



# SOUND SAVINGS: PRESERVING AUDIO COLLECTIONS



## Proceedings from the Symposium

### THE LIBRARY OF CONGRESS DIGITAL AUDIO PRESERVATION PROTOTYPING PROJECT\*

#### **Carl Fleischhauer**

Project Coordinator  
Office of Strategic Initiatives  
Library of Congress

#### **Introduction**

The Digital Audio Preservation Prototyping Project was established at the Library of Congress for several reasons. The underlying motive—not always visible in our presentations—is that the time has come to change our approach to reformatting recorded sound collections, for reasons I will outline in a moment. The surface motive, the trigger to action, is the planned move in 2005 by the Library's Motion Picture, Broadcasting, and Recorded Sound Division to a new facility in Culpeper, Virginia. Substantial funding for the new National Audio-Visual Conservation Center comes from David Woodley Packard (the son of David Packard, co-founder of the Hewlett-Packard Corporation) and the Packard Humanities Institute.

The project has embraced sample collections from two Library of Congress divisions: the Motion Picture, Broadcasting, and Recorded Sound Division (M/B/RS) and the American Folklife Center (AFC). In another talk at the Sound Savings conference, the archivist Michael Taft described the AFC Save Our Sounds effort, which is allied with the prototyping effort. Overall, the project's focus has been on reformatting sound recordings, with an eye on moving into video. We want to reach some useful conclusions next year, in time to apply the lessons in the new building.

The prevalent practice for reformatting audio and video from the 1960s and 1970s into the 1990s has been "copy to analog magnetic tape." We see four reasons to change. First, there is the matter of media life expectancy. Magnetic tape (analog or digital) will not last as long as the archetypal media used for reformatting: microfilm. Second, there is the issue of quality loss as a result of making the copy. Analog-to-analog copying introduces what is called generation loss. This is tolerable with microfilm, when the time between re-reformatting is long. But with audiotape the time between re-reformatting is relatively short, and the adverse effects are troubling. Third, there is the problem of device and media obsolescence. We are seeing a virtual cessation of manufacturing of analog-tape media and analog-tape recording devices. Finally, the digital era is here, and we need to engage it, and not just to serve reformatting. The next

generation of content to reach our institutions will be digital to begin with, and its preservation for the long term will depend upon techniques similar or identical to those we establish to sustain digitally reformatted content.

The prototyping project has also been motivated by the desire to model new ways to provide access to researchers. The production of digital masters makes it relatively efficient to produce service copies, e.g., compressed copies that can be accessed in our secure local area network or streaming copies for the Web. At the Library of Congress, copyright considerations and consideration of the prerogatives of folk communities mean that we must limit access to much of our recorded sound collection, i.e., many items cannot be placed on the public Web. But after the collections have moved to Culpeper, we want our reformatted content to continue to be accessible in reading rooms on Capitol Hill, and the digital service copies that we place in the Library's secure storage systems will help us accomplish that goal. We are also exploring ways to provide access more widely, perhaps to remote sites, legal circumstances permitting.

One of the advantages of digital-file reformatting is the ability to reproduce an entire object. For example, here is a description of our digital reproduction of a sound recording made by the U.S. Marine Corps in the Pacific during World War II. The 1945 original was recorded on Amertape Recording Film, sprocketed 35mm film that ran through a recorder that cut grooves in the surface. Within a year or so, the Marine Corps copied the film to 16-inch transcription discs. These have since deteriorated but they were used as the source for our audio. (We hope to go back to the film at some point.) The digital copy provides access not only to the audio but also to images of the film box, the disc labels, and a content log sheet that had been packed with the film. This virtual package is presented in an interface that permits a researcher to play the audio, zoom in on the images, and examine detailed technical metadata.

The preservation approach we are exploring has at its core a *digital object* or *information package* that includes bitstreams, i.e., the files that contain the audio and images, and metadata. These packages will be managed in digital *repositories*, sophisticated versions of the computer storage systems we are using today. CDs or DVDs will not be used to store the content. It is worth saying that content management—what happens inside the repository—has at its heart a paradox. Digital content depends on specific information technology systems to keep it alive and to render it for users. But information technology systems are inherently obsolescent and will be replaced in relatively short time periods and thus our content must also be system independent. At any given moment, content lives on *this* media—disks in this server, for example—and is sustained by *this* information technology system, but the content must transcend the lifespan of any given media and system.

Our preservation explorations have wrestled with four issues: selecting the target format for reformatting, determining the quality of the reformatted copy, shaping the information package and the importance of metadata, and analyzing longevity in a “media-less” environment.

### **Selecting the Target Format**

The first issue concerns the choice of bitstream structure and file type. This entails striking a balance between six factors:

**Disclosure:** Are specifications and tools for validating technical integrity accessible to those creating and sustaining digital content? Preservation depends upon understanding how the information is represented as bits and bytes in digital files.

•  
**Adoption:** Is this format already used by the primary creators, disseminators, or users of information resources? If a format is widely adopted, it is less likely to become obsolete rapidly, and tools for migration and emulation are more likely to emerge.

•  
**Transparency:** Is the digital representation or encoding open to direct analysis with basic tools? Digital formats in which the underlying information is represented simply and directly will be easier to migrate to new formats. Encryption and compression inhibit transparency.

•  
**Self-documentation:** In part, this is about the package inside the package. Does the file format include metadata that explains how to render the data as usable information or understand its context? Self-documenting formats are likely to be easier to sustain over long periods and less vulnerable to catastrophe than ones that are separated from key metadata.

•  
“**Fidelity**” or support for high resolution: Does the format “hold” high resolution audio?

•  
**Sound field support:** Does the format represent stereo and even surround sound?

What formats have we selected? For our audio masters, our bitstream choice is pulse code modulated (PCM) sampling, uncompressed. This is the type of bitstream used on audio compact disks and it meets the transparency test. The file format we use is WAVE, from Microsoft, and it meets the adoption, disclosure, and fidelity tests. By the way, we feel that the “PCM-ness” of the bitstream is more important than the “WAVE-ness” of the file; Macintosh users put their PCM bitstreams into AIFF files to equal effect. We have not yet begun using what is called the Broadcast WAVE Format, which would get a higher score on the self-documentation test than ordinary WAVE. Meanwhile, we are curious about one-bit-deep formats like the DSD structure on SONY’s Super Audio Compact Disk (SACD) but this bitstream structure is not yet widely adopted. DSD also gets occasional negative write-ups in the trade press, so we are taking a wait and see position. Since our reformatting is limited to mono and stereo material for the moment, we can afford to put off addressing the matter of surround sound. For audio service files, we use WAVE at lower resolution and MP3 compressed files.

For our image masters, our bitstream choice is bit-mapped or raster, also uncompressed. The file format we use is TIFF, another industry standard, originally from Aldus and now from Adobe. Here, the “bit-mapped-ness” of the bitstream is more important than the “TIFF-ness” of the file. For image service files, we use JPEGs and expect to switch to JPEG 2000 after this new format has been more widely adopted.

### **The Quality of the Reformatted Copy**

The second central issue has been the subject of several interesting and instructive discussions by LC staff working on the prototyping project. Our talk revolved around questions like, “What

does high resolution mean?” and “Why should we seek it?” In the end, our decision-making turned on some unexpected factors, some of which are beyond the reach of science and objective measurement.

With sound, the analysis of resolution starts with considerations of sampling frequency, measured in cycles or kilocycles per second. Roughly speaking, digitizing audio means taking the analog waveform and representing it as a large number of points or dots—connect the dots and you have your waveform back. The more dots, the better you can redraw your sound wave; the more dots, the better you can represent the fine parts of the curve that represent high frequency sounds. This parameter can be compared to spatial resolution for images. A digital image consists of row upon row of picture elements, pixels for short, often called “dots.” The higher the number of pixels, the higher the spatial resolution.

The second key parameter is bit depth, which audio engineers sometimes call “word length.” With audio, the more data you have for each sample—the longer the word, so to say—the more accurate the position of the sample in terms of amplitude. Greater bit depth gives you a lower noise floor and lets you represent a greater dynamic range, which can be especially helpful when transferring, say, field recordings made in hard-to-control circumstances. Compact disks usually have 16 bits (2 bytes) per sample, while many professional recording systems offer 24 bits (3 bytes). The imaging analogy is that an image 24 bits per pixel can reproduce more colors than 8- or 16-bit representations and thus offers the possibility of greater color fidelity.

Everyone is convinced that it is a good idea to digitize audio at 24 bits per sample. Keen ears can hear the difference and, although we have not done so, one could exploit test signals to compare distortion and noise. And it was in the discussion of bit depth that one of the “unmeasurable” factors was articulated: “You want a cushion of extra data,” the engineers said, “just to protect you when you copy items with a wide or varying dynamic range, or to give you elbow room to fix things later in the event that an operator doesn’t do a perfect job.”

I have heard an analogous argument regarding imaging, especially when reformatting photographic negatives. The proposal is to make a preservation master image with a “flat” (low contrast) contrast curve and 12- or 16-bits-per-channel instead of the customary 8. Then a future user could manipulate the image to restore it or for a desired aesthetic effect, and resave it at 8 bits deep. The outcome of this process would be an image with a full set of tones at the 8-bit depth, i.e., the histogram for the new 8-bit image would be free of gaps. In contrast, if you started with an 8-bit-per-channel master, manipulated it, and then resaved the copy at 8 bits, the resulting copy image would lack some tones, i.e., the histogram would have gaps.

In contrast to the consensus we reached regarding the desirability of greater bit depth for sound recordings, our conversations about sampling frequency revealed differences of opinion. Some of us on the administrative side imagined that the starter question would be: “What is the range of sound frequencies that we might expect in this original item?” Our idea was that we would set the frequency range of the digital copy to more or less match the frequency range inherent in the original item. What frequencies have been captured, for example, on a 78 rpm disc from the acoustic era? From 8-10 kilocycles per second? The usual rule for digital sampling is to work at twice the highest frequency you want to reproduce. Therefore something on the order of 20 kilocycles per second should capture the full range of frequencies on the original 78. Or to take

another example, suppose a collector used an analog Nagra tape recorder to record folk music at 7.5 ips with a Neuman condenser microphone. What is the highest frequency tone that we might expect to hear when the tape is played back? Most engineers would say that such a recording system is not likely to capture much sound with frequencies above 14 or 18 kilocycles per second. Thus if we digitally sample at 44 or 48 kilocycles, we ought to capture the full range of frequencies.

The engineers, however, did not want to work at 44 or 48 kilocycles, to say nothing of 20. They advocated 96, with some eyeing 192. The argument here—and really this argument takes both higher sampling frequency and greater bit depth into account—largely concerns factors that pertain to practical production matters or “downstream” possibilities, and which are therefore not very susceptible to objective testing. The following paraphrases capture some of the dialog:

•  
“There may be hard-to-hear harmonics that you won’t want to lose.”

•  
“Copies with less noise and less distortion can more successfully be restored when you come back to them.”

•  
“In the future we’ll have even better enhancement tools and post-processing, so save as much information as you can.”

•  
“What if you need extra data to support certain types of resource discovery?”

The last bullet refers to what is called “low-level” features by the working group developing MPEG-7, an emerging metadata standard associated with moving image work. In the world of sound, low-level data would be used to support “find this melody” queries, for processes that produce transcripts of spoken word content, or to support a system executing the famous query “find me more like this one.”

Thus the analysis of the inherent fidelity of the original did not provide the steering effect some of us had expected. Meanwhile, we also made test copies at higher and lower resolution and asked, “Can you hear the difference?” But these informal A-B tests also fell short of being conclusive. One engineer has proposed carrying out some empirical tests on post-processing actions to confirm the idea that the restoration of a recording (e.g., careful cleanup for publication as a CD) would be more successful if the master was at high rather than moderate resolution. But we have not yet carried out such an experiment.

The outcome is that our team generally works at the upper limit of available technology. We produce most of our audio masters at 96 kilocycles and 24-bit word length. At this time, we make two service copies: first, a down-sampled WAVE file at compact-disc specifications: 44.1 kilocycles and 16-bit words, and second, an MP3 file that is very handy in our local area network. Meanwhile, we produce images of accompanying matter, like disc labels, tape boxes, and documents. The master images are at 300 pixels per inch (ppi), with a tonal resolution of 24 bits per pixel.

Our project development has highlighted two additional topics that have to do with reproduction

quality. The first has to do with practices, including the use of professional equipment and professional workers. On the equipment side, one key device is the analog-to-digital converter, the device that actually samples the analog waveform and spits out the bits. Professional converters are generally external to the computer workstation (or digital audio workstation) and are superior to and more costly than “prosumer” a-to-d devices, often installed as a card in the desktop computer. We avoid cleanup tools when making masters. And for mono discs in our collections, we copy using a stereo cartridge to allow for future processes to “find the best groove wall.”

On the human side, digitizing requires professional skills in both the digital and analog realms. A professional worker must not only be conversant with a-to-d converters and workstations, but must also be a master of the art and science of playing back originals to the best effect, no mean task when you confront instantaneous discs, cylinders, wires, and sticky tapes. In the new center at Culpeper, we see these professionals as our supervisors, contract overseers, and as experts who perform the most difficult work.

As we plan the future, we would like to include apprentice workers in the team, as well as outsource certain types of material. We have so many items in need of reformatting, that we are seeking ways to increase efficiency. Elements that we hope will accomplish this include sorting originals by “transfer efficiency” category, that is, by putting groups together that have the same technical characteristics. We would like to find and employ expert systems (automated tools) to help us judge quality or at least spot anomalies to inspect later. For some categories, we want to experiment with having a single operator copy two or three items at once. I will note that some interesting high-volume production tools are emerging from the PRESTO (Preservation Technology for European Broadcast Archives) project organized by broadcasters in Europe <<http://presto.joanneum.ac.at/index.asp>>. At the same time, our team has been very interested to learn about Carl Haber and Vitaliy Fadeyev’s cutting edge experiments at the Lawrence Berkeley National Laboratory to use high-resolution imaging to recover sound from discs and cylinders <<http://www-cdf.lbl.gov/~av/>>.

The second additional topic pertaining to high resolution concerns the role of objective measurement. In imaging, this is related to the use of targets and, in audio, to standardized sets of tones. The outputs from targets or tone sets permit you to measure the performance of the equipment used to produce an image or an audio file, and the setup or adjustment of that equipment. The measurement of targets and tones does not help you evaluate actual “content” images or sounds directly.

In library and archival reformatting circles, the development of imaging targets is farther along than practices for using audio tone sets. I participated in an image-related contracting activity at the Library in 1995 and, at that time, the appropriate targets, the availability of measuring tools, and ideas about how to interpret the outcomes were not at all mature. Recently, experts have wrestled with what are called *performance measures* for digital imaging systems. You can’t necessarily believe your scanner when it says 300 ppi, we are told. Instead, we should measure what actually comes through the system. For example, use modulation transfer function (MTF) as a yardstick for delivered spatial resolution. But the process of implementing performance measures for imaging has not yet reached its conclusion. My impression is that the investigators working on this are not ready to say what the MTF pass-fail points ought to be for, say, a system

used to digitally reproduce a typical 8x10-inch negative.

On the audio side, our work group has made sound recordings of the standard ITU test sequences known as CCITT 0.33. There is one for mono and one for stereo, and both are 28-second-long series of tones developed to test satellite broadcast transmissions. With appropriate measuring equipment, recordings of the tones can be used to determine the frequency response, distortion, and signal-to-noise ratio produced in a given recording system. We have looked at the numbers but we are not yet ready to say where the pass-fail points ought to be for the equipment we might use. The recording industry may have more sophisticated or more appropriate performance measures, not well known in our circles, but I am sure that those of us working on the problem in the archive and library community will get smarter (or better informed) with time.

### **Shaping the Information Package and the Importance of Metadata**

The third central issue concerns the *information package*, a complex multipart entity. As noted earlier, the package's data takes the form of audio, video, or image bitstreams, while its metadata represents a familiar trio from digital library planning: descriptive, administrative, and structural.

In our prototyping project, our main descriptive metadata is for the object as a whole, and is often a copy of a MARC (MACHine Readable Cataloging) record in the Library of Congress central catalog. The copy is massaged to create a MODS XML record <<http://www.loc.gov/standards/mods/>>. But our complex objects often benefit from additional descriptive metadata for individual parts, e.g., song titles, artists for disc sides or cuts, or names associated with a particular element within a digital package. The descriptive metadata for these elements are encoded in what MODS calls *related items*, a kind of "record within the record." One type is called a *constituent related item*, and this fits our case very nicely.

Our administrative metadata is extensive. For example, we include a persistent identifier and ownership information, meant here not in the copyright sense but rather to identify the party responsible for managing this digital object. We include information about the source item and any conservation treatment that may be applied, data about the processes used to create the digital copy (sometimes called *digital provenance data*), and technical details about the file we have created. In the latter two categories, we have made use of sets of data elements under discussion by working groups within the Audio Engineering Society; our versions of these data sets are linked from this Web page: <<http://lcweb.loc.gov/rr/mopic/avprot/metsmenu2.html>>. Meanwhile, we have not made a big effort to collect true rights data but we do categorize objects to permit access management at the Library.

Structural metadata records the relationships between parts of objects. For example, when we reformat a long-playing record boxed set, we produce sound files for all of the disc sides, as well as images of the labels, the box, and the pages in the accompanying booklet. Thus our digital reproduction will include several dozen files and these are documented in the structural metadata. In the interface for end users, this metadata supports the presentation of the package and enables the user to navigate the various parts of the digital reproduction.

Although we have not implemented this in our prototyping, we know that there is a need for an additional category of metadata to support long-term preservation. This category is described in a helpful report from the Research Library Group and OCLC (Online Computer Library Center) titled *Preservation Metadata and OAIS Information Model* <<http://www.oclc.research/pmwg/>>. Examples of digital preservation metadata include “fixity” information, e.g., checksums to monitor file changes; pointers to documentation for file formats; and pointers to documentation of the environment required to render files.

We are encoding all of the metadata using the emerging Metadata Encoding and Transmission Standard (METS) <<http://www.loc.gov/standards/mets/>>. We worry about the extent of metadata that we wish to capture and count on the pressures of actual production to have a winnowing effect. Meanwhile, it is critical to continue the development of tools to automate the creation of metadata, especially administrative metadata.

### Longevity in “Media-less” Environment

The fourth central issue highlights the importance of *keeping* digital copies, a need that rivals and may even surpass the need to *make* the copies in the first place. This is where the repository comes in, a topic of discussion rather than a point of action for our prototyping project. Regarding the repository, our project and the planning for the new National Audio-Visual Conservation Center intersect with Library-wide digital planning—including a repository—being carried out by new National Digital Information Infrastructure Preservation Program (NDIIPP) <<http://www.digitalpreservation.gov/>>. We anticipate that the design for the Library’s repository will be consistent with the important NASA Open Archival Information System (OAIS) reference model, now an ISO standard <<http://ssdoo.gsfc.nasa.gov/nost/isoas/>>.

The OAIS model is the source of our *packaging* jargon. The model articulates a content life cycle in which a producer sends a *submission information package* to the repository, where it is ingested and reshaped to make an *archival information package*, suitable for long-term management. When an end user, called a consumer in the model, requests a version of the object for viewing or listening, the repository reshapes the content into a *dissemination information package* for presentation. We anticipate that the center at Culpeper will play the role of producer, preparing submission information packages for the Library’s repository.

During this period when the Library’s repository is under development, we place our carefully named files in UNIX file systems established in the Library’s storage area network. Although less sophisticated than the planned repository, the storage area network has an active backup system in place, a system that has sustained the eight million files from our American Memory program for several years. We keep planning improvements in our practices. For example, we now segregate our masters and service files so that a higher level of protection can be applied to the masters. For now, the METS metadata is stored as individual XML files. In effect, we are storing virtual information packages, “ready to submit.”

There is a policy implication here. Keeping digital content requires a significant information technology infrastructure, meaning both systems and people. That may be fine for larger organizations but what about smaller or independent libraries and archives? Small sound archives are clearly not in a position to mount this level of IT infrastructure. What are they to do?

There are two dimensions to this issue. Some future-oriented discussions in the NDIIPP context have suggested that there should be many libraries and archives—thought of as those who organize, catalog, and provide access to content—served by a *few* repositories—the keepers of the bits. This suggestion raised follow-up questions: How might such a many-few structure be established? Who would pay for what?

As these longer-term policy questions are being considered, there are pressing questions for today. Is there a suitable holding action for *keeping* digital content? For audio, would it be a good idea for small archives to store their files on multiple CD-Rs or DVD-Rs, or to write to data tape, as an interim solution? Ought one work in a hybrid manner, digital and analog, in spite of the extra cost? There are no authoritative answers for these difficult questions and this has impaired our ability to provide our colleagues with definitive answers.

\* Portions of this paper have been taken from a talk presented at the 2003 Preservation Conference at the National Archives and Records Administration <[http://www.archives.gov/preservation/conferences/papers\\_2003/fleischauer.html](http://www.archives.gov/preservation/conferences/papers_2003/fleischauer.html)>. This paper represents work carried out for a federal government agency and is not protected by copyright.

---

© The  
Association of  
Research  
Libraries  
**Maintained**  
by: [ARL Web](#)  
[Administrator](#)  
Site Design  
Consultant:  
[Webster](#)  
[Media](#) **Last**  
**Modified:**  
**December 15,**  
**2003**

[Proceedings](#)  
[Home](#)