulary_id/135.html.

FORMAT

Description

A field that indicates the specific type of audio, video, or film asset, such as “cylinder” or “VHS.”

Rationale

Format will allow you to organize and group like assets when planning for digitization. In addition, knowing the types and quantities of formats in your collection will allow you to gather more accurate digitization cost estimates, making budgeting and fundraising easier.

Procedure

You may have personal experience with many of the formats in your collection; however, some formats will likely be unfamiliar. Throughout its 150-year history, audiovisual recording has spanned well over 100 different formats, making format identification sometimes difficult. The following resources will help with format identification.

- National Archives and Records Administration, *How Do I Identify Audio Formats?*
<https://www.archives.gov/preservation/formats/audio-identify-formats.html>
- National Archives and Records Administration, *How Do I Identify Video Formats?*
<https://www.archives.gov/preservation/formats/video-identify-formats.html>
- National Archives and Records Administration, *How Do I Identify Motion Picture Film Formats?*
<https://www.archives.gov/preservation/formats/motion-picture-film-identify-formats.html>
- Association of Recorded Sound Collections, *Guide for Audio Preservation, Chapter 2, Audio Formats: Characteristics and Deterioration*
<https://www.clir.org/pubs/reports/pub164/pub164.pdf>

- Mona Jimenez and Liss Platt, *Videotape Identification and Assessment Guide*. Texas Commission on the Arts, 2004

<http://www.arts.texas.gov/wp-content/uploads/2012/04/video.pdf>

- Timothy Vitale and Paul Messier, *Video Format Identification Guide* on the *Video Preservation Website*, 2007
http://videopreservation.conservations-us.org/vid_id/index.html

- Film Preservation Guide, Chapter 2, *Understanding Film and How it Decays*
https://www.filmpreservation.org/userfiles/image/PDFs/fpg_2.pdf

- Preservation Self-Assessment Program, *Collection ID Guide*
<https://psap.library.illinois.edu/collection-id-guide>

Once you’ve identified the format, record it in your inventory. You may also fill in the Media Type once identification has been accomplished. Use a controlled vocabulary, such as an internal list or PBCore’s instantiationPhysical list: http://metadata.registry.org/concept/list/vocabulary_id/145.html.

TITLE

Description

A free text field containing the title of the asset.

Rationale

Depending on the level of description your items have associated with them from earlier cataloging efforts, the title is often your best clue about the content. The title will likely play a role in the prioritization of your assets for digitization in the future.

Procedure

If a catalog record already exists, take the title from there. If a record does not exist, make your best guess as to the title based on whatever is on the asset itself. This may consist of a professionally printed label where the title is clear or a handwritten label where some guessing is necessary. If no label exists, simply title the asset “untitled” or “unknown.”

Optional fields

COLLECTION NAME

Description

A field denoting the asset's parent collection, if any.

Rationale

If your institution has many collections, it will be helpful to have the parent collection recorded. This will allow you to group items by intellectual content, which will assist with prioritization.

Procedure

Collection names are often recorded on boxes or other containers. Or perhaps your assets are grouped by collection in a particular location. Collection names can be difficult to assign to audiovisual materials as they are often part of institutional knowledge and not necessarily recorded or easily interpreted. Use a controlled vocabulary based on the collection names already in use at your institution.

DESCRIPTION

Description

A free text field for any contextual information recorded on the asset or its container or gleaned from other sources such as donor agreements.

Rationale

Recording as much easily-discernible data as possible while you are handling your assets is a smart strategy. Recording any printed or written information that is on the asset now will save you time later.

Procedure

Record any printed or written information on the asset that does not fall into one of the other fields listed here. Use a separator, such as a semicolon, to separate different lines or ideas.

DATE

Description

A field noting the date the asset itself was created.

Rationale

The age of an asset can be helpful in determining its risk of degradation and therefore prioritization.

Procedure

Date can be a tricky data point to determine as there can be multiple dates associated with one asset. One tape could have been used to record on multiple days. A tape could have been a duplicate of another asset, meaning it was created on one date, but the content is from another date. Do your best to record the date that makes the most sense to you and record any additional dates in the Description field.⁹ Use a consistent date format, such as yyyy-mm-dd.

GENERATION

Description

A field that defines the relationship between original material and copies.

Rationale

The Generation field will allow you to identify duplicates in your collection and will aid in prioritization.

Procedure

Before an asset is reformatted, its generation can be determined only by using the label information present on the object. Unfortunately, there is no other way to determine an asset's generation. Your organization should compile a list of relevant generations (e.g., Production Master, Access Copy, Sub-Master, Dub, etc.) and use that controlled vocabulary or use PBCore's instantiation-Generations list: http://metadataregistry.org/concept/list/page/1/vocabulary_id/147.html.

⁹ Should your collection be one where it makes sense to record multiple dates for each asset, you may wish to explore PBCore for further guidance on date fields: <http://pbcore.org>. The instantiationDate element may prove useful: <http://pbcore.org/pbcoreinstantiation/instantiationdate>.

PART**Description**

A free text field that notes if the asset is a piece of a larger work. For example, if a full-length film is broken up into four reels, you might complete this field as “Reel 1 of 4.”

Rationale

Part is helpful in determining if your institution holds the entirety of a particular work, and it will allow you to keep parts grouped together when prioritizing for digitization.

Procedure

Like Generation, Part can only be determined by the asset’s label information.

COMMERCIAL OR UNIQUE/RARE**Description**

A field indicating if an asset is commercial or unique in nature.

Rationale

This field is extremely important in prioritization. See “Section 3.2: Content Value” below for more information on prioritization.

Procedure

Usually, identifying whether or not an asset is a commercially released item is fairly straightforward. Is it in a professionally produced package? Is there a recording label or production studio listed on the label? Is there a barcode printed on the label or a matrix number? If yes, then it’s a commercial item.

Assets with handwritten labels can usually be assumed to be unique. The exception to this would be if it is a home copy of a commercially released recording (think of recording songs off the radio onto cassette). Do your best to identify whether assets are commercial or unique; when in doubt, assume they are unique.

Identifying commercially released records that are rare (e.g., short runs, out of print) is a bit trickier and will require

subject expertise or research. You will have to determine if this is a priority for your organization. Use a controlled vocabulary of consistent internal terms.

COPYRIGHT/RESTRICTIONS**Description**

A free text field that explains the terms defining an asset’s availability for re-use.

Rationale

Restrictions on making an asset freely available to the public will play a role in prioritization.

Procedure

Restrictions on usage may be recorded in an existing catalog record or in an asset’s donor agreement. Record any information you have on copyright status and restrictions here. If no information is known regarding an asset’s status, record that as well.

CONDITION**Description**

A field recording the physical condition of the object.

Rationale

Assets suffering from irreversible degradation may be prioritized for digitization. Some methods of degradation can be contagious; infected assets should be quarantined.

Procedure

Some modes of degradation are easily identifiable just by inspecting an asset, and these modes are the ones we wish to record here. Deep knowledge of recording formats and their chemistry is not required, just an observant eye. For more information, review “Chapter 1: Introduction to AV Preservation Challenges.” Here are some red flags to look out for:

- **Mold:** Mold can affect any format in your collection. Keep an eye out for fuzzy or speckled debris. Active mold may spread and will continue to degrade your assets. It is best to quarantine any assets infected with mold.

- **Odor:** Odd smells are indicative of some sort of chemical reaction and may indicate that an asset is actively degrading. Film that smells specifically of vinegar should be quarantined from other film.
- **Broken or cracked discs or cylinders:** Depending on the break, this may make them unplayable by traditional means.
- **Palmitic acid:** A white, oily sheen on discs indicates they are in the early stages of delamination. Once delamination begins, the lacquer will crack and flake off, rendering the disc unplayable.
- **Crazing:** The next step in delamination of discs. A network of fine cracks manifests on the discs. Once the lacquer flakes off, the disc will not be playable.
- **Long play tapes:** Long play tapes are less of a condition issue and more of a reality of the original format. They are thinner than tapes of shorter durations and are therefore prone to stretching or snapping on playback.
- **Broken carriers:** Cracked cassette cases, broken reels, etc.
- **Unspooled tape:** The tape is off of its reel or out of its cassette, possibly crumpled and tangled. This will lead to mechanical problems on playback and potentially a loss of content.
- **Spoking:** Spoking is most easily identified on open reel tape and film (Image 2.1). As magnetic tape or film shrinks, the tension in the reel builds, forcing the tape/film to “bend” in a regular pattern to help relieve that tension. These assets may require some preservation work prior to digitization or the use of specialized equipment.

This field can be either free text or a controlled vocabulary.

For more information on condition issues affecting audiovisual materials, visit the following resources:

- FACET: The Field Audio Collection Evaluation Tool, Format Characteristics and Preservation Problems Version 1.0
http://www.dlib.indiana.edu/projects/sounddirections/facet/facet_formats.pdf

IMAGE 2.1

Physical condition: A film reel that exhibits signs of spoking.

Credit: AVP



- Association of Recorded Sound Collections, *Guide for Audio Preservation, Chapter 2, Audio Formats: Characteristics and Deterioration*
<https://www.clir.org/pubs/reports/pub164/pub164.pdf>
- Film Preservation Guide, Chapter 2, Section 2.6, *Common Types of Decay and Damage*
https://www.filmpreservation.org/userfiles/image/PDFs/fpg_2.pdf

Inventory Tools

How do you best capture this item level information? Determine whether your institution has any internal tools or catalog systems and if they will work for your audiovisual materials. If the answer is no on either count, here are some free, open source options that may work for you.

AVCC

AVCC is an abbreviation for AudioVisual Collaborative Cataloging, which is a free, open source web application developed by AVPreserve and funded by Library of Congress, METRO, and AVPreserve. AVCC was developed to enable collaborative, efficient item-level cataloging of audiovisual collections. The application incorporates built-in reporting on collection statistics, digital storage calculations, shipping manifests, and other data critical to prioritizing

and planning preservation work with audiovisual materials.

AVCC establishes a minimal set of required and recommended fields that provide basic intellectual control enabling quantification, planning, and management of collections. The focus of AVCC is two-fold: to uncover hidden collections via record creation and to support preservation reformatting in order to enable access to the content itself. For more information, visit <https://www.avpreserve.com/tools/avcc>.

PSAP

The Preservation Self-Assessment Program, or PSAP, is a free online tool developed by the University of Illinois at Urbana-Champaign. PSAP was developed to help collection managers prioritize efforts to improve the conditions of their collections. Meant for institutions with few-to-no preservation or conservation staff, PSAP is designed to be simple and easy to use for those with little preservation training.

PSAP aims to:

- Support targeted preservation assessments of paper documents, books and bound items, photographic and image materials, audiovisual materials, and non-composite museum objects made of ceramic, glass, stone, or metal.
- Perform item- and collection-level assessments.
- Provide textual and image-based educational resources to aid in the identification of different types of materials and their preservation challenges.
- Address factors of storage and display, applicable to situations from open exhibitions to closed archives.

For more information, visit <https://psap.library.illinois.edu>.

AV COMPASS

AV Compass is a free, online suite of tools developed by the Bay Area Video Coalition with support from the Andrew W. Mellon Foundation. Intended for use by individuals and organizations alike, AV Compass features step-by-step educational videos, PDF guides, and a tool for creating inventories.

AV Compass includes:

- Instructional guides and eleven short videos that walk a user through the assessment.
- Overview of preservation concepts.
- Directions on how create and implement a preservation plan.
- A free tool to create an inventory of your collection, which you can export and send to collecting archives and preservationists.

For more information, visit <http://www.avcompass.bavc.org>.

SPREADSHEETS

Spreadsheets are a perfectly adequate option for creating an item-level inventory for some institutions. They are familiar to most staff and easy to use. Simply add whichever fields you plan to populate during your inventory and get started. Be sure to create and use controlled vocabularies where applicable—consistency is key.

SECTION 2: SELECTION FOR DIGITIZATION

The Mission Statement as a Guide

The process of selecting assets for digitization will be much more efficient and precise if you approach the issue with your mission statement in mind. Your archive should have a mission statement, approved by the head(s) of your organization. This statement should define the authority of the archivist within the organization and the parameters of the archival program.¹⁰ It is these parameters that will influence your decisions with regards to the selection of audiovisual materials for digitization.

Some questions to consider when selecting items for digitization include:

- Do you serve a specific audience's needs?
- What materials will be most useful to that audience?
- Is there a record management lifecycle in place at your organization?

Do you need to worry about destroying certain materials?

Assets that don't fit within the archive's mandate set forth by your organization may not be considered preservation-worthy. Some assets may even need to be destroyed according to existing record retention policies. Referring back to your mission statement and associated policies will help you determine which assets are likely candidates for digitization.

Don't have a mission statement? Write one! Operating your archive without a mission statement puts the health of your archive in jeopardy. A strong mission statement informs your audience about the principles on which you base your existence, justifies your work to governing bodies, and provides a clear direction for your decision-making.

Goals for the Collection

Closely related to the mission statement, any specific goals you have in mind for your archive will come into play when selecting items for digitization. Are there any initiatives you are hoping to launch in the future? Are there items in your collection that align with those initiatives? The selection process will be much more effective and useful if you are thinking about your mission, your goals, and what you need to meet those goals both in the short term and in the long term.

Selection of Best Copy

Archives commonly have multiple versions of the same content. This could be the result of previous preservation reformatting efforts or could be due to production processes.

LEGACY PRESERVATION AND VIEWING COPIES

These duplicates will be identical copies of original audiovisual materials either on the same format or a newer format. It has been common practice to make "viewing" copies of archival materials in order to lessen the wear and tear on the original when screening items for patrons. Or, in an effort to battle obsolescence, content would be migrated to a newer format (e.g., 1 inch open reel videotape to Betacam). No matter the motivation, these copies will almost always be inferior to the original. The exception will be if the original has become so far degraded over time that the newer copy becomes the only viable means of transferring the content. The identification of "best copy," or the best copy of an asset currently in existence, will be important in the preservation of collections like this.

PRODUCTION PROCESSES

If your organization holds any production-oriented collections in its holdings, chances are you will have a combination of production elements (i.e., working copies and component parts that went into creating the final edited master) as well as final edited masters. You could even have several different "final" versions of a work. This would be a good

¹⁰ See SAA's "Museum Archives Guidelines" for an adaptable example. <https://www2.archivists.org/groups/museum-archives-section/museum-archives-guidelines>.

time to reference your mission statement and goals. Most archives find that the final edited masters will be of greatest interest and value to their organization and patrons; there will not be strong use cases for maintaining or digitizing the production elements.

However, some archives do have use cases for maintaining and digitizing all production elements in their holdings. These archives might be affiliated with an active production studio that may wish to maintain the option of using archival footage in reissues or new work. Or, these archives may be part of an educational institution to which the process of creating audiovisual media is of great interest and importance to its students and faculty. Whatever the case, if this is your mission, you must plan to digitize and preserve all production elements in your collection in addition to final edited masters.

SELECTING BEST COPIES WITHIN A BUDGET

If you find yourself in one of the scenarios outlined above and you are able to weed certain items from your collections (whether they are viewing copies or production elements), then your long term burden of preservation has been lightened, potentially significantly. However, the identification of best copies, elements, and final master copies for deselection and selection will require investigation and resources. Materials will need to be viewed and comparisons made in order to determine which copy is now the best quality copy. This may either be handled internally or through a vendor. If performed internally, you will need to procure the appropriate equipment, supplies, and staffing to screen your assets. If performed through a vendor, materials should be grouped by production title prior to sending to a vendor for selection of best copy.

In an ideal world, we would be able to view each asset and make an informed decision as to which copy is the best quality and which version we wish to digitize for preservation. However, whether you do this in-house or with a vendor, it will add to the cost of your digitization project. Rather than viewing or listening to each asset, your budget may require you to make some assumptions regarding best copy. For example, it is usually a good choice to select the original master for preservation, as any copies made in the past will be of inferior quality. The reality is that in some cases, the original master will be more degraded than a copy, which is why we wish we could

watch each asset to make a decision. Making an assumption is the only way some of us will be able to afford to preserve the assets we have in our collections; just make sure you think through the possible scenarios and make a well-informed decision. You can choose to select a few items that are of extreme importance to your organization for analysis by a vendor to select best copy.

SECTION 3: PRIORITIZATION FOR DIGITIZATION

After you decide which of your audiovisual assets must be digitized, you must prioritize them. Chances are you may not have the budget, resources, or staffing required to digitize all of your assets at once. Therefore, having a planned order and timeline in mind will be helpful for planning purposes. There are four main issues to address when prioritizing items for digitization: technical needs, cost to digitize, content value, and use value. These will be discussed in the following subsections.

Technical Needs

Physical audiovisual materials face multiple risk factors that render them partly or completely inaccessible in a relatively short period of time. As discussed in Chapter 1, plastics, metals, dyes, and other chemical materials are inherently unstable and reactive to their environment. The triggers and timeline for chemical degradation vary from format to format as the chemical makeup varies.

In addition, the technological dependencies of audiovisual materials limit the time that materials are accessible. Throughout its 150 year history, audiovisual recording has spanned well over 100 different formats. The majority of these formats are not interoperable and rely on their own proprietary technology for playback. Once a manufacturer discontinues production of a given format, obsolescence begins, leading to the loss of expertise, parts, equipment, and documentation and eventually resulting in inaccessibility of content. Age is not the only factor in obsolescence. Sometimes formats are very long-lived, such as VHS or LP, or they are extremely short-lived, such as MII or DAT. Long-lived formats will stave off obsolescence longer than short-lived ones, but all physical audiovisual formats are at risk of becoming obsolete.

MEDIASCORE

In 2015, Indiana University developed and released an application called MediaSCORE,¹¹ which enables a detailed analysis of degradation and obsolescence risk factors for most physical audio and video formats. In order to function properly as a prioritization tool, MediaSCORE's developers created a prioritization scoring system based, in part, on a format's risk of obsolescence and degradation. Higher scores indicate a higher level of risk. Lower scores indicate a lower level of risk. Should you choose to use MediaSCORE as a prioritization tool, you will have the opportunity to enter a wide variety of item specific factors, including age, brand, and condition, all of which will have an effect on an item's level of risk. However, using even the base score alone provides you with a great start on prioritizing your items based on technical needs.

The following list of audiovisual assets ranked by technical risk was created by considering the format's MediaSCORE base score as well as professional experience of practitioners in the field.

A few notes on the preceding ranking:

- This list only considers each format's general level of obsolescence and degradation; it should not be the only factor you use in your institution for prioritization. Content value, use value, and cost to digitize, as well as any item-specific preservation problems will need to be considered as well.¹²
- This is not a comprehensive list of audiovisual formats. Should you encounter a format in your collection that is not on this list, it was most likely not a widely used format and is likely obsolete. You should rank it higher (i.e. with a lower number) on this list.

Film is ranked lowest on this list as numbers 39, 40, 41, and 42. Film is on a different degradation and obsolescence trajectory than audio and video formats. If stored under proper conditions,¹³ it will remain viable for years to come. If film is not stored under proper conditions, it should be moved higher in the ranking.

TABLE 2.1

Prioritization of audiovisual formats based on technical needs

Rank	Format
1	2-inch open reel video
2	Lacquer disc
3	½-inch open reel video
4	MII
5	D1, D3
6	U-matic
7	PCM 1600, 1610, 1630 (U-matic)
8	DAT (DDS)
9	PCM-F1 (VHS, Betamax)
10	D2
11	1 inch open reel video
12	8mm family (video)
13	DTRS/DA-88 (Hi-8)
14	ADAT (VHS)
15	MiniDisc
16	Betamax
17	DV family
18	VHS, S-VHS, VHS-C
19	DVCPRO
20	Cylinder
21	Metal disc
22	Wire reel
23	¼-inch open reel audio
24	Microcassette, minicassette
25	Compact cassette
26	8-track
27	Betacam, BetacamSP, BetacamSX
28	Digital Betacam
29	D5
30	HDCAM
31	XDCAM
32	Laser disc
33	CD, CD-R, CD-RW
34	DVD, DVD-R, DVD-RW
35	Pressed 78 rpm disc
36	Pressed 45 rpm disc
37	Blu-ray
38	Pressed LP disc
39	Color acetate film
40	Black and white acetate film
41	Color polyester film
42	Black and white polyester film

¹¹ <https://github.com/IUMDPI/MediaSCORE>

¹² See "Chapter 1: Care and Handling" for item-specific preservation problems.

¹³ See Chapter 1, "Sections 4: Film" and "Section 8: Additional Resources" for more information about storing film.

Content Value

An assessment of an asset's content value will help you prioritize based on the historical or research importance of the content. Due to budget and timeline constraints, you will probably not be able to digitize everything in your collections before it is no longer viable; therefore, it makes sense to prioritize based on the value of an asset's content. In many cases, content-based advocacy catches an administrator's or donor's attention more than technical needs. However, a prioritization plan should consider both.

Answers to the following questions will help you to prioritize based on content:

- Is this commercial/published or unique/rare content?
- Is this content currently under copyright that restricts availability and re-use?
- Is this duplicate content of other items in the collection?
- What role does this play in the collection's origin and in the current research value of the collection? (e.g., Is it production material of less value than the final master?)
- What is your archive's role in custodianship (e.g., long-term preservation or access only)?

Based on these questions, unless content value dictates otherwise, a content-oriented prioritization scheme would be outlined as:

1. Prioritize unique/rare recordings over commercial/published recordings.
2. Prioritize recordings not under copyright over those under copyright.
3. Commercial/published works should be digitized if they are considered to be of continuing research value to the collection. These items may be digitized at a lower resolution if it is determined that:
 - a. other institutions or original rights holders are also caring for them,
 - b. the content is easily replaceable through other means, or

- c. there is limited value in long-term preservation, but there is more immediate value in exhibition or educational use.

Exact or lower quality duplicates, such as published materials or dubs from a master source, should not be digitized. Duplicate content derived from production processes should be prioritized for digitization based upon the research priorities of the collection.

Use Value

This consideration will be driven primarily by research requests or in support of your institution's activities such as events, educational use, and web publication. As such, it should be a more strategic consideration used to promote the work of your institution and the depth of its collections. Essentially, use value considers how the content of an asset relates back to the mission of your institution. Here, you will want to ask yourself the following questions:

- Does this content support my mission?
- Does the content fit within my collecting policy?
- Does this content serve my core audience?

If you answered yes to these questions, then you should treat that asset as a high priority.

Cost to Digitize

Based on the various technical issues discussed in "Section 3.1: Technical Needs," different formats will cost different amounts to digitize. This could be due to the scarcity of playback machines or the rate of chemical degradation of the object. No matter the reason, the result is that it is vastly more costly to digitize one format versus another. For example, digitizing a 2 inch open reel videotape can cost as much as ten times the amount of digitizing a compact cassette.

You will have to take into account all factors—technical issues, content, use value, and cost—to decide if it makes more sense to digitize one 2-inch open reel videotape or 10 compact cassettes. There is no right or wrong answer here, as long as you think through all the factors and decide what is best for your institution.

Finding a Balance

Prioritization of your institution's assets for digitization requires a careful balance of technical needs, content value, use value, and cost. No one area should necessarily dominate your decision making process.¹⁴ How each area is weighed will be unique to your institution. Hold discussions with interested stakeholders to determine a course of action that is right for your institution and collections. That being said, in some ways technical needs are still paramount. The reality is that not all of the audiovisual assets in your collection will be viable more than 10 years from now. If there is a chance you will wish to preserve the content of certain formats, regardless of the value of the content, then it is a good idea to digitize those sooner rather than later.

CONCLUSION

The inventory and assessment of collections is a critical process that will have a lasting impact on an audiovisual media collection at every subsequent stage of reformatting and preservation planning. When creating an inventory, the entry of the required fields: Unique ID, Location, Media Type, Format, and Title will provide an institution with the basic level of knowledge needed to determine which of its materials need to be reformatted first. Optional fields such as Collection Name, Generation, Description, and Condition, among others in this chapter, will further enhance an institution's decision-making abilities.

There are helpful tools that can streamline the inventory process, ranging from third party software that is selected and downloaded online to spreadsheets that come installed on nearly every computer. Whatever the chosen tool, once the information is gathered, a repository must make careful choices when selecting materials for digitization. These decisions should take into account the institution's mission statement, the materials' content value, and the technical needs of the collection. As discussed in Section 3, the application MediaSCORE can help make these technical needs clear.

The inventory and assessment of audiovisual media underpins and supports all future projects for the collections, making the time spent reviewing and describing your materials worthwhile. In the long term, these steps will help you develop metadata for digital audiovisual collections, which makes conducting effective salvage efforts easier. In the short term, an institution will be prepared to draft a Statement of Work and submit a Request for Proposal to a vendor while applying resources efficiently. These are the key components of pursuing a reformatting project and are discussed further in "Chapter 3: Planning, Preparing, and Implementing Reformatting Projects."

¹⁴ Unless "digitize everything" is something you can afford to do, in which case you should prioritize based on technical needs.

CHAPTER 3

PLANNING, PREPARING, AND IMPLEMENTING REFORMATTING PROJECTS

by Chris Lacinak, Founder and President, AVPreserve and edited by NEDCC staff

AT THE HEART OF PLANNING AND IMPLEMENTING a successful reformatting project is clear and strong communication. Whether this work is performed in-house or with a vendor, there is very little difference in the steps that should take place. Due to this fact, and for the sake of simplicity, this chapter will refer to “clients” and “vendors” to mean both internal digitization programs and contracted vendor services. The client is the department, unit, or organization with the need to reformat audiovisual media in their holdings. The vendor is the internal department or an external organization that will perform the reformatting.

Solid communication begins with the client clearly articulating their needs and specifications. This goes beyond specifying the input (e.g. Open reel audiotape, VHS) and output (e.g. 24-bit 96 kHz WAVE file, uncompressed Quick-Time file). When only given these two reference points, too many assumptions are left to chance; each one of these assumptions creates an opportunity for disconnects and miscommunication, greatly increasing the likelihood of the vendor delivering an end-product that does not actually meet the client’s needs.

The need for specificity and detailed communication is particularly important with the reformatting of content stored on legacy audiovisual media, such as Digital Betacam, VHS, audiocassette, and audio open reel. Most organizations do not have the equipment or expertise to view the media they hold in their collections and therefore have no point

of reference for the quality of the moving image and sound recorded on their assets. What is typically known is that it is normal and expected for there to be challenges, audiovisual artifacts, and certain aesthetic characteristics when reproducing content recorded on legacy media. This is due to a combination of obsolescence, degradation, and the fact that legacy technologies generally have specifications, which were typically high-performance at the time of their release but are much lower quality when compared to current technology. How does a client know that a poor-looking or -sounding transfer is due to the original recording quality, degradation, or an error on behalf of the vendor? The answer in most cases is that they almost never know, but this does not render the client powerless or subject to blindly trusting the vendor.

Clients can greatly mitigate risk by using specifications to ensure preservation-oriented workflows and practices are utilized and quality assurance and control measures are incorporated. A great place to document these expectations and specifications is in a statement of work (SoW), a document that is typically incorporated into a request for proposal (RFP) and details all of the requirements of a project with regard to standards, practices, protocols, timeline, and technological specifications. Using an SoW will also help clients fairly perform comparative analysis between vendor proposals. When there are few or no specifications included as part of an RFP, there is no way to identify reasons for differences in vendor proposals,

such as wildly varying pricing, or to understand how the differences relate to vendor workflows and practices.

Many organizations do not have the expertise in-house to draft a detailed SoW for a preservation-oriented reformatting project. This chapter will provide an SoW outline and set of discussion points along with sample SoWs and specifications. These should not be used in a “copy-and-paste” fashion. This will not serve the client, the client’s true needs, the vendor, or the spirit of strong and clear communication. They should be used for establishing a terminology and a framework that organizations can use as a foundation for research, conversations with vendors, and ultimately creating their own SoW.

While solid communication begins with the client, it does not end with the client. Vendors have a critical role to play in this regard as well. Vendor communication begins with responding to a SoW, asking questions, offering suggestions, and engaging in a conversation. These points of feedback can be used by the client as the basis of discussions held with other vendors and for performing further research within the client’s own organization. Vendor communication also consists of keeping in touch with the client at logical points throughout the project, although the expectation for communication from the vendor should be made clear within the client’s SoW.

Too many times, the project ends when a vendor sends the final deliverables of a project to a client and no quality control is performed by the client to ensure that there are no audio or video quality issues or that specifications regarding file naming conventions, organizational conventions, and metadata were met by the vendor. In these cases, a question is raised about the value of creating a set of digitization specifications for a vendor if the client is not going to verify that they have been met and that the deliverable meets their expectations and needs. While the vendor should perform their own quality assurance, the client should perform quality control in addition to this. Mistakes happen, even from well-intended, expert vendors. The client shares part of the responsibility and burden to ensure that the final deliverable meets the specifications outlined in the SoW. For every specification provided to a vendor, the client should have a quality control protocol that verifies that the specification is met.

When issues arise, they may or may not be vendor errors. There needs to be communication between the client and vendor regarding identified issues in order to see them through to a mutually agreed upon resolution, and any necessary rework must be performed and put through the same quality control process.

When solid planning and preparation is followed by strong and clear communication between the client and vendor, the risk of failure or disappointment is greatly reduced. There is an old saying that good fences make for good neighbors. Consider a detailed SoW to be a good fence. It leaves little question about the parameters of the agreement. Some may feel that this inserts a formality or rigidity in the client-vendor relationship that disallows a more friendly and convivial relationship. To the contrary, like good fences, clear and explicit statements of work and specifications allow both clients and vendors to relax, knowing what’s expected of them, fostering a more positive and cooperative relationship.

SECTION 1: THE DIGITIZATION SIGNAL PATH

While the exact technologies and workflows vary from project to project, it is helpful to understand the general setup and process that a vendor will use with your AV materials in the reformatting process in order to determine your internal priorities and specifications. The diagram below shows a very generic signal path with associated AV inputs and outputs. This diagram assumes that the original item has already been inspected and any necessary physical repair and stabilization has already been performed. This diagram also only speaks to AV and not to metadata.

This diagram is extremely simplified in order to show the high-level, salient components of a digitization signal path. An original physical item is played back on a machine. This machine must be calibrated and aligned to the original item in order to achieve a faithful reproduction of the original recording. The signal processor inserted in this diagram could be any number of different devices depending on what the original item format is. For audio discs this may be a preamp and equalizer. For video formats this may be

a time base corrector. For audio open reel this may be noise reduction decoders. Principally, signal processors in a preservation reformatting signal path are not about enhancing the signal. They are devices that are necessary in order to reproduce the original recording with integrity. If they were not present in the signal path, a faithful reproduction could not be achieved. In some cases, their absence may mean that no reproduction could even be achieved.

The output of the signal processor(s) is fed into the analog to digital converter (ADC), responsible for transforming the signal from its analog source to a digital approximation. The quality of the ADC is critically important in the ability to create a digital copy with high precision and integrity. There are video ADCs and audio ADCs. The digital stream of bits created by the ADC is routed to an encoder which takes this raw digital stream of data and converts, maps, and packages it into a file format or wrapper. When this is written to the hard drive as a file, the Preservation Master file is produced as a result. In order to create derivative files such as Mezzanine and Access Copy files, a transcoder is needed to decode the Preservation Master file, thereby creating a data stream that can then be used to convert, map, and package the AV into the appropriate specifications for the Mezzanine and Access Copy file formats.

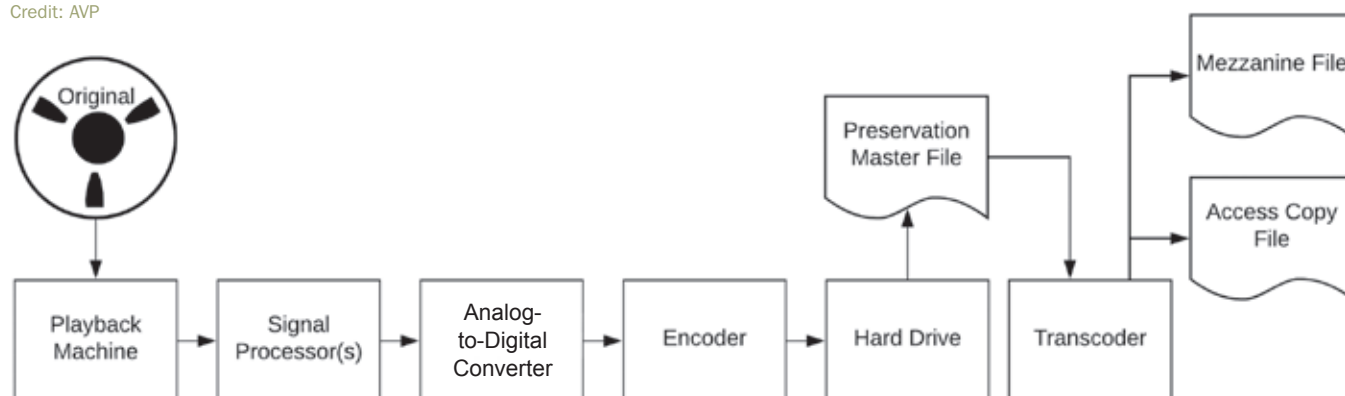
SECTION 2: THE REQUEST FOR PROPOSAL (RFP) PROCESS

For organizations working with either an internal or external vendor, releasing a request for proposals (RFP) is the best way to identify the best fit for the project and establish good communication early. If the RFP process is new to you, thinking of it like hiring an employee can be helpful—many of the steps are similar. There should be a vetting and selection working group established for this purpose, just like a hiring committee would be gathered to evaluate candidates for a job opening. A “job description” should be released and advertised in the form of a written SoW. Vendor proposals should be reviewed, and vendors should be interviewed by the committee, just as you would with a candidate.

In smaller organizations, the SoW is often the entire RFP. In larger organizations, particularly when there are procurement departments, there are often additional administrative, contractual, and legal sections, and the SoW is one part of the RFP. In this case, once the SoW is finalized it is turned over to the people in charge of procurement and typically goes through review with a focus on contracting and legal aspects. In cases in which there are a number of people involved in the process on the client side, it is important

FIGURE 3.1
A Diagram of the Digitization Signal Path

Credit: AVP



before the RFP goes out for all stakeholders to meet, review the SoW, and confirm the timeline to make sure that all parties are on the same page. This should include the primary drafters of the SoW, procurement, legal, IT, and any other stakeholders the project will either depend on or directly impact. Once everyone is in sync and all dates in the timeline are updated and current, the RFP can be released. This may be done by sending the RFP to selected vendors, posting it publicly, or both. In some cases, law or organizational policy may dictate whether an RFP must be posted publicly.

After an RFP is released, it is customary to provide a period in which vendors can ask clarifying questions about the RFP, your organization, or the project. Publishing specific information about how vendors can contact you with questions in the RFP, or even setting up specific conference calls, can keep this process moving quickly.

Once the RFP has been issued, vendor questions have been posed and responded to, and the final proposals are in from vendors, the vetting and selection process begins. Each member of the working group should review each proposal and document comments and questions. Some organizations choose to use a scoring system, rating vendors in specified categories. If this is the case, documenting these scores should be done as part of the review process as well. Once everyone has reviewed the proposals, the working group should convene to discuss. If there are any vendors that are definitely seen as a bad fit, they may be removed from the process. All questions for each vendor that is moving forward in the process should be consolidated and either a meeting should be scheduled with each of the vendors to ask questions or an email should be sent asking for written responses along with a deadline. Once all responses are received, the working group should reconvene to discuss and select a final vendor. Keep in mind that selecting multiple vendors is also a possibility. For very large projects, it might make sense to split between vendors either to spread the load or to work with vendors that are particularly strong with a given set of formats. Keep in mind that the client may need to revisit the vendor candidate pool if the pilot or project goes poorly with a selected vendor. All vendors should be notified of the decision promptly regardless of whether they were selected or not.

SECTION 3: INTERNAL PROJECT PLANNING

Conducting a successful reformatting project requires thoughtful planning that addresses three major questions:

1. What are the human, technology, and financial resources required within the client organization?
2. What are the roles each client staff member will play throughout the project?
3. How will the client organization handle the digital files and original analog materials once they are returned?

Answering these questions before the project gets underway will help the reformatting process run smoothly and avoid issues that may arise from having an incomplete understanding of available funds, team roles, and steps to undertake when receiving digital files. The following subsections provide guidelines to consider when making decisions regarding the reformatting project. Existing literature on effective project planning and management is rich and addresses many of the concerns relevant to AV reformatting projects. Where appropriate, this textbook provides excerpts from important works on project planning.

Committing Staff and Resources

Available staff and funds will invariably play a critical role in determining the scope of the project and the pace at which it is completed. This makes it important to identify available funds and staff hours as soon as possible when planning a reformatting project so that expected goals are realistic and that the projected timeline is manageable. However, identifying these resources is only the first half of the first step of project planning. The second half is allocating resources and delegating roles to staff effectively. A valuable reference that addresses both of these topics is the *Handbook for Digital Projects: A Management Tool for Preservation and Access*.¹⁵ Specifically, a section in Chapter 3, “Project Management: Creating a Plan of Work and Budget,” discusses approaches to committing staff and resources in great depth. And while this reference places a focus on scanning and microfilming, its advice regarding project planning remains true and applicable to audiovisual reformatting.

15 Maxine K. Sitte, ed. *Handbook for Digital Projects: A Management Tool for Preservation and Access*, 1st ed. Andover, MA: NEDCC, 2000.
<https://www.nedcc.org/assets/media/documents/dman.pdf>

Project Team Activities

PROJECT MANAGEMENT

Having a project manager on the client side is critical for a smooth and successful project. The project manager oversees all parts of the project, including, but not limited to: vendor interfacing; managing the client staff; monitoring scheduling and timeline; keeping track of quality control issue resolution and rework status; coordinating with client stakeholders; troubleshooting issues that arise; and proactively working to anticipate and identify prospective issues. Whether this resource is internal to the organization or a third-party provider, it is a critical role that must be filled.

RFP DRAFTING AND PROPOSAL VETTING

The client will be responsible for drafting the RFP, identifying vendor candidates, vetting proposals submitted by vendors, and ultimately selecting the vendor or vendors with which they wish to enter into a contract. This process should be driven by the project manager with the inclusion of other key stakeholders.

CONTRACTING

Once a vendor or vendors have been selected, a formal contract should be written. If the client has contracting, procurement, and/or legal resources on staff they will likely be responsible for this step. If not, the vendor can be asked to provide a draft agreement, but the client should be sure to have a legal advisor review before signing.

THE PILOT PROJECT

A pilot project is a project phase in which a sampling of items is selected to run through the entire process in order to identify and resolve any problem areas or gaps in the processes, protocols, specifications, or understanding. This allows both the client and the vendor to identify and resolve any issues in the process and specifications before launching into full production mode. A pilot phase is highly recommended, especially the first time a project is being performed, the first time a project with a new set of specifications is being performed, or whenever a client and vendor are working together for the first time. Usually it is best to select a sampling of items that will represent

the variances and test as many aspects of the SoW as possible.

There really is no “right” size for a pilot project. Generally, the size is based on a combination of factors including the amount of time it will take the vendor to complete, the number of items needed to represent the variables and diversity of the project, and the number of items seen as sufficient to serve as a proper test. The more complex the specifications and the larger the project, the more important a pilot phase is.

In cases where a pilot project proves very challenging, or in cases of very large projects, it may make sense to perform multiple pilot projects. In the former case this will help ensure that the process is fully ironed out before launching into full production mode. In the latter this may allow a ramping up period where each pilot is successively larger before hitting full throughput.

This is particularly important in the case of large projects where items are being digitized quickly, errors propagate at scale, and much time, energy, and good will can be expended in dealing with identifying and resolving errors. It is much easier to do things right the first time than to have to go back and fix things. This is almost always true, but it’s particularly true with larger-scale projects.

Pilot phases tend to be challenging because they often reveal more issues to be resolved than anticipated and resolving those issues tends to take collaboration between parties, which takes time. Essentially, the pilot process takes time, and it is important to allow for an appropriate amount of time. And as frustrating as it can be to iron out details when everyone wants to get working in full production mode, the up-front hard work pays major dividends throughout the rest of the project. Rework and re-performing quality control on a small number of items at the beginning of the project is far less painful for all parties than doing the same on a large number of items at the end of the project.

PREPARING, PACKING, SHIPPING, AND RECEIVING

Depending on the size of the project and how the original items are stored, this step could take a significant amount of time. Client staff are responsible for carefully packing up all assets to be sent to the vendor(s) and creating

a detailed manifest for each box and for each shipment. Packing and shipping recommendations by carrier type were outlined in Chapter 1, “Section 6: Shipping of Media Carriers.” In some cases, there may be work to do before items can be packed. Some organizations have requirements regarding the minimum amount of processing that must be performed on items before they can be sent for digitization. These requirements are usually focused on registering items with the organization’s system of record, assigning or capturing an identifier, and perhaps capturing one or more additional metadata fields. If these requirements exist, then completing them can take a lot of resources and time. This should be identified and accounted for as a process that precedes packing and shipping. Also remember to account for the fact that you will be receiving and reshelving these materials, too. Additionally, make sure to account for all staff time as well as the costs for materials and supplies for all of these processes.

QUALITY CONTROL

Upon receipt of digitization deliverables from the vendor, client staff must perform quality control. These measures may include testing that ensures that file specifications align with the requested target specifications, checking file naming conventions, making sure all files are accounted for, validating any XML files, and performing qualitative tests such as playing back files to check for audiovisual quality issues. The quantitative quality control may be performed with less expert, but well-trained staff. The qualitative quality control process benefits greatly from using staff that have audio knowledge and are able to identify, interpret, and articulate audio quality issues. The quality control process will take significant staff time and should not be overlooked when preparing timelines and budgets. It is frequently the case that client organizations do not adequately staff this role, resulting in a project bottleneck that can put the success of the project at risk. Aside from human resources for quality control, it is critical that client organizations use the appropriate specialized equipment and environments that quality control requires. Performing quality control using sub-par equipment or while in a poor environment can hinder quality assurance. For AV quality control it is important that clients have, or procure, equipment and use an environment that keeps the following in mind:

- The computer used for quality control must have sufficient processing power, memory, and bandwidth for working with audiovisual files at scale and running the necessary applications and processes.
- The computer used for quality control must have installed, and be able to run, all necessary specialized applications for video, audio, metadata, and file processing.
- The computer used for quality control must have a high bandwidth connection to any storage locations that it will be reading files from or writing files to.
- The audio digital to analog converter used must be of high quality. Audio playback must not be performed using the computer’s internal commodity hardware sound card.
- The audio reproducers (e.g. audio monitors, headphones) must be high quality, capable of reproducing the full audio frequency spectrum with precision, and intended for professional audio use.
- If using headphones for quality control, they should be very comfortable for the user to wear for extended periods. They should be over the ear, high-quality headphones created for professional audio use.
- If using audio monitors for quality control, the room in which quality control is being performed should be acoustically transparent. Rooms may be acoustically treated to achieve this as necessary. The acoustics of the room must not add noise that makes it difficult to perform critical listening (e.g. excessive reflections from close hard surfaces, reverberation).
- The environment that the user is in must be free of noise (e.g. humans, machines, HVAC).
- When performing video quality control, video computer monitors used must be high-quality monitors offering precision video reproduction (e.g. color, contrast) used in professional critical viewing environments.
- CRT video monitors should be used for viewing any files that result from originals that utilize interlace scanning. (This applies to all analog legacy video formats and some more modern digital formats.)

- For critical viewing of video, the user should have total control over the lighting in the room.

Failure to perform good quality control not only stands to waste time and money on the process itself, but also puts the goal of preservation at risk and all of the finances dedicated to the project. If there are concerns about the ability to perform this work, in-house clients may consider hiring a third party to perform quality control for them.

Receiving Digital Copies and Physical Originals

When digitization takes place in a preservation context, it would be remiss for the client organization not to prepare for the eventual return of the digitized files. “Chapter 4: Managing Digital Audiovisual Collections” addresses this topic and provides guidance on the necessary protocols and infrastructure. Organizations should consider this in advance of creating significant quantities of files; they should create plans and be prepared to accommodate both the ingest of media and metadata files into access and preservation systems as well as the long-term storage and management of these files.

MANAGING PHYSICAL ORIGINALS

After the physical assets have been digitized, it may be tempting to assume that they are no longer necessary. However, the physical assets still play a role in the overarching preservation strategy. As long as the original assets are able to be reproduced, they offer risk mitigation against loss of digital files, poorly performed reformatting transfers, and future technological advancements that may bring improved reformatting capabilities that the client may want to keep open as an option. Follow the recommendations outlined in Chapter 1 in order to maintain these backups for as long as possible.

SECTION 4: DRAFTING A STATEMENT OF WORK

An SoW, as defined for the purpose of this text, is a document that provides context to help the reader understand the goals and intent of the project, followed by details of the scope, process, timeline, communication protocols, technical specifications, and a request for pricing. In larger

organizations an SoW may make up one part of a request for proposal (RFP) that also includes legal and contractual information. In smaller organizations it may make up the heart of the RFP with few other components.

The process of drafting an SoW and issuing RFPs provides an opportunity to learn from others, ranging from organizations that have recent experience to share about performing similar work to vendors who are willing to engage and share their expertise and broad experience across a diverse customer base. This section will provide a framework, reference points, and a vocabulary that can be used to prompt conversation with others as you create an SoW. Before issuing your RFP, ask colleagues to review and critique it in order to help confirm and/or shape decisions. If the SoW will be used with a digitization lab that is internal to the client organization, meet with the digitization lab team throughout the SoW drafting process to seek feedback and input. If issuing an RFP for external service providers, the Q&A portion of the process can be a great point of reflection if vendors ask questions that point out unforeseen issues or identify gaps that may have been missed. The author of this chapter worked for a reformatting service provider serving archives between 1999 and 2006, created and taught a graduate level course on this topic for three years, has drafted over 25 SoWs for clients of AVPreserve since 2006, and has managed several of those projects through to completion. Even still, every new SoW and every new implementation brings new insights, lessons learned, and ways to improve and do better the next time. There is always more to learn and ways to continue improving, so use this document as a starting point, but also be sure to use the process as a platform for learning and refining.

There are a number of pieces of information that are important to convey in any SoW, for the sake of both the client and the vendor. The headings below may serve as an outline for an SoW or may simply serve as specific points that are incorporated into an SoW. The thoughts and considerations supporting each heading will provide the information needed to best decide how to incorporate these points into an SoW that is designed to serve a specific organization and set of circumstances.

Before delving into the details below, it may be helpful to review the diagram of the digitization signal path in Section 1 in order to gain perspective on what follows

the vendor selection and pilot project phases of a larger project. This diagram illustrates how many of the items discussed below tie together within the larger process.

Sections of a Statement of Work

ABOUT THE CLIENT AND PROJECT

This section provides the vendor with some context about the client organization and project so they have a basic understanding of who the client is and what their goals are. This serves two important functions. First, it provides frames of reference to ensure understanding and to help avoid misinterpretation based on incorrect assumptions. Second, it allows the vendor to engage as a collaborator more easily because they have the information that will help them to offer suggestions or provide alternative paths to take while reaching the stated goals, objectives, and outcomes.

BRIEF DESCRIPTION

In this section, provide the vendor with the scope of the materials to be digitized. Typically, this would be a breakdown of formats by quantity and duration. If duration of the recorded content is unknown, use the maximum media duration of an asset (e.g. 60 minutes for a 60-minute audiocassette) as a safe outside estimate. If media duration is used, it would be wise to point this out to the vendor so that they know the logic behind these numbers. It is best to provide format information at the variant level. For instance, if the collection contains Betacam, Betacam SP, Betacam SX, and Digital Betacam, do not put all of these formats under “Betacam.” List each of them separately and provide the quantity and total duration for each.

This section points out the importance of performing an inventory of the materials to be digitized prior to drafting the SoW. Depending on the number of items that are candidates for digitization and the size of the client’s budget, there may be the need to prioritize items for selection as well. Read more about performing inventories and selecting for digitization in “Chapter 2: Inventory and Assessment.”

TIMELINE

It is easy to overlook the number of steps that must take place in the contracting and implementation process as well as the number of parties that may need to be involved. Documenting and communicating each step with its associated due dates allows all parties to work more cooperatively and in synchronization. The act of thinking through the process in this level of detail helps to create a realistic timeline that will avoid unpleasant surprises down the road. The timeline below offers an example of what the steps might include. Your timeline may look different depending on the contracting process.

- **RFP distributed:** yyyy-mm-dd
- **Bidder questions:** yyyy-mm-dd
- **Client responses** (Q&A will be provided to all respondents): yyyy-mm-dd
- **Proposals due:** yyyy-mm-dd
- **Award:** yyyy-mm-dd
- **Contracting completed:** yyyy-mm-dd
- **Pilot assets shipped to vendor:** yyyy-mm-dd
- **Pilot phase completed by:** yyyy-mm-dd
- **Media shipments begin:** yyyy-mm-dd
- **Ship files to Client:** To be delivered on a monthly basis for quality control and approval
- **Client QC completed:** To be completed xx days after delivery on a regular basis
- **Rework completed:** Within xx days of completion of project
- **Ship source objects to Client:** Within 30 days of completion of Client approval of rework
- **Project completed by:** yyyy-mm-dd

There are two phases in this process that are most likely to cause timeline issues and problems between clients and vendors. It is important to recognize them as part of the process and to allow sufficient time for them. These are the pilot and rework phases.

The pilot phase is discussed in its own section below. The rework phase, the phase in which vendor mistakes are corrected, is often overlooked entirely. All parties would prefer to think that the project will be smooth sailing and any necessary rework will be negligible at most. Sometimes this is true, but major problems arise when rework hasn't been factored into the timeline at all, especially when everyone is up against a hard deadline, and the vendor delivers right at the end of said deadline. Note that performing a pilot phase will mitigate the amount of rework that is necessary, but it is still important to plan for rework nonetheless.

Aside from performing a pilot phase, another way to avoid rework is to stage the deliverables so that quality control is performed at multiple periods throughout the project instead of all at once at the end of the project. When quality control is performed earlier in the process, it offers opportunities to identify and correct any issues as work proceeds. Also, it allows for more measured staff allocation for both the client and vendor, which is a much better option than performing quality control and rework all at once, usually when there is little time left in the project, a situation which is not conducive to performing good quality control or rework. Spreading quality control and rework out over the course of the project lessens the chance of ending up with unpleasant surprises at the end. One final note on rework is that it quickly leads to the potential for challenging project management if it isn't identified as an auxiliary process to the main digitization effort. It is important to establish a process between the client and vendor on how rework will be managed, tracked, and reported throughout the project.

CLIENT POINTS OF CONTACT

This section should identify points of contact and associated information on the client side for project management, technical matters, administration, contracting, and invoicing. Not all points of contact have to be named in the SoW. Some can be identified as the project evolves, though it is best to provide the information up front if it is known.

COMMUNICATION PROTOCOL

This section defines how the client and vendor will communicate, at what frequencies, and at what point in

the timeline. There are eight points of communication that should be considered and addressed in this section.

1. Shipment of materials from client to vendor

The client should identify the information and means of delivery that will precede and/or accompany the shipment of materials to the vendor so that the vendor can plan and prepare accordingly for the management of this information upon receipt. This typically consists of two documents.

The first document is a shipping manifest that the vendor will reconcile the received materials against. There should be a manifest for each box that is printed and then both included inside and affixed outside of the box. Each box should have a box ID that is documented in a master shipping manifest and that is used for internal tracking purposes within the client organization. The box ID should be clearly placed on each box.

The information on the shipping manifest should include item level identifiers and/or label information that will allow the vendor to quickly identify the item on the manifest when unpacking the boxes. If this proves difficult based on the current labeling and/or identifying mechanisms on the items being shipped, then applying new identifiers to the items should be considered. In addition to including printed copies of the shipping manifests in each box, it is also a good idea to send a digital version of the master shipping manifest containing the total number of boxes, the box ID for each box sent, and a list of all of the items in each box.

The second document is a Source Metadata document, which includes information that will be helpful to the vendor as well as metadata that the client would like to persist through the digitization process and that the vendor should incorporate into the deliverables, potentially to remain as part of an archival package. This might include unique identifiers, title information, rights information, or any type of descriptive or administrative information that makes sense to keep as part of the archival package. The Source Metadata should be provided as a digital file, typically in spreadsheet form or XML. This may serve as an electronic backup to the packing manifest as well.

If the client is responsible for shipping materials to the vendor, this section should also identify how far in

advance the client will send notification of shipment and the Source Metadata document. If the vendor is responsible for transport of the materials, then this section should define when and how the client and vendor will be in communication about pickup to ensure a smooth and coordinated shipment.

Note that this may happen in batches for larger projects as opposed to shipping all items at once for smaller projects.

2. Verification of receipt at the box level by vendor

This should identify how long the vendor has to notify the client that they have received the materials and to offer a box-level reconciliation confirmation, as well as the means of communication and which client point of contact(s) to notify.

3. Verification of receipt at the item level by vendor

This should identify how long the vendor has to notify the client that they have received the materials and offer an item-level reconciliation confirmation. It should also stipulate the expected means of communication and which client point(s) of contact to notify.

4. Routine status updates

The need for status updates throughout the process will be dependent on the size of the project. Projects that will be finished in a few weeks may not need any status updates. Projects that last longer than a month will benefit from them. Projects that last for multiple months require them for successful completion.

This section should identify the frequency and method of status updates. The method may be meetings, emails, updating of shared cloud-based documents, or some combination of these. This section should also lay out the expectation of what information will be provided in status updates. If the method chosen is meetings, this section might also provide a generic agenda for these meetings, typically including a status report on progress from the vendor, reporting on projected progress over the next period, discussion of any outstanding issues, questions from either party, and any administrative or financial matters.

In larger projects, it is often advantageous for the client and vendor to have shared access to a project management spreadsheet (e.g. Google Sheets) or application where each party can see the overall status of the project via summary information in real time, as well as status at the item level where it is possible to track and communicate about each item through the process. For instance, the vendor might document issues encountered in inspection and place them here for the client to have insight into the issue and make a decision about how the vendor should proceed. Or the client might document their quality control findings and the vendor might document their response and/or rework status. This document or application may also specify the unique identifier for the storage media that the generated files are being delivered on in order to make it easier for the client to manage quality control and receipt of final deliverables. If such a system is used, it is important to identify who will be responsible for populating which fields and at which point in the process. Having such a mechanism for project management and immediate transparent access to the project status by all parties promotes tighter coordination and collaboration as well as overall project success.

5. Shipment of digital files from vendor to client

This should identify how far in advance and who the vendor will notify within the client organization before shipping the deliverables for quality control. Note that the vendor should maintain possession of the original materials until quality control has been fully completed in case the vendor needs access to the items for inspection or rework.

This section should also identify the information that the client expects the vendor to send in advance of and/or along with the shipment. At minimum this should be a shipping manifest detailing the number of boxes being sent, the box ID for each box being sent, the number of items being sent, and the item ID for each item being sent. This should be included in print form inside and outside of each box, and a master shipping manifest should be sent electronically. The electronic version might also include a manifest of the contents of each item being sent.

The client should identify the turnaround time on notifying the vendor of receipt at the box and item level to ensure closing the loop. Note that this may happen in batches for larger projects as opposed to shipping all items at once for smaller projects.

6. Quality control findings and resolution

This should identify how quality control will be managed and tracked throughout the process. Quality control consists of the client performing quality control, client documentation and reporting of findings to vendor, discussion, and investigation of any identified issues by vendor, rework as necessary, and re-performance of quality control as necessary.

This should identify the method of documenting and tracking this information, who is responsible for documenting which parts, and the turnaround time for each part of the process.

Properly managing this auxiliary process can be confusing and is often difficult in the midst of a digitization project. Establishing the protocols and methods of communication here will go a long way in making this process as smooth as possible. To reiterate, having a shared project management spreadsheet or application that incorporates quality control documentation and management is advised.

7. Shipment of original materials from vendor to client

This should identify how far in advance the vendor will notify the client before shipping the original materials, the means of notification, and who should be notified. It should also identify the information that the client expects the vendor to send in advance of and/or along with the shipment. At minimum this should be a shipping manifest detailing the number of boxes being sent, the box ID for each box being sent, the number of items being sent, and the identifying information for each item being sent within each box. This should be included in print form inside and outside of each box, and a master shipping manifest should be sent electronically.

The client should identify the turnaround time on notifying the vendor of receipt at the box and item level to ensure closing the loop. Note that this may happen in batches for larger projects as opposed to shipping all items at once for smaller projects.

8. Notification of ability to delete information from vendor systems

Before a vendor deletes client information from their systems, there should be explicit agreement from the client. This ensures that all quality control issues are resolved and that the client has time to ensure that all digital information is secured, with multiple copies stored in multiple locations, and any ingest routines are completed. Obviously, this needs to happen in a reasonable amount of time, as it places a burden on the vendor to maintain this data on their systems. Therefore, this section should not only identify the means of communication but also the anticipated turnaround time in order to allow the vendor to plan accordingly.

SHIPPING

The shipping section should be used to discuss any aspects related to the required method, speed, protocols, care and handling standards, material specifications, and insurance requirements related to packing and shipping. This is a good place to identify who is responsible for providing packing materials, performing packing and shipping, and paying for shipping. Note that if the vendor is responsible for any of the labor or costs associated with packing and shipping, be sure to make it clear that they need to include this in their proposal and pricing.

PILOT PROJECT

This section should lay out the parameters of the pilot project, including the size of the sample to be included, the formats that will be selected, the conditions under which a second pilot project may be performed, and any other relevant details. The size of the sample should be based on the size and variety of the formats selected for digitization in the SoW. In general, the pilot project will be most useful if it includes a sampling of each of the formats the client plans to reformat. This will give the client and the vendor the opportunity to test the specifications outlined in the SoW for each type of material. A second pilot project will be warranted if errors due to the reformatting process are identified or if specifications the client outlined in the SoW are found to be unsuitable for the goals of the project.

DEFINITIONS

One of the largest communication issues that tend to cause problems is how different people apply different definitions and assumptions, resulting in miscommunication and misunderstanding. This section should be used to define any terms that are commonly interpreted in different ways or that the client is using in a way that may not be obvious to others.

CARE, HANDLING, AND STORAGE

The care, handling, and storage of materials is a common point of differentiation between vendors that are preservation focused and those that are not. This section provides an opportunity to convey requirements and request proof of adequate care, handling, and storage from vendors.

Relevant aspects to raise here in the context of care, handling, and storage include:

- staff experience;
- company policies, protocols, and practices;
- storage and facility environmental specifications; and
- security.

MEDIA TREATMENT AND PREPARATION

Vendors have differing approaches to media treatment and preparation, such as cleaning, baking, and repair. These differences can have significant cost implications and may represent differences of opinion with the client as well. This section should be used to communicate any specifications that the client feels strongly about so that all vendors are referencing the same specifications in their pricing. In the absence of preferences from the client, this is a good place to ask the vendor to provide information about their standard media treatment and preparation protocols in order to provide more insight into their practices and to help you understand pricing variances between vendors. Note that this is also an area that may separate preservation-oriented vendors from non-preservation oriented vendors.

REFORMATTING

Similar to the section on media preparation and treatment, the reformatting section is a place to document any client preferences and/or to request information on vendor standard protocols. There are four main areas that are generally addressed in this section:

1. Reproduction Setup

Referring to the actions performed to the equipment and signal path immediately before beginning the transfer, reproduction setup is an important cost driver and differentiator for analog sources. Typically, this speaks to calibration and alignment processes performed on the playback equipment in order to help ensure a faithful reproduction of the original recording when playing back the source item. Different media types and formats have different reproduction setup routines, and this section should speak to each specific one. If there are no client preferences, inquiring about the vendor's standard protocols for this will allow comparisons between vendor proposals and pricing.

2. Signal Path

The path that the audio and/or video signal follows from the output of the playback device to the input of the recording device is the signal path. The quality of the devices in the signal path, the cable and interfaces used, and the way in which the signal path is constructed has a significant impact on the quality of the transfer. Stating preferences and/or gaining insights into a vendor's signal path is a useful piece of information for evaluation.

3. Image and Sound Processing

Best practices for preservation have dictated for decades that "flat" transfers be made when creating a preservation master. "Flat" transfers are unaltered digital versions of the content on the analog media. The intent of this approach is clear: not altering the signal in any way that detracts from a faithful reproduction of the original recording. However, the interpretation and application of "flat" differs from format to format and from person to person. It is important to forego the assumption that everyone has the same interpretation of "flat" and instead to be specific about what types of image and sound processing are and are not acceptable.

Inquiring about the vendor's standard protocol and opinion on this matter is appropriate here.

It may be the case that mezzanine or access copies (see section below for more details) have different specifications for image and sound processing. Some organizations want enhanced copies for access while others prefer to enhance only specific pieces of content, on demand and as needed. Any preferences should be stated here.

4. Destination File Format Specifications

Whereas the other subsections of the reformatting section offer some latitude insofar as the client may ask the vendor for input or for an opinion, the specifications for destination file formats does not. This is where the client must document in as much detail as possible the target digital formats that they want to receive. These may differ based on media types and format, and in these cases, it is important to provide details for each distinct set of specifications and to clearly identify which media types and formats they apply to.

There are three common destination file types, each of which has a distinct role to play:

- a. **Preservation Master**—The primary role of the Preservation Master is to provide an authentic reproduction of the original recording that enables a path to sustainability and long-term access.
- b. **Mezzanine**—The primary role of the Mezzanine is to offer a copy that is ready for use in standard production workflows and systems. There is more than one use case which suggest there can be a routine need for a high-quality version (e.g. for editing or broadcast). While a Preservation Master is in a format and resolution that makes it cumbersome to work with, and an access copy may not be of sufficient quality, the Mezzanine-level file fulfills the need. Mezzanine files are not always needed and should be created only when there is a clear and frequent need. An alternative to creating Mezzanine copies of all files is to create them on-demand from a copy of the Preservation Master file. In cases where the

need for Mezzanine level files is infrequent, this is often the best option.

- c. **Access Copy**—The primary role of the Access Copy is to provide a file that can be reviewed using standard commodity hardware and non-specialized software and infrastructure in order to maximize accessibility of the content.

Depending on the media type and format, file format specifications for each of these file types might include: wrapper specifications, codec specifications,¹⁶ bit depth, sample rate, fourcc code, color space, chroma subsampling, compression algorithm, scan type, pixel height and width, aspect ratio, recording standard, number of audio channels, handling of closed caption information, handling of timecode information, bit rate, optimization for streaming vs download, and endianness.

For video sources it is especially important to specify that all channels of audio be digitized as part of the recording. Video recordings may have anywhere from a single channel to multiple channels of audio depending on the format and how it was recorded. Not all playback devices will necessarily play all channels of audio on a videotape. For formats such as Betacam SP, there are decks that can record and playback four channels, and there are decks that can record and playback two channels. The latter deck is less expensive, so it is not uncommon for vendors to have two-channel decks. When a videotape with four channels of audio is played in a deck that supports only two channels of audio, it may not be possible to identify that there are additional channels present on the tape. The same is true when playing a videotape with timecode in a deck that does not support timecode. Unfortunately, lack of awareness of these issues over the past decades has almost certainly resulted in the loss of a great deal of audio and timecode information. Ideally, the presence of all audio channels and timecode should be identified and reproduced in the digitization process.

An aspect of file format specifications that is not obvious is that, even with all of the above-mentioned details, different encoders using the same exact specifications might construct files differently. If a large

16 Joshua Ranger, "A Primer on Codecs for Moving Image and Sound Archives." 2010. <https://www.avpreserve.com/a-primer-on-codecs-for-moving-image-and-sound-archives>

number of files resulting from a digitization effort share the same specifications but are constructed in multiple ways, this inconsistency can lead to challenges in the performance of quality control, variable results in compatibility with software applications, and eventually, challenges with migration and transformation.

One way to mitigate this issue when working with a single vendor is to utilize MediaInfo¹⁷ or a similar application to create a profile for each target file format and to utilize this profile as the lower level specification moving forward to ensure consistency. This can also be useful with multiple vendors as a conformance point that each vendor must meet to help ensure as few inconsistencies as possible.

It is true that any archive that is involved in generating, acquiring, and preserving digital files will have to learn to work with the inherent variability and diversity that comes along in this domain. While this is the reality, it is also true that consistency is a friend to preservation, and where we have the ability to insert our control to gain consistency we should.

HEAD AND TAIL CONTENT

This section speaks to how to handle the head and tail of content, common parlance for the beginning and end of a program or piece of media. A range of challenges may be present, and it is important to communicate how the vendor should handle each of these scenarios.

The first scenario that occurs is that there may be content at the beginning and/or end of program material that may or may not be relevant or related to the content of interest for preservation. For instance, the beginning of a piece of media may have minutes of black or silence before there is any program material. Or 20 minutes of program material might be recorded onto a 60-minute piece of media, leaving 40 minutes of black or silence at the end. Or there may be bars, tone, countdowns, or slates before the program material. In almost all cases, no organizations are interested in a vendor recording extended periods of black or silence. Usually organizations want to maintain bars, tones, slates, and the like for the preservation master. It is important to let the vendor know how they should handle these issues when encountered. Note that it is

common for an SoW to state how many seconds of black or silence a vendor should record at the end of a program before stopping the recording. However, it is important to include in the SoW that the vendor is responsible for ensuring that there is no additional content recorded after a period of black or silence to avoid missing content in the digitization process.

The above speaks to how to handle the creation of the preservation master. Specifications used for the preservation master may not apply to mezzanine or access copies. Depending on how these lower-resolution target formats will be used and by whom, the client may choose to eliminate anything that is not content of interest (e.g. bars, tone, slate).

REFERENCE FILES

As discussed in the introduction, one common and significant challenge when working on digitization of legacy media is the issue of not having a reference of quality for the original recording due to lack proper equipment, expertise, and labor resources. Therefore, when a digital file is returned from a vendor it is difficult to assess the quality of the work performed. There may be indicators of possible issues, but it is rare that qualitative review yields evidence of obvious errors on behalf of the vendor. Qualitative indicators that are identified in quality control often lead to an exchange with the vendor and further investigation that may confirm whether or not an issue is a vendor error, part of the original recording, or perhaps due to degradation. However, this confirmation is based on the review and feedback of the vendor. The client has little control to make their own assessment and judgment on the matter. This is simply a reality of doing this work, and because clients have little control when it comes to evaluating the output, it is necessary to have control over the processes and practices. This is the reason for the great level of detail and consideration given to these matters throughout an SoW.

Another way to mitigate the risks associated with not having a reference for quality of the original recording is to provide a known reference signal (e.g. bars, tone, black, silence) and have the vendor input it to the same system(s) that will be used to digitize client materials. The vendor is asked to replace the output signal of the playback device

17 <https://mediaarea.net/en/MediaInfo>

used to reproduce client content with the output of a signal generator that outputs the specified reference signals. The specified reference signals are then routed through the exact same signal path and captured in the exact same way that the client's audio or video signal will be routed and captured. This provides some insight to the client on the quality of the signal path and systems that are being used by the vendor. The client can utilize test and measurement software such as waveform monitors, vectorscopes, oscilloscope, audio level meters, and other tools to assess the quality of the recorded signal. Because the input signal is a known reference, variances between the input signal and the recorded output signal can be identified. In essence, it is letting the client know what impact the vendor system is having on the signal, which speaks directly to the quality of the signal path and system, as well as the degree to which it is in good operating condition and calibrated.

This is not a catch-all by any means. This is something that is done periodically, likely under the careful attention of the vendor, and may not be representative of the many tens, hundreds, or thousands of hours that this same system may be put to use for digitizing original materials. It also only speaks to the part of the system that follows the playback device, which means that it is not speaking to the quality and condition of the playback device, or the quality of the setup, calibration, and alignment performed by the operator. However, while it is not perfect, it is still far better than having no insights or reference points. This section of the SoW should provide detailed specifications for the reference signals to be used, the period of time that each should be played, and the precise order in which they should occur. It should also identify how frequently and under what circumstances to create reference files.

Typically, a client will ask for a reference signal per setup, or distinct signal path, being used for digitization. For instance, if the vendor is using four Betacam SP decks and digitization stations and two VHS decks and digitization stations, then this would yield six reference files. A distinct signal path may also be defined by the equipment that is used in between the playback deck and capture station.

When the reference file is digitized, the specifications used should be the same as the specifications for the

preservation master. It is also recommended that the client request that mezzanine and access copy files be made using the same method, hardware, software, and protocols that will be used with the client's original materials. This provides an opportunity to assess both qualitative and quantitative aspects of the derivative creation process as well, utilizing a known reference signal.

For small projects, it may be the case that only one set of reference files per setup needs to be created for the entire project. For longer term projects, it is advisable to identify the frequency with which reference files should be generated. This may be at the batch level if the project is being split into batches, but it could also be weekly or monthly. Whatever the case, it is wise to have periodic snapshots of the vendor's system, as opposed to just one, to ensure that the equipment is in good condition and that the system is properly calibrated and aligned for the duration of the project.

The above approach works well for audio and video, but film is somewhat trickier. For one, there are no widely adopted test films to fulfill this purpose. While test films with different test patterns and signals have been produced from time to time, there is no standard test film that has emerged in the way that test recordings became the norm for audio and video, or even for still image digitization, which has widely-adopted test targets. At this point there is no common test protocol that is used in film digitization projects that is similar to what is discussed above. Creating test films for use on film scanners would be a resource-intensive endeavor, which is certainly a contributing factor for why they are not in wide use.

On the plus side for film, there is often much greater insight into the quality of the original film. Because a client can see the images on a film and can see if a film is deformed or scratched, the disconnect that exists for audio and video is diminished for film.

Finally, in addition to the items mentioned above regarding the contents of this section of the SoW, the client should specify organizational and file naming conventions for the reference files and any associated embedded or external metadata to be delivered along with them.

DIRECTORY STRUCTURE AND FILE NAMING

Digitization projects produce files of all types (e.g. media files, metadata files), and the client must provide specifications for the vendor on how to name and organize them. When establishing file naming conventions for a collection, most people think in terms of newly-derived files reformatted from other sources. In reality, there will be more and more born digital content that already has filenames deposited with archives. In some cases, this content can be renamed to fit the archive's naming structure with no loss of information, but in other cases, such as with MXF files, the inherited naming structure refers to complex file and directory structures that must be maintained in order to preserve the entire content. Naming structures should be flexible enough to recreate any necessary naming conventions.

There are multiple questions that should be answered before drafting organizational and file naming specifications.

- **Are there existing organizational and naming conventions?**

Many organizations already have organizational and naming conventions that can be adopted or adapted. Some organizations are unhappy with their current organizational and naming conventions, so this may be an opportunity to address that. If you choose to depart from existing conventions, be sure to maintain awareness of any internal systems and processes that may have a dependency on the existing organizational or naming conventions and update accordingly. If there are no current conventions, then the opportunity exists to start from a blank slate and create a new convention without concern for inconsistency or incompatibility with internal systems and processes. For tips on file naming, see below.

- **What is the most persistent and pervasive identifier in use in the client organization?**

Naming conventions for directories and files are often based on the identifiers for an object. It is also common for organizations to have multiple identifiers for items. When selecting one to use, give thought to which of these is the most persistent and pervasive, both historically and moving forward. Are any systems or processes dependent on any particular identifier? Sometimes organizations have no identifiers or find that existing identifiers are unsuitable. In this case it may

be appropriate to establish a new identifier scheme and assign new identifiers to items; however, if there are existing identifiers, check to see if they will, in fact, work. Do not rush into creating a new identifier scheme without first diligently thinking it through, as the implications are far reaching, and ending up with too many identifiers can be a problem in and of itself.

- **Will this be the SIP or AIP, or will this organization and file naming be an interim specification?**

SIP and AIP are terms from the Open Archival Information System (OAIS) that stand for “submission information package” and “archival information package” respectively. Although these terms have specific meaning within the context of OAIS, a SIP can roughly be thought of as the components (e.g. metadata and media files) that make up the submission to the archive for longer term storage. These components may undergo further processing as part of the ingest routine before establishing an AIP. An AIP can roughly be thought of as the components that make up the collection of items that the archive will manage over the long term.

Some organizations create specifications for deliverables of a digitization project that equate to a SIP. Once the deliverables are received from a vendor, the client will put them through their own internal processes to generate an AIP. These processes may end up altering the organization and naming of the directories and files delivered from the vendor. In this case it is important that the directories and files are organized and named in way that is aligned with internal systems and processes and that makes the ingest and processing as efficient as possible.

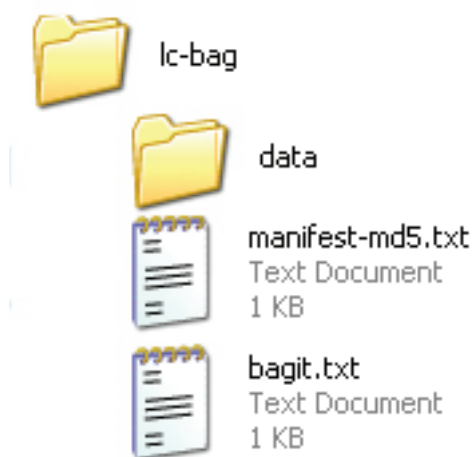
Other organizations create specifications for deliverables of a digitization project that equate to an AIP. They ask the vendor to organize and name their directories and files in such a way that that no further reorganization or naming is necessary to create the AIP. This approach requires consideration of more than just organization and naming of directories and filenames; it requires consideration of the metadata deliverables and parsing through which fields will be produced by the vendor and also of which fields must be provided by the client. (See information about the Source Metadata document under Communication Protocol, above.)

- **To bag or not to bag?**

BagIt is a file packaging format standard developed by the Library of Congress and the California Digital Library and maintained by the Internet Engineering Task Force. BagIt was originally developed as a means of sending and receiving files in a “bag” that would provide built-in data integrity verification for the recipient of the data. In other words, it allows someone who is receiving files to verify that all of the files sent to them are present and have not been altered since the sender created the bag. Another use case that exploits features of the BagIt specification is longer term storage in archives. Routine file attendance and data integrity verification are activities of a digital archive, and some organizations choose to use bags as an encapsulated AIP.

FIGURE 3.1
Bag File Structure

Credit: <https://blogs.loc.gov/thesignal/files/2015/03/bags.jpg>



Whether the vendor’s deliverables are being treated as a SIP or an AIP by the client, bags may prove useful. If being treated as a SIP, a bag can be used to verify file attendance and data integrity by incorporating a bag validation tool into the ingest routine. If being treated as an AIP then file attendance and data integrity can be validated upon receipt, and on an ongoing basis.

At its core, the BagIt specification provides file attendance and data integrity validation through the creation of checksums for every file and a manifest that documents all files in a bag. If the client chooses not to use bags in the deliverables, then it will be important to provide

an alternate specification for the provision of manifests and checksums from the vendor so that file attendance and data integrity can be validated upon receipt by the client.

- **Film**

As is the case throughout this document, film proves to be somewhat of a different animal. Digitized audio and video produce a single file from a single original item, with some exceptions for multiple sides of an item or very long recordings. With film, the most widely adopted preservation master formats produce a file for each frame of the film. This results in many thousands of files from a single film, often stored in a directory that essentially serves as the “wrapper.” Each of the frame-level files are named using a common base name with sequential modifiers appended. These names cannot be modified because players are dependent on the file names in order to play them back properly. The directory is typically named with only the base file name. Some organizations choose to specify that a bag be created from the parent directory so that there is some encapsulation of the directory that provides an automated way to validate file attendance and data integrity at the level of the film.

In addition to being an organizational and file naming consideration, the deliverable of the preservation master for film has logistical implications. Because the structure is different, it may take additional planning when determining how the files will be managed and used and whether or not the difference in structure will have implications.

METADATA

Metadata generation is a byproduct of any digitization project. Whether and how the generated metadata is captured and delivered is up to the client and must be specified. Metadata specifications for projects generally come in two types: external and embedded.

External

External metadata is delivered as a sidecar file, usually either in XML, CSV (spreadsheet format), or a combination of the two. If using a spreadsheet format, CSV is advisable over any proprietary format because it is a transparent

and widely adopted format for data exchange that can be opened in any text editor or spreadsheet application.¹⁸

In addition to specifying the format for delivery, the client must specify the fields, vocabularies, and structure of the metadata being delivered. The types of information encompassed in external metadata deliverables include technical, administrative, preservation, and descriptive metadata.¹⁹ Utilization of standards helps avoid reinventing the wheel and saves some time and effort. Sometimes organizations will adopt a standard as a whole, and other times organizations will choose to use standards in part, selecting

fields or sections along with their associated vocabularies from one or more standards. Organizations often have their own fields and vocabularies that they want included in the metadata deliverable as well. Some organizations may request a single master metadata deliverable while others may ask for multiple metadata file deliverables. There is no single right way to come up with a metadata specification. The right way will be whatever proves to be functional within the client organization. Here are some considerations that will help guide an organization to an answer:

FIVE TIPS FOR FILE NAMING

1. Beware of filename dependency

Filenames are not actually part of the file but rather are part of the file system and are therefore not dependably persistent over time and across systems. Instead, the Unique ID (UID) assigned to the object should be the constant identifier used to track and maintain the provenance of the file. The UID may end up being the same as the filename, but regardless, be sure to embed the UID inside the file in an appropriate and documented place.

2. Do not overthink filenames

Whether the filename is a randomly generated value or not, be systematic. Think, “Is this logical? Can I spell out the rules easily enough to do batch renaming?” In trying to create the perfectly contained and expressed filename or UID structure, there is a strong temptation to overthink to the point that they become non-systematic or too idiosyncratic to be logically parsed. If a naming structure is not systematic enough to have a piece of software perform a series of logical renaming steps, then manual renaming and retyping of filenames may be required at some point in the future.

3. Do not use filenames as database records

Filenames are not the place to cram in a bunch of descriptive and structural information. That’s what databases are for! All we require from a filename and ID is that they act as a link to the database record for that unique object. Trying to cram excessive descriptive information into a filename creates unwieldy names and is often futile because of how often conditions or conventions change and new scenarios come up over time. Having filenames that are tied too closely to specific scenarios creates inflexible structures that require non-systematic revision when situations change, which creates the predicament described in tip #2.

4. Keep filenames machine-readable

There is often an urge to make a file naming structure decodable by humans, but it also needs to be decodable by computers. Avoid characters that are not URL compatible, that require escape characters, or that are reserved by operating systems. Limit options to numbers, letters, periods, and underscores.

5. Consider existing filenames

As noted above, the inherited filenames structure of born-digital content may refer to complex file and directory structures that must be maintained in order to preserve the entire content.

18 See “Managing Data in Spreadsheets” for a tutorial on how to avoid accidentally corrupting data when managing it within spreadsheets. <https://www.avpreserve.com/papers-and-presentations/managing-data-in-spreadsheets-tutorial>

19 “Chapter 4: Managing Digital Audiovisual Collections” goes into these types of metadata further.

How will the metadata deliverables be used?

Considering the questions and points posed below will help determine:

- The format(s) to specify in the SoW
- The fields, vocabularies, and structure to specify in the SoW
- The number and types of external metadata deliverables to specify in the SoW

Questions and Considerations:

- Will deliverables be used as the AIP and kept in packages for long-term storage as-is upon receipt from the vendor?
- Will metadata be ingested or imported into systems? If so, what are the specifications for import into those systems?
- Is the client able to transform the metadata deliverables to meet any system import specifications as necessary, or do they need the vendor's deliverables to conform with the system import specifications?
- Is it easier to send the vendor pre-existing metadata to have them incorporate into their deliverables along with the metadata they are generating, so either (a) the metadata can be ingested in a streamlined way all at once instead of combining things together after the deliverables are received, or (b) the vendor is delivering a package that meets the client's AIP specification? The client may even want to ask for external metadata deliverables that are included as part of an AIP in addition to metadata deliverables that are delivered outside of an AIP that can be used for ingest into systems. If planning on delivering metadata to the vendor for incorporation into their deliverable, the client should specify this as part of the SoW and include it as part of the Source Metadata document(s) sent to the vendor.

What information is desired for capture?

Considering the questions and points posed below will help determine:

- Who will populate which fields?
- When will those fields be populated?

Questions and Considerations:

- In order to put some shape around an answer to this question, standards and metadata used in existing internal systems can serve as a guide. This might also be a good time to start capturing fields that are needed but have not been historically captured. Some standards that are typically used or referenced include:
 - PBCore (<http://pbcore.org>)
 - reVTMD (<https://www.avpreserve.com/products/avi-metaedit-revtmd>)
 - AES 57 (<http://www.aes.org/publications/standards/search.cfm?docID=84>)
 - PREMIS (<https://www.loc.gov/standards/premis>)
- Once there is a list of fields and associated vocabularies, think about which fields it makes sense for the vendor to capture. Of the fields that remain, which metadata already exists in some form? If the balance of fields contains required fields, there will need to be a determination of where and how those will be populated.
- When deciding which fields it makes sense for the vendor to populate, a typical approach is to have them capture any identified fields that would be permanently lost if they weren't captured at the time of digitization (e.g. playback devices, settings used for the digitization process, transfer operator, reproduction issues) and any fields that would be more efficiently captured by the vendor than by the client (e.g. original item specifications, duration).
- Many organizations inquire about whether or not they should ask vendors to generate descriptive metadata, such as describing the content of a recording during the transfer process. Most of the time this does not make sense for two reasons: 1) Vendors are not subject matter experts in the content being digitized and are likely only able to provide the most generic descriptive metadata, and 2) digitization workflows have shifted dramatically toward high throughput workflows that do not lend themselves to systematically capturing descriptive metadata. This question is usually based on the perception that because the vendor is already watching the material, they can easily document what they are seeing. In this way it appears to be an easy add-on or a potential byproduct of the digitization process; however, this is

not usually the case. Descriptive metadata generation should be considered as a separate service with its own set of specifications and special systems and expertise.

When and how will metadata deliverables be validated?

Considering the points posed below will help determine:

- The number and types of external metadata deliverables to specify
- When and how quality control will be performed on these deliverables

Validation is the process of vetting external metadata deliverables to make sure that they conform to the defined fields, vocabularies, and structure that make up the specification. One advantage to using standards in whole and having metadata delivered as independent files is that there is often an XML Schema Definition, or XSD, for a given standard. An XSD provides the rules against which an instance of a document claiming to be in conformance with a given standard can be validated. This may be useful not only to determine that the metadata received meets the specified standards, but also for potential future use cases in which a standard for data exchange, migration, import, or transformation to an organization or system is required or useful.

If such potential future use cases are less of a priority, then validation may be performed in another way. The client might choose to create a local XSD file to validate against as part of quality control or ingest, or the systems importing the data might perform validation as part of the import process.

Embedded

Embedded metadata can be most simply defined as metadata that is stored inside the same file, or container, that stores the audiovisual signal to which that metadata refers.

In many ways, one can think of embedded metadata as the file-based domain's equivalent of the physical domain's labels, annotations, and written documentation stored inside of material housing, or even as "in-program" annotations such as audio and video slates at the head of a recording.²⁰

Every file format has distinct embedded metadata specifications and fields. For instance, there are different options for embedding metadata in WAVE²¹ files than there are in MP3 files.²² Embedded metadata is what enables the display of information such as artist and title in applications that play back media files. The primary goal of embedding metadata for the purpose of preservation should be to identify:

- the object in an instance where it is dissociated from its external metadata
- the holding organization
- the data source that holds information about the object
- the copyright status

The Federal Agencies Digitization Guidelines Initiative (FADGI) published guidelines for the use of embedded metadata in WAVE files,²³ which may be adapted to other formats accordingly.

There are a few additional considerations to keep in mind when it comes to creating embedded metadata specifications.

- Digital files that are acquired by an organization, rather than created through digitization, likely will have existing embedded metadata that was generated by people, software, and/or hardware prior to acquisition. In the interest of maintaining the authenticity of the original object, these files should undergo a different process with regard to embedded metadata in order to maintain the authenticity of the original object.

20 Joshua, Ranger. "Embedded Metadata in WAVE files: a look inside issues and tools." 2014. <http://www.avpreserve.com/papers-and-presentations/embedded-metadata-in-wave-files-a-look-inside-issues-and-tools>

21 Federal Agencies Audio Visual Digitization Working Group. *Task 5.4: Assess Options for Embedding Metadata in WAVE Files and Plan the Audio Metadata File Header Tool Development Project: Assessment Report and Initial Recommendations*. http://www.digitizationguidelines.gov/audio-visual/documents/AVPS_Audio_Metadata_Overview_090612.pdf

22 "ID3 tags are the audio file data standard for MP3 files in active use by software and hardware developers around the world." <http://id3.org>

23 Federal Agencies Audio-Visual Working Group. *Embedding Metadata in Digital Audio Files: Guideline for Federal Agency Use of Broadcast WAVE Files*. http://www.digitizationguidelines.gov/audio-visual/documents/Embed_Guideline_20120423.pdf

- While some file formats/wrappers have robust options and toolsets for reading and writing embedded metadata, others are lacking in this regard. For instance, MXF and DPX files, common target formats for video and film, present significant challenges. In cases like these, side-car files with the embedded metadata equivalent may be created as an interim solution, or the choice may be made not to have an embedded metadata deliverable for these files.
- When creating an embedded metadata specification, consider how most applications manifest and present this information. The choice of where to store the metadata within the file may be influenced by how accessible it is to users across an array of applications and tools.
- Embedded metadata can be fragile and may be accidentally erased and/or augmented if not handled in a considered way.²⁴

QUALITY ASSURANCE AND CONTROL

Vendor Quality Assurance

Quality assurance is actually addressed at several points preceding this section in the SoW, but this is the place to bring about clarity on two fronts. The first is to request details about the vendor's quality assurance protocols in their proposal. The second is to be explicit about protocols and practices required to perform the statement of work.

This is particularly important because differences in quality assurance protocols and practices are often the biggest driver of variation in pricing offered by vendors. It is important both to clearly understand the details of their quality assurance plan to gain insights into differences that may manifest in pricing and to make sure that each vendor provides a proposal that meets the core quality assurance requirements. This enables more of an apples-to-apples comparison.

Client Quality Control

In the same way that it doesn't make sense to create a rule that can't be enforced, there is little sense in creating

a specification that is not checked for compliance. The client and vendor are partners in ensuring success, and it is the obligation of the client to perform quality control in order to live up to their side of the partnership. There will be quality issues in every project; this is simply a reality. Having a reasonable number of errors does not mean that the vendor did a bad job. On the flip side of that statement, if the vendor did fail to perform quality work or there was a misunderstanding and misapplication of a specification and subsequently quality control is not performed, then the client is complicit in the poor outcome. It does not matter if a client and vendor have been working together for many years using the same set of specifications; quality control should be performed diligently and routinely for all work performed. With this in mind, it is important for the client to have a documented quality control protocol accompanying each SoW to ensure that it is performed comprehensively and consistently. At a high level, this protocol should provide a check for each specification and requirement in the SoW.

Quality Control Check Categories

The checks to be performed can be categorized as quantitative or qualitative and automated or manual. The categories that a given check falls into will determine the best method and approach for implementing it.

- **Automated Quantitative:** These are checks that are performed and reported with reliability based on logic applied by a software application. For example, checking a client-specified embedded metadata field in a file may be performed using an application such as MDQC.²⁵ Checking that a particular copyright statement is present in an embedded copyright field or checking that the sample rate of all audio preservation masters is 96 kHz are quantitative checks that can be performed in an automated way with a definitively reliable result.
- **Automated Qualitative:** These are checks that use algorithms to perform the job in an automated fashion and that require subjective judgment calls by humans to increase the reliability of reporting. For example, automated tools that identify and report on audio and

24 Association of Recorded Sound Collections (ARSC) Technical Committee. *A Study of Embedded Metadata Support in Audio Recording Software: Summary of Findings and Conclusions*. <https://www.avpreserve.com/tools/a-study-of-embedded-metadata-support-in-audio-recording-software-summary-of-findings-and-conclusions>

25 Metadata Quality Control (MDQC), <https://www.avpreserve.com/tools/mdqc>

video errors fall into this category. These types of systems may be useful for creating more efficient workflows, but they will over- and under-report issues and require human review.

- **Manual Quantitative:** These are quantitative checks for which either (a) there is no logic that can be implemented in an automated way or (b) the resources required to develop an automated routine are not justified by the level of manual effort required to do the same task. For instance, checking that the title field contains the correct title based on the original item or reviewing the file naming and organizational conventions of deliverables from a vendor fall into this category. Depending on the circumstances, these checks may or may not be able to be automated, or they may be able to be automated but there are so many conditions in the logic that it would be overly complex and resource-intensive to build the appropriate tool.
- **Manual Qualitative:** These are checks that require the subjective judgment of a human. These can be used in combination with automated qualitative checks or used alone. An example of the former is the human review of a report generated by an automated qualitative tool that identifies issues in the audio and video. An example of the latter is ensuring that the qualitative differences between the preservation master, mezzanine copy, and access copy for a given original item are within acceptable limits of variable quality.

Quality Control Resource Planning

Depending on the specifications used, the available data, and the tool set at hand, a quality control protocol consisting of a combination of these checks should be documented. Once the appropriate categories and associated methods are identified for each check, it is necessary to analyze the allocation of human and machine resources needed to make sure that the time it takes to implement the quality control protocol aligns with available resources.

Some checks will be computer- and labor-intensive, while others will require very few resources and can be performed quickly. For those that are quick and require few resources, it makes sense to perform them 100% of the time. On the other hand, resource-intensive checks will

require scaling down the percentage of files checked based on the available resources.

Consider the scenario where a staff member has 20 hours per week dedicated to quality control for deliverables coming in at a pace of 250 digitized items per week. Let us say that performing the audio and video quality checks takes five minutes per original item. At 250 items per week, this activity alone will require just over 20 hours. With all of the other quality control checks that must be performed, as well as the handling of media and other administrative tasks, it is not feasible to spend all 20 hours solely on the audio and video quality check. A solution to this is to perform the check on a sampling of materials. Say that after adding up the time that all of the other quality control checks and administrative tasks will take, there are only four hours that can be allocated toward checking of audio and video quality. At five minutes per original item, this means that there is enough time to perform this check on 48 items, or 19%.

Sampling is a widely adopted and perfectly acceptable approach, but it is important to keep a couple of things in mind. The first is that if the error rate of the sample set is high and/or there are consistent errors found, the sample size should be increased to track down the severity and extent of the issue. The second is that the higher the sample size, the greater the confidence and the lower the margin of error. While scaling and sampling can be used within reason, it does have its limits, and it is important for a successful outcome to provide sufficient staff for the performance of quality control. It is also important that the staff performing quality control has the appropriate expertise necessary to make subjective judgment calls.

The above text mentions human resources, and it is equally necessary to analyze and plan for the compute resources that are required to implement the quality control protocol. Like the human resource calculations in the previous example, the approach here is to identify the length of time it will take for computers to perform their job. This will make sure that there are no bottlenecks that may become barriers to the performance of tasks by staff or slow down the required throughput to meet the demand. For computer-intensive processes, the options may be to utilize sampling for certain processes or to increase the compute resources to scale accordingly.

Quality control documentation and rework

Documenting and resolving quality control issues can be confusing and challenging. It is advisable to have a shared system (e.g. Google Sheets or a custom software application) that allows both the client and the vendor to comment on quality control issues in real time, where there is no likelihood of version control issues, and where all current data is immediately available to all parties at the item level. Having an explicit quality control status is useful so all parties can see if an issue is pending resolution, requires rework, or has been resolved. Be sure that any rework goes through the quality control process again; do not make the assumption that rework must be correct. Also, consider updating data tape, such as LTO tapes, or hard drives that resulted in quality control issues and the need to perform rework.

Reference Files

There is a preceding section of the SoW that specifies reference files that the vendor will record through their system and that will give the client insight into the performance of the vendor systems. In order to check these reference files, it is necessary to procure software that will analyze these files. This likely includes software with video waveform and vectorscope displays and audio level meter and frequency analysis displays. Using this software to determine the extent to which the reference files meet standard targets will provide the client with the information they need to either identify potential issues or confirm the health of the vendor's systems.

Financial Planning for Quality Control

Implicit in the above discussion of human and compute resources is the need to plan and budget for resources within the client organization. Budgeting the appropriate amount of time for properly-skilled staff, as well as for specialized equipment and software, to perform quality control tasks is critical to ensuring the success of a digitization project. These should be planned for and included in the project budget from the outset.

DELIVERY

This section of the SoW defines several aspects of the delivery of media from the vendor to the client. Note that the file naming and organizational conventions

are integrally tied to this section. Specifications of the relationship between the vendor deliverable and the AIP and whether to bag or not should be incorporated here.

Media

The first item to specify here is the media and formatting that the vendor should use when sending the deliverables to the client. This is usually hard drives and/or data tape. Be sure to specify lower level requirements regarding file systems (e.g. NTFS, HFS+, LTF5) to ensure that data is able to be accessed and worked with upon receipt. Otherwise, specifications for delivery on hard drive is straightforward.

Delivery on data tape requires consideration of the following factors:

- *Hardware and software*

Assuming that the client is venturing into data tape for the first time or they are going to upgrade data tape formats, the client should first look within their own organization to identify the existing hardware and software infrastructure. If it uses current technology that makes good sense from a preservation perspective, determine if it is possible to share these resources and take advantage of the organization's investment. Sometimes there may be an overlap in technology but a big disconnect in policy, or there may be no additional capacity or bandwidth. For instance, it is often the case that an organization's IT department uses data tape for backup purposes. However, these solutions often utilize proprietary software (see next bullet point), have limited retention policies that don't align with preservation, and generally manage the data using very different workflows and practices when compared to a preservation environment. This is not always the case though, and sometimes there can be resource sharing that utilizes separate policies, practices, and management.

If there is no opportunity to take advantage of existing infrastructure within the client organization, or it does not make sense, then it is necessary to budget for the appropriate hardware and software to be able to access the data on the data tapes. The client must be sure to source this equipment and budget for it from the outset.

- *Proprietary versus open*

Writing and reading data to data tape requires an intermediate application. Until recently, these intermediate applications were all proprietary applications that utilized their own methodology for writing and reading data. This reliance on a proprietary third-party application to access data represents an increased risk to the preservation of that data. Within the past few years, LTFS (Linear Tape File System) was introduced for the LTO (Linear Tape Open) data tape format. LTO and LTFS are both open standards resulting from a cooperative effort across manufacturers to agree on and publish a set of standards for both a data tape format and the methodology for writing and reading data to the data tape format. This mitigates the increased risk of using a proprietary third-party methodology for accessing data on data tape. This may or may not be reason enough to select LTO and LTFS as the specification, but this general point should be thoroughly thought through as part of the specification process.

- *Implications to quality control protocol*

As opposed to hard drives which are instantly accessible, data tape, despite the claims of manufacturers, simply takes longer to work with. It often requires copying data from tape to a local drive in order to perform the quality control checks. This can add significant machine and labor time. In some cases, the bottlenecks that this creates may dictate the details of the quality control protocol and require more sampling than may be considered ideal. An alternative approach to scaling down the amount of quality control is to request that some items be delivered on both hard drive and LTO to enable faster quality control and ingest of materials.

- *Library system versus storage on shelves*

Data tape devices range from single tape drives to tape library systems that can hold many tapes. The former is fully manual and results in storing data tapes on shelves. This has some advantages, but it falls short of being able to implement a digital storage solution that meets the requirements of a digital preservation environment. The latter better meets the demands of a digital preservation environment because the automated nature of a tape library enables working across a large set of tapes, and this approximates the functionality found when working with disk-based servers and network

attached storage devices. The downside is that the cost and complexity is much greater compared to a single drive approach.

- *Migration*

Data tape formats have a limited lifespan. Some have roadmaps that define how frequently updated versions will come out and how much backward compatibility will be supported. Regardless of delivery media and storage choice, however, migration is a reality, and the greater amount of effort required with data tape requires additional planning. Thought should be given to when the migration should occur, the level of effort involved, and the costs to procure new hardware and staff or to outsource the effort.

Shipment guidelines

In addition to specifying the delivery media, this section should specify how the media is packed, shipped, and identified at the box level and media level. Specifications may include conventions for identifiers and where and how they should be applied. Furthermore, there should be specifications for the metadata that accompanies the media delivery, identifying the contents of the delivery media at the item level. This may be included in a shipment manifest, in the project management documentation, or both.

This section of the SoW should also specify when and how the original items should be shipped back to the client. Typically, originals are not shipped back until all quality control and rework is complete for all of those items. They are also sent separately from the delivery media to mitigate the risk of loss during the delivery process.

This is also a good place to reiterate who is responsible for performing and paying for the shipment and any shipment protocols or requirements.

Protocol for deletion

The client must be explicit about expectations for when and how the vendor will delete the client's data from their systems. There is a balance to be aware of here between the burden placed on the vendor who must store large quantities of data on their systems and the need of the client to thoroughly work through the quality control process and make sure that the data is safely on client systems before the vendor deletes their data. The turnaround time for client quality control and vendor rework is defined in

the timeline section of the SoW, but this is a good place to reiterate that timeline and to add explicit instructions that the vendor should not delete any client materials until given written approval to do so.

Importance of an Internal Ingest Protocol

Once quality control is complete, rework performed, and final deliverables received for a given set of items, there will be an ingest process. The details of the ingest process will depend on the given organization and may range from simply placing LTO tapes onto shelves and documenting the location to copying the data off and running it through extensive processes in order to prepare and deposit the data into digital preservation repositories and populate access systems. Whatever the case, think through the details of the ingest process. Consider the specifications of the deliverables and the delivery media and think through the handling of the media, the processing of any data, and the updating and population of systems. Document these details and analyze and calculate the human and compute resources required and the associated timeline. There may need to be meetings with others in the client organization or third party storage/system providers to figure out the details of ingest and to resolve any prospective issues. The quality control and ingest protocols go hand in hand and are a critical part of internal documentation.

Staffing

Aside from quality control and ingest of delivered items, there will be effort associated with receiving, reconciling, and reshelving the original items once they are received. Be sure to plan for this and staff the effort adequately.

RFP RESPONSE

In this section, the client communicates to prospective bidders how they should respond to the RFP. The SoW has many specifications and requests for specific pieces of information from the vendor. Providing a checklist for the vendor along with a general comment about how they should respond and what should be included in their proposal is useful for all parties.

PRICING INFORMATION

If the client does not provide specific direction on how bidders should provide pricing information, it becomes nearly

impossible to perform comparative analysis. It is also the case that there are unknowns on the client side regarding the quantity of items that will be digitized (this is usually dependent on the pricing that comes back from the vendors) and the program duration of the content on the media being digitized; therefore, it is valuable for the client to provide a structure for pricing information that is consistent across vendors and that builds-in variables that provide a better understanding of pricing under different scenarios. This may be accomplished by building a range of durations and quantities into the client-specified reporting structure for pricing. Any specified reporting structure should include costs for media, supplies, shipping, and any supplementary services that are needed.

QUESTIONS

The SoW is a detailed document with many parts that have had a great deal of consideration put into them. It is in the best interest of all parties to allow the vendor the opportunity to ask questions about the SoW. This will help them better understand the SoW and put together a representative proposal, or it may expose flaws in the SoW that need to be corrected. The number and quality of the questions asked also provides an indication of how thoroughly the vendor has read the SoW.

This section should reiterate the date by which any questions from vendors are due, who they should be delivered to, how they should be delivered, and the timeline and method for responses.

Many organizations choose to share all questions (anonymized) and all answers with all participating bidders. Others choose to simply respond privately to the inquiring bidder. Regardless of approach, it should be documented here so the vendors will know who will see their questions and the associated responses.

Examples of Statements of Work

The following guides offer examples of much of what is discussed above, and they can serve as references for developing an SoW. Note that the title of these documents uses the phrase “Request for Proposal.” In the context of this chapter we would refer to these as Statements of Work.

- Guide to Developing a Request for Proposal for the Digitization of Audio. <https://www.avpreserve.com/papers-and-presentations/guide-to-developing-a-request-for-proposal-for-the-digitization-of-audio>
- Guide to Developing a Request for Proposal for the Digitization of Video (and More). <https://www.avpreserve.com/papers-and-presentations/guide-to-developing-a-request-for-proposal-for-the-digitization-of-video-and-more>
- Digitizing Motion Picture Film: Exploration of the Issues and Sample SOW. http://digitizationguidelines.gov/guidelines/FilmScan_PWS-SOW_20160418.pdf

SECTION 5: POST-DIGITIZATION

The text above discusses the importance of creating an ingest protocol and coordinating with other stakeholders in the client organization (and perhaps partner organizations) who are depended upon or directly impacted. Also discussed as part of the ingest protocol was walking through the details of populating client systems with media and metadata. All of these are critical elements of properly planning for post-digitization. One aspect not yet covered is how to calculate storage requirements for each of the target systems, in total and over time. This will be an important piece of data for IT infrastructure planning, for budgeting for the delivery media in the project, and for logistical planning for quality control, ingest, and longer-term storage.

The basis for calculating storage requirements will be an inventory of items selected for digitization. Populate a spreadsheet with the following columns:

- Format
 - Quantity
 - Estimated average duration (use media duration if program duration is unknown)
 - Estimated total duration (quantity x estimated average duration)
 - Preservation master file size (GB/TB per hour/min x estimated total duration)
 - Mezzanine copy file size (GB/TB per hour/min x estimated total duration)
 - Access copy file size (GB/TB per hour/min x estimated total duration)
- Total file size (PM file size + Mezz copy file size + Access copy file size)

Summing across all formats will provide the total required storage capacity. If the capacity of the delivery media is known, then the quantity of media to purchase can be calculated, and the client can plan for what will be received.

Knowing total storage capacity is helpful, but a lower level piece of information that is more useful when planning for IT storage infrastructure is how that data will be produced over time. This is much more true for large projects than small projects where everything will come in at once.

To calculate the storage capacity growth over time, the client organization will need to have either specified or have a sense of the frequency and quantity of batches being delivered from the vendor. Note that this same information is also useful for planning for quality control and ingest staffing. Once the frequency and quantities (and possibly formats) of deliveries are known, the spreadsheet created to calculate total storage capacity can be used to calculate storage capacity growth over time. Furthermore, if it is known which target formats will populate which target systems, this calculation can be performed at the target system level.

CONCLUSION

Managing a successful reformatting project requires effective communication, a basic knowledge of AV digitization techniques, and a deep understanding of how to work with a vendor. Creating a Request for Proposal and a Statement of Work informed by a detailed inventory and guided by the walkthrough detailed in this chapter will help to ensure that project goals are understood by all and that files are properly delivered to the institution. Additionally, a process for quality control, rework, and ingest should be in place to make sure that returned files are vetted and effectively stored in the repository.

It will be the responsibility of the institution to establish a sustainable preservation program to ensure that the delivered files are well cared for long into the future. This requires planning, policies, a storage infrastructure, active management, and metadata. These topics as well as basic concepts in digital preservation are described in “Chapter 4: Managing Digital Audiovisual Collections.”

CHAPTER 4

MANAGING DIGITAL AUDIOVISUAL COLLECTIONS

by Amy Rudersdorf, Senior Consultant, AVPreserve

THE FUNDAMENTAL REQUIREMENTS OF DIGITAL preservation are threefold: (1) maintain the bits, or building blocks, of the digital files; (2) maintain the content of the file (the movie in the video file, the song in the audio file) so that it is accessible and understandable; and (3) preserve both the bits and content for as long as necessary. This last requirement demands more than a technological solution. Even more than physical audiovisual collections, digital collections require holistic management to ensure their long-term preservation and access—and organizational factors are critical. Incorporating the organization’s goals and objectives is essential to ensuring support for a sustainable digital preservation program. Without high-level organizational buy-in, ongoing resources (staffing, funding) together with technology (hardware, software, storage)—essential aspects of a sustainable digital preservation program—are not guaranteed. Digital audiovisual collections management involves planning in order to establish policies and standards-based practice, so that staff understand their roles and so that technology can be best utilized for an institution’s digital collections.

The world in which digital collections live is one of constant change. For audiovisual files, carriers (storage media) and wrappers and codecs (together, the file format) must each be addressed separately when thinking about best practices for long-term management. Change can occur at any level. For example, a digital storage system may need to be replaced even if the file format within it is stable.

Likewise, a file format may need to be migrated to a newer version, while storage remains unchanged. In the digital world, where every bit and byte matters, keeping in mind the relationships between file, carrier, and storage will help you best manage your collections.

This chapter will include an overview of the approaches, standards, and considerations for an institution beginning to manage their digital collections. The beginning of the chapter focuses on the organizational infrastructure that supports sustainability of digital collections. Organizations must consider factors that make long-term preservation possible, such as risk management, preservation management, standards and guidelines, policies and planning, organizational infrastructure, and planning and phasing. The second half of the chapter describes the activities and technology required for digital preservation. These include storage (redundancy, media, and geographic diversity), monitoring (fixity), information security, metadata, refreshing and migration.

When establishing a digital preservation program, keep in mind that not all of the activities in this chapter need to be put in place at once. For some institutions, implementation of a complete set of digital preservation policies, for example, may not occur right away. Building momentum within the institution might be the first step, which will ultimately lead to the adoption of policies, committed financial support, and deployment of technologies. The recommendations herein should be considered the

fundamental components of a sustainable preservation program. With continued evolution and improvement, the adoption of these plans, policies, and services will support the goal of long-term preservation and access to your digital audiovisual collections.

SECTION 1: DIGITAL PRESERVATION CONCEPTS

Definitions

“**Digital preservation** is the active management of digital content over time to ensure ongoing access.”²⁶ It is an integral part of a larger process of **curation**, which consists of the activities across the content lifecycle described throughout this textbook: selection and appraisal, description, ongoing care and management, long-term access, and/or deaccessioning/disposal.

Without **active management**, which includes many of the activities outlined in this chapter, digital assets and their associated content are at risk, potentially through media failure, human error, inaccessibility due to format obsolescence, or barriers to discoverability due to poor metadata. The functions of digital preservation reduce these risks and together help to ensure that content remains accessible over time.

Risk Management

At the heart of a preservation strategy is **risk management**. The opportunities for loss or damage to digital collections is inherent in providing access to users, who may include digital collections managers, IT staff, students, the public, and others. A preservation strategy can be an effective way to mitigate these risks to the greatest extent possible.

The nature of risks is varied and may be human-generated, mechanical, or natural. The human risks to technology may be purposeful (file formats are not selected for **migration**, metadata is not captured), nefarious (viruses, cyber-attacks), or accidental (deletions, misfiling or misnaming files). Organizational risks include insufficient

planning and policies, which lead to a loss of or lack of sustainable funding to support trained staff and/or appropriate technologies. Risks may also be mechanical, such as when files change at the bit level without human awareness or media and storage fail. Risks may also come from nature; floods and fires can destroy electronic media on which files are stored.

Over time, risks evolve based on the organization and its resources as well as industry-wide technical changes. As risks change, how institutions identify, respond, and monitor them must change, too. Successful preservation strategies must be flexible, yet cautious, to be able to react to risk effectively. Through planning and management, risks can be mitigated as they emerge.

Preservation Management

Due to the varied nature of risks to digital assets, it is important that digital preservation be approached from a programmatic standpoint with administrative support that makes preservation a priority. One approach to thinking about preservation management is illustrated by the concept of the “three-legged stool,” in which the organizational infrastructure, technological infrastructure, and resources all have equal footing to create a stable digital preservation program. Nancy McGovern’s *Digital Preservation Management: Implementing Short-term Strategies for Long-term Problems* describes these three interlocking structures as follows:

1. *Organizational Infrastructure* includes the policies, procedures, practices, and people—the elements that any programmatic area needs to thrive but that are specialized to address digital preservation requirements.
2. *Technological Infrastructure* consists of the requisite equipment, software, hardware, secure environment, and skills to establish and maintain the digital preservation program. It anticipates and responds wisely to changing technology.
3. A sustainable *Resources Framework* addresses the requisite startup, ongoing, and contingency funding to enable and sustain the digital preservation program.²⁷

26 Library of Congress, “What is Digital Preservation?” <http://www.digitalpreservation.gov/about>

27 Anne R. Kenney, et al. “Digital Preservation Management: Implementing Short-term Strategies for Long-term Problems.” <https://dpworkshop.org/dpm-eng/conclusion.html>

Determining the “what” (organizational infrastructure), the “how” (technological infrastructure), and the “how much” (resources framework) helps to identify an institution’s needs for a sustainable digital preservation program.

Standards

Many industries employ standards to make certain they comply with accepted practice, ensure the safety of their customers and employees, and provide a foundation upon which new technologies can be built. Two international standards documents serve as the cornerstone for the management of digital collections, guiding institutions in the development of sustainable preservation programs and serving as benchmarks for institutions that are maintaining preservation management technologies. These are:

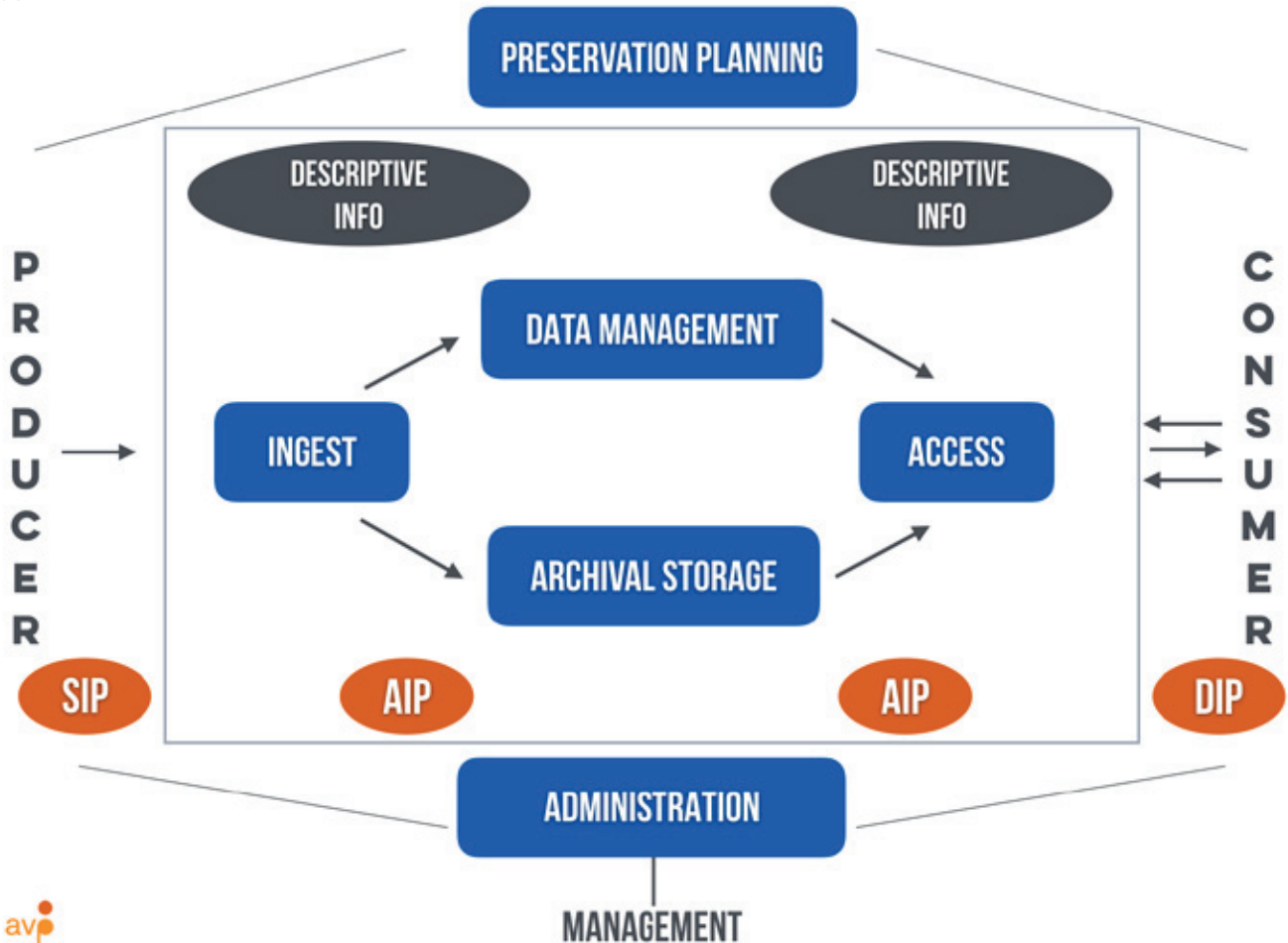
- *ISO 14721:2012, Space data and information transfer systems—Open archival information system (OAIS)—Reference model*, commonly referred to as the OAIS Reference Model, or OAIS, and;
- *ISO 16363:2012, Space data and information transfer systems—Audit and certification of trustworthy digital repositories.*

OAIS REFERENCE MODEL

The Open Archival Information System (OAIS) reference model is a conceptual framework for an archival system dedicated to preserving and maintaining access to digital assets over the long-term. Because it is a conceptual model, it is not inherently prescriptive. Instead of specifying technology, staffing, or resources, it provides guidance

FIGURE 4.1
Open Archival Information System (OAIS) Reference Model

Credit: AVP



about best practice for building a sustainable preservation environment. OAIS states that archival storage is necessary but does not specify how that storage will be implemented, the technology required, or how many staff are needed to maintain it. Likewise, access is a component of the OAIS model, although how access is provided will be unique to each organization. Each component represents a process within a system but not the specific resources and technology required to support that process.

The framework takes into account producers (creators) of content (and their embodiment as file-based assets and data that will be preserved); the system (technology, workflows) in which the content is preserved; the administration, management, and preservation planning structure that administers the program; and the consumers (users) that will use the content at some point in the future. Content is packaged in different formats throughout its lifecycle (reflected in the orange ovals in the graphic on page 61):

- Submission Information Packages (SIPs), created in preparation for **ingest** (submission into the archival system)
- Archival Information Packages (AIPs), the content that is managed and stored over time, which may include the SIP contents and additional data created by the system
- Dissemination Information Packages (DIPs), the content made sharable for users, typically a subset of the AIP

The value of OAIS is that it provides a model for functions that should occur in a preservation environment and the types of content that must be managed over time. It is an example of a holistic approach to digital preservation, taking into account not only the technology but also the people, resources, and organization as well. Compare this conceptualization to the three-legged stool analogy above.

AUDIT AND CERTIFICATION OF TRUSTWORTHY DIGITAL REPOSITORIES

ISO 16363 is the international standard that describes the characteristics of a trustworthy digital repository (TDR). It includes categories of metrics that identify the individual components that together comprise a TDR. The OAIS framework is the basis for the TDR. The expectation

is that any TDR will embody the OAIS model, including a robust submission and ingest process, an archival storage and data management system, and an access component—all overseen by a fully developed organizational and management structure that ensures programmatic longevity and stability. See the figure above to see how these facets interact. Digital objects in ISO 16363 are referred to as “information packages,” as in the OAIS standard. Other vocabulary from OAIS appear in the standard as well.

While the standard was developed to provide a framework for certifying a digital preservation program as “trustworthy,” the reality is that the great majority of institutions will probably not attain certification. Instead, many organizations use the standard’s 109 metrics to guide development and growth of their program and archive and to focus energy and resources on areas for improvement.

STANDARDS-BASED DECISION MAKING

Standards offer comprehensive guidance on the making of highly functional and robust digital preservation environments. However, implementing these standards can be a major undertaking, requiring significant resources and cooperation from numerous stakeholders. If an organization lacks the resources to build a local system that meets these standards, there are a number of alternatives available to it, but these standards should still be consulted to provide guidance in choosing the right option.

Look for partnership organizations. Many, often larger, organizations have built their digital preservation infrastructure with standards in mind. Partnering with them to deposit into their archive might be a more sustainable option than building your own. Review a partner organization’s documentation and talk to its staff to understand their approach to digital preservation and whether it complies with OAIS and TDR.

Consider consortial or partnership organizations. Organizations like DPN, MetaArchive, and others have built communities around their digital preservation services. They work together to ensure that these services are built within the ISO framework.

Purchase preservation storage services. DuraCloud and Preservica are just two of several digital preservation products built with guidance from the ISO standards.

No matter which direction an organization takes—building its own preservation system or looking outward for services—the ISO standards should be kept in mind. Whatever the approach, standards-based decision making will help to build a robust and sustainable digital preservation program.

Resources

Definitions

- Association for Library Collections and Technical Services. “Definition of Digital Preservation.” <http://www.ala.org/alcts/resources/preserv/defdigpres0408>
- Digital Curation Centre. “What is Digital Curation?” <https://dcc.ac.uk/about/digital-curation>
- *Digital Preservation Handbook*. “Digital Preservation.” <https://www.dpconline.org/handbook/glossary#D>
- Library of Congress. “About Digital Preservation.” <http://www.digitalpreservation.gov/about>

Standards

- ISO 14721:2012, Space data and information transfer systems—Open archival information system (OAIS)—Reference model. <http://public.ccsds.org/publications/archive/650x0m2.pdf>
- ISO 16363:2012, Space data and information transfer systems—Audit and certification of trustworthy digital repositories. <https://public.ccsds.org/pubs/652x0m1.pdf>

Guidelines

- Data Seal of Approval. <https://www.coretrustseal.org/about/history/data-seal-of-approval-synopsis-2008-2018>
- “Levels of Digital Preservation.” National Digital Stewardship Alliance (NDSA). <http://nds.org/activities/levels-of-digital-preservation>

SECTION 2: ORGANIZATIONAL INFRASTRUCTURE

A key facet of a sustainable preservation program is the organizational infrastructure that supports it. This includes all the elements that any programmatic undertaking needs to thrive, including planning, policies, funding, procedures, stakeholders, and decision makers, all of which must be tailored to address digital preservation requirements.²⁸ Organizational infrastructure is one of the three main components of the international standard described above, *ISO 16363: Audit and Certification of Trustworthy Digital Repositories*, which defines programmatic **governance** structures (i.e. dedicated leadership and oversight functions), preservation policies, and sustained support for staffing and funding as fundamental.

To establish organizational support for a digital preservation program, planning and documentation are essential. Documenting the current state of the collections and the future goals for preservation will help cultivate organizational buy-in for a digital preservation program.

Preservation Planning

The first step in developing a sustainable preservation program, or in other words, achieving the balance of the three-legged stool, is planning. Preservation planning can be defined as “a process by which the general and specific needs for the care of collections are determined, priorities are established, and resources for implementation are identified.”²⁹ Digital preservation plans differ amongst institutions, but generally they support the following objectives:

- Organizational commitment to the preservation of and continued access to digital collections through ongoing financial and resource support
- Authenticity of digital collections (ensuring files are trustworthy representations of their original content)
- Environmental controls on the physical media on which digital collections are held

28 Anne R. Kenney, et al. “Digital Preservation Management: Implementing Short-term Strategies for Long-term Problems.” <https://dpworkshop.org/dpm-eng/conclusion.html>

29 NEDCC, “Preservation Leaflet 1.1: What is Preservation Planning?” <https://www.nedcc.org/free-resources/preservation-leaflets/1.-planning-and-prioritizing/1.1-what-is-preservation-planning>

- Monitoring collections and addressing errors or changes as necessary
- Migrating digital collections as software and/or hardware becomes obsolete³⁰

Digital preservation plans can be brief (1 to 2 pages), concisely communicating these objectives through one- or two-sentence statements. Many institutions make their digital preservation plans available online, and each is unique to the organization, its structure, and its goals. These online publications are invaluable resources for institutions establishing new preservation programs. Understanding how others address their commitment, needs, and resource allocation can inform decision making at your own institution. The good news is that you do not have to write a digital preservation plan from scratch; when you find one that you feel meets your needs, consider adapting it to your local conditions.

Preservation Policies

A digital preservation policy is the framework around which a digital preservation strategy is developed at an institution. It is integral to documenting the institution's commitment to digital preservation services, identifying at a high level what digital content is in scope of the digital preservation policy (in this case, digital audiovisual collections) and providing a vision for moving these strategies into action. As noted above, many institutions make their digital preservation policies available online, and templates are available as well.

A digital preservation policy document may include the following types of information:

- Mission and vision statements
- Definitions of terms
- References to other policies (collections, preservation of analog collections, etc.)
- Preservation strategy
- Process workflows
- Definitions of preservation levels (bit-level preservation versus full preservation of the files as they exist today)
- Definitions of acceptable preservation formats and metadata standards
- Overview of storage and technical infrastructure
- Definitions of communities (producers of content, users of content, etc.)

Policies and plans should be approved by high-level administration to ensure that administrators understand the goals of a digital preservation program and to solidify their ongoing support to sustain the program. Policies also facilitate buy-in from colleagues and staff as they integrate the digital preservation program as a key function of the organization.

Resources

Plans and Policies: Samples and Templates

- Bishoff, Liz. "Digital Preservation Plan: Ensuring long term access and authenticity of digital collections." *Information Standards Quarterly* 22:2, 2010. https://groups.niso.org/apps/group_public/download.php/4250/FE_Bishoff_Digital_Preservation_Plan_isqv22no2.pdf
- Dartmouth College Library Digital Preservation Policy, 2015. https://www.dartmouth.edu/~library/preservation/docs/dartmouth_digital_preservation_policy.pdf
- de Jong, Annemieke. *Digital Preservation Sound and Vision: Policy, Standards and Procedure*, 2016. <http://publications.beeldengeluid.nl/pub/388>
- MetaArchive Cooperative. "Digital Preservation Policy Template." 2010. https://metaarchive.org/wp-content/uploads/2017/03/ma_dp_policy_template.pdf
- NEDCC. "NEDCC Digital Preservation Policy Template." <https://www.nedcc.org/assets/media/documents/SoDAExerciseToolkit.pdf>

30 Liz Bishoff, "Digital Preservation Plan: Ensuring long term access and authenticity of digital collections." https://groups.niso.org/apps/group_public/download.php/4250/FE_Bishoff_Digital_Preservation_Plan_isqv22no2.pdf

SECTION 3: STORAGE INFRASTRUCTURE

In the digital environment, how digital content is stored is paramount. As noted above, standards such as the *ISO 14721: Open Archival Information System (OAIS)* and *ISO 16363: Audit and Certification of Trustworthy Digital Repositories (TDR)* provide a framework for how repositories should be structured and managed and what actions should be taken on the digital content within them. For example, best practice suggests that digital content be stored on “active” servers that are backed up and managed with preservation in mind. Storage on fixed devices, such as DVDs or external hard drives that are not monitored or backed up, do not meet the requirements of a TDR. And, they have high failure rates—the bits rot and the media inevitably degrades, which can mean catastrophic loss to collections stored on them.³¹

The often-heard “storage is cheap” aphorism is true when it comes to per-bit storage costs at scale. However, the reality for many institutions, especially small and mid-sized organizations, is different. The cost of digital collections storage can be significant, especially for audiovisual collections that are terabytes or petabytes in size. For this reason, cost often plays a significant role when an institution selects the type of storage and determines how it is managed. The good news is that there are less-expensive options that still allow for an acceptable level of maintenance of digital collections. Understanding what is available, the associated costs, and the risks involved in the various options is key to choosing the best solution for an institution.

At a high level, storage options can be broken down into two categories: local or on-premise storage, and cloud or outsourced storage. Based on an institution’s requirements, technical infrastructure, and resources, one or both options may be feasible. Decisions about what type of storage works best for an institution’s needs should be influenced by such factors as:

- **The level of reliability or “uptime” required.** Do you need immediate access to your digital content or can there be delays of minutes or hours in retrieving it?
- **The number and types of users that need access to it.** Who will take responsibility for managing the digital content—digital collections managers only, the entire archives staff, or someone else? Will a version of the content also be publicly accessible?
- **Types and amount of digital content.** How much storage do you need? At what rate will it grow?
- **Redundancy.** Is an institution capable of safely managing two or more copies of its digital content locally, or must it rely on cloud storage?

Considering these issues alongside best practices such as those in the NDSA Levels of Digital Preservation (see “Section 1.5: Resources”), levels of effort required, and the resources in place to support them will help an institution identify the best storage options for its situation.

In the rest of this section, **storage media**, **storage architectures**, and **storage capacity** are detailed in an effort to provide practical guidance on best practices.

Storage Architectures

Online, nearline, and offline are terms used to describe different types of storage architectures. These terms speak to the ease and immediacy with which data can be accessed as well as the varying costs and scalability of storage.

Online: In this context, online means that the data is immediately available to users on a storage system. Servers that host an institution’s networked drives are examples of online storage systems. This is the fastest, but also the most costly, of the three architectures. It is also the most common. Examples of online storage include flash and spinning disk, both described below in “Section 3.2: Storage Media.”

Nearline: In this case, digital content is available to users with some lag time, which can be a few seconds to a minute or longer. It is automated and networked in the same way that online storage is, but the media is different, typically a magnetic tape library. (Magnetic tape is described below in “Section 3.2: Storage Media.”) This tends to be

31 “Seagate slapped with a class action lawsuit over hard drive failure rates,” PCWorld.com, February 2, 2016. <http://www.pcworld.com/article/3028981/storage/seagate-slapped-with-a-class-action-lawsuit-over-hard-drive-failure-rates.html>

an option used by larger institutions that have the resources to diversify their storage architectures.

Offline: Here, digital content is stored on a piece of media that requires a human to connect it to a computer in order to access the data on it. The most common offline media type used for digital preservation is magnetic tape. Offline storage is often used to backup digital content for long periods of time. This kind of storage architecture is cost effective, but it takes time to access because it is not connected to a network. For digital preservation, offline storage is often used for the third-copy backup, or disaster recovery copy, of digital content.

Why use one architecture over another for audiovisual content? There are a number of reasons, including the following:

- **Cost.** Offline storage tends to be the cheapest, but it has drawbacks in that it is not actively managed by automated digital preservation processes like fixity monitoring.
- **Immediacy.** Online storage has the quickest response time – typically content is immediately available when needed. However, it may not be necessary for all content to be accessed immediately. Second or third copies of digital files are often stored on nearline servers or offline Linear Tape Open (LTO) tape because they will not be accessed regularly and because **latency** is higher.
- **Bandwidth constraints.** Larger files, such as high-resolution video, take time to transfer over networks. If quick access to high resolution files is required, the cloud might not be an ideal solution because it requires transfer over the internet, meaning that the available bandwidth in and out of your facility provides additional constraints.
- **Scale.** The scale of audiovisual collections can be massive. It may not be cost effective to store all digital content on online storage. As long as two copies are actively managed using online or nearline architectures, other copies can be stored offline (e.g. on LTO tape), which tends to be the most cost-effective method of storage.

Many institutions consider a hybrid on-premise and cloud solution for their storage architecture. Cloud storage vendors have options for both online and nearline storage that, together with local storage, may provide an institution with a redundant and secure approach to managing its digital content.

Storage Media

There are a variety of storage media options available to digital collections managers. Some are widely accepted as preservation-appropriate, while others are recognized as problematic due to their susceptibility to failure and obsolescence.

Removable media, such as portable hard drives, portable flash drives, or CDs and DVDs, are not considered viable as part of an overall preservation strategy. These media are highly susceptible to failure from degradation of the components that comprise them. The reality is that files are often stored on or burned to them and then the media itself is filed away or stored in boxes and not actively monitored. These media are also resource intensive to monitor for errors; human intervention is required to plug in a flash drive or play DVDs to identify errors in the files stored on them. For audio or video collections, this can mean hundreds or more of DVDs to monitor. If errors are identified on a DVD, it is typically the entire media object that has

FIGURE 4.2
Examples of removable storage media in use from 1970 to 2010, including 8" floppy disk, 5.25" floppy disk, 3.5" floppy disk, cassette tape, 8mm tape, CD, DVD, ZX Microdrive, SDHC card, CompactFlash card, and USB disk.

Credit: https://commons.wikimedia.org/wiki/File:Forty_years_of_Removable_Storage.jpg



failed, which leads to complete data loss or cost-intensive data recovery procedures that are out of reach for some institutions.

Two predominant storage media that are considered good options for digital preservation management are referred to colloquially as “spinning disks” and “magnetic tape.” This is by no means an exhaustive list, but it provides some idea of the options available for the purposes of managing digital content.

Spinning disk storage that is part of a networked storage environment is commonly used in digital preservation environments. Spinning disk storage has quick response times and allows for active monitoring, such as fixity checking, to take place. This type of storage is often highest in cost because the media is expensive, the servers are always on and must be maintained in an environmentally controlled and secure area, and staff must be available to keep the servers up and running.

Magnetic data tape, most commonly LTO tape, is typically used either for nearline or offline storage (described in “Section 3.1: Storage Architectures”). Magnetic tape media is less expensive than spinning disk or other low latency storage options, and the cost of managing it over time is greatly reduced, especially for offline storage. Like other removable media, its mediated nature slows access

FIGURE 4.3

The inside of a spinning disk hard drive. Note how the head extends over the disk similar to the way a stylus rests on top of a vinyl record.

Credit: AVP



FIGURE 4.4

An LTO data tape cartridge.

Credit: <https://commons.wikimedia.org/wiki/File:LTO2-cart-purple.jpg>



and preservation activities such as active fixity monitoring. However, magnetic tape is much more reliable and far less prone to failure than portable drives and optical disc media such as CDs and DVDs. Magnetic data tape can be stored in what is known as a tape robot, which can provide some automation for access and preservation activities.

Media that enables digital collections managers to actively monitor the health of their collections is always the best choice when deciding on storage options. Luckily, this type of storage also tends to be the most prevalent. No matter what choice you select for storage, though, always backup your data at least once and ideally twice (three copies total), or more.

Options for Storing Digital Collections

A major factor in deciding what type of storage to choose for digital content management, beyond what your institution already has in place, is cost. And when determining cost, it is important to take all of the costs of managing storage into account. The total cost of ownership (TCO) considers all of the media, labor, and overhead costs that go into installation, ongoing management, and even migration from one storage option to another at some point in the future. Of all the costs, ongoing management is the highest, so institutions more frequently consider cloud storage as a way to alleviate the day-to-day costs and responsibilities for storing digital content. Whether cloud storage is the answer depends on each institution’s local organizational, resource, and technology infrastructure.

LOCAL STORAGE

Storage offerings are as diverse as the institutions that they serve. They may be online only, or some combination of online, nearline, and offline. In all cases, there are associated costs to managing the servers and media on which digital content is stored. Staffing, facilities, and on-going management of—and upgrades to—technology must all be factored into the costs of maintaining storage locally (i.e. “on premise”). Digital collections managers should develop strong relationships with IT staff who manage storage at their institution, so that they can work together to build the best local storage environment possible for the digital content they wish to preserve over time.

CLOUD STORAGE

Cloud storage is a service model in which digital content is maintained, managed, backed up remotely, and made available to users over the internet. Examples of cloud storage include Amazon S3, Amazon Glacier, and Google Cloud Storage.

Cloud providers offer different services, features, and performance levels based on costs and the intended market. A few considerations when assessing cloud storage options are:

- **Latency.** How quickly does the system respond to requests for access to a digital file?
- **Geographic diversity.** Will your data be stored in one location or backed up to multiple locations?
- **Security.** What services are in place to ensure your data is safe?
- **Disaster recovery.** What happens if systems fail?
- **Exit path policies.** How difficult is it to get your data out, either in chunks or as a whole?
- **Costs.** What are the costs to upload data into the cloud? What are the ongoing service costs? What does it cost to download your data or exit the service entirely?

COMPARISON OF CLOUD AND LOCAL STORAGE

Before deciding on one solution over another, a comparison of the features of each, in relation to the need for long-term management of digital collections, should be undertaken. Some considerations are listed in the accompanying table.

Each type of storage has its own financial and organizational implications, and each institution will need to weigh the factors above to come up with a solution that best suits their needs. In some cases, it will not be an either/or decision but a solution that uses both types of storage to their best effect for the institution’s unique situation.

For example, one institution might have a mandate to maintain all collections, whether digital or physical, onsite. In this case, they may opt for a local-only storage solution. Another institution might not have the infrastructure and staff to manage collections onsite, due to costs or personnel restrictions, and may opt instead for cloud storage (from Amazon, Microsoft, Google, etc.) or even a provider like Preservica or DuraCloud that offers a set of preservation services in addition to cloud storage. And, as is more and more often the case, an organization might opt for a hybrid approach. In this case, they may choose to keep a single online copy on local storage so they have quick access to files when they need them. Secondary and tertiary copies may be stored locally on online or nearline storage or in the cloud. Often, yet another copy is stored on magnetic data tape (such as LTOs) in a different geographic location. These second and third backup copies tend to be versions of files that do not need to be accessed readily except for periodic fixity checks. This hybrid approach is an excellent way of (a) alleviating single points of technology failure by distributing content across storage solutions and (b) distributing content across geographically diverse locations.

TABLE 4.1**Comparison of cloud and local storage.**

Credit: AVP

	Cloud	Local
Cost	<p>Change in cost of services over time are unknown.</p> <p>Pricing is frequently akin to early cell phone plans; i.e. there are lots of unknowns until you're "in."</p> <p>Pay as you go. Only pay for what you use.</p> <p>Amount of admin required is typically unknown up front.</p>	<p>Most storage systems last 5–7 years. The cost of replacement must be taken into consideration.</p> <p>A significant portion of costs are upfront to pay for new technology, although ongoing costs for staffing and facilities is also a factor.</p>
Staffing	<p>Requires some staff to configure options, troubleshoot with technical support, and coordinate efforts. This is less of a staffing burden than local storage.</p>	<p>Requires dedicated staffing to manage infrastructure and users.</p>
Support	<p>Support will be different depending on the service provider. Because the user bases tend to be large, generalized services such as knowledge bases or FAQ pages are available.</p>	<p>Support is dependent on the IT staff responsible for managing the storage environment.</p>
Exit Path	<p>Many cloud storage plans make it cheap to upload content but very expensive to download it. Different cost models exist, and it is important to consider them carefully.</p>	<p>There is a clear exit path that is straight forward, although it requires more logistical planning and coordination on the part of the IT staff and the digital collections manager.</p>
Scalability	<p>It is relatively easy to increase the amount of storage you need—it is cost dependent.</p>	<p>Typically, scalable but may take more staff time and financial and computing resources to grow storage capacity.</p>
Forward Looking	<p>Storage and computing in general are trending toward the cloud.</p>	<p>It may pay off to take a "wait and see" approach with cloud storage, so you have more time to understand the true nature of cloud storage and computing as it matures and is tested over time.</p>
Sustainability	<p>Pay-as-you-go provides for more predictable financial planning over a longer period of time but requires continual investment.</p> <p>If funding stops or goes away, there are few/no options for what to do with the stored content.</p>	<p>Ongoing funding is required, and when technology must be updated, there are short-term capital cost increases.</p> <p>If funding stops or goes away suddenly, the infrastructure exists to buy time while you come up with alternatives.</p>

Resources**General**

- Griffith, Eric. "What Is Cloud Computing?" *PCMag.com*, May 3, 2016. <http://www.pcmag.com/article2/0,2817,2372163,00.asp>
- "Nearline storage." Wikipedia. https://en.wikipedia.org/wiki/Nearline_storage
- Newman, Jared. "Seagate slapped with a class action lawsuit over hard drive failure rates." *PCWorld.com*, February 2, 2016. <http://www.pcworld.com/article/>

[3028981/storage/seagate-slapped-with-a-class-action-lawsuit-over-hard-drive-failure-rates.html](http://www.pcworld.com/article/3028981/storage/seagate-slapped-with-a-class-action-lawsuit-over-hard-drive-failure-rates.html)

- "Storage." *Digital Preservation Handbook*. <https://www.dpconline.org/handbook/organisational-activities/storage>

Standards

- ISO 14721:2012, Space data and information transfer systems—Open archival information system (OAIS)—Reference model. <http://public.ccsds.org/publications/archive/650x0m2.pdf>

- ISO 16363:2012, Space data and information transfer systems—Audit and certification of trustworthy digital repositories. <https://public.ccsds.org/pubs/652x0m1.pdf>

SECTION 4: ACTIVE MANAGEMENT

Digital preservation requires active management to ensure that digital assets persist over time. Compared to analog counterparts that may need little beyond stable environmental conditions, digital content requires constant monitoring for changes and mitigation should changes arise.

Change can occur in a variety of ways. Files may be updated or altered intentionally for good or for nefarious purposes. Natural disaster, hardware and software failure, **bit rot**, and human error can all affect the stability of digital content, as well. A key component of active management is an awareness of the threats posed by change and employing systems to reduce the chances that change can occur.

This section describes some of the key functions needed to actively and fully manage digital content for preservation.

Redundancy and Geographic Separation

Maintaining multiple copies of digital content in different geographic locations is a fundamental practice of preservation, whether in the physical domain (more than one institution may preserve copies of the same film) or in the digital domain (the same audio file may be stored on servers in Los Angeles and New York). It also means ensuring that your digital content is stored on more than one type of media; for example, spinning disk and data tape.

In the digital realm, it is ideal to maintain three copies of your digital content, stored in different geographic locations, on different types of media, and maintained in such a way that the copies are always the same.³² This ensures that if something happens to one copy in one location or on one type of media, at least two other unchanged copies of the digital content persist.

PRESERVATION DATA BACKUPS

Most of us have accidentally deleted a file that our IT department was able to restore, either in its entirety or in an earlier state based on when they last performed a backup of the server on which the file was stored. Active data backups involve copying actively used production files, often on a daily basis, for the purpose of short-term retention while the files are in use. These backups might be saved for a week or a month, but after a period of time they are overwritten by new backups. Access copies of digital audiovisual content—the files that you use and share on a frequent basis—are typically backed up in this way.

Preservation data backups are slightly different. Instead of files in active use, the best-quality versions of the digital files that are in a finished, inactive state (often referred to as “preservation masters”) are copied from a primary storage location to a secondary (and ideally, tertiary) storage location. All three copies of replicated data are typically composed of identical “packages” that contain the digitized preservation masters, as well as preservation metadata to help identify and use the files when, for example, the access copies are no longer viable and must be replaced. The preservation masters are not meant to be accessed in the near term, however. The length of time that these preservation packages are backed up is dependent on their retention requirements. In the case of digital audiovisual content, the retention schedule is often “for as long as possible.”

GEOGRAPHIC DISTRIBUTION

Ensure that copies are in geographically disparate storage locations/systems to decrease the likelihood of loss due to localized disaster or service interruption. For example, if your data center is located along the coastal United States, it is best to store duplicate copies of your digital content on servers in another region where hurricanes are not a concern. Or, for example, a university may store digital content on one type of media in campus buildings, on premise, and on another type of media in a satellite location, off site, 20 or 30 miles away. It may also maintain a third copy of the digital content even further away to provide geographic diversity, or even in the cloud (see Section 3).

32 Megan Phillips, et al. “The NDSA Levels of Digital Preservation: An Explanation and Uses,” www.digitalpreservation.gov/ndsaworking_groups/documents/NDSA_Levels_Archiving_2013.pdf

Ideally, storage devices will allow preservation packages to be actively monitored so that errors in data can be identified and repaired. How that identification and repair happens is described in the next section.

File Fixity and Data Integrity

In the context of digital preservation, fixity describes the unchanged state or “fixed-ness” of a digital file. Fixity monitoring can identify whether a file has changed for any number of reasons, such as human error, hardware failure, or bit rot, so that action can be taken to repair it.

ZEROS AND ONES

Digital files are, at their most basic level, made up of a series of zeros and ones.³³ Each number is a “bit.” A file is essentially a very long list of bits stored in a particular order that is different from the order of every other digital file, and this enables a computer to access and play it. A small change, such as turning a single zero into a one, changes the makeup of the entire file, sometimes in catastrophic ways.

Keeping track of all of the bits is hard. Files can be thousands or millions of zeros and ones long, and it is important to ensure that files are stable, or “fixed,” so that they can be accessed. Instead of keeping track of the bits themselves, we can check the files’ fixity by watching or listening to each of them on a regular basis, but that is time consuming and unrealistic in large collections. Luckily, there is another method for monitoring fixity to ensure that if a file does change, we catch that change and can repair it.

MONITORING FIXITY WITH CHECKSUMS

To make monitoring files for changes to the bits easier than keeping track of millions of zeros and ones, we use shorter, alphanumeric strings that reflect the uniqueness of every digital file. These strings are called checksums and are generated by a program that reads the zeros and ones of a file and creates a unique string of characters to represent them. This checksum becomes that file’s signature and will remain the same as long as the bits do not change.

If a file is changed, even in seemingly insignificant ways (you may not even be able to tell just by playing a file), a completely different checksum will be produced by the checksum generator.

Checksums are valuable for several reasons. For example, they can be used to authenticate a file. If a file is the official version of a video, it can be authenticated by first creating a checksum signature and then running the checksum program later to be sure that the signature has not changed. Or, if a file is being deposited in an archive, a checksum may be produced before deposit and then after deposit. This one-time test authenticates that the file is what the depositor understands it to be and that it was not corrupted upon ingest.

One of the greatest values of checksums is their use in monitoring file fixity, which tests for fixedness or stability of the bits over time. It is important to note that while different files have different checksum signatures, exact copies of files will have the same signature. As long as a file doesn’t change, it will always have the same checksum signature as other identical copies. That means files can

FIGURE 4.5

Image of a list of checksums, whose signatures are on the left, and their associated files. The algorithm used to create the checksums in this example is called SHA256, although there are others, such as MD5 and SHA1, in wide use as well.

Credit: AVR

```
99 00:00:00 99 99
2016-08-18 15:43:57
||-||0
sha256
3305a6ec078c2627991e86b7385459dbf43e983bee35de6cc13dd7c26e44c5ff /Users/.../photos/choc_pudding.jpg 93656899
218a3e8add730d479bb492c3ef86f1c10084e311c2c79267c0f706b48c33ed8b /Users/.../photos/green_beans.jpg 93656846
24bd9d27a52a11c5edd4a4bc7d75edbb2724cecbf44541e0ebfcb7c34df0484a /Users/.../photos/ice_cream.png 93656879
f55a999e834bd9c6660bbfd676f65860ebefb2d9f008ad5469a1f311cb631653 /Users/.../photos/potato_salad.jpg 93656713
```

33 Computers use binary code (zeros and ones) to process information.

be monitored for change using a tool that checks fixity on an ongoing basis and repairs files when checksums do not match by replacing them with another unchanged copy of that file. Monitoring fixity over time (e.g. once every month, six months, or a year) allows you to identify changes and replace corrupt files with an unchanged copy.

Finally, while checksums are the primary mechanism for monitoring fixity at the bit level, they can also be used to monitor file attendance, or identifying if a file is new (the checksum signature has never been produced before), removed (a checksum is missing from a list), or moved (the checksum appears with files in another location). Tracking and reporting on file attendance is a fundamental component of digital collection management and fixity.

THE FIXITY TOOL

Some institutions have sophisticated systems and workflows that automate the monitoring of fixity and file attendance in their digital collections; however, many do not. A good place to start with fixity and tracking checksum signatures, if technology systems are not available to automate this process, is by maintaining an inventory in spreadsheet software that lists the files in your collections, their locations on your servers, and their associated checksums and the dates that they were produced. Over time, the inventory will help you identify changes at both the bit and file level so that you can find and repair the errors in your collections.

There are a variety of tools for creating and verifying checksums. Some of these include:

- ExactFile (Windows, <http://www.exactfile.com>). Calculates a variety of checksums.
- FastSum (Windows, <http://www.fastsum.com>). Calculates MD5 checksums.
- HashMyFiles (Windows, http://www.nirsoft.net/utis/hash_my_files.html). Calculates SHA1 and MD5 checksums.

In addition, all operating systems have built-in checksum generation and validation functionality; however, using them requires access via a command-line user interface.

An institution might also consider the free and open-source tool Fixity.³⁴ Fixity was created with smaller and/or lesser-resourced organizations in mind and is a simple application that enables automated checksum production and file attendance monitoring and reporting. Fixity scans a folder or file directory and creates a manifest of the files including their file paths and checksums, against which a regular comparative analysis can be run. Fixity monitors file integrity through generation and validation of checksums and file attendance through monitoring and reporting on new, missing, moved, and renamed files. Fixity emails a report to the user that documents flagged items along with the reason for a flag, such as that a file has been moved to a new location in the directory, has been edited, or has failed a checksum comparison for other reasons.

Information security

The ultimate goal of preservation is to ensure that collections remain unchanged and accessible over time. Information security protocols help to minimize accidental or nefarious changes by users and the public and to track how files are changed, who changed them, and when the alterations were made. Protocols help ensure that authenticity is maintained, and when it isn't, that a change is documented so that an organization can act on that change by, for example, replacing corrupted files with backup copies. This section describes some of the ways that information security is used to maintain the authenticity of digital collections.

USER PERMISSIONS

In a digital preservation environment, organizations must control which users are accessing and manipulating data. Some users may have access to view files, while others may have controls over where files reside, their formats, and who can access them. Creating, assigning, logging, and managing permissions and restrictions are critical to mitigating the risk of intentional or unintentional data corruption and misuse of content. Many preservation management systems make permissions management easy. However, when data management happens manually or outside of a management system, then setting access permissions on directories on networked servers can be a

34 <http://www.avpreserve.com/tools/fixity>

TABLE 4.2

An example of how digital preservation access permissions can be applied to users within an organization. Permissions will differ for each organization, and it is important that staff responsible for digital preservation have a voice in how they are applied.

Credit: AVP

Staff position	Access Permission on Storage Device			
	Read	Write	Move	Delete
Digital preservation staff	X	X	X	X
IT support staff	X	X	X	
Other internal users	X			

TABLE 4.3

An example of a manually constructed audit trail that includes user, file, action, and date.

Credit: AVP

User	File	Action	Date
Jane Smith	f10d6b9e.wav	Moved file from Directory 1 to Directory 2	2016-09-07

good way to create a secure space to store digital collections. IT staff can usually help implement these strategies.

Computer systems offer varying levels of permissions that enable or restrict access to digital collections. Four types of permissions are read, write, move, and delete. These functions are more or less what they sound like, although the concepts of “read” and “write” may not be entirely obvious. In this context, “read” access refers to the ability to view files in a directory without being able to edit or delete the files. “Write” access means that a user can add files to a directory or edit files within it. The levels of permissions are cascading, starting with the lowest level of access—none—and ending with the greatest—“delete.” This means that a user with “delete” access also has “move,” “write,” and “read” access. Conversely, a user may have only “read” access or no access at all. If you are able to set permissions on directories that contain preservation copies of digital files, then consider doing so, but with caution. Always be sure that an administrator has full access to the collection, so that if the digital content needs to move, migrate, or be monitored for fixity, there are no restrictions on doing so.

As with all digital preservation activities, it is important to document decisions about permissions. It is one thing to assign permissions, but having information at hand that documents who has been assigned which permissions will

help identify whom to contact when diagnosing data errors. Logging access and internal actions taken on digital collections is equally important, although challenging, in an environment where the logging is not an automated process. Logs, or audit trails, enable digital collections administrators to audit actions taken and track changes to a user and date, which can be valuable when trying to understand when and where errors have occurred. Although by no means impossible, these activities can be time consuming in a manual workflow. An example of a manually constructed audit trail might include the following notations (Table 4.3).

The good news is that many preservation management systems automate this work. Whatever preservation tools you have at your disposal, always remember to do the best you can do now, so that you’re ready to do more when the resources and infrastructure enable it.

PROTECTING FROM EXTERNAL THREATS

Every organization needs to be concerned not only with controlling access to digital collections from within, but also with protecting assets from external threats. This is particularly important for storage devices such as servers that are connected via networks and to the internet. Controlling access to these devices with good password and username practices is imperative. Passwords should be

unique, high-strength (upper and lowercase letters, numbers, and symbols), long (12–15 characters is recommended),³⁵ and changed routinely.

Operating systems, when not updated routinely, can be another potential risk to a network. For example, at the time of this writing in 2017, Windows XP runs on 7% of computers across the globe, although support for it from Microsoft officially ended in 2014.³⁶ This means that security patches, created in response to active virus threats and other nefarious software, are rarely produced. This puts these computers and the networks on which they run at risk of security breaches. Taking a proactive approach to keeping operating systems up to date decreases the chances of data breaches.

VIRUS SCANNING

Another important component of information security is monitoring digital collections for viruses and other corrupting malware. Virus scanning should be performed on files that are being brought into a digital preservation environment from external sources to avoid infecting existing files and systems. If you are performing digitization internally and have full control over the files being created, this is less of an issue. It becomes a greater concern when accepting files from external sources, such as donors or even other units within an organization. Virus scanning should be performed on all external data transfers into the environment and then routinely in the digital preservation environment as an added precautionary measure. One approach is to have a dedicated “clean” computer that is not connected to a network or the internet. Files can be virus tested on this machine before transferring them to networked storage. This may be especially useful to institutions where IT support is not readily available.

A good idea is to create a folder that is specifically used to store newly acquired files to prepare them for ingest into your preservation environment. The folder should not be directly connected to your preservation environment. If you cannot check the files on the media on which they are delivered, transfer new files directly into this folder and

immediately perform virus checking on them. That way, if corruption is identified, the files cannot infect your existing digital collections.

There are many low-cost virus-scanning software options available on the market today. If you have an IT department, talk to them about virus scanning before you purchase your own software. They might have options available that are already in use in your organization.

Resources

- AVPreserve. “Fixity.” <https://www.avpreserve.com/tools/fixity>
- De Stefano, Paula, et al. “What is Fixity, and When Should I be Checking It?” *Checking your Digital Content: An NDSA Publication*, 2014. <http://www.digitalpreservation.gov/documents/NDSA-Fixity-Guidance-Report-final100214.pdf>
- “Information Security.” *Digital Preservation Handbook*. <https://www.dpconline.org/handbook/technical-solutions-and-tools/information-security>
- “Storage.” *Digital Preservation Handbook*. <https://www.dpconline.org/handbook/organisational-activities/storage>

SECTION 5: METADATA

We recommend that you read this section together with “Metadata,” in Chapter 3 (page 49), to get a complete understanding of the types of metadata pertinent to audiovisual collections, how metadata is generated, and standards for capture.

Metadata is the information about a digital file that allows us to understand, use, manage, and preserve it. Without it, we would not know about the file (e.g. the title, who created it, and on what date), what the file is (e.g. the wrapper or codec in use, data rate, pixel dimensions,

35 Brian Barrett, “7 Password Experts on How to Lock Down your Online Security.” Wired.com, May 5, 2016, <https://www.wired.com/2016/05/password-tips-experts>.

36 Nicole Becher, “Patching is Hard, and Windows XP Will Die Harder,” http://www.slate.com/articles/technology/future_tense/2017/05/patching_is_hard_and_windows_xp_will_die_harder.html

duration, etc.), how it relates to other files (e.g. part one of three), and how it has been monitored over its lifetime (e.g. fixity checks). Without the appropriate metadata, a file becomes inaccessible and unusable, ultimately losing its value.

Metadata is produced at various times during a file's lifespan. Descriptive and technical metadata are often captured at the point of creation, such as during the digitization process. Preservation metadata, on the other hand, is generated on an ongoing basis to log actions performed on the file, including such activities as transferring it from one storage environment to another, performing fixity checks, or migrating from one format to another. Logging preservation metadata over time creates an audit trail that ensures a file remains authentic and accessible for the long-term.

Types of Metadata

Each type of metadata plays a different role that, along with a digital file, makes up the information package that ensures long-term access. It is useful, then, to describe the different types of metadata and how they affect a digital file.

DESCRIPTIVE METADATA

Descriptive metadata is the information about a file or files that enables identification and discovery. Descriptive metadata includes the title, creator, date of creation, and keywords that document the subject of the file's content.

STRUCTURAL METADATA

Structural metadata is the information that designates how a set of files relate to one another, such as songs on a CD or how the parts of a single file are structured. For example, it could tell you that a file comprises a container file with one video, two audio, and two subtitles tracks.

ADMINISTRATIVE METADATA

Administrative metadata includes information about how to manage a digital file and track its process history. This ranges from rights metadata, which indicates who owns or holds copyright for a file and how it can be used and accessed, to **technical** and **preservation metadata**, which are described in detail below.

While all forms of metadata help provide long-term access to digital collections, technical metadata has particular significance for audiovisual content and preservation metadata is key to ensuring that digital content can be managed over time.

TECHNICAL METADATA

Technical metadata captures the essence of a digital file. It is the technical information that describes how a file functions and that enables a computer to understand it at the bit level, so that it can be played back in a way that is useful for a viewer or listener. Technical metadata includes information such as wrapper, codec, compression, and aspect ratios. Often, technical information is embedded in the file itself and is read directly by compatible software and hardware. For preservation purposes, technical metadata is extracted and stored outside the file so that as formats obsolesce and compatibility fades, a file's basic structure is understood and can be migrated to a new format that allows it to be accessed. Many audiovisual metadata schemes rely on technical metadata—extracted from embedded metadata—as one method for maintaining the usability of digital content over time.

PRESERVATION METADATA

Preservation metadata is the information necessary to support the management and long-term accessibility and usability of an object. It tracks the processes that are necessary to manage a file in a digital environment over time, including: monitoring fixity and performing any repairs that are identified during fixity checks, auditing logs to identify when and who has interacted with an object, monitoring obsolescence information, and documenting provenance information to support the authenticity of an object. Examples of preservation metadata include checksums, storage locations, and records of process activities and dates (for example, that a file is moved from one location to another and the date that the move occurred).

Embedded Metadata

Embedded metadata is the information that is stored within a file that also stores the content to which the metadata refers. For example, a WAVE file contains both the music and the technical information to play the file. Embedded metadata can also include descriptive information,

as in mp3 files, which enables display of an artist, album, and title in applications that play them.³⁷ Thinking about it another way, embedded metadata is the digital equivalent of physical labels, annotations, and written documentation stored inside a material housing or the video slates at the head of a recording.

Embedding information about the holding organization (the data source that holds information about the object) and the copyright status also helps to identify the file if it becomes disassociated with the metadata that is part of its information package. The Federal Agencies Digitization Guidelines Initiative (FADGI) is a set of published guidelines for digitization processes and offers guidance for the use of embedded metadata in WAVE files.³⁸ For example, the guidelines offer recommendations about how to store embedded metadata in WAVE files that result from the digitization process.

EXTRACTION TOOLS

Because of the embedded nature of much technical metadata, tools have been developed to automate the extraction of this information from the files in which it is held. Two of those are:

- **FITS**, <https://projects.iq.harvard.edu/fits>, is a command-line tool developed by Harvard University Library that identifies, validates, and extracts technical metadata for digital objects, including some audiovisual formats. The metadata is exported into an XML file.
- **MedialInfo**, <https://mediaarea.net/en/MedialInfo>, is an open-source program that extracts technical metadata about media assets and exports it into a variety of formats including txt, EBUCore, PBCore, and reVTMD, which are all described below. It works with a variety of audio and video formats and has a GUI interface, so command-line knowledge is not necessary. It is available for many operating systems.

Standards, Schemas, and Guidelines

Metadata standards, schema, and guidelines are invaluable to the creation, management, and sharing of information. They tell us how and why certain metadata should be captured, enabling us to easily understand metadata created by others and minimizing the obstacles of sharing information between systems. Metadata can be stored in Excel spreadsheets, as XML files, or in databases such as content management systems and institutional repositories, as well as in other formats. However metadata is stored, using standards to create and structure it will make it more broadly understood and interoperable.

The standards and guidelines briefly described below are just a few of the most recognized and recommended for the management of audiovisual collections.

EBUCORE

EBUCore is based on the Dublin Core standard and adapted to broadcast media. It is a descriptive and technical metadata schema developed and maintained by EBU, the largest professional association of broadcasters in the world. EBUCore captures the minimum information needed to describe radio and television content.

Link to more information about EBUCore:
<https://tech.ebu.ch/docs/tech/tech3293.pdf>

Link to the EBUCore metadata specification:
<https://tech.ebu.ch/MetadataEbuCore>

FADGI (FEDERAL AGENCIES DIGITIZATION GUIDELINES INITIATIVE)

Begun in 2007, this is a collaborative effort by US federal agencies to define common technical guidelines, methods, and practices for digitizing historical content and the capture of technical metadata. The focus of the audiovisual working group, in particular, is to identify, establish, and disseminate information about standards and practices for the digital reformatting of historical and cultural audiovisual materials by federal agencies, although the guidelines

³⁷ <http://id3.org>

³⁸ Federal Agencies Audio-Visual Working Group, "Embedding Metadata in Digital Audio Files: Guideline for Federal Agency use of Broadcast WAVE Files," version 2. http://www.digitizationguidelines.gov/audio-visual/documents/Embed_Guideline_20120423.pdf

have seen broad use beyond the US government as well. The effort covers sound recordings, video recordings, motion picture film, and born-digital content.

Link to more information about the FADGI audiovisual working group and its guidelines: <http://www.digitization-guidelines.gov/audio-visual>

METS (METADATA ENCODING AND TRANSMISSION STANDARD)

The METS schema is a standard for encoding descriptive, administrative, and structural metadata about objects within a digital library, expressed using XML. METS provides an XML document format for encoding metadata necessary for both the management of digital library objects within a repository and the exchange of such objects between repositories or between repositories and their users. METS is a Digital Library Federation initiative that is maintained by the Library of Congress.

Link to more information about METS: <https://www.loc.gov/standards/mets/mets-home.html>

PBCORE (PUBLIC BROADCASTING METADATA DICTIONARY PROJECT)

PBCore is a metadata schema designed for sound and moving images. It can be used as a guideline for cataloging the descriptive and administrative information about audiovisual content. It can also act as an exchange mechanism to share information between institutions or applications, and much more. PBCore expands on the Dublin Core standard. It was created by the US public broadcasting community and is maintained by WGBH in Boston.

Link to more information about PBCore: <http://pbcore.org>

PREMIS (PRESERVATION METADATA: IMPLEMENTATION STRATEGIES)

The PREMIS Data Dictionary for Preservation Metadata is the international standard for metadata to support the preservation of digital objects and ensure their long-term usability. PREMIS is a comprehensive, practical resource

for implementing preservation metadata in digital archiving systems.³⁹ It is maintained as a standard by the Library of Congress.

Link to more information about PREMIS: <http://www.loc.gov/standards/premis>

REVTMD

reVTMD is an XML schema tailored to include fields that address the creation and long-term management of reformatted videos, especially for the cultural heritage community. It is a concise subset of the large array of technical metadata available for digital media, structured in a way to make it highly usable for accessing and managing all types of video files. The captureHistory section is especially helpful in capturing process history for preservation purposes. reVTMD was developed by the US National Archives and Records Administration in collaboration with AVPreserve.

Link to more information about reVTMD: <https://www.weareavp.com/tag/revtmd>

Link to the reVTMD XML schema: <https://www.archives.gov/preservation/products/reVTMD.xsd>

Resources

- Baca, Murtha, ed. *Introduction to Metadata*, 3rd edition. <http://www.getty.edu/publications/intrometadata>
- Greisinger, Peggy. "A brief overview of metadata for audiovisual materials," November 6, 2014. <http://ndsr.nycdigital.org/a-brief-overview-of-metadata-for-audiovisual-materials>
- International Association of Sound and Audiovisual Archives. "Metadata." In *Task Force to establish Selection Criteria*. <http://www.iasa-web.org/task-force/7-metadata>
- Zeng, Marcia L. *Metadata Basics*. <http://metadataetc.org/metadatabasics>

39 Library of Congress, PREMIS Data Dictionary for Preservation Metadata, <http://www.loc.gov/standards/premis>

SECTION 6: PLANNING FOR OBSOLESCENCE

As this chapter has made clear, managing digital content requires a great deal of planning to ensure that the three-legged stool of organization, resources, and technology remains balanced over time. And, like so much related to digital preservation, this planning cannot stop; it is a long-term commitment to long-term access of digital content. Planning today takes into consideration the needs of tomorrow to ensure that wrappers, codecs, and the media on which digital objects are stored do not obsolesce—and if they do, that there are established pathways for moving them to new formats and media so that they remain accessible over time. (For more information on reformatting, see Chapter 3.) It's not only formats that obsolesce but also the media on which they are stored and the systems that are required to read or play them. If, for example, an mp3 music file is on a CD ROM but you no longer have a player with which to listen to it, the viability of the file format is only one concern; finding a way to read the media is another complicating factor.

Obsolescence monitoring

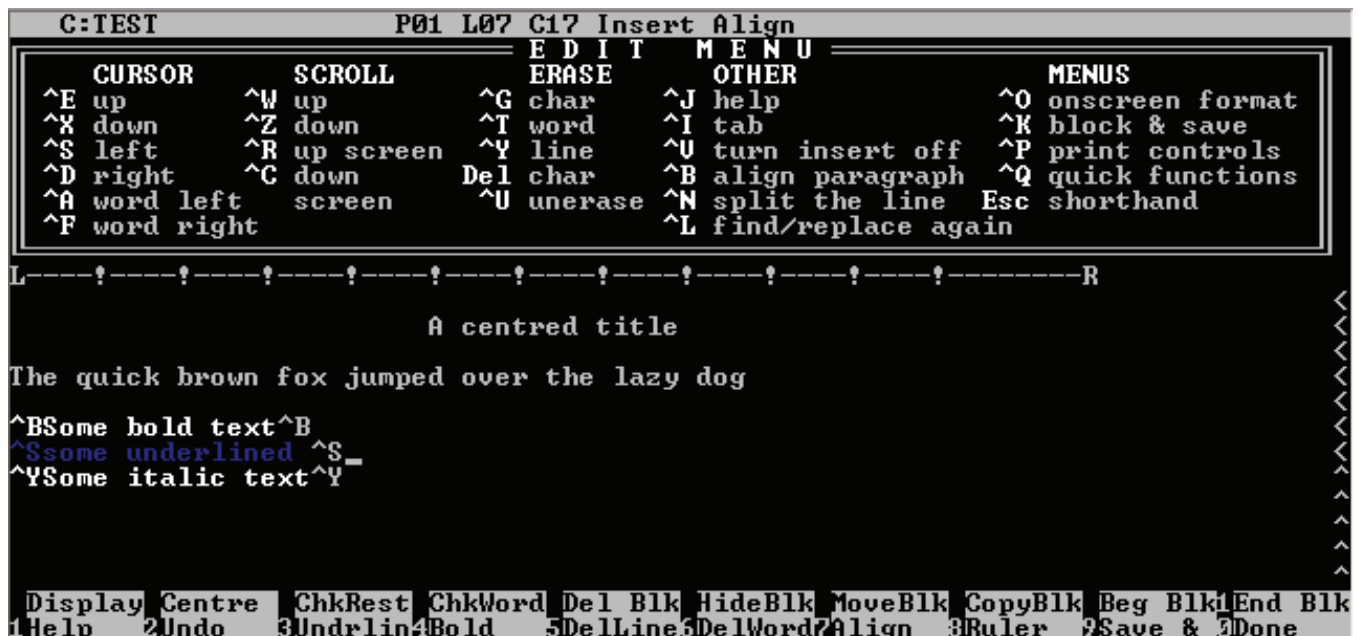
A fundamental practice of digital preservation is monitoring wrappers, codecs, and media to ensure they remain usable and that they have not become obsolete. Obsolescence can happen at the file or program level (such as Real Media and WordStar) or with the media on which the file is stored (such as Zip drives). When this happens, the digital content becomes unreadable or inaccessible.

Before obsolescence happens, steps can be taken to transition to current storage options or formats, such as Microsoft Word. Keeping track of changes in technology, called “obsolescence monitoring,” involves maintaining awareness of file formats, software, and systems that are ubiquitous (which have less of a chance of becoming obsolete) and, more importantly, awareness of more specialized or less-used formats, which have a greater chance of obsolescence. Monitoring can be accomplished without specialized technologies simply by keeping up with the changing landscape; however, technology-based watch systems are available. Paired with proactive planning, these systems can help ensure that digital content does not become obsolete.

FIGURE 4.6

An LTO data tape cartridge. A screen shot of WordStar 4.0.

Credit: AVP.



The good news is that obsolescence tends to happen slowly. Routinely reviewing files, conferring with colleagues, and learning about industry-standard formats are some of the best strategies for ongoing obsolescence monitoring.

Refreshing and Migration

The ultimate goal of digital preservation is the ongoing accessibility of digital content. Think of “content” in this context as you would the moving image on a reel of film. The physical object will change when the film is replaced by a reformatted version stored on digital media, but the content—the images and sound and the order in which they are played—remain the same. Digital audiovisual content must be able to transcend software and hardware changes over time, so choose non-proprietary or “open” formats when possible, such as Broadcast WAV for audio. The media on which digital files are stored will also become obsolete or unstable with age—usually more frequently than file formats do—and those files will need to move to more up-to-date media. Some estimates put the lifespan of hard disk drives at three to five years; others put their median lifespan at six years or more.⁴⁰ When the risk of loss of digital content due to the threat of obsolescence becomes too great, we look to refreshing and migration.

Refreshing refers to the approach of transferring files and/or data from one media, server, or system to another. This may consist of moving files from an aging server to a new one or shifting metadata from an obsolete database to a more widely used system. Great care must be taken to ensure that all data is transferred in a “lossless” way and that the integrity of the content is verified after the move. Error checking could include running fixity checks on files moved from one server to another, looking for changes or missing files, or reviewing a significant sample of metadata records (10% or more). The frequency with which refreshing happens depends on the technology; data on servers should be refreshed at least every three to five years. Shifts in metadata databases are dependent on the ongoing maintenance of the system in which they are stored.

Migration refers to the transfer of the content and meta-data from one audiovisual format, such as a wrapper and/or codec, into another. Migration is necessary when the wrapper/codec or the software used to read the format becomes less ubiquitous and is on the verge of becoming inaccessible. Migration may consist of transferring audio from one wrapper to another without changing the codec, or it may consist of transcoding the audio and placing it in a new wrapper. Real Video, a proprietary video file and encoding format, used to be relatively common for video streaming on the web in the late 1990s and early 2000s. Today the ubiquity of formats like MP4 with h.264 encoding have rendered Real Video practically obsolete. If care is taken to ensure that video and audio are captured at high resolution using stable and open or uncompressed formats, the need to use migration for preservation will be less likely. However, migration, or perhaps the creation of new derivatives from preservation masters, might be required as web trends shift over time.

Because obsolescence can happen at many levels (codec→wrapper→content management system→storage media), refreshing and migration plans must consider all of these: hardware (e.g. servers), software (e.g. video platforms, codecs), and databases (e.g. digital asset management systems [DAMS], collection management systems [CMS]).

Initially selecting the codecs and wrappers, systems, and media with the greatest longevity and openness is ideal and means that refreshing and migration are not immediate concerns. While refreshing and migration are inevitable, being able to postpone that need means you can focus on other aspects of digital preservation management. That isn’t to say that planning should not happen. Identifying funding that will be required when new hardware must be purchased or when staff are hired to complete the process of migration and refreshing helps to future-proof (and disaster proof) digital collections.

40 Brian Beach, “How long do disk drives last?” Nov. 12, 2013. <https://www.backblaze.com/blog/how-long-do-disk-drives-last>

Resources

- “Digital Preservation Strategies.” In *Digital Preservation Management: Implementing Short-Term Strategies for Long-Term Solutions*. <https://dpworkshop.org/dpm-eng/terminology/strategies.html>
- “Obsolescence: File Formats and Software.” In *Digital Preservation Management: Implementing Short-Term Strategies for Long-Term Solutions*. <https://dpworkshop.org/dpm-eng/oldmedia/index.html>
- *Sustainability of Digital Formats: Planning for Library of Congress Collections*. <http://www.digitalpreservation.gov/formats/intro/intro.shtml>

SECTION 7: PRIORITIZATION AND PHASING

Approaching digital preservation as a whole can be intimidating and overwhelming. Prioritizing approaches based on a digital preservation plan can assuage those feelings. Instead of thinking about the “foreverness” of digital preservation, consider it in periods of five- or 10-year increments. At the end of each period you have the option to make a decision about what to do in the following five or ten years. You can choose to do nothing, do the minimal amount necessary to maintain the option to decide again later, or you might choose to pursue a more robust solution, in whole or in part.

The aphorism “perfect is the enemy of good” is a useful way of thinking about prioritizing and phasing digital preservation activities. Doing something now—for example, creating an inventory of your collections, developing a collection policy, or making a backup copy of your digital content—is better than waiting for the perfect technological solution along with the resources and organizational framework to support it.

So, how to prioritize? Consider the following factors:

- **Impact.** Small steps can have great influence. Proper file naming, good metadata, and the use of open format wrappers and codecs are a few strategies that will have a significant impact on the longevity of your digital

collections. They do not require complex systems and expensive technology to implement but they make digital content more findable, understandable, and usable over time.

- **Feasibility.** The reality is that some digital preservation practices require technology and resources that not all institutions can staff or fund. Other activities may be too time consuming for one person to undertake. Figure out what you can manage right now, and do it. As your capacity and experience grows, you might find that you gain support for more technologies and resources, making more complex activities more feasible.
- **Urgency.** An important consideration in making sure you are effectively using resources—especially if they are limited—is identifying what needs to be done right away. For example, moving digital content off of fragile or obsolescing media such as CDs and onto actively managed servers before data is lost may be of greatest urgency for your institution. Identify problems that, if you do not address them, will become more significant. Also, if opportunities arise, such as a one-time source of funding, be prepared to take advantage of them in a timely manner. Of all the factors, urgency should take priority, especially if digital content is at immediate risk.

Being flexible, putting what you know into practice, and taking a proactive approach today will establish a foundation that makes implementation and adoption of new technologies and programmatic preservation strategies possible and easier in the future.

NDSA Levels of Digital Preservation

Sometimes it is hard to know where to start. With so many standards and guidelines, beginning to address the challenge of long-term, active management of your digital content can be challenging. The National Digital Stewardship Alliance (NDSA) had this challenge in mind when it developed the “Levels of Digital Preservation.”

The “Levels of Digital Preservation” are a tiered set of recommendations for how organizations can begin to build or enhance their digital preservation activities.⁴¹ The Levels (LoDP) are meant to be an easy-to-use set of guidelines

41 Levels of Digital Preservation. <http://ndsa.org/activities/levels-of-digital-preservation>.

for those beginning to think about digital preservation, as well as for those with established programs that are ready to take the next step and enhance services. The focus of LoDP is on the content in digital collections and the infrastructure in place to manage it. LoDP is not designed to assess the overall readiness of digital preservation programs, like ISO 16363 does (see “Section 1.4: Standards”), because it specifically addresses technology and not issues related to organizational infrastructure.

The LoDP guidelines are organized into five functional areas that are at the heart of digital preservation technology systems and which are addressed in more detail in “Section 4: Active Management:”

- storage and geographic location
- file fixity and data integrity
- information security
- metadata
- file formats

TABLE 4.4
National Digital Stewardship Alliance’s Levels of Digital Preservation (2013)

Credit: NDSA

	Level 1 (Protect your data)	Level 2 (Know your data)	Level 3 (Monitor your data)	Level 4 (Repair your data)
Storage and Geographic Location	<ul style="list-style-type: none"> • Two complete copies that are not collocated • For data on heterogeneous media (optical discs, hard drives, etc.) get the content off the medium and into your storage system 	<ul style="list-style-type: none"> • At least three complete copies • At least one copy in a different geographic location • Document your storage system(s) and storage media and what you need to use them 	<ul style="list-style-type: none"> • At least one copy in a geographic location with a different disaster threat • Obsolescence monitoring process for your storage system(s) and media 	<ul style="list-style-type: none"> • At least three copies in geographic locations with different disaster threats • Have a comprehensive plan in place that will keep files and metadata on currently accessible media or systems
File Fixity and Data Integrity	<ul style="list-style-type: none"> • Check file fixity on ingest if it has been provided with the content • Create fixity info if it wasn't provided with the content 	<ul style="list-style-type: none"> • Check fixity on all ingests • Use write-blockers when working with original media • Virus-check high risk content 	<ul style="list-style-type: none"> • Check fixity of content at fixed intervals • Maintain logs of fixity info; supply audit on demand • Ability to detect corrupt data • Virus-check all content 	<ul style="list-style-type: none"> • Check fixity of all content in response to specific events or activities • Ability to replace/repair corrupted data • Ensure no one person has write access to all copies
Information Security	<ul style="list-style-type: none"> • Identify who has read, write, move and delete authorization to individual files • Restrict who has those authorizations to individual files 	<ul style="list-style-type: none"> • Document access restrictions for content 	<ul style="list-style-type: none"> • Maintain logs of who performed what actions on files, including deletions and preservation actions 	<ul style="list-style-type: none"> • Perform audit of logs
Metadata	<ul style="list-style-type: none"> • Inventory of content and its storage location • Ensure backup and non-collocation of inventory 	<ul style="list-style-type: none"> • Store administrative metadata • Store transformative metadata and log events 	<ul style="list-style-type: none"> • Store standard technical and descriptive metadata 	<ul style="list-style-type: none"> • Store standard preservation metadata
File Formats	<ul style="list-style-type: none"> • When you can give input into the creation of digital files encourage use of a limited set of known open formats and codecs 	<ul style="list-style-type: none"> • Inventory of file formats in use 	<ul style="list-style-type: none"> • Monitor file format obsolescence issues 	<ul style="list-style-type: none"> • Perform format migrations, emulation and similar activities as needed

Each functional area is gauged against a set of criteria that help an institution identify their own level of digital preservation readiness. The four levels are progressive; the requirements in the first level are building blocks for levels two through four. Although level one is the foundation on which the other levels are built, sometimes an institution's readiness in another level is higher than it is in level one.

The important thing, and what LoDP can offer to a new or growing digital preservation program, is the ability to prioritize the needs of an organization, effectively identifying what the organization is capable of attaining at the present moment. This means that the organization can be at a readiness level of three without having fulfilled all of the requirements of level two in the same category. After level one is addressed, the organization can shift its focus to filling in missing requirements in level two and progressing onward to level four. The move towards a programmatic solution should be continuous, fluid, and flexible with the understanding that a preservation program is being built that can withstand future organizational changes.

Resources

- Hodges, Patricia, et al. "The Five Organizational Stages of Digital Preservation." *Digital Libraries: A Vision for the 21st Century: A Festschrift in Honor of Wendy Lougee on the Occasion of her Departure from the University of Michigan*, 2003. <https://quod.lib.umich.edu/s/spobooks/bbv9812.0001.001/1:11/-digital-libraries-a-vision-for-the-21st-century?rgn=div1;view=fulltext>
- National Digital Stewardship Association (NDSA). "Levels of Digital Preservation." http://ndsa.org/documents/Levels_v1.pdf
- NEDCC. "Preservation Leaflet 1.4: Considerations for Prioritization." <https://www.nedcc.org/free-resources/preservation-leaflets/1.-planning-and-prioritizing/1.4-considerations-for-prioritizing>

CONCLUSION

Digital preservation is a multi-faceted endeavor that requires institutional planning and policies, a suitable storage infrastructure, active management, and well thought-out and consistently applied metadata. Careful planning and thoughtful policies will provide a program with concrete, achievable goals and will help to ensure the organizational commitment needed to make a program sustainable over time. A storage infrastructure that addresses the needs of the institution while accounting for the costs involved will promote sustainability as well, and the options described in this chapter will empower organizations to make a thoroughly informed decision. However, without vigilant monitoring and management, the security and fixity of files cannot be assured, no matter how well designed the storage is for digital collections. Metadata that complies with standards and is informed by guidelines will facilitate this monitoring and keep digital audiovisual collections accessible into the future. These components, combined with planned migration of collections to address obsolescence, constitute the makings of an effective digital preservation program for audiovisual collections.

This series of programmatic elements can at first appear daunting to those at repositories seeking to reformat and care for audiovisual media. Developing a sense of your institution's priorities will make phasing these components into a preservation program possible. The creation of these priorities should take into account their potential impact, feasibility, and urgency. The NDSA Levels of Digital Preservation can provide a helpful framework to further assist in this decision making.

Because a preservation program is so dependent on the unique context of an institution, there is no single correct way to create a program that is sustainable over time. Using the information outlined in this chapter will best prepare you to chart a course that meets your institution's specific needs at a pace that takes into account the resources available to you. In "Chapter 5: Disaster Preparedness and Response," you will learn about the considerations, strategies, and tips that will help you to secure and salvage your collections when faced with an emergency and to ensure that your developing preservation program is ready for the unexpected.

CHAPTER 5

DISASTER PREPAREDNESS
AND RESPONSE

by Kara Van Malssen, Partner and Senior Consultant, AVPreserve

THIS CHAPTER OUTLINES THE FUNDAMENTALS OF disaster prevention, preparedness, and recovery for audiovisual collections. The contents of this chapter address issues of disaster preparedness and response for physical collections—discs, reels, cassettes, and film—that occupy shelves and boxes. However, one thread that will run throughout this chapter is the necessity of performing reformatting and digital preservation for disaster preparedness purposes. As this textbook makes clear, transforming the content of physical media into digital files is critical as analog equipment and expertise rapidly disappears and demands for accessing digital content increase. Disaster prevention and recovery are yet another justification for digitization and proper digital preservation. If managed correctly, through proper backup and geographic separation, recovery of digital data from an offsite copy will be a relatively painless and lossless process compared to the difficult recovery and virtual guarantee of a certain degree of loss in the event that a flood or fire affects physical media.

Collection managers should prioritize digitization and digital collections management for valuable materials. In the meantime, taking measures to protect collections will go a long way toward ensuring that their contents will be accessible over the long term.

Another tenet of this chapter is that disaster preparedness does not have a stopping point or an ending point. Once one risk is addressed, another can emerge. In order

to be effective, disaster preparedness must be an ongoing, integrated part of collection management.

SECTION 1: DISASTER PREVENTION AND MITIGATION

Preventative measures will help mitigate the chances that a disaster will occur and/or will minimize the effects when one does occur. Proper building and collection security, repairing faulty or exposed wiring, and installing storm shutters are all examples of actions that will reduce the likelihood of theft, electrical fires, and storm damage. Good storage, staff training, and collection knowledge can prevent a host of potential disasters that audiovisual archives may face, especially long-term disasters where damage builds over time, such as continual mishandling and tape deterioration.

The activities outlined here and in “Section 2: Disaster Planning” are not mutually exclusive, nor should they be thought of as sequential. Both preventative actions and preparedness measures can be tackled in parallel. Start with a risk assessment to identify priority areas, and begin to address both preventative measures, such as fixing a leaky roof, while simultaneously performing preparedness actions, such as gathering supplies in case a typhoon threatens to make the damage worse before roof repairs are complete.

Assess and Reduce Risk

Disasters come in all shapes and sizes. A leaking pipe that goes unnoticed for a few months and that results in mold growth is not the same as a catastrophic fire or Category 5 hurricane. Some damage might not be evident immediately, and performing a risk assessment survey will help you prepare for the threats most likely to affect your collection.

The Canadian Conservation Institute (CCI) and the International Centre for the Preservation and Restoration of Cultural Property (ICCRPM) have developed a useful model for managing and reducing risks to cultural collections using a 5-phase risk management cycle.⁴² This simple structure, outlined below, provides a framework for managing risks to collections of audio, video, film, or other formats. By following the ICCROM model, collection managers can address and subsequently eliminate large and small-scale disaster risks to their collections.

ESTABLISH THE CONTEXT

Perform a valuation of the collection by looking at the organizational context. Are all items valued equally? Do you place higher value on original recordings than you do on commercial recordings? Value may be determined by national significance, high historical or cultural value, or how well an item fits the institution's mission. Some items may be assigned high value because there are no duplicates of that content.

Appraising audiovisual materials is a subjective task that requires collection and subject matter expertise; in large institutions it will likely require the input of many staff members. In small institutions or individual collections there may only be one person who can perform appraisal. At minimum, determine a rough sense of the priority items or collections. Identification of priority items will be relevant to preparedness planning and will be critical during a disaster recovery operation.

IDENTIFY RISKS

Identify risks to your region, city, building, and collection. Review the history of disasters. Look at possible risks

and think through prospective scenarios that may cause damage. This information will guide how you prepare for an emergency. Sometimes the biggest threats aren't obvious, so be sure to carefully look at all levels of risk. For example, you might not think much about the aging electrical system in the building, but this can actually be a significant fire hazard. Risks may include rare events (flood, fire, earthquake, war) and cumulative events (water leaks). Write simple descriptions of potential scenarios that illustrate what could happen in an actual emergency in order to document risks and bring them to life.

ANALYZE RISKS

Categorize risks by their frequency and their likely impact. For each one, determine how often the event is likely to occur and how much value will be lost in individual items as well as in the collection as a whole. For example, scratches on an LP caused by a collapsed shelving unit will not likely damage the entire disc and in many circumstances can be repaired. However, the same collapsed shelf could permanently destroy wax cylinders and other fragile formats, thereby affecting a larger percentage of the value and content of each impacted item. Depending on the number of affected items and the total size of the collection, the percentage of the collection affected could be insignificant or very large.

PRIORITIZE RISKS

Prioritize risks based on your analysis. Rate risks according to the probability of occurrence and the level of impact. For example, is a flood likely to happen in your area in the next ten years? Is it likely to cause major damage to collections? If so, this should be a high priority. Is poor labeling and tracking likely to cause important items in the collection to become lost within the next year? If so, this is also a high priority. Risks that are less likely to occur frequently or will have less of an impact are comparatively lower priorities.

TREAT RISKS

Take steps to minimize or reduce identified risks. While you may want to address the biggest risks first, these may

42 *The ABC Method: a risk management approach to the preservation of cultural heritage* <https://www.canada.ca/en/conservation-institute/services/risk-management-heritage-collections/abc-method-risk-management-approach.html>

also be the most challenging. Don't let the biggest tasks prevent you from taking steps to treat the lower-priority risks, especially if these are relatively easy to resolve.

Mitigate Risk

This section highlights areas that typically need attention to reduce risk to collections. This is not an exhaustive list, but it describes solutions for common risks, including those that might be overlooked.

BUILDING STRUCTURE AND SYSTEMS

The building that houses a collection is the first and at times the only line of defense against disasters. Yet, quite often the building itself or weaknesses within it pose the greatest threat to audiovisual collections. Leaking pipes and electrical fires are often the origin of disasters.

When possible, work with the building's maintenance and facilities staff to identify potential hazards and correct them. Incorporate preventative actions into maintenance scheduling for your collection and building.⁴³ This includes activities such as regular cleaning of the work and storage areas, inspection of the facilities, and maintenance of the plumbing. If the collection is at a very small institution, conduct a thorough inspection of the building. Seek help from professionals as well as friends or acquaintances who have experience with building construction, renovation, or maintenance.

Take steps to ensure that the building can protect collections from external threats. For example, if you are in a hurricane-prone area, fit shutters and update windows to code. If your area is frequently affected by dust storms, you may need to use heavy curtains or seal cracks in walls and windows.

For the safety of collections and staff, fire detection and suppression systems should be installed if they are not already.⁴⁴ There are a variety of systems to choose from, including fire extinguishers, wet and dry sprinklers, gas suppression systems, and smoke and heat detectors. Equipment should be regularly inspected and maintained.

In particular, extinguishers require replacement and gas suppression systems require recharging at specific intervals. All staff members should know where detection and suppression systems are located in the building and should be trained in their operation.

PROPER STORAGE

These important points will help reduce the effects of any disaster on audiovisual materials and will increase the chances of a complete recovery of content in the event of damage.

- **Keep tapes rewound.** If a tape is damaged, the exposed area may need to be removed. It would be much better to lose the leader at the beginning rather than important content in the middle of the tape.
- **Store collections off of the floor.** Do not store master materials in a basement or directly under a roof.
- **Store cassette tapes with spine up.** This will help maintain proper tape pack distribution and will shed falling water from above in case of sprinkler activation

FIGURE 5.1
Cassettes stored properly, with their spines up.

Credit: AVP



⁴³ For an extensive checklist of housekeeping and maintenance activities, see Johanna Wellseiser and Jude Scott, *An Ounce of Prevention: Integrated Disaster Planning for Libraries, Archives, and Record Centres* (2nd ed., Lanham, MD, The Scarecrow Press, 2002), pp. 53-61.

⁴⁴ Not only is this important to ensure the safety of personnel and collections, but keep in mind that many granting agencies require that institutions have these systems in place before they will fund preservation projects.

or roof leaks. Ideally tapes should be stored with the spine label facing up.

- **Be sure all items are in some sort of enclosure.** Plastic cases and containers are preferred. The enclosure will be the final point of defense before the carrier itself is damaged.
- **Strive for proper climate control.** Storing media in a climate-controlled environment will greatly increase the life expectancy of your media by reducing the risk of long-term damage, such as binder degradation, vinegar syndrome in acetate open reel tapes and film, and mold growth. This is not always possible, but at minimum don't store valuable materials in areas where climate fluctuates, such as attics. Try to keep them in an insulated environment, and avoid low lying storage areas such as basements, which are more flood prone.
- **Keep a complete inventory of all materials off-site.** In a disaster, it is likely that databases and other electronic records will be unavailable. Having a paper inventory with identifiers will help enormously with identification, prioritization, and recovery.
- **Reformat physical collections as soon as possible.** Obsolescence and physical disasters pose enormous threats to the longevity of audiovisual content. Reformating is a step toward protecting recordings from these threats. Reformatted copies are easier to geographically distribute and are also more widely accessible.
- **Back up and geographically separate digital collections.** Ensure there is at least one geographically separate copy of all digital content. Geographic separation means that the backup is far enough away that immediate risks to the primary copy are not threats to the second copy. Further is always better, but as far away as possible, at minimum in another building, is better than keeping backups in close proximity to primary copies.

COLLECTION PROFILING

Collection managers who have been through a disaster say that knowing your collection is the single most

important factor to successful recovery. Collections that are unfamiliar, unprocessed, or have no identification are almost impossible to prioritize for recovery. Having a broad overview of the collection will help identify how many items need to be stored in specific environmental conditions, will enable classification by vulnerability to water or fire, and will identify the number of items that may need to be treated by an external service provider in an emergency.

Start by completing a collection profile that documents the number and format of items in the collection. Begin with broad categories, and, when feasible, further categorize into specific audiovisual formats (analog video cassettes, digital video cassettes, film, open reel audio, etc.).

INVENTORY & LABELING

Subsequently, work on creating or updating collection records and labels.⁴⁵ Adequate identification will be critical in a disaster. Recovering audiovisual materials is an expensive and time-consuming process. If you are unable to identify badly damaged items, you may spend time cleaning or sending something to a lab that is not of high priority or that has an undamaged duplicate stored elsewhere, potentially neglecting the items that are of high value or are irreplaceable. Furthermore, labs will need information about the items they are recovering both in order to give accurate pricing and to perform the correct procedures for the carrier.

At minimum, item labels should include a title or brief description and an identifier. Identifiers should correspond to those in your inventory. Ideally, item labels would also include a total running time and record date.

The inventory can have more detail than the label and, in this way, acts as a cross reference. Include the format, date, associated collection, and description. For formats such as 1/4" open reel, it is helpful to note on the label whether the tape is full, half, or quarter track and the recording speed. Inventories should include the role/generation of the item as well: Is it unique? Is it a duplicate? Is there a master recording stored elsewhere? Is this just an access copy?

45 See "Chapter 2: Inventory and Assessment" for more information.

DEACCESSIONING

In a disaster scenario, materials tend to get knocked over, mixed together, and removed from their original storage locations. Sometimes, media are separated from their containers due to the force of water or impact, and labels can become smudged or detached altogether. As a result, it can be difficult to identify which items are which, and furthermore, which are the most important. In these situations, it becomes painfully clear that precious time may be spent recovering low- or no-value items at the expense of the highest priority materials. Deaccessioning can go a long way toward ensuring that only the important items receive the critical care they need in a recovery operation.

CARE AND HANDLING

Small disasters are often the result of mishandling. These disasters can be prevented by ensuring that staff and caretakers are well-trained in handling collections with care, are observant of problems, and are able to record accurate information about an item's condition.

To guarantee these types of problems are mitigated, staff working with collections should have the following minimal training:

- **Format identification:** All staff should be familiar with formats found in the collection. At a minimum, this means being able to distinguish different types of media (e.g., Betacam SP versus Digital Betacam, LP versus shellac, Mini DV versus DAT).
- **Handling:** This includes proper transport, inspection techniques, machine threading, rewinding, etc.
- **Condition reporting:** Up-to-date, accurate records about collection items are a critical resource in a disaster and will help with prioritizing recovery decisions.
- **Storage preparation:** Ensure that staff members rewind tapes and secure film and open reel tape ends before an item is returned to storage.

SECTION 2: DISASTER PLANNING

Often we think that the goal of disaster planning is the creation of a written disaster plan. In fact, the value of

disaster planning is in the process itself—training, familiarization, and practice—so that when a disaster does strike, all staff or collection stewards are prepared. Disaster plans in written form are incredibly valuable. However, these only become effective when collection caretakers participate in their creation and maintenance and are trained in their use. Disaster plans are excellent reference guides; they contain important contact information and emergency response procedures. Effective disaster planning, however, is an ongoing process, one that lives in the minds of those responsible for collections. For individual collectors and artists in particular, a written plan is not nearly as important as taking appropriate steps to be prepared in an emergency.

Disaster preparedness planning should focus on two key goals:

1. Knowing how and when to react to an emergency warning; i.e. ensuring that the appropriate policies and procedures are put in motion when such warnings are issued.
2. Enabling effective response in case of damage.

The key to disaster preparedness is for plans to be quick and efficient to implement in an emergency. This is best accomplished by testing, revising, and practicing the plan regularly. For example, an institution could designate a week every six months to be “disaster week.” During this time, a group of people who have designated responsibility for disaster preparedness planning will review the plan, revise it as needed, check and re-stock response supplies, train staff (especially new staff hired since the last “disaster week”), and conduct drills.

Preparedness Policies and Procedures

Not all emergencies are preceded by warnings, but when they are, preparedness plans should be put into action to help prevent an emergency from turning into a disaster. Time is of the essence. A well-researched and tested disaster preparedness plan will allow you to take swift action in the short window of time available.

Disaster preparedness actions are by nature very contextual. Hurricanes and typhoons, dust storms, flash floods, tornadoes, fires, and civil unrest warnings require very different preparation steps. Localized threats, arising in

scenarios when pipes have aged or electrical wiring is faulty and the building itself is the threat, require a different set of preparedness steps. Even when the threat is shared, the way it is addressed depends on the location of the building, the material of structures, and where and how collections are stored.

When creating a disaster plan for your organization, start by considering each line of defense for the collection. Think of these as layers as those of a Russian matryoshka, or nesting, doll: the outermost layer provides protection for the next, which protects the one inside of it, and so on. Start with the outermost layer, typically the building or the structures that surround the building. Whenever possible, collaborate with people who know the structure well, such as the building manager or owner. Research and talk to others in the region who have buildings of a similar structure and material to find out what types of precautions they take when emergency warnings are issued. Factor in structural vulnerabilities that may need to be addressed. For instance, are sandbags needed to reduce the chance of floodwaters from entering the building?

Beyond just the building, additional defensive lines may need to be strengthened. Depending on the anticipated risks, emergency procedures may include covering shelves and boxes with plastic sheeting, moving collections off the floor and onto higher shelves, moving collections from one room to another, or other procedures.

All individuals responsible for the collection should be involved in establishing preparedness procedures, led by identified members of the “disaster team.” The team should work together to determine when preparedness procedures should go into effect at the point that warnings are issued, bearing in mind the individuals’ personal needs. If people need time to secure their own homes or evacuate the area, this should be accommodated in the preparedness procedures.

Staff Training and Simulation

Once a disaster strikes, the availability of individuals to help with recovery will vary greatly depending on their personal circumstances. Therefore, as much as possible, preparedness should be inclusive of anyone who might be available to help in a disaster. A first step in getting staff

thinking about how to react in a disaster is to conduct training exercises. Simulation training is highly recommended as it provides an effective way for people to become familiar with preparedness actions and the process of disaster recovery. It also provides some sense of the confusion and anxiety that arise during disasters, while simultaneously surfacing ways that the constraints and urgency of a recovery scenario can influence disaster preparedness steps.

Two types of staff training should be conducted:

1. **Disaster preparedness drills:** These drills prepare staff to react when a disaster warning has been issued. The more rehearsed these are, the more efficiently they can be enacted when the time comes.
2. **Disaster recovery training:** This training places staff in a post-emergency situation and allows them to work through the organizational factors involved in a recovery operation. It also allows them to practice salvaging collection items.

Simulated disaster recovery training involves setting up a small, isolated “collection” of materials that have been affected by some type of disaster. The collection should include a mix of materials that represent items of different “value” and types, including paper artifacts alongside audiovisual formats. Since nearly all disasters involve water, this is a good damage agent to use during recovery simulations.

The experience should force participants to think through the disaster recovery process and inform the preparedness process by raising questions such as: Who is our insurance company? How do you prioritize when media items are unlabeled? Are these commercial LPs really valuable? How do you clean a ¼” open reel tape? The training should be guided by an expert but largely leave the operation of the recovery and the discovery of lessons-learned to the participants. Training activities include: ensuring the area is safe to enter, damage site survey and documentation, establishing roles and responsibilities, gathering and managing supplies, triage and prioritization, handling and workflow, documentation, and cleaning and drying. The outcomes of this type of training will likely reveal steps that can be factored into preparedness planning procedures.

Equipment and Supplies

In an area-wide emergency, supplies will be hard to come by, and at first you will likely only be able to work with what is immediately available. Even in the case of a small, localized disaster, such as a burst pipe, it will be much easier to protect collections from damage if basic supplies such as plastic sheeting are on hand. Keeping a stock of emergency supplies will go a long way toward effective recovery.

Use your knowledge of the area and the building (see Section 1 of this chapter) to know which supplies and equipment will be essential after a disaster. For example, if electricity is frequently interrupted, an uninterruptible power supply (UPS) or a small power generator might be a wise investment. If your area is earthquake prone, having emergency lighting or at minimum an adequate supply of flashlights and batteries will be critical.

Keep emergency supplies in a watertight plastic container and in an easily accessible place. This may be near building entrances or in your car. It doesn't hurt to have caches in multiple places. A list of basic supplies you might want to have on hand includes:

- A few gallons of distilled water
- Nitrile gloves (latex and powder free)
- N95 face masks
- Trash bags: large and small
- Tape: paper, masking, duct, and strong (e.g., Gorilla) tape if available
- Plastic sheeting (to cover doorway entrances in advance of flooding, to cover surfaces during recovery)
- Rolls of paper (to cover surfaces for drying)
- Paper towels (lots and lots)
- Towels
- Felt markers (Sharpies), pencils, pens
- Flat trays or bins (for moving items)
- Flashlights, lanterns, headlamps, and batteries
- Notepads and clipboards
- Microfiber towels or other lint free cloths
- Isopropyl alcohol
- Cotton buds (Q-tips)

- Buckets
- First aid kit

Prioritization and Identification

Think about which items in the collection will be a priority to evacuate or recover if they are damaged in a disaster. This is the point at which the appraisal process is very important (see Section 1). Make sure these items are well identified. Consider storing priority collections in a separate area that can be easily reached in the dark in case of a power outage. Organizations have used various approaches to ensuring these items can be quickly found and retrieved in an emergency, including locations identified on building diagrams and glow-in-the-dark stickers on shelves.

Disaster Plan Documentation

A disaster plan will be an invaluable resource in the event of an emergency. The written plan is a quick reference for telephone numbers and email addresses of staff and

FIGURE 5.2

A sample disaster supply kit.

Credit: Kenneth Spencer Research Library, University of Kansas
<https://blogs.lib.ku.edu/spencer/tag/disaster-kit>



external resources (e.g. recovery labs). It also contains information about preparedness steps for the building and collections, including floor plans, the location of emergency exits, shut off valves, electrical breakers and outlets, and priority materials, and the details of salvage procedures.

A written disaster plan does not need to be overly detailed and should not be too long. It should be a well-organized resource and reference in the event of an emergency—not a verbose book that must be read from cover to cover for effective response, salvage, and recovery. Sections of the plan should be tabbed so they can be easily located in a time of crisis. Lists, diagrams, and bulleted, bold text will be most useful in an emergency. Make sure the plan is clear and easy to use. Additionally, it should be available in print and electronic form.

A disaster plan should include:

- **Pre-disaster action steps and evacuation instructions.** Outline the steps to be taken to protect people and collections in the event of a hurricane, typhoon, forest fire, or other emergency for which warnings are issued. Detail evacuation procedures and the location of alarms. Make sure this information is easy to find and very simple.
- **Internal communication information.** Collect the following from all staff: all telephone numbers (home, mobile, partner's mobile, etc.), email addresses (including personal email in case work servers go down), Twitter and Facebook handles, etc. Don't assume that phone networks (landline and cellular) or the internet will be available and reliable in the aftermath of a major disaster. Some methods of communication may work while others don't, and this may change over the course of several days. As part of your planning, ensure that you have access to and are familiar with alternative options such as SMS text messaging and Twitter. You might consider creating a Facebook Group that can be used to keep everyone informed in times of emergency. This will enable updates to be sent out to all staff members simultaneously.
- **Service provider list and contact information.** Include insurance companies, labs, recovery experts, full recovery services, conservators, roofers, plumbers, electricians, transportation and storage services, rental facilities,

drying facilities, the police and fire station, and other local emergency management response agencies.

- **Partner institutions and professional networks.** As part of the disaster planning process, talk to sister institutions or other professionals in your area that may be able to help in an emergency. Come to an agreement about what services or assistance may be provided, such as storing evacuated priority collections. Be sure to offer reciprocal assistance.
- **Building floor plan.** Include the location of water shutoff valves, electrical switches, disaster equipment and supplies, and all collections.
- **Priority collections list and location.** Identify these items in the floorplan as well as in a separate inventory.
- **Response structure and job assignments.** Identify who is in charge of the response effort and include their contact information. Include a backup in case the first person is not available. Be sure anyone put in charge has the authority to spend funds. List names and contact information of staff members who have been trained for various aspects of disaster response and recovery and who are familiar with the disaster plan.
- **Basic salvage instructions.** Instructions should be relevant to the materials in the collection.

Most importantly, be sure the disaster plan is reviewed and updated regularly and that staff training is conducted periodically.

SECTION 3: FIRST RESPONSE STEPS

Preventative actions and preparedness steps will likely eliminate or reduce any damage to collections. However, sometimes even the most well-prepared collections can be affected. The goal of this section is to provide a basic overview of first steps to take to salvage valuable recordings. It does not go into detail about recovery or restoration. Recovery is a highly specialized task best left in the hands of experts. Even if you are not able to send tapes to a lab right away, taking proper steps to recover collections will buy time while funds are raised. Nonetheless, the earlier experts can be contacted, the better.

This section briefly walks through basic recovery steps, focusing on both human safety and the reduction of further risk to collections during the recovery process. Recovery itself is full of risk; the likelihood of mishandling, losing/dissociating materials from labels or cases, lack of documentation, and slow response leading to mold growth or other damage, is dramatically increased. Seeing your valuable collection lying under a pile of debris or submerged under water induces a sense of panic. Being well prepared will help alleviate permanent loss of content.

Recovery Tips

DON'T GIVE UP HOPE

In “Magnetic Tapes Can Survive Flood Exposure,” Peter Brothers notes that it is often assumed that water damaged tapes are ruined and unsalvageable. In fact, this is often not the case. Even tapes that have been submerged for extended periods of time have been recovered by experts. Brothers writes that no matter how bad they may look, “most wet tapes can now be saved and restored, if they are treated properly.”⁴⁶

CALL THE EXPERTS AND AUTHORITIES AS EARLY AS POSSIBLE

The disaster plan should contain contact information for authorities and experts, including insurance companies, disaster recovery services (for clearing water out of the building), labs and conservation professionals, local and federal disaster recovery agencies, etc. As soon as a damaged item is identified, contact these groups. They can help you determine what next steps to take, including whether or not you should attempt to begin recovery or instead wait for help.

SAFETY FIRST

Human safety should always trump the desire to get in and rescue valuable recordings. The first step in recovery must always be ensuring that the area is safe to enter. Live wires, contaminated standing water, and damaged structures can pose enormous risk to humans. Have the

building inspected by an authority or expert and cleared for entry before proceeding and handling media.

STOP AND/OR MINIMIZE DAMAGE

Do what you can to reduce risks to people and collections if the threat is ongoing. Shut off valves as well as electrical and climate systems. Cover collections with plastic sheeting if water or debris is falling. Move collections out of the hazardous space as quickly and safely as possible.

ACT QUICKLY BUT RESPONSIBLY

Disaster recovery literature for collecting institutions often names 72 hours as the time window during which materials must be rescued in order to be fully recovered. While salvaging media within this timeframe is ideal, it is not always possible. Entire areas of a building may be cordoned off for days or even weeks due to hazardous conditions. Even after you gain entry to the space, it might be several hours or days before a recovery plan can be put in place. It is important that salvage be conducted quickly but carefully at this stage. In some cases, more damage may be done the longer the media sits under a pile of rubble or under water, but in other situations this may not be the case. Mishandling and dissociation of media and containers are some of the biggest risks at this stage. Create a plan of action to avoid these threats.

DON'T ATTEMPT TO PLAY WET MEDIA

Do not, under any circumstances, attempt to play wet or contaminated media. This can lead to damage of both the media item itself as well as the playback equipment, which are both valuable resources. Wet or contaminated media will need to be cleaned in distilled water and dried before any content recovery can be attempted.

IDENTIFY DAMAGE AGENTS

It is important to identify the types of contaminants that may be affecting media items. If collections are submerged in water, attempt to identify what types of contaminants may be in the water: salt, chlorine, sewage, etc. This will help determine what recovery actions need to be taken.

46 Peter Brothers, “Magnetic Tapes Can Survive Flood Exposure,” <http://www.specsbros.com/disaster-recovery-magnetic-tapes-can-survive-flood-exposure.html>.

Items submerged in saltwater, for instance, must be cleaned in distilled water as soon as possible, as the salt is highly corrosive and will quickly damage any metal parts (rollers, layers in optical discs, etc.).

CAREFULLY DESIGN A RECOVERY OPERATION—DON'T JUST ACT

Before diving into salvage and recovery, develop a comprehensive plan. At this point, an awareness of the potential risks is key. The plan should factor in:

- **Space:** Identify a clean and well-ventilated space for cleaning and drying. Remember that drying space will have to accommodate media items as well as their cases, covers, labels, and inserts. Ensure that surface area is sufficient.
- **Supplies:** Make a list of your supply needs (such as gloves, cleaning supplies, flat surfaces for transport, paper towels, etc.), assemble what is available around you, and send someone out to find the rest. In an area-wide disaster, this can be particularly challenging due to regional demand for recovery supplies. Be creative, and start working on this as quickly as possible.
- **Roles and responsibilities:** Many institutions have reported that having “too many cooks in the kitchen” doesn’t work in an emergency situation. Identify at least one person who has the authority to spend funds. Quickly identify a coordinator who can establish needed roles and begin to fill these. These roles will change as the operation progresses: people who start by moving media to the recovery space may become responsible for cleaning or documentation later on. Necessary roles will likely include: coordination, documentation, cleaning, transport, security, and external communication. Some roles will require multiple people.
- **Documentation:** Documentation is perhaps the most important aspect of the recovery. It starts with documentation of the disaster area. Photographic as well as written documentation of the damage will be critical for insurance claims.⁴⁷ Ensure that damage to the building

and collections is thoroughly documented. Next, procedures should be well-documented and accessible to everyone participating in the operation. Finally, documentation must become ingrained in all aspects of the recovery: which cassette goes with which insert, what day and time drying started, and the names and contact information of the day’s volunteers. A lack of documentation, particularly the associations between media items and their cases/labels, is one of the biggest risks to successful recovery of audiovisual items.

- **Training & knowledge transfer:** Everyone who participates in the recovery should be trained in the specifics of the workflow and procedures. In a situation in which staff and volunteers will be coming and going according to their availability, ensure that procedural knowledge is passed between people and documented as they cycle through.

TRIAGE AND PRIORITIZE

Separate wet from dry items, and separate items by degree of damage. Attempt to identify the most valuable items and make these the first priority for salvage. The Association of Moving Image Archivists (AMIA) website provides a comprehensive list of triage steps.⁴⁸

FIGURE 5.3
An example of documentation: a U-matic tape and its case with matching and clearly-identifiable labels.

Credit: AVP.



47 Be sure you have approval from your insurance provider before attempting to move or salvage collections. In some cases, an adjuster may need to come survey the disaster site first.

48 See “Disaster Recovery—First Actions for Film, Tape, and Discs,” *amianet.org*, <https://amianet.org/wp-content/uploads/Resource-Disaster-Recover-First-Actions.pdf>

Media-Specific Salvage

This section offers strategies for salvaging damaged audiovisual items. It does not cover all imaginable types of damage, but instead focuses on the most common such as water, contaminants, and debris. The strategies are limited to formats commonly found in audiovisual collections.

These instructions are limited to salvage and stabilization only. They are not equivalent to recovery or restoration. These steps are intended to stop or slow down ongoing damage and buy you time before transferring contents off of the media item. Following cleaning and drying, media items may still need to be sent to a lab for full restoration and transfer.

ALL ITEMS

1. Remove media from containers, cases, or sleeves.
2. Remove wet inserts from cases.
3. Ensure that ALL pieces of the media item are labeled with a common identifier so that they can be brought back together after drying. Discard any containers that can easily be replaced (e.g. CD jewel cases). If containers have to be cleaned, be very careful not to smear or remove label information.
4. Cleaning should be done using distilled water. Tap water and even filtered water should not be used for cleaning as the mineral contents can be very harmful to the media.
5. All items that are cleaned in water should be left to dry for at least 48 hours before being placed back inside containers—longer if the relative humidity is high in the drying area or if the item is severely waterlogged.

OPTICAL DISCS (CD, DVD)

1. Read notes for All Items above.
2. Do not freeze optical media.
3. Contaminated or water-damaged discs should be rinsed in clean distilled water. Do not submerge discs that are not already wet or have not been compromised through debris or contamination.
4. Using a lint-free (e.g. microfiber) towel, dry the data side of the disc by wiping from the center out in a sun-ray motion.

5. If there is any residue remaining, clean using a Q-tip with a solution of 1/3 isopropyl and 2/3 distilled water.
6. Blot the label side; wiping may remove or smudge labels.
7. Ideally, dry discs in new, clean jewel cases with the data side down and the jewel case open like a book, standing upright. If jewel cases are not available, lay media flat on a clean, dry surface with the label side down.
8. If needed, label with a felt tip marker on the (usually clear) inner plastic ring of the disc.

ANALOG TAPE (VHS, U-MATIC, BETACAM, AUDIO CASSETTES, OPEN REEL AUDIO)

1. Read notes for All Items above.
2. Do not freeze magnetic tape.
3. Do not attempt to rewind wet or damaged tapes.
4. Even when a tape is fully submerged, it is likely that only the exposed parts have been compromised. Taking tapes apart, unwinding, or unspooling will in most cases do more damage. Do not take these actions unless advised by an expert.
5. Contaminated or water-damaged tapes should be rinsed in clean, distilled water. Carefully ensuring that the tape does not unspool in the water, submerge it briefly, giving the tape a slight shake. Dispel dirty water into a separate bin or bucket so as to not further contaminate the water.
6. Remove any residue on the outside of the cassette using a Q-tip with a solution of 1/3 isopropyl and 2/3 distilled water, taking care not to smudge or smear the label.
7. Lay upright to dry with the exposed portion of the tape facing up. For cassettes, prop open the lid and hold in place with a Q-tip.

DIGITAL TAPE (MINI DV, DVC PRO, DVCAM, DIGITAL BETACAM, DAT)

1. Read notes for All Items above.
2. Do not freeze magnetic tape.
3. Do not attempt to rewind wet or damaged tapes.
4. Do not submerge in water under any circumstances.

5. Taking tapes apart, unwinding, or unspooling will in most cases do more damage. Do not take these actions unless advised by an expert. Even when a tape is fully submerged, it is likely that only the exposed parts have been compromised.
6. Clean the outside of the cassette using a Q-tip with a solution of 1/3 isopropyl and 2/3 distilled water.
7. Lay upright to dry with the exposed portion of the tape facing up.

FILM (8MM, 16MM, 35MM)⁴⁹

1. Read notes for All Items above.
2. Do not unwind the film. The film may have stuck together, and unwinding could cause further damage.
3. Do not try to dry the film until you have been instructed by an expert on the best way to do so.
4. If the film is not wet, do not submerge it.
5. Rinse the film in clean water to remove debris.
6. If you have access to a freezer, place the film in a plastic bag, remove as much air as possible, and seal the bag. A supermarket bag will suffice for this.
7. If you don't have access to a freezer, place in a bucket of cool water. Change the water daily for up to two weeks until you can get the film to an expert

LACQUER DISCS

1. Read notes for All Items above.
2. Do not submerge in water under any circumstances.
3. If wet, dry off immediately, laying the disc on a clean, dry, flat surface and using a soft, non-shedding, non-abrasive cloth.
4. Avoid flexing the disc. Lacquer discs may have a glass base that can break. Flexing may also promote delamination if there are already issues with the disc.
5. If packing, place disc in a sleeve and pack with clean, flat cardboard spacers in between each disc. Pack vertically and snug, making sure there is no lateral movement but not so tight that it is stressing the disc.

SHELLAC DISCS

1. Read notes for All Items above.
2. Do not submerge in water.
3. Shellac can be cleaned in a solution of distilled water and a few drops of mild dishwashing detergent.
4. Using a microfiber or other lint-free cloth, wipe discs using a circular motion following the direction of the grooves.
5. Rinse in clean, distilled water.
6. Wipe again in a circular motion with a dry lint-free cloth.
7. Lay flat to dry.
8. Place in a clean sleeve.

VINYL DISCS

1. Read notes for All Items above.
2. Vinyl can be cleaned in a solution of distilled water and a few drops of mild dishwashing detergent.
3. Using a microfiber or other lint-free cloth, wipe discs using a circular motion following the direction of the grooves.
4. Rinse in clean, distilled water.
5. Wipe again in a circular motion with a dry lint-free cloth.
6. Lay flat to dry.
7. Place in a clean sleeve.

WAX CYLINDERS

1. Read notes for All Items above.
2. Do not submerge in water.
3. Gently dry with a non-linting, non-abrasive cloth. Too much pressure may crack the cylinder or alter the grooves of a soft wax cylinder.

49 Adapted from the National Film and Sound Archive of Australia, "First Aid for Water Damage." <https://www.nfsa.gov.au/first-aid-water-damage>

CONCLUSION

Small, localized emergencies and large-scale disasters continue with alarming frequency and impact as the climate changes. Additionally, the potential for mistakes, negligence, and crime is ever present. Disaster preparedness must be a concern for all individuals and institutions that create or collect content with long-term value. Though disasters may be relatively infrequent, their impact can be devastating, with the potential for total loss of material. Even the simplest response or recovery plan can be highly effective if it is practiced and understood by all stakeholders. Doing something is always better than doing nothing when it comes to emergency preparedness.

This chapter has outlined many emergency preparedness and response basics that have been tried, tested, and improved by those experienced with disaster response and recovery of audio and other media over the years. However, these are only general guidelines. Always remember that the cause, circumstances, and context of each disaster will vary greatly. Guides like this one provide a general playbook but can't answer every question. Again, it is critical to have contact details for emergency response agencies (e.g. FEMA), experts, insurance, and others, so that these groups can be contacted as early as possible and guide you through the dos and don'ts of your particular situation.

Finally, as has been noted in several places in this chapter, the digitization of valuable audiovisual content and proper management of digital collections are two of the most important disaster preparedness steps you can take. A well-managed digital archive with proper intellectual control, backup, and geographic separation will always fare better in a disaster than will unique analog materials. An added benefit is that digitization, already necessary for most analog formats today, will best position the collection for long-term preservation and improved access.

SELECTED BIBLIOGRAPHY

American Institute for Conservation, 2017. Field Guide to Emergency Response. Available for purchase from <http://store.conservation-us.org/site/index.php?app=ecom&ns=prodshow&ref=FAIC-2>

AMIAnet.org. Disaster Recovery – First Actions for Film, Tape, and Discs, <https://amianet.org/wp-content/uploads/Resource-Disaster-Recover-First-Actions.pdf>

AVPreserve, 2012. Disaster Response Information and Assistance. Available from <http://www.avpreserve.com/blog/disaster-response-information-assistance>. Includes links to audiovisual specific resources, support/assistance contacts

Brothers, Peter (n.d.). Magnetic Tapes Can Survive Flood Exposure. <http://www.specsbros.com/disaster-recovery-magnetic-tapes-can-survive-flood-exposure.html>

FEMA (n.d.). Heritage Emergency National Task Force website. <https://www.fema.gov/media-library/assets/documents/113297>. Links and downloads for emergency response funding, recovery procedures, guidelines, etc.

Metropolitan New York Library Association, 2012. Disaster Recovery Resources. Available as an archived webpage at <https://web.archive.org/web/20151022195348/http://metro.org/articles/disaster-recovery-resources>. Includes links to disaster assistance hotlines, resource pages, grants and funding.

National Film and Sound Archive of Australia (n.d.). First aid for fire damage. <http://www.nfsa.gov.au/preservation/care/first-aid-fire-damaged-audiovisual-materials>

National Film and Sound Archive of Australia (n.d.). First aid for water damage. <http://www.nfsa.gov.au/preservation/care/stabilising-audiovisual-after-floods>

GLOSSARY

Active management: The performance of consistent and ongoing digital preservation activities (e.g. fixity and validation) to ensure a digital file's continued access for as long as necessary.

Artifact: Anomalies during visual or aural representations of recordings.

Audit trail: The information associated with a digital file that tracks the transactional history of it from the point of capture or ingest to know whether it has been managed without change to the bits that make it up and according to relevant policies and standards.

Authenticity: The quality of being genuine and free from tampering and is typically inferred from internal and external evidence, including its physical characteristics, structure, content, and context.⁵⁰ Trustworthiness.

Back coat: Layer added to some magnetic tape to help support the magnetic recording layer. The back coat reduces tape friction, dissipates static charge, and reduces tape distortion.

Binder system: System through which magnetic particles are held by a binder to a substrate layer.

Bit rot: The corruption, loss, or decay of bits, the building blocks of digital files.

Carrier type: Refers to the physical carrier of the AV material. Examples of carrier type include reels and cassettes.

Checksums: Alphanumeric strings that reflect the uniqueness of every digital file.

Curation: The activities that are performed on a digital file throughout its lifecycle, including selection and appraisal, description, ongoing care and management, long-term access, and/or deaccessioning/disposal.

Degradation: The process in which the quality or integrity of an object is destroyed over time.

Delamination: In disc media, the process that causes layers to separate from the support base.

Digital preservation: The active management of digital content over time to ensure ongoing access.⁵¹ It is an integral part of curation (see definition above).

Digitization: The representation of an object, image, sound, moving image, or document by generating a series of numbers that describe a discrete set of its points.

File attendance: Ensuring that there are no missing or unexpectedly present files in a given location.

Fixity: File fixity refers to the property of a digital file being fixed, or unchanged. Fixity checking is the process of verifying that a digital object has not been altered or corrupted.⁵²

Governance: In the informational sense, governance is the set of structures, policies, procedures, processes, and controls implemented to manage information at an enterprise level, supporting an organization's immediate and future regulatory, legal, risk, environmental and operational requirements.⁵³

50 "Authenticity." Glossary of Archival and Records Terminology. Society of American Archivists. <http://www2.archivists.org/glossary/terms/a/authenticity>

51 "About." Digital Preservation. Library of Congress. <http://www.digitalpreservation.gov/about>

52 Wikipedia. File Fixity. https://en.wikipedia.org/wiki/File_Fixity. Accessed December 29, 2016.

53 Wikipedia. Information governance. https://en.wikipedia.org/wiki/Information_governance

Ingest: The process by which digital files and their associated metadata (called a Submission Information Package, or SIP) is deposited or submitted into a digital repository.

Latency: In computer networking, latency is the time interval between the request for information, such as a digital file, and the retrieval or display of that file to the user by the system.

Machine transport: Playback equipment.

Mandrel: A cylindrical rod placed through a cylinder and used to rotate it for playback.

Media type: AV materials are classified as audio, video, or film during the cataloging and inventory processes.

Metal evaporated tape process: Process in which magnetic particles are vaporized from a solid and deposited onto a substrate layer.

Migration: Converting from one format to another format considered to be of greater stability.

Obsolescence: The state of being which occurs when an object or practice is no longer wanted or used. Usually occurs when a new technology supersedes the old.

Preservation planning: A process by which the general and specific needs for the care of collections are determined, priorities are established, and resources for implementation are identified.

Refreshing: Copying information content from one storage media to the same storage media.⁵⁴

Reproduction method: Method in which a recorded signal is played back from a physical media object.

Risk management: The systematic control of losses or damages, including the analysis of threats, implementation of measures to minimize such risks, and implementing recovery programs.⁵⁵

RPM: Rotations per minute. Used to indicate recording speed for discs and cylinders.

Sidecar file: A file that is stored next to the AV file in the same directory.

Signal path: The route that an audio signal travels from source to output. This may be within a single device (CD to speaker within a stereo system) or within a workflow (original audio recording to reformatted digital file).

Slipping: Tape pack problem in which either single strands or groups of strands are misaligned and migrate to rest against the edge of the flange. May cause edge damage to the tape or film.

Splice: When two ends of a tape or film are joined together using specially formulated splicing tape.

Sticky shed syndrome: A condition resulting from the deterioration of the binder in magnetic tape that results in gummy residues on tape heads during playback.⁵⁶

Storage architecture: The computing and network infrastructure required to store digital files.

Storage capacity: The amount of data a storage device can hold, often measured in gigabytes (GB), terabytes (TB), and petabytes (PB).

Storage media: Devices on which data is stored. These include computer hard disks, optical disk drives, USB drives and other external hard drives, DVDs, and magnetic data storage tapes.

Stylus: A hard point following a groove in a phonograph record and transmitting the recorded sound for reproduction.

Substrate: The backing film needed to support the magnetic recording layer of a magnetic tape.

Tails out: A method for winding tape onto a reel where the end of the tape is on the outside.

54 Digital Preservation Coalition Digital Preservation Handbook Glossary <https://dpconline.org/handbook/glossary#R>

55 "Risk Management." Glossary of Archival and Records Terminology. Society of American Archivists. <https://www2.archivists.org/glossary/terms/r/risk-management>

56 "Sticky Shed Syndrome." Glossary of Archival and Records Terminology. Society of American Archivists. <https://www2.archivists.org/glossary/terms/s/sticky-shed-syndrome>

