

# ARSC Technical Committee

## Preservation of Archival Sound Recordings

*Version 1, April 2009*

This document contains the Association for Recorded Sound Collections Technical Committee's (ARSC TC) recommendations for preserving sound recordings, and more specifically for transferring sound recordings to an archival digital format, based on the state of the art and the technology in 2009.<sup>1</sup> This statement updates the ARSC/AAA (Associated Audio Archives) recommended practices that guided audio preservation work starting in the late 1980s. This is not intended to offer specific procedures or best practices; for that level of detail we refer you to the list of audio preservation resources that accompanies this document. Rather, this document provides general guidance using, as much as possible, language accessible to a non-expert audience, although those with more experience in preservation matters may also find it useful.

The ARSC TC endorses the work of the International Association of Sound and Audiovisual Archives (IASA) Technical Committee as published in IASA documents *TC 03 (The Safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy)* and *TC 04 (Guidelines on the Production and Preservation of Digital Audio Objects)* and directs interested parties to those works. We add our perspectives here on the general principles in the IASA documents along with some additional thoughts of our own. We hope this document will provide useful guidance to public and private archives, libraries, museums and other collections that include recordings of speech, music, ambient sounds and other audio.

### General Principles and Guidelines

After more than a hundred years, it is no longer practical to make archival analog copies of audio recordings to preserve their content. High-quality recording equipment (and the pool of replacement parts needed to keep existing equipment working) and tape are increasingly difficult to obtain; within the next decade or less, it is likely that finding recorders and tape will become even more of a problem.

Storage of archival audio as digital files is now the better approach for several reasons:

- Analog copies lose quality with each generation, while a copy of an uncompressed digital audio file can be identical to the digital original
- High-quality audio tape for archival copies is often prohibitively expensive

- Many physical digital carriers such as Digital Audio Tape are obsolete; others such as CD and DVD are expensive and difficult to manage for long-term preservation
- Archives can and should move digital files to new storage media using automated processes as and when data storage technology advances without the need to re-create each digital file; migrating and duplicating analog recordings to new media is a more costly, labor-intensive process
- Archiving audio recordings as digital files makes possible easier and expanded access to their content

Therefore, we believe that the time has come for archives and other collections to transfer their analog holdings to digital files. We can no longer assume that researchers will be able to gain access to the contents of an analog archive thirty or more years from now. Analog sound carriers continue to degrade, some quickly and catastrophically. Obsolescence—of playback machines, technical expertise, tools, and audio formats—makes it more difficult and expensive to retrieve deteriorating media. To ensure that audio content survives, some holdings must be digitized relatively soon. Please refer to the ARSC Technical Committee’s list of “Essential Resources for Audio Preservation” to help identify media most at risk.

It is important to devise a plan before starting any preservation project. The plan should include:

- A survey of holdings to determine which important materials are most at risk
- A decision on whether to use a vendor or in-house resources (or a combination of both) and what facilities and staff training are necessary
- A long-term preservation storage structure for the digital files as an independent system or as part of the institutional IT infrastructure
- Funding for all of the above

It is still important to preserve and maintain the original analog master recordings in order to allow future generations of archivists and audio technicians to apply new methods that may extract better quality and/or additional information compared to today’s technology. In addition, the originals form part of the historic record and keeping the originals should be part of a complete backup strategy. For specific details about storage requirements of analog recordings, see the resource list cited above.

Today, most new sound recordings are produced as digital originals, in a wide variety of formats, on a wide range of storage media. Over time, equipment that can play many of these digital formats and media will also become difficult to find. Finding equipment for some older digital formats is already a challenge. We must assume that both recording technology and digital storage methods will continue to advance and change, and those changes will, if anything, come even more rapidly than the advances of the last century. Today, an important part of our job as conservators of recorded sound is to make things as easy as possible for our colleagues and successors fifty to two hundred years from now.

## Metadata -- Related Data and Other Information

Neither analog nor digital recordings exist in a vacuum. Printed or handwritten notes and other documents, photos and other images, and published reviews are often as important as the recordings themselves. In addition, technical information about both the recording and the digital transfer is a key part of the complete preservation package.

Metadata – literally “data about data” – is defined by the National Information Standards Organization as “structured information that describes, explains, locates or otherwise makes it easier to retrieve, use or manage an information resource.” Within the context of a sound archive, access to metadata and other documentation will be essential when the next generation of archivists tries to recover and understand the contents and context of an archived collection of sound recordings.

The digital copy should be a complete surrogate for the original recording, including all the supplemental information (notes, labels, images of the box, sleeve or other container and all available technical information) that accompanies the original carrier.

## Digital Copies

It is essential to transfer physical items (e.g., analog and digital original master recordings) to one or more common digital file formats, using procedures that embody these principles:

- The archival master (also called preservation master) digital files must be as accurate copies of the originals as possible; any compromise in the transfer will remain in the files forever (or until somebody goes back to the original masters, if they are still playable)
- The transfer process should not add any artifacts or subjective changes. The initial digital transfer to archival masters is not the place for file compression or other data reduction such as MP3 encoding, change in dynamic range, or noise reduction that is not already encoded in the original signal. However, it is appropriate to decode recordings that have been encoded using noise reduction systems such as Dolby or dbx if an accurate assessment has been made of the type of noise reduction used and the performance of the decoder. If unsure, two copies may be stored in the digital archive: one in “raw” form without decoding, and a second copy with decoding applied, each clearly marked.
- The digital archival master file format should be widely used, self-documenting, with publicly accessible specifications, and it should not depend on a specific operating system or type of equipment (see below, “Making Digital Archival Masters”)
- The archival storage media should also be widely used, and it should not depend on proprietary equipment
- The digital audio file should be part of a larger preservation package that includes related metadata as well as pointers to associated digital text and image files (or the files themselves) and physical artifacts

- Digital archival master files require unique persistent identifiers, a role that may be played at least temporarily by a consistent descriptive file-naming scheme. This identifier should also appear in a database or spreadsheet to associate files with metadata.

## Transfers from Analog Masters

The process of converting a recording from an analog master to a digital copy is a critical point in preserving a recording. If the transfer is done properly, it may never be necessary to return to the analog original. Therefore, it is essential to understand both the original technology used to create the recording, and the process of digital conversion.

### ***Playback***

Do not cut corners on playback equipment and methods. Every analog medium requires a specific set of tools and methods to prepare and play a recording; quarter-inch tape, cassettes, wire recordings, broadcast transcription discs, dictation media, cylinders and other historical media hold recorded sound, but each requires very different playback equipment. In all cases, experience and expertise with specific media and equipment are critical for the best possible playback.

However, a few general rules apply:

- Transfer from the original masters whenever possible. If the originals are lost, destroyed, or damaged, use the best available copies. Sometimes “best” is not self-evident; in that case, consider transferring all logical candidates and comparing the sound quality of the digital copies. Always include a transfer of the original analog master (if it is available), even if other copies sound better or are in better shape, as proof of its condition.
- Whenever it is appropriate, clean or physically restore recordings before transfer to obtain the best results
- Use the best quality playback equipment available, and keep the equipment in good operating condition
- Always adjust the playback machine for each recording to obtain the most accurate reproduction of the original recording
- If you do not understand how a playback device works or how to align it, get help from somebody who does
- Before you try to play a recording, test and evaluate all the equipment in the signal chain
- The goal of preservation is to present an *unaltered* digital representation of the analog recording for use by future generations. Do not introduce any audio processing, compression, normalization or “sweetening” during your digital transfer – you can do that later on a secondary digital copy; it is best to presume that the next generation of audio engineers might develop a better way to clean up noise and distortion, so we should not force them to undo our less accurate processing. Always

monitor the transfer through high-quality headphones or monitor speakers. If you hear problems (such as drop-outs, skips, or off-speed playback), stop well after the event, apply the appropriate treatment to the media or adjustment to the playback device and attempt to re-transfer the offending section with sufficient overlap to provide clear context for editing – unless it is more efficient to re-transfer the entire item. If you detect mechanical playback problems, include detailed notes that might help a future transfer technician solve them.

- Use expert help to repair, clean, and/or transfer fragile or difficult media
- Exercise quality control during and after the transfer as problems or defects in the first digital copy will carry over to all future versions

In general, it is important to remember that making an archival master copy is not the same as creating an original recording, preparing it for re-issue, or listening to a recording in your living room or classroom. As stated in the ARSC/AAA guidelines from the late 1980's, the sound archivist's job is to preserve history, not to re-write it or "improve" it. At this stage, your primary goal is to provide the most accurate possible copy of the original. You are creating raw material for other archivists, media producers and researchers, each of whom will have different requirements; let them worry about reducing noise or applying other processing to meet their own specific needs.

## ***Metadata***

Keep detailed notes about the digital transfer, including the date and location, the make and models of the equipment used, the settings on the equipment, the signal path, any technical adjustments, the full names of the people operating the equipment, and as many additional technical details as possible about the source recording, the analog playback and the digital copy. This information should become part of the metadata that accompanies the digital audio file.

## ***Making Digital Archival Masters***

Archival master digital audio files for archival storage should use a widely available uncompressed format. The *de facto* standards for archival master audio file formats are WAV and Broadcast Wave Format. In practice, the Broadcast Wave Format is the better choice because it provides a specific, defined location within the file itself for metadata about the content, ownership, source recording, and digitizing signal chain associated with the digital file, as well as a unique source identifier (USID). It also provides a timestamp to allow related files to be sequenced in the correct order.

A bit depth of 24 bits per channel for preservation transfers is generally accepted, and we strongly recommend this. Sample rates are somewhat controversial; there is some disagreement in our field about the minimum acceptable values. The *de facto* standard sample rate for archival master digital copies, which we recommend, is 96 kHz. Lower sample rates may be acceptable under certain circumstances but this decision should be made by experts and fully justified. When in doubt, use 96 kHz.

The analog-to-digital converter is an extremely important link in the process of making digital copies. The best A-to-D converters for archival transfers are the ones designed for

“professional” applications rather than consumer-quality sound cards and audio processors on computer motherboards. The consumer-grade sound devices and on-board audio processors found in many desktop and laptop computers are often not accurate or quiet enough for archival transfers.

Best practices documents recommend using an external analog-to-digital converter located outside of the computer (most often connected to the computer through USB, FireWire or AES/EBU interfaces). Like every other element of the signal chain, the converter should be of the highest quality that the project budget can afford. Note that not every external converter can provide acceptable results and some experts believe that a few excellent converters that are installed in a computer’s expansion slots may provide acceptable performance. We recommend seeking expert help in addressing this issue.

## Transfers from Digital Originals

Many recordings created in the last ten or twenty years were produced as digital originals on tape, compact discs, MiniDiscs and other media, or directly to hard drives and flash media. In most cases, it is good practice to save the original master files on the original media, along with archival copies on archival storage media. However, if the master was created on a reusable storage device (such as flash media or a portable MP3 recorder), it is acceptable to preserve a bit-for-bit copy on other, more durable media as your “master.”

For archival preservation and storage of existing digital files, keep a copy of the original file, and a copy converted to an uncompressed WAV or Broadcast Wave file (unless the original file is already in one of those formats). Even if the original was created in a compressed file format such as MP3, Ogg Vorbis, or a file created from streaming media, archival master copies of digital recordings should be uncompressed Broadcast Wave files. While migration from the original file format to WAV or Broadcast Wave will eventually be necessary for long-term sustainability, it is not simple, and it is acceptable to delay such migration as long as a close watch is kept on the relevant technologies for impending obsolescence.

In general, archival master copies of digital files should use the same bit depth and sample rate as the originals.

If the original digital masters are in a proprietary format for which the playback equipment does not provide a digital output, treat the recordings like analog masters and create new archival master digital copies following the practices outlined above.

## Storing the Digital Files

Storing digital files for preservation presents two separate requirements: protected interim, local storage before files are ready for transfer to the long-term solution, and long-term storage for preservation. Long-term storage demands a careful strategy to ensure that the content remains viable over a very long period of time, even if it moves to different storage media. This strategy requires a paradigm shift from traditional audio archiving: the *content* is the target for preservation rather than the carrier. The owners of an archive should expect to migrate content from carrier to carrier as needed. A few basic storage principles include:

- Generate local backup copies of all files as soon as possible after creation
- Store backups on a separate device or separate media from the original. Make at least two, or preferably three sets of archival master digital files of preserved content and store them in different locations, and possibly on different types of media.
- Use a method to verify data integrity such as checksums for all files to be retained over the long term
- Monitor the storage media for errors and degradation over time
- Monitor the lifecycle of digital file formats, and adopt new ones when the previous standard becomes obsolete

Digital preservation requires ongoing active management of the archive including data integrity checking, evaluating obsolescence issues, and planning for the next migration. It is no longer possible to put storage media on the shelf and forget about them.

No single type of storage media is perfect for long-term storage of digital files. Each has its advantages and disadvantages. Briefly:

- Data tapes are inexpensive, reliable, and robust but the tape drives required to produce them are relatively expensive up-front and they require specific knowledge to integrate into a data storage system
- Hard drives are familiar, easy to use, and they provide random access, but as mechanical devices, they are not as reliable or robust as data tape over long periods of time
- RAID arrays remove some of the disadvantages of individual hard drives while they provide larger capacities than individual drives. They are extremely easy to use, especially when supplied as network-attached storage, but still should be considered mechanical devices that may fail at some point in time. A single RAID array is not adequate for long-term storage.

Optical discs require special mention as unsafe preservation storage media. Using this format requires costly test equipment to check the quality of both blank media and digital recordings. Even if this equipment is already available to an archive, the cost of storage media may be higher than the other formats, except in very small volume – individual CDs or DVDs are not expensive, but the cost-per-megabyte is substantially greater than other media types.

Recordable audio CDs are generally not appropriate for archival storage. In addition to the other problems with optical discs, the format cannot hold files at higher bit depth and sample rate than 16 bit, 44.1 kHz, which is far less than the *de facto* 24 bit, 96 kHz standard. In addition, migrating an archive stored on optical discs is expensive, because it requires much human intervention. Further information on the use of optical discs may be found in the resource guide.

One place where audio CDs may be the only workable option is in very small archives of oral history cassettes where tapes are unstable and/or at serious risk and funding constraints force a do-it-yourself preservation program. This is a temporary measure, and these audio CDs should be transferred soon to other media and ingested into a larger managed digital repository.

Digital Audio Tapes (DATs) are also unsuitable for archival storage. Numerous reports of playback problems suggest that DATs are unreliable for long-term storage because they often develop dropouts and mechanical problems. Further, the format is obsolete — no new DAT machines have been produced since late 2005 and spare parts for playback equipment will become difficult or impossible to obtain in the near future.

Many institutions are developing mass storage systems for digital files. Some are also adding preservation services to create what is often called a trusted digital repository (TDR). This type of repository uses mass storage systems and a set of stringent criteria for preserving and providing access to digital files. In such a system, data integrity is checked regularly, the relationships among files are preserved using metadata documents and metadata systems, changes to files are tracked, and content is migrated automatically when file formats become obsolete. Many archivists believe that storage in an actively managed TDR provides the strongest possibility that content will remain accessible over the long term. Several organizations are developing processes to certify TDR's.

For archives lacking the expertise or resources to create and manage the physical aspects of a digital preservation system, procuring the services of a TDR to care for preservation files may be the answer. This may be done through partnership with an academic institution or purchasing services from a vendor, assuming that TDR's will eventually be developed by commercial operations. Whether the trusted digital repository is a commercial, educational, or other type of institution, it is vital for the client archivist to develop and maintain a strong working relationship with the IT staff of the repository to ensure that the archive's interests are well served.

Regardless of storage strategy, long-term digital preservation requires a strong commitment from archive owners that provides for periodic data migration and sustained system support. The owner of every archive should expect to revisit these issues on a timely basis.

## **Documentation**

A digital archive must include a catalog that identifies its contents and provides all the technical information that future users might need, in order to retrieve and use the contents of the archive. It should also contain links to digital metadata files and pointers to physical artifacts related to each digital audio file.

## **Conclusion**

International standards and best practices for audio preservation are largely in place. Costs associated with the storage of archival master audio files and associated metadata have and will continue to drop to the point where they are less of a barrier. While the combination of carrier degradation and format obsolescence makes digitization in the near-term a necessity, it is now possible to undertake enduring digital preservation work in earnest.



Our window of opportunity is open, and it is our generation's task and responsibility to convert analog holdings to digital assets, and to provide the next generation of archivists with sustainable content.

---

<sup>1</sup> The issues addressed in this document were discussed and debated by the full ARSC Technical Committee. Lead writer/editor for the document: John Ross. Additional writing/editing: Mike Casey. Contributions from: Peter Alyea, Marcos Sueiro Bal, George Blood, Adrian Cosentini, Bruce Gordon, Eric Jacobs, Richard Hess, Chris Lacinak, Brad McCoy, John Narus, David Nolan, Doug Pomeroy, Marie O'Connell, Tommy Sjöberg, John Spencer, and Charles Van Winkle.