Media Preservation and Digitization Principles

Mike Casey

Director of Technical Operations, Audio/Video
Media Digitization and Preservation Initiative
Indiana University

Version 1.0
3/17/2022

Acknowledgements

This document was reviewed by the practitioners listed below. Their experience and insights produced a number of ideas and suggestions that were incorporated into the text. The resulting work is deeper, richer, and more accurate because of their efforts. The author, however, remains responsible for any inadvertent additions, omissions, or misinterpretations.

Natalie Rose Cassaniti, Cultural Collections Officer, Parramatta Heritage Centre, City of Parramatta, New South Wales, Australia

Jon Dunn, Assistant Dean for Library Technologies at the Indiana University Libraries

Patrick Feaster, Media Preservation Specialist, Media Digitization and Preservation Initiative, Indiana University

Hannah Frost, Assistant Director, Digital Library Services, Stanford Libraries

Chris Lacinak, Founder and President, AVP weareavp.com

Will Prentice, Training & Dissemination Manager, Unlocking Our Sound Heritage project, British Library

Mike Shallcross, Digital Preservation Librarian at the Indiana University Libraries

Kimberly Tarr, Head, Media Preservation Unit, New York University Libraries

The idea to develop a set of media preservation principles was hatched in conversation with Chris Lacinak over Indian food in Oslo, Norway. Chris and I were attending the 2010 Joint Technical Symposium and our conversation ranged freely that evening, inspired by some of the conference presentations from earlier in the day. Thanks to Chris for working on this at its beginning and at its end.

Contents

1 Introduction

The time period in which preservation action for media holdings is both possible and feasible is short. This fundamental problem is particularly acute for analog and physical digital recordings, which are afflicted by what we might call environmental challenges. They are actively degrading, some catastrophically. They are subject to rapidly advancing obsolescence that results in the increasing scarcity of playback machines, parts, and expertise, among other issues. They receive inadequate resources for the preservation tasks needed. This environment of degradation, obsolescence, and insufficient resources is deadly, and it is widely thought that it will make the preservation of large holdings of audio and video recordings unaffordable in the not-so-distant future. These impediments are also not foreign to the class of recordings generally known as 'born digital.'

Despite the environmental challenges, media preservationists are often charged with safeguarding holdings of great historical, cultural, aesthetic, financial, community, and documentary value and significance. At the very least, the recordings they shepherd are significant enough to justify resources to save them for the future. The media preservationist has an obligation to the recording, its owner or custodian, future users, colleagues, the subject(s) of its content, and the public as a whole to make appropriate choices that support long-term preservation. Their work is one of the key contributors to the success or failure of the process by which the content on media recordings is transmitted through time.

2 Scope

To respond to the issues outlined above, a set of general principles is needed to guide the development and implementation of preservation strategies and policies so that efficient, accurate, sustainable, and enduring work is supported. This media preservation principles document is narrowly focused on the preservation of media (sometimes called audio-visual or AV) recordings, including the act of digitization that is, in many cases, central to their survival. The document does not address the larger and broader concerns fundamental to archiving media recordings that include a range of topics pertinent to archives, libraries, and other types of organizations, as well as private individuals, that have taken up this charge.[1] It is not a code of ethics that governs morality and defines acceptable conduct for an organization or an individual.[2] Rather, it is a statement of principles, with some associated policies and practices, in which the principles themselves are brought to the forefront for consideration.[3]

[1] For philosophies relating to archiving audiovisual recordings, the classic work is: Ray Edmondson, *Audiovisual Archiving Philosophy and Principles (Third Edition)* (Bangkok: UNESCO Bangkok, 2016). https://bangkok.unesco.org/content/audiovisual-archiving-philosophy-and-principles

[2] See, for example, the IASA Code of Ethics at https://www.iasa-web.org/code-ethics

[3] Another document that provides principles in this space is: IASA Technical Committee, The Safeguarding of the Audiovisual Heritage: Ethics, Principles and Preservation Strategy, Co-Edited by Will Prentice and Lars Gaustad.

This document begins with a statement of principles for media preservation programs. Intended to be interpreted broadly, 'programs' include operations in libraries and archives in many settings such as universities, broadcasting companies, and museums; the activities of private collectors, cultural organizations, and local history initiatives; and the protective efforts of corporations, sports teams, and religious organizations, among others. These programs may be formally established, or they may act on a more informal basis. What they have in common is a desire to safeguard audio and video holdings for future use.

The principles herein were initially developed for parts of Indiana University's (IU) Media Digitization and Preservation Initiative (MDPI) or its planning forerunner known as the Media Preservation Initiative. They originally addressed concerns and issues specific to the IU project but are expanded and generalized in this document to apply to a range of media preservation programs and media digitization operations.

In this document, the term 'media' refers specifically to audio and video recordings. While many, or even most, of the principles may be relevant for motion picture film holdings, they were not developed from the considerable film experience at IU.

Some of the principles include policies that in several cases, but not all, represent those enacted at IU for the audio and video parts of MDPI, which focused on the digitization of analog recordings and the transfer to digital files of physical digital recordings. Other policies were created in this document for the sake of illustration. This is far from a complete set of policies and there are any number of other areas in which an organization may need to establish policies to guide work. Some of the policies in this document may be appropriate for some entities engaged in preserving media while inappropriate for others. Policies, in contrast to principles, are not universal or unchanging and may differ from organization to organization.

Following the exploration of principles for media preservation programs is a look at principles for media digitization operations that are focused specifically on digitizing analog recordings and transferring physical digital recordings to digital files for long-term preservation. Digitization operations may be attached to organizational programs, private collectors or groups, or the services offered by a vendor. This section includes specific digitization practices derived from the set of media digitization principles and policies. Many of the practices explored in this section are in active use at IU for audio and video digitization. However, neither the collection of policies nor the practices discussed in the document should be considered comprehensive.

Version 4, 2017 (= Standards, Recommended Practices and Strategies, IASA-TC 03). International Association of Sound and Audiovisual Archives. www.iasa-web.org/tc03/ethics-principles-preservation-strategy

3 Definition of Terms

A **principle** is a fundamental or general truth, concept, or value that is a guide for behavior or evaluation. Principles guide and/or govern the development of specific policies or practices employed within the space that they represent. Principles are unchanging and universal in nature and exist independently of context. While principles may not provide direct answers to the specific problems and conundrums that arise in daily work, they help frame appropriate questions and provide perspective and a way of thinking for use in the decision-making process. They can be used as touchstones, returned to again and again for general, baseline guidance on what is considered true in our field.

A **first principle** is a foundational assumption that cannot be deduced from any other assumption. It represents the highest possible degree of generalization and is the most important reason for doing or believing something. Its truth is self-evident.

Principles by themselves are not enough to establish actions. They must give rise to policies that determine actions to be taken in particular situations. Policies are enacted using specific practices. This gives us what we might call the three P's of preservation planning and performance: principles, policies, and practices.

A **policy** is a set of ideas or statements that are used as a basis for making decisions about how a goal is to be reached. A policy may also be thought of as a definite course of action selected from alternatives. Policies are implemented by practices (also called procedures). Policies are often constructed in the form of "when faced with x, we do y." For example, when we receive moldy tapes, our policy is to isolate them, so they do not harm staff or contaminate other recordings.

A **practice** is a regular action or procedure that is performed in response to a specific situation. It is the detailed steps by which something is done and is often presented step-by-step. To expand on the above example, it may be the archive's practice to isolate moldy tapes by enclosing them in a tightly sealed box that is stored in a room dedicated for this purpose. This is the specific practice that enacts the above policy. Practices may vary by context and may be different from one organization to another, even though they are responding to the same policy or principle. More than one practice may be a defensible response to any given policy or principle.[4]

[4] Examples of best practices documents in our field include: IASA Technical Committee, *Guidelines on the Production and Preservation of Digital Audio Objects*, ed. by Kevin Bradley. Second edition 2009. (= Standards, Recommended Practices and Strategies, IASA-TC 04). www.iasa-web.org/tc04/audio-preservation and Mike Casey and Bruce Gordon, *Sound Directions: Best Practices for Audio Preservation*, 2007. Available at https://hdl.handle.net/2022/27339

It can be difficult to differentiate principles, policies, and practices. One tactic that may help is to observe your reaction to reading or encountering one of these. For example, reading a principle naturally elicits a response along the lines of "that's right," "that is true" (or not true, if you disagree with it) or "that makes sense."

In contrast, reading a policy typically evokes a response such as: "here's what we need to do," "we should do that" (or not do that if you disagree.) A practice may provoke a response that is more along the lines of "here is what we are specifically doing."

To aid reference and discussion, the policies and practices are numbered under the Principle to which they are attached. For example, the second policy under Principle 14 is numbered Policy 14.2. Practices are further numbered under the policy to which they are attached. For example, the second practice under the same policy is numbered Practice 14.2.2. (Principle 14/the second policy under Principle 14/the second practice under the second policy under Principle 14)

Below is a concrete example.[5]

Principle 7: Accuracy, Faithful Reproduction, and Integrity

The products of preservation work must be as accurate as possible, representing the source recordings faithfully and with the highest level of integrity.

Policy 7.1: An azimuth adjustment is performed for all open reel audio tapes and all audio cassettes prior to digitization in order to capture as much of the high frequency information on the tape as possible.

Practice 7.1.1: The digitization engineer takes the appropriate azimuth tool, inserts it into the slot, and turns it in one direction and then the other, all the while listening for maximum high frequency content. This procedure is done by listening. The setting that provides maximum high frequency content is used for playback.

Additional Terms

For the sake of clarity, below are definitions of additional key terms used in this document.

[5] Will Prentice suggests another way to understand the differences: "When a new staff member begins work and is unfamiliar with the context of the role, we might invoke *practices*, that is, simply tell them what to do. When that person becomes more familiar with their role, they may begin to ask the reason for certain practices, in which case we invoke *policies* to explain why we do things in a certain way. Inevitably, however, they will encounter a situation where even the policies seem insufficient or ambiguous, perhaps when encountering edge cases or exceptions. In this situation we invoke *principles*, in order to ensure the consistency of any policies or practices that result."

Media—this word has several meanings. For our purposes, media refers to physically stored content, specifically the materials that hold data in any form.[6] Audio and video recordings are the media forms, types, or formats relevant to this document.

Media holder—any entity that possesses audio and video and is engaged in trying to preserve it—archives, libraries, universities, broadcasters, private collectors, artists, corporations, production facilities, etc.

Time-based—something (a recording, for the purposes of this document) that has duration and unfolds or 'plays' over time. Time-based media include audio, video, and motion picture film.

Content—the target for preservation. Sometimes called essence, content is the material that is recorded on the physical media or into a digital file.

Carrier—typically refers to the physical media object that contains the content. For example, an open reel audio tape or a VHS videotape.

Accuracy—this term is used in several of the principles. The British Dictionary provides the following definition, which aligns closely with the intended meaning in this document: faithful measurement or representation of the truth; correctness; precision.[7]

Faithful—this term is also used in several of the principles and is a close synonym to accuracy. Faithful may be defined as adhering or true to fact, a standard, or an original.[8]

It is also necessary to understand the various types of audio and video recordings which are popularly classified as analog, physical digital, or born digital.

Analog recording—the recorded signal is a continuous physical representation of the original sound and/or light waves. Analog formats include audio cassettes, VHS, LPs, U-matic video, and others.

Digital recording—represents the original sound or light waves as a series of discrete pulses that may be interpreted as binary numbers (0's and 1's). The term physical digital refers to recordings in which the audio and/or video content is tied to the physical format and carrier. For these recordings, the digital data on the physical carrier must be transferred to digital files for long-term preservation. Examples include Digital Audio Tape (DAT) and Digital Betacam (DigiBeta), which is a video format.

[6] See, for example, https://sites.google.com/site/multimediadreamwiki/1--introduction or https://www.yourdictionary.com/media
[7] https://www.dictionary.com/browse/accuracy
[8] See https://www.dictionary.com/browse/faithful

Born digital recording—is often used to identify recordings that originate as digital data in a digital file. Note, however, that physical digital recordings may also (or may not) be 'born digital' although they are not digital files.[9]

There is a more precise way to identify and define the various types of recordings by focusing on the signals/data that are stored on the media and how they may be read and captured.[10] That gives us the following categories:

- **Analog baseband**--signals derived from analog formats via Composite, S-Video, Component outputs, and/or analog audio outputs.
- **Digital baseband**--signals derived from digital formats via SDI and HD-SDI. DigiBeta is an example.
- **Digital native**--data recorded onto media as DV or MPEG or PCM based encodings, where the DV, MPEG, or PCM encoding can be extracted in its native form. MiniDV and audio CD are examples.
- **Digital file-based**—data recorded onto media as a file consisting of a wrapper and codec that can be retrieved from the recording media in its native form.

In this document, the more popular terms are used to make things simpler and more understandable. Therefore, analog corresponds to analog baseband, physical digital is digital baseband or digital native, and born digital refers to digital file-based.

[9] The physical digital recording itself (a DAT, for example) is born *digital* but the recorded content may be a copy of an analog recording, so the underlying content is not *born* digital.
[10] This content was provided by Chris Lacinak.

4 First Principles

First Principle A. Inherent Value

Time-based media content has inherent value as a primary information source for future uses.

In much the same way as books, manuscripts, or photographs, time-based media may hold information that is defined as important for future use. Individuals, organizations, and societies have produced or collected large quantities of audio and video recordings that are considered valuable to both themselves and others. These recordings may, for example, have the potential to contribute to research, support instruction, or heighten aesthetic experience. They may have business or financial value and may be for sale. They may be defined as carrying historical or cultural value. There are numerous ways in which value is assigned or attached to recordings by human actors interested in their future use. It is beyond the scope of this document to analyze the process of conferring value or the specific value judgments made. Instead, it reminds us that time-based media recordings join other forms, formats, or information sources in providing data that may be of value into the future.

First Principle B. Protecting value

Recordings that are considered valuable must be protected from loss.

The central task for media preservationists is to take steps to protect or preserve time-based media recordings that are defined as valuable so that they may be used in the future.

First Principle C. Separation of Content from the Carrier

The content captured on a media recording via the recording process, not the carrier or the physical object itself, is the most valuable target for preservation.

Virtually no audio or video file formats or physical carriers are expected to remain playable in the foreseeable future due to degradation and obsolescence issues. The only way to preserve their content over the long term is to transfer to digital files while it is still feasible to do so. For many reasons, digitization is the methodology used to preserve content from analog audio and video carriers. Some organizations may also prioritize the survival of the original carrier (if the content is tied to a physical object) if it possesses artifactual value. However, in most cases the survival of the content necessarily takes precedence.

5 Program-Level Guiding Media Preservation Principles

Below is a set of general guiding principles for media preservation programs as they develop preservation strategies and policies, specify infrastructure, and undertake planning activities. These are adapted, revised, and expanded from the Indiana University publication *Meeting the Challenge of Media Preservation: Strategies and Solutions*.[11]

Each section below is arranged in the following manner: title of the principle, statement of the principle, discussion, policies emerging from the principle, practices emerging from a policy, and specific examples. These sections are labeled in Principle 2, below, as an example.

Principle 1: Taking Action

Active degradation and the rapidly advancing obsolescence of audio and video recordings require immediate, and ongoing, preservation action.

Audio and video content may be lost if action is not taken now. In most cases, doing something with the resources available is better than doing nothing, understanding that compromises should never be greater than necessary and must be well understood. That is, determine what represents the best that you can do at this time and ensure that all work meets that standard. Document areas that can be upgraded in the future with the application of additional resources.

Note that perfection may be virtually impossible to achieve. For example, equipment—including playback machines, analog-to-digital converters, preamps, and others—is designed and manufactured to attain different levels of quality and has different price points. An organization, vendor, or a private collector may not be able to afford the most expensive (and, ostensibly, the most accurate) equipment on the market. This does not necessarily mean that a defensible digitization workflow cannot be developed. In addition, playback machines for analog and physical digital recordings are typically old and obsolete. They may be difficult to obtain and repair and may no longer function at their highest levels. Choices may be limited.

Finally, those engaged in preserving audio and video may not have the resources to collect and/or generate as much metadata as they would like. Again, this does not necessarily preclude undertaking solid preservation work.

The goal is not perfection, which is unachievable, but making informed and justifiable choices. It may be useful for institutions to draw upon in-house or outside technical expertise to come to an understanding of the strengths and weaknesses of their choices. This will help in assessing whether a defensible workflow that meets its needs and requirements is possible.

[11] This document is available at https://hdl.handle.net/2022/14135

[title] **Principle 2: Long Time Horizon**

[statement] Media preservation requires a commitment to the long-term.

[discussion] Although difficult and inherently inexact, it is necessary to anticipate the impact of our decisions not just five or ten years from now, but thirty to fifty years into the future. Even better would be anticipating the consequences of a decision in 200 or 300 years, although it is hardly possible to accurately envision the circumstances and context that will surround media preservation that far into the future. This does not mean developing a detailed preservation strategy that covers 300 years, or even 30 years. Rather, it involves researching longer-term implications as part of any decision-making process in areas concerning preservation. It also supports prioritizing favorable long-term outcomes over favorable short-term outcomes when they are in direct competition. Finally, this principle suggests that certain long-term financial and staff resources will be necessary for the digital preservation of selected content.

In any case, long-term preservation is not a one-time endeavor, but an ongoing set of strategies actively applied throughout a preservation system over a very long period of time.

Policy 2.1: All preservation-related decisions will consider not just the short- and medium-term, but also long-term implications and consequences. Preservation decisions will support strategies and infrastructure that make possible long-term preservation.

[example] Example: Using a lossy compressed format such as MPEG2 that conforms to the IMX/D10 specification for video preservation master files may cost less to store and provide short-term convenience. However, this format may not have the resolution required for demanding use cases. It is also not widely adopted for preservation and likely to become obsolete in the next decade or two. In this case, an uncompressed or mathematically lossless video preservation master file would better support the preservation of content over a long period of time. This is an example of prioritizing favorable long-term outcomes over favorable short-term outcomes.

Principle 3: Timeliness

Media preservation requires timely intervention.

Timely may be defined as occurring at a suitable or appropriate time, happening at the best possible moment, or well-timed.[12] A 'timely' intervention (such as digitization) is one that occurs at the most appropriate, or the chosen, time for the object of the intervention, whether that is sooner or later. Holding off an intervention until some point in the future may be just as timely as taking action now, depending on the goals and priorities defined by the preserving organization.

[12] See Dictionary.com https://www.dictionary.com/browse/timely?s=t and the Cambridge dictionary https://dictionary.cambridge.org/us/dictionary/english/timely

It is also true that the opportunity to preserve some media objects that are highly threatened due to active degradation, obsolescence, or poor storage conditions may be lost if preservation is not undertaken in the near-term. A prompt intervention inspired by these threats may well be timely, but whether or when it is actually carried out depends on priority, the principle discussed next.[13]

Principle 4: Priority

Media preservation actions are taken in order of priority.

If the principle of timeliness suggests when an action (such as digitization) should take place with a group of recordings, then priority provides the criteria by which the *specific* time to take action is calculated. It is widely acknowledged that resources for preservation are scarce. Many believe that not everything will be preserved in time due to lack of resources, ongoing degradation, and advancing obsolescence. If this is true, it is incumbent upon practitioners to select recordings that represent the highest priority to their organization for preservation treatment.

Data for use in determining priority may come first from an assessment of value, as it is obviously not wise to expend resources on recordings with no or low value to the holder. Further data may be generated by an evaluation of condition that includes an identification of preservation problems, as well as an analysis of obsolescence issues. Together, these factor into a calculation of the level of threat to the viability of a recording or collection of recordings. Curators, archive managers, and other experts may use their in-depth knowledge and experience to bring other aspects into this mix for consideration.

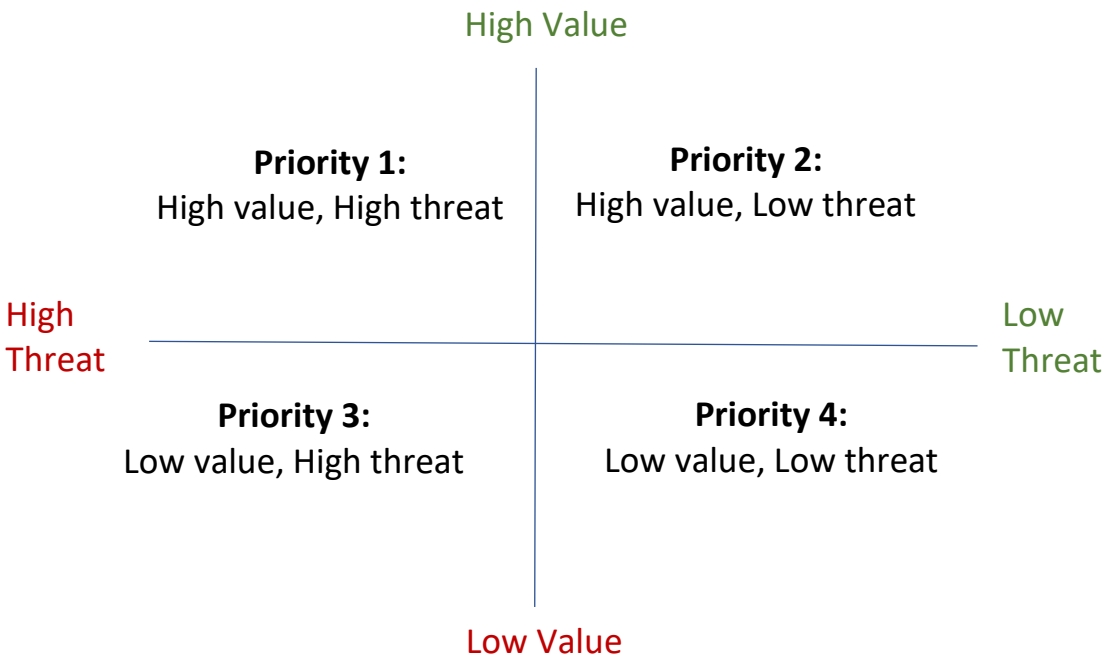
Policy 4.1: Digitization prioritization decisions will take into account the need for timely intervention by using an analysis of value, degradation, obsolescence, and storage conditions for the format or collection in question to determine their priority.

Practice 4.1.1: Staff use the criteria outlined in the MediaSCORE and MediaRIVERS[14] prioritization applications to generate data and develop an analysis of value, degradation, obsolescence, and storage conditions for use in prioritization decisions.

Example: Parts of the prioritization process may be conceptualized as an exercise in risk management, with risk defined as the probability of a loss. The loss that we are concerned with is, of course, that high-value audio and video content will not be preserved in time. One of the classic risk management responses is to find ways to reduce risk. Within the prioritization endeavor, risk may be reduced by placing recordings higher on the prioritization list, which increases the odds that preservation action will be taken before loss occurs.

[13] Ideas that led to the linking of the principles on Timeliness and Priority came from Mike Shallcross and Chris Lacinak.
[14] See https://github.com/IUMDPI/MediaSCORE

Here is one way in which recordings may be categorized after assimilating the data from the various analyses described above. In this example, value trumps threat.[15] [16]



[15] https://www.weareavp.com/wp-content/uploads/2015/03/IU_mscore_mrivers_guide.pdf

[16] The idea of treating the prioritization process as an exercise in risk management and the original version of the quadrant diagram were developed by Chris Lacinak.

This is a useful high-level, relatively simple, categorization scheme. In real life, it becomes much more difficult to sort what is typically a large number of recordings that fall into a medium value or medium threat range in order to construct meaningful priorities.

Consider, for example, two collections in the Priority 2 and Priority 3 quadrants. A collection of lacquer discs that are beginning to delaminate are identified as a very high threat because the format is actively degrading and content will soon be lost. If this collection is rated low in value, it would fall into the Priority 3 category. Compare that with a collection of open reel audiotapes that is considered high value, but lands in Priority 2 because the threat is low. Do you digitize the tapes first, understanding that content from the lacquers may be lost? Or do you digitize the lacquers first to safeguard the content and let the higher value collection wait, understanding that there are often no guarantees of future preservation funding? There is no right answer to this dilemma. Ultimately, the decision will rely upon the judgment of the stewards of these collections, based on the data collected for the above-mentioned analyses.

Principle 5: Primacy of the Unique and Original

Unique and/or original items receive highest priority.

A unique item is the only version that exists. A unique recording is often, but not necessarily, an original. Sometimes original recordings are lost or un-discoverable and a later generation copy is thus unique. Originals are the items created by the recording device and represent the first version made.

Original items are more highly valued than copies for technical reasons as they typically have advantages in fidelity. Copies made in the analog domain suffer from generation loss, which is the result of noise and bandwidth issues in the analog equipment used to make the copy. Unique recordings are more highly valued for research purposes than items also held by other organizations as they enable the holding institution to provide access that cannot be provided elsewhere. This raises the profile and reputation of the holding institution.

However, just because a recording is unique does not necessarily mean that it is valuable. The content matters, as does the interest of the end user.[17] For example, a unique recording of fourth graders reading reports on the U.S. Civil War may not be considered highly valuable to most people (although it may be valuable to the students or their families.) However, a unique recording of the leading scholars of the day giving presentations on the Civil War may be considered of very high value to an archive that specializes in this topic.

There are complications. It may be more costly to expend resources researching the provenance of a difficult recording or collection then to simply digitize it. Also, for items held in broadcast collections, an edited master may be a second-generation tape, but it may also be

[17] This paragraph developed from comments by Will Prentice.

the only version that contains the specific content that was aired. It is also possible that the original tapes used to construct the edited master no longer exist as original recordings were often re-used for other recording sessions.

The above analysis of unique and original recordings is most clearly drawn when applied to media holdings that are bound to a physical carrier, particularly those that were recorded in the field. This analysis breaks down when born digital file-based recordings are brought into the mix. What can be said about a recording that was created as a digital file and can be easily copied with no loss of fidelity? What is the original? What is unique? Within this paradigm, these distinctions lose meaning. Archivists have therefore expanded their notions of unique and original, especially as they inform value, beyond the physical object. For example, uniqueness may be more usefully found in the content on the recording rather than the recording itself. It may lie more easily in the aggregations of recordings or in the processes that produce recordings. The concept of the original may no longer fit content that is recorded initially as digital files. Its usefulness may be severely limited when applied to these types of recordings.[18]

Policy 5.1: When necessary and feasible, research the provenance of a recording to determine if it is unique and/or original.

Principle 6: Digitize or Transfer Once

Due to time and resource constraints, and the very large number of recordings in need of preservation, it is highly desirable to digitize analog recordings or transfer physical digital recordings to digital files just once.

The cost of digitization, the large number of media objects that require preservation intervention, and degradation issues may preclude returning to source recordings a second time for preservation transfer. Although media holders sometimes find it necessary to digitize for access purposes only, these workflows often produce lower resolution copies that may not support all future use cases. In this situation, it may become necessary, although highly undesirable and perhaps impossible, to undertake digitization a second time to produce suitable digital objects for both preservation and long-term access.

Policy 6.1: Preservation decisions for analog recordings will support a “digitize once” philosophy that strives to make the current effort the last playback of the source recording.

[18] The ideas in this paragraph are adapted from James M. O’Toole, *On the Idea of Uniqueness*, American Archivist, Volume 57, Fall 1994.

Principle 7: Accuracy, Faithful Reproduction, and Integrity

The products of preservation work must be as accurate as possible, representing the source recordings faithfully and with the highest level of integrity.

The value of media recordings is often dependent—at least in part—on integrity, accuracy, and quality. For example, a broadcaster may have specific fidelity or quality requirements that must be met before a recording is included in a production for broadcast. If playback during digitization is not handled accurately, it may result in a lower quality representation that is not suitable for broadcast. As another example, a news reporter may need to trust that the content on the source recording was captured in its entirety—every last second—when drawing conclusions about an event that was recorded. Finally, a music researcher may need to rely on the integrity of the digitization engineer's work to accurately analyze musical variables such as dynamic range and timbre, which can easily be altered by the use of inappropriate digitization procedures.

While the term 'accuracy' is regularly used to evaluate the act of digitization itself, it also applies to other parts of the digitization endeavor such as pre-digitization preparation or metadata generation and collection. Preparation of an open reel audio tape or an audiocassette, for example, includes an analysis of variables such as track configuration and noise reduction, which must be accurate for faithful reproduction. Also, collecting inaccurate metadata can lead a user to make faulty judgements about the recorded performance. A classic example is misinterpreting a crack in a cylinder recording as a drum. Metadata from the digitization engineer identifying and describing the crack would prevent this mistake.

Policy 7.1: Preservation and access decisions will consider the implications of concepts such as faithful reproduction, accuracy, integrity, and completeness.

Policy 7.2: Digitization workflows will be chosen and/or designed to produce digital objects that represent source recordings as faithfully, accurately, and with as much integrity as possible for use into the future.

Example: For a digital file to accurately represent the signal on a source open reel tape, the azimuth of the tape machine playback head must be aligned to the orientation of the signal on the tape. This pre-digitization adjustment enables the capture of as much high frequency content as is available from the source tape. If this adjustment is not done, and if the signal on the tape remains misaligned with the playback head, the digital file will contain less high frequency information than is potentially available from the source tape. This results in a less accurate, lower integrity representation of the source recording. It has not been faithfully reproduced. Note that performing this adjustment purposely misaligns the playback head. In its perfectly aligned state, it cannot provide the highest integrity representation of the signal on the tape, unless that signal is also aligned perfectly.

Example: It is sometimes stated that a recording is best reproduced using equipment from the time period in which it was created. A variant of this is the idea that the best results come from playing the recording on the same brand and model with which it was originally recorded. This is almost never true. More accurate results can be obtained from using modern, high-quality equipment that is able to extract more of the available signal. Older equipment was often able to record more information than it could extract. Using higher quality machines results in a representation of the source recording that has greater integrity.

Principle 8: Standards and Best Practices

Standards and best practices help media preservation programs ensure that preservation work is high quality, sustainable, interoperable, accurate, and consistent.

Following standards and best practices greatly decreases risk over the long term. Aligning with others using the same software, hardware, procedures, or technologies creates the opportunity to pool collective knowledge and leverage existing resources. With more people doing the same thing, there is a greater body of experience to draw from and more tools available to carry out the work efficiently and effectively. Problems are more likely to be resolved and the cost to any given organization lower. When a path to the next best thing is desired or needed, it is more likely to be developed for a large number of organizations using the same standard than for small groups each pursuing different procedures.

In some cases, standards or best practices may not represent the best choice for a specific project. For example, an organization may not have the resources to implement a metadata standard that is particularly complex or detailed. Some deviations from standards and best practices are intentional due to disagreements with standard practice or attempts to pioneer new practices. Other deviations are considered temporary, used until the organization can obtain the funding, expertise, equipment or whatever is needed to meet a standard or best practice. Non-standard choices, or partial implementations of a standard or best practice, are documented for future understanding by persons with the appropriate authority and expertise.[19]

Policy 8.1: Preservation decisions, services, workflows, and procedures will adhere to the greatest extent possible to internationally recognized standards and best practices in areas where they exist.

Policy 8.2: Written documentation of the choices made along with appropriate reasoning is provided when preservation decisions, services, workflows, and procedures deviate from, or do not make use of, standards and best practices. Any changes to policies or practices are documented in a change log.

[19] The points in this paragraph come from Hannah Frost and Chris Lacinak.

Principle 9: Preservation and Access

Preservation and access are interdependent and equally important for media collections with value for future use.

Both preservation and access are central to the mission of entities that seek to preserve media content. It is often stated that an item is not truly preserved if it is not accessible. While generally true, in any given situation one or the other may take precedence as their relationship is not always one of equal opposites.[20] Failing to optimize access will generally *not* harm preservation but failing to optimize preservation *will* harm access.[21]

For example, a researcher wishes to hear the content of a collection of lacquer discs. Because this format is already rapidly degrading, playback will only hasten its deterioration. Therefore, playback requests are withheld until a preservation-quality copy can be made. In this situation, preservation takes precedence over the demands of access.

Access can also drive preservation decisions. It is possible to transfer a media object to a format that makes long-term access more difficult. For example, digitizing an analog videotape to a lossy compressed format may result in compromised access. Managing this content over time will require decoding from one lossy codec and encoding into another multiple times. This may result in the accumulation of compression artifacts. In addition, available resolution from a lossy compressed format may not be enough to support the needs of some end users. In this situation, the need to guarantee the highest quality access over a long period of time may force a change in the digitization plan.

Policy 9.1: Access to media objects is provided in such a way that long-term preservation is not endangered or compromised. Whenever possible, preserving a media object utilizes procedures that will not compromise future access and usability.

Principle 10: Knowledge and Expertise

Successful media preservation requires knowledge and expertise from a range of disciplines.

Media preservation ultimately relies on expertise brought to the table by a number of disciplines. This includes digital preservation, which may be defined as "all of the processes and activities required to maintain access to digital information for as long as it is needed."[22] While it is beyond the scope of this document to detail digital preservation principles and best practices, it is important to understand that digital preservation best practices in areas such as

[20] Here, opposite is defined as situated in corresponding positions with relation to an intervening line, space, or thing. See dictionary.com for this definition at https://www.dictionary.com/browse/opposite?s=t
[21] The ideas in this paragraph come from Will Prentice.
[22] See the Yale University website at https://guides.library.yale.edu/digitalpreservation

long-term storage must be engaged for the fruits of our media preservation labors, including digitization, to be sustainable over time.

Organizations will have different ways of acquiring and making use of the knowledge and expertise necessary for media preservation. For example, some may hire staff with the required expertise while others may make use of consultants or service providers in specific areas. In very small operations, a staff member may develop or acquire expertise in more than one area to create a successful program.

Example: To be fully functional, the preservation system at Indiana University for audio and video needs to draw upon expertise in audio and video engineering, digital preservation, metadata librarianship, subject matter curation, library cataloging, copyright and intellectual property issues, storage of digital files for long-term preservation, network administration, server administration, software development, and preservation administration, among others. These areas of expertise must not only be available but effectively integrated across the entire preservation ecosystem for preservation to be successful.

Principle 11: Efficiency

Due to time and resource constraints, media preservation actions must be delivered as efficiently as possible.

It is imperative to develop highly efficient workflows to enable successful preservation of audio and video holdings in time and within available resources. This is particularly true in light of the large number of analog recordings worldwide that require digitization, the limited resources available to nearly all media holders, and the limited window of opportunity available for digitization.

An efficient approach may include the use of parallel transfer (digitization) workflows, where one operator digitizes multiple recordings at the same time. Like 1:1 workflows (one engineer digitizing one recording at a time), parallel transfer operations must engage international standards and best practices. Digitizing organizations may also wish to consider the overall value of specific holdings as determined in the prioritization process. Especially for content requiring less intervention or considered of lower value, high efficiency workflows may be the most appropriate choice.

The process of assessing efficiency may also extend to an examination of workflows for automation and batch procedures, particularly post-digitization. Efficiency can further be addressed in quality control workflows, description workflows including the repurposing of metadata, and others.

Policy 11.1: All analog recordings are evaluated for their suitability for high-efficiency digitization workflows. This evaluation will consider issues around fragility, presence of

preservation problems, uniformity of source materials, and unusual technical characteristics in making this determination.

Policy 11.2: Parts of the workflow that precede or follow digitization such as preparation of recordings, metadata collection, quality control, creation of deliverables, and others are evaluated for efficiency.

Example: The IU Music Library holds some 38,000 open reel tape recordings of student and faculty recitals and concerts dating from the 1950s. While curatorial staff tell us that there are recordings of a number of prominent classical and jazz musicians interspersed within this collection, the majority of items are judged to be of moderate value. That is, they are valuable enough to justify digitization and long-term preservation but are not considered highly valuable. In addition, some of the most valuable tapes were digitized as part of an earlier project. Digitizing this collection using highly efficient parallel transfer workflows may entail a greater risk than using 1:1 workflows, but IU is willing to accept this risk rather than incurring the much higher cost of using less efficient 1:1 workflows. Calculations indicate that it would take 27.86 years to digitize this collection using a single 1:1 workflow versus 7.43 years for one parallel transfer workflow that digitizes four recordings simultaneously. This is one situation where a principle that addresses efficiency may be successfully invoked to help make this decision.

Principle 12: Redundancy

Managed multiple copies decrease the risk of loss by lessening the dependency on any single copy. [23]

All analog and physical digital media carriers are endangered by either degradation or format obsolescence or both. In addition, the technology necessary to play back these carriers is potentially a source of wear, if not damage. For these reasons, media archiving has always relied upon reformatting as a basic preservation and access strategy to mitigate the risk of loss of content. This was true in the analog domain where for many years backup analog copies were produced for researcher access and for preservation. It is also true in the digital domain where both the act of digitizing analog recordings and the incorporation of 'born digital' recordings may result in multiple copies of preservation master files and digital derivatives to serve both preservation and access. In accordance with digital preservation best practices,[24] in order to best mitigate the risk of intentional or unintentional loss, multiple copies of digital files

[23] The term 'managed,' as used here, refers to a process by which copies are tracked and evaluated over time. Multiple copies by themselves are not useful unless you know how many you have, where they are, and what condition they are in. The addition of these ideas was suggested by Hannah Frost.

[24] Digital preservation is defined above as part of the principle on knowledge and expertise (Principle 10). There are a number of good documents that address digital preservation best practices. For starters, see Digital Preservation Coalition. (2015). *Digital Preservation Handbook* (2nd ed.). https://www.dpconline.org/handbook and Bodleian Libraries, "Introduction to Digital Preservation" (2020) https://libguides.bodleian.ox.ac.uk/digitalpreservation/home

are needed. These are best stored across different storage environments, IT management structures, and geographic regions.

6 Media Digitization Principles

The following principles, along with the practices outlined below, are specific to the work of digitizing analog audio and video recordings. This digitization may be undertaken by vendors and/or by organizations and individuals digitizing their own holdings. Whereas the principles up to this point are written for any organization or individual engaged in a formal or informal program to preserve media recordings, the ones that follow are intended for those actually preparing and playing recordings and converting them to digital files. Note that the digitization function as defined here includes preparatory work such as cleaning or repairing recordings. Note as well that the digitization function is considered to include staff members of an organization who are managing digitization done by a vendor. These managers must participate in the technical decisions made by the vendor that may be governed by the principles below as well as the needs and desires of the organization that holds the recordings.

The list of digitization principles that follows is not (and could hardly be) considered exhaustive.

Principle 13: Beneficial and Harmful Results

Preservation is best served by weighing the potential benefits of an action against the risk of harm.

Decisions that invoke this principle are best made, if possible, by a group rather than an individual. This minimizes the risk of relying on a single point of failure or on personal idiosyncrasy and increases the odds that sound judgement will be applied to the issue at hand.[25]

The statement, "First, do no harm," is often quoted by media preservation practitioners as a general principle to follow in working with media recordings.[26] However, all interventions with media recordings carry at least some risk of harm, and this risk may be necessary to meet the goals/principles of faithful reproduction and integrity. This may be justified since the target of preservation is usually the essence/content and not the media object that carries the content. This is explored under First Principle C at the beginning of this document.

It is tempting to cite principles used by document and object conservators to guide interventions for the sake of preservation. In the conservation field, objects are typically not treated if there is a risk of harm. In this case, 'preventative conservation' would be the appropriate solution – stabilizing the environment in which the object is housed. The last few

[25] Comments from Will Prentice led to the development of this paragraph.

[26] Typically associated with the medical profession, it is actually not part of the Hippocratic Oath as is widely thought. See http://www.lynnwebstermd.com/first-do-no-harm/ and https://www.health.harvard.edu/blog/first-do-no-harm-201510138421

generations of conservators have realized that any type of treatment changes an object, if even at a microscopic level, and that an assessment of potential harm must take this into account.[27]

However, there are two big differences between conservators and media preservationists in this area: first, because of ongoing degradation and rapidly increasing obsolescence, the content on media recordings is not expected to survive or be affordably accessible for much longer. This fact makes choices around interventions even more challenging because actions must be evaluated within the context of soon losing the opportunity to preserve content partially or even completely.[28] In contrast, conservators usually work with objects that are expected to remain intact and accessible for a longer period of time. Second, the objects themselves often have artifactual value and must be preserved through conservation work.

With media recordings, the carrier may well have enduring value: iconic or cultural if the original recording was particularly important, or scholarly if some aspect (e.g., handwriting, damage) not easily captured in the description can reveal some additional detail of the provenance of the content. However, this is usually not the case. The carriers are most often not considered valuable. It is the content or essence that is the primary target for preservation.

If we restrict our actions to those that have no risk of harm, then we will not be able to work effectively, efficiently or in a timely manner to preserve content. This directly violates Principle 11 above and, depending on the specific circumstances, may conflict with Principle 7, among others.

A more nuanced and useful guide in this area may be the philosophy or doctrine of Double Effect.[29] This doctrine states that it is permissible to cause a harm as a side effect (a "double effect") of creating a good or beneficial result. For our purposes, it is important to separate harm to the media carrier and harm to the content (sometimes called essence) on the carrier. The beneficial result may be the recovery of the content on the carrier and its subsequent preservation.[30]

Here are some of the tenets of the doctrine of Double Effect, adapted for relevancy to media preservation work:

- The harmful result may not be willed but may be permitted.
- Practitioners should strive to attain the beneficial result without the harmful result if possible or, at least, to minimize the harmful result.

[27] See, for example, Alison Richmond and Alison Bracker, *Introduction* in Alison Richmond and Alison Bracker editors, Conservation: Principles, Dilemmas and Uncomfortable Truths (Routledge, New York) pages xvi-xvii. Also in the same work, Elizabeth Pye, *Archaeological Conservation: Scientific Practice or Social Process,* 136, 141,
[28] Thanks to Natalie Rose Cassaniti for pointing this out and for the resources on conservation.
[29] See https://plato.stanford.edu/entries/double-effect/
[30] This idea is from Chris Lacinak.

- A proportionality condition must be met: the beneficial result must be sufficiently desirable to compensate for allowing the harmful result. In other words, the extent of the harmful result must be adequately offset by the magnitude of the beneficial result.
- The harmful result does not directly produce the beneficial result.

A cost-benefit analysis may be useful for evaluating the potential harmful and beneficial results of a proposed intervention and for reaching the best decision. Note that it may be necessary to evaluate the potential harm to the carrier through our actions, against the potential harm to the content through inaction, since degradation and obsolescence issues are likely to intensify over time.[31]

Policy 13.1: Decisions on procedures or courses of action that carry the risk of harm are made by a group of people with skills, training, and/or competencies relevant to the pertinent issues.

Example: Sticky shed syndrome tapes are unplayable and, as a result, the content on these tapes is inaccessible. However, there is a temporary remedy. These tapes are routinely baked, which enables them to be played and thus digitized. There is no hard science on how baking impacts a tape, and it may well cause harm to the carrier. Yet, the practice is ubiquitous because it results in stabilization of the media, enabling the faithful reproduction and capture of the original recording. We have generations of anecdotal evidence that baking (usually) works quite well. Invoking proportionality, we might argue that any harm is likely to be subtle (minor loss of high frequency content, for example, although this is purely conjectural) and offset by a benefit that is quite large—the ability to play, capture, and preserve previously unplayable content. The content of many thousands of sticky shed tapes may now be used thanks to this technique. In this situation, it appears that the benefits outweigh the potential risks.

Example: Cleaning videotape is sometimes necessary to achieve optimal playback. However, the cleaning procedure introduces risk of damaging the tape. For example, some cleaning machines make use of burnishing blades that scrape loose debris from the tape. While most engineers have not experienced problems, some have, and there is debate in the field on the relative safety of this type of machine. If cleaning leads to significantly improved playback, increasing the accuracy with which the recording is represented in the preservation master file, what is the point at which the risk moves from acceptable to unacceptable? Is it one in 100 tapes are damaged, or one in 1,000?

Example: Sometimes there are difficult choices that must be made. For example, in the early 1990s, before disc scanning and imaging methodologies were on the horizon, a delaminating lacquer disc left archivists with a no-win situation: play the disc (assuming it was playable) understanding that it would deteriorate significantly and perhaps completely as a result of playback attempts. Alternatively, place the disc back in storage understanding that it would continue to deteriorate to the point where it cannot be played within a relatively short period.

[31] This idea is from Will Prentice.

The potential beneficial result of the first choice is the recovery of some or all the disc's contents. The potential harmful result (the 'double effect') is that playback renders the disc unplayable in the future. Proportionality may be explored by assessing how much of the content is likely to be recovered vs. how likely is it that the disc will be further damaged, given the disc's current level of damage. Minimization of the harmful result may be attempted by manipulating disc playback variables such as tracking force, anti-skate, speed, and stylus size.

The potential beneficial result of the second choice is preserving as much of the disc as possible for a non-contact playback methodology, which is now available. The potential harmful result is that the disc further delaminates in storage and becomes unplayable. A cost-benefit analysis may include researching the availability and affordability of non-contact methodologies now and in the near future. The calculus may also hinge on whether we think it is useful or even possible or feasible to save the disc for a potential future playback.

Principle 14: Accuracy and Completeness

Many future uses of digitized media recordings require that preservation master files and metadata documents represent source recordings as accurately and completely as possible.

This is similar to Principle 7 above in the section on preservation program principles. It is a key principle for both a program and for specific digitization work and bears recasting here.

The primary goal of media digitization for preservation is to produce a representation of the source recording that is as accurate and complete a reproduction as possible for use by future generations. This representation is in the form of a digital file called a preservation master file that functions as the primary surrogate for the target content. The result is a copy that looks or sounds as close to the source recording as possible, including all its inherent imperfections and flaws.

Achieving this goal requires digitization engineers to adhere to specific procedures and for the digitizing organization to implement specific practices in the areas of accuracy and completeness as described below.

Choices of analog-to-digital converter, sample rate, bit depth, and playback techniques may all impact the accuracy of the conversion to the digital domain. Choices relating to metadata collection may affect the integrity of the digital representation of source recordings. Rules governing the creation of both preservation master files and production master files have an impact on the accuracy, integrity, and completeness of the digital files. This includes the choice of the file format itself.

Policy 14.1: Develop and/or adopt digitization workflows, procedures, equipment, personnel, and space that support the creation of digital files that represent the source recording as accurately as possible.

Practice 14.1.1: Professional equipment is used for digitization.

The overall objective is to transfer the full dynamic range and frequency response of the source recording to the preservation master file. For this reason, professional-level equipment that is as high quality as can be reasonably afforded is used for digitization. For certain formats or for specific technical characteristics, professional-level equipment may not exist. For example, professional VHS machines were never built to support the long-play (LP or EP) mode used to record some tapes. Also, there are no professional-level open reel audio machines that support tapes recorded at 1.875 ips and 0.9375 ips. In these cases, the highest quality consumer or prosumer machines are used.

Practice 14.1.2: Digitization that requires critical listening and viewing takes place in studios designed as critical monitoring spaces.

Audio engineers must be able to hear accurately the signal from the playback machine uninfluenced by the room so that they may make appropriate adjustments when preparing a recording for digitization. These adjustments include, for example, selecting the size and shape of the stylus that will yield the lowest amount of noise and distortion during digitization of a disc recording. Likewise, video engineers must be able to accurately view the video signal, which means that lighting and paint color specifications must be met. Digitization rooms are designed by studio designers to meet these critical functions.

An important exception are rooms that house high-throughput (parallel transfer) workflows. These workflows depend less on careful critical listening and viewing to reach their objectives.

Practice 14.1.3: No adjustments are made to any device in the signal chain during digitization.

Example: An audio engineer digitizing an open reel tape encounters a section of the tape in which the levels are exceedingly low. Should the engineer boost the level during playback to create a preservation master file that will make it easier for researchers to hear and understand the content? Doing so produces a file that is not as faithful or accurate a representation of the source recording as possible. Boosting the level during recording changes the dynamic range on the tape and the loudness of one performance relative to others. These are variables studied by some researchers. Therefore, audio gain (volume) is not raised or lowered during digitization regardless of how high or low the signal reaches. If the signal level is considered problematic, digitization is started again after levels are reset. Similarly, video levels (luma and chroma) are not raised or lowered during digitization.

Adjustments to the playback device that optimize the contact point between machine and recording (for example, a tape playback head or a disc stylus) may be made during digitization. This would include, for example, adjusting tracking for some video formats or manipulating the turntable tonearm's headshell and stylus to achieve more accurate tracking of a disc.

Practice 14.1.4: No denoising, color correction, or other signal restoration procedures designed to improve the listenability or viewability of the content are made in the preservation master file.

Signal restoration procedures are inherently arbitrary and result in the alteration of the audio or video signal. While they are not appropriate for preservation master files, they may well be useful in a derivative file to enhance use of the content. Such work is clearly documented for both future users and preservation administrators.

Practice 14.1.5: Physical restoration and stabilization procedures such as cleaning, baking, or repairing that are designed to optimize the physical and mechanical part of the playback process to achieve a faithful reproduction are permitted.

Practice 14.1.6: Use uncompressed or mathematically lossless formats for preservation master files created from the digitization of analog recordings.

For example, Indiana University uses the Broadcast Wave Format (BWF) for audio preservation master files and the FFv1 format with a Matroska wrapper for video preservation master files. BWF is uncompressed while FFv1 is mathematically lossless.

Lossy compressed files are permanently data-reduced files. This permanent data reduction does not meet the objective of creating an accurate representation of the source recording with the greatest possible fidelity for use into the future. It is also true that future conversions of lossy compressed files to new formats—which involve decoding from one lossy codec and encoding into another—will lead to an accumulation of errors that will result in increasing diminishment of quality over time. Finally, the resolution of lossy compressed files may not be high enough to support some uses now and into the future.

Policy 14.2: Develop and/or adopt digitization workflows and procedures that support the creation of digital files that represent the source recording as completely as possible.

Practice 14.2.1: Preservation master files are not edited.

Even content that is clearly unintended by the creator of the recording is kept in the preservation master file so that it is a complete representation of the source recording. Note that unintended content may be of value to users. Given that source recordings will not be accessible in the future, this content will be lost if not preserved in the file. If extraordinary circumstances warrant the removal of content (for example, restricted Native American ceremonies), this may be done in a derivative file at the direction of collection curatorial staff.

Practice 14.2.2: Gaps (sections with no apparent recorded signal) in the middle of source recordings are either retained in the preservation master file or partially removed and documented.

Example: An audio engineer digitizing a cassette tape encounters a five-minute gap on the tape where there is nothing recorded. Should the engineer digitize the silence?

There are several ways to think about this dilemma. A gap in the middle of a tape, while not uncommon with field recordings, represents an atypical and/or unexpected condition. A strict interpretation of the principles in this document suggests that unrecorded silence in the middle of the source recording, with content beginning again afterwards, be retained in the preservation master file to represent the source recording accurately and completely. However, it could be argued that, if the recording machine was turned off, there was no intention of recording, and the existence of a gap could be documented by metadata.

Some practitioners feel that sustainability and environmental issues should override the relevant principles. They argue that one minute (or five minutes) of unrecorded tape may not seem like much data but, when aggregated over collections and archives, can amount to a huge amount of storage (with its attendant energy usage and costs) to maintain 'nothing.'[32]

If the organization's policy is to *not* preserve gaps on the source recording in the digital file, the following additional policies and practices must be enacted:

- Some amount of the unrecorded tape must be retained in the file to signify that a gap was present and to prevent a mistaken interpretation that the content was continuous.
- The duration of the gap on the source recording should be documented in metadata.
- If the gap is intentionally meaningful to the recording, it should be preserved in its entirety. Performance art or production elements, for example, may contain unrecorded sections that meet this standard.
- If the gap is the result of possible malicious tampering, it should be preserved in its entirety. For example, if content was erased to conceal something.
- The gap must be played in its entirety during digitization to confirm that there is no content present.

Unrecorded (blank) tape after the end of content, when the recording machine has clearly been taken out of record, is typically not retained. Unrecorded tape at the beginning of content, before the recording machine was placed in record, is also not retained. There is less concern about these two situations as they are commonly found and even expected with tape formats. They are usually not documented.

Unmodulated but visible grooves (grooves with no recorded signal) at the beginning or end of a disc are digitized so that future users may know with certainty that there is no content in this area of the disc. If the grooves are visible to the naked eye there is the expectation that there is recorded content, even though it is not always possible to tell if the grooves were actually modulated and content exists. The temptation is to wonder if something was missed. The best

[32] This idea is from Kim Tarr.

way to avoid retrieving the source disc to find out is to play through the unmodulated grooves and preserve this in the digital file.

There is also a parallel that can be drawn with book digitization projects, many of which (but not all) do not scan pages that are blank. Usually, metadata is added that documents the number of pages that were blank and not scanned.[33]

Practice 14.2.3: Verify the beginning and the end of content, using leader tape as needed to enable the capture of recorded material at the extreme edges of the tape.

Digitization engineers take steps to make certain that all content on a recording, including content at the very beginning and the very end, is captured in the preservation master file. These steps include such things as the use of leader tape as needed at the beginning and/or end of a source open reel audio tape, as well as basic procedures like starting the record machine first and the playback machine second at the start of digitization.

Principle 15: Arbitrary Judgment Calls

Workflow components that rely upon personal opinion or interpretation represent potential weak links in the preservation chain and require mitigation and/or additional analysis and documentation.

Some aspects of digitization are inescapably arbitrary, that is, contingent solely upon a single person's discretion, interpretation, or opinion on what is best. Selecting styli, setting audio and video levels, performing an azimuth adjustment, and choosing playback equipment are all examples of technical actions that rely on the professional judgment of a person or persons. The choices made are open to interpretation and there may reasonably be disagreement between two professionals with similar skills and training. Given that no one is perfect, these judgment calls may not produce products that are 100% accurate.

In some situations, it may be desirable to preserve multiple versions from digitization to leave open the option of re-doing the procedure that was a judgment call. For example, digitization of audio cassettes encoded with Dolby noise reduction can be undertaken in such a way that it produces a preservation master file with no noise reduction applied and, at the same time, a preservation master-intermediate file with the engineer's choice of a noise reduction setting that sounds best to him or her. This is described in more detail below.

Policy 15.1: Staff with appropriate technical training and knowledge of preservation goals and practices are used to make judgment calls in the workflow.

[33] See, for example: https://www.aaa.si.edu/documentation/digitizing-entire-collections-chapter-4-for-scanning-technicians and https://memory.loc.gov/ammem/arendthtml/build.html#bpvc This analogy was suggested by Patrick Feaster.

Example: Performing an azimuth adjustment on an audio tape (described in Principle 7, above) is somewhat arbitrary and, if not performed well, may result in high frequency content that was not recovered. This risk can be reduced by using trained personnel with critical listening skills to perform this procedure. It may also be mitigated by targeted QC that compares the digital file to the source tape with the azimuth adjusted.

Policy 15.2: Discs requiring a playback equalization curve for which the exact curve used during recording is unknown, are digitized both with and without a curve.

Example: Selection of a playback equalization curve often represents a professional judgment call for non-RIAA discs. Any given transfer engineer's selection may or may not be accurate in the evaluation of others. In order to maintain maximum flexibility into the future, discs may be transferred both with and without a curve at the same time in one pass with both files preserved. The version without the curve may be used at any point in the future with whatever playback curve is desired. The version with the curve represents both the way the disc was intended to be reproduced as well as a more listenable version for researchers.

Policy 15.3: Audiocassettes encoded with Dolby B or C noise reduction are digitized both with and without decoding.

Example: Audiocassettes recorded in the field may be encoded with Dolby but not marked as such. They may also be marked inaccurately as to the type of Dolby used. It is also thought that cassette degradation may result in reduced levels on tape and/or other problems that result in the Dolby system not working very well. In these cases, the choice of which Dolby to use or even whether to use it at all during playback can be highly arbitrary. For this reason, a preservation master file is created without applying any noise reduction and, at the same time, a preservation master-intermediate file is generated with the engineer's choice of a noise reduction setting that sounds best to him or her. This provides options to future staff or users who may disagree with the choice that was made.

Principle 16: Trust

The workflows and equipment used in media digitization operations engaged in preservation work, and the products of these operations, cannot be trusted by themselves to meet established specifications.

Everyone makes mistakes and any operation created or guided by human beings will make mistakes. Therefore, systems and programs must be developed to increase the likelihood that the products of digitization *will* meet specifications. In addition, systems and programs must be developed to make certain that the output of a digitization workflow actually *does* meet specifications. This is classic quality assurance and quality control. The notion of trust, however, also extends to the performance of the digitization signal chain, which must be regularly verified so that the products of digitization meet an organization's specifications. It further

includes the equipment, space, and personnel used for digitization. These are all part of quality assurance.

Policy 16.1: Use a quality control (QC) program in which the products produced by a digitization operation are checked for adherence to the specifications of the digitizing organization.

Policy 16.2: Use a quality assurance (QA) program with procedures, policies, and decisions that increase the probability that the digitization system will produce products that meet the specifications of the digitizing organization. QA is a proactive process that consists of taking steps in advance to ensure that the product will meet the specification.

Policy 16.3: The digitization signal chain is tested and measured, and its performance verified, once per [whatever unit of time is needed for the specific operation. Week? Month? Other?] This includes system integrity checking, calibration and alignment of equipment, and other measures designed to verify the integrity of playback and digitization.

Policy 16.4: Use audio engineers, video engineers, and digitization operators to perform digitization. Engineers digitize recordings that require 1-1 attention. Digitization operators digitize recordings in parallel transfer workflows or in other situations where highly developed critical monitoring or technical skills are not as needed.

Personnel with specific skills, experience, and aptitudes perform digitization. This facilitates trust and contributes to quality assurance.

Audio and video engineers are trained in critical listening and/or viewing and have wide technical experience. These traits are essential for handling the more complex, delicate, and subtle digitization challenges. Usually, this is because the recordings are part of a fragile format, diagnosed with a preservation problem, or have other issues that call for the critical listening or viewing expertise, technical skills, and experience with deteriorating legacy recordings that an engineer provides.

Digitization operators are typically less experienced and less skilled than engineers. They may have an aptitude for technical work and receive specific training in digitization on the job. While critical monitoring and technical skills are certainly important for parallel transfer operations run by operators, the objectives of these workflows may be a little different than the 1:1 digitization workflows handled by engineers.

Policy 16.5: Use digitization engineers who are well-trained in the specifics of their craft.

Audio and video engineers trained in production work are often not skilled in the tasks and procedures required in a preservation workflow. They may not have experience manipulating a deteriorating recording on a legacy format using an old and obsolete playback machine. Specific training in digitization for preservation is provided to digitization engineers. This may be provided in-house if the expertise exists or outside of the organization if that is where it resides.

Policy 16.6: Digitization is as close to 100% attended as reasonably possible for 1:1 transfers.

Practice 16.6.1: Engineers remain in their studios during digitization, focusing on the task at hand.

Transfer engineers listen to or view as much of the content during digitization as is reasonably possible in order to verify optimal playback and digitization. Audio and video monitors are used, and playback machine transports are monitored, during the entirety of the digitization process.

Media Preservation and Digitization Principles

First Principles

First Principle A. Inherent value

Time-based media content has inherent value as a primary information source for future uses.

First Principle B. Protecting value

Recordings that are considered valuable must be protected from loss.

First Principle C. Separation of Content from the Carrier

The content captured on a media recording via the recording process, not the carrier or the physical object itself, is the most valuable target for preservation.

Program-Level Guiding Media Preservation Principles

Principle 1: Taking Action

Active degradation and the rapidly advancing obsolescence of audio and video recordings require immediate, and ongoing, preservation action.

Principle 2: Long Time Horizon

Media preservation requires a commitment to the long-term.

Principle 3: Timeliness

Media preservation requires timely intervention.

Principle 4: Priority

Media preservation actions are taken in order of priority.

Principle 5: Primacy of the Unique and Original

Unique and/or original items receive highest priority.

Principle 6: Digitize or Transfer Once

Due to time and resource constraints, and the very large number of recordings in need of preservation, it is highly desirable to digitize analog recordings or transfer physical digital recordings to digital files just once.

Principle 7: Accuracy, Faithful Reproduction, and Integrity

The products of preservation work must be as accurate as possible, representing the source recordings faithfully and with the highest level of integrity.

Principle 8: Standards and Best Practices

Standards and best practices help media preservation programs ensure that preservation work is high quality, sustainable, interoperable, accurate, and consistent.

Principle 9: Preservation and Access

Preservation and access are interdependent and equally important for media collections with value for future use.

Principle 10: Knowledge and Expertise

Successful media preservation requires knowledge and expertise from a range of disciplines.

Principle 11: Efficiency

Due to time and resource constraints, media preservation actions must be delivered as efficiently as possible.

Principle 12: Redundancy

Managed multiple copies decrease the risk of loss by lessening the dependency on any single copy.

Media Digitization Principles

Principle 13: Beneficial and Harmful Results

Preservation is best served by weighing the potential benefits of an action against the risk of harm.

Principle 14: Accuracy and Completeness

Many future uses of digitized media recordings require that preservation master files and metadata documents represent source recordings as accurately and completely as possible.

Principle 15: Arbitrary Judgment Calls

Workflow components that rely upon personal opinion or interpretation represent potential weak links in the preservation chain and require mitigation and/or additional analysis and documentation.

Principle 16: Trust

The workflows and equipment used in media digitization operations engaged in preservation work, and the products of these operations, cannot be trusted by themselves to meet established specifications.

Media Digitization Practices

Below are digitization practices discussed in the text above. They are compiled here to make them easier to view and evaluate together. This is not a comprehensive set of practices, nor is it intended as a set of best practices. Instead, it is a list of practices that serve to illustrate the text in this document.

Practice 14.1.1: Professional equipment is used for digitization.

Practice 14.1.2: Digitization that requires critical listening and viewing takes place in studios designed as critical monitoring spaces.

Practice 14.1.3: No adjustments are made to any device in the signal chain during digitization.

Practice 14.1.4: No denoising, color correction, or other signal restoration procedures designed to improve the listenability or viewability of the content are made in the preservation master file.

Practice 14.1.5: Physical restoration and stabilization procedures such as cleaning, baking, or repairing that are designed to optimize the physical and mechanical part of the playback process to achieve a faithful reproduction are permitted.

Practice 14.1.6: Use uncompressed or mathematically lossless formats for preservation master files created from the digitization of analog recordings.

Practice 14.2.1: Preservation master files are not edited.

Practice 14.2.2: Gaps (sections with no apparent recorded signal) in the middle of source recordings are either retained in the preservation master file or partially removed and documented.

Practice 14.2.3: Verify the beginning and the end of content, using leader tape as needed to enable the capture of recorded material at the extreme edges of the tape.

Practice 16.6.1: Engineers remain in their studios during digitization, focusing on the task at hand.