Paper delivered at the
Preservation Conference: Digital Technology vs. Analog Technology
National Archives at College Park
March 27, 2003

Carl Fleischhauer Project Coordinator
Office of Strategic Initiatives
Library of Congress

1. Reformatting Audio and Video and the Motivations for a New Approach

The prevalent practice for reformatting audio and video from the 1960s and 1970s into the 1990s was, in brief, "copy to analog magnetic tape." (Meanwhile, at the Library, we retain and store the originals.) Some nuances were added in the 1990s. For video, for example, we at the Library have been making a pair of copies: one on digital videotape and one on analog videotape (two kinds of Betacam). And although we ourselves do not follow this approach, in recent years other archives have been digitally reformatting their audio materials to compact disk, sometimes in the CD-audio format, sometimes using the disks as containers for audio files, typically WAVE files.

These nuances mean that our conference title is a bit of an oversimplification and, indeed, the parameters for our current Library of Congress explorations of digital reformatting concern not only *analog* as compared to *digital (*as target formats) but also *media-dependent* as compared to *media-less.* What has motivated us to undertake these explorations in our prototyping project? What factors seemed to us to be shortcomings of the prevalent practices I just mentioned, practices that have tilted toward analog (until very recently) and are generally media dependent?

We have identified five factors. First, there is the matter of media life expectancy. Magnetic tape (analog or digital) will not last as long as media like microfilm. And sometimes we have worked with specific media that didn't behave as advertised. The Library has audio preservation copies we made in the 1970s on analog tape that now suffers from what is called sticky-shed syndrome; thus these deteriorating preservation copies must themselves be reformatted. And on the digital disk side, we all worry about the life expectancy of writeable media.

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 might be tolerable with, say, microfilm, when the time between re-reformatting is long. But with audio and video tape the time between re-reformatting is relatively short, and the adverse effects are especially troubling. Meanwhile, some video digital tape formats--Digital Betacam for example--actually conceal modest compression in the way the signal is laid down, which presumably would have a visible effect after enough re-copies have been made, although the SONY corporation insists that hundreds (dozens?) of acts of copying could occur before you saw anything.

Third, there is the problem of device and media obsolescence. On the audio side, we are seeing a virtual cessation of manufacturing of analog-tape media and analog-tape recording devices. On the video side, this takes a little different form, what we might call format obsolescence, and it plagues both analog and digital. The video signal--what goes thru the cable between devices and out to a display monitor--is very standardized. But the way the signal is actually laid on a tape tends to be proprietary and varies from system to system. And these tape systems tend to have a lifespan measured in terms of from one to three decades. For example, the U-matic 3/4-inch videotape system was very prominent in professional circles in the 1970s and 1980s; today, we are beginning to encounter difficulties in finding machines and blank tape.

There are proposals to address this issue, including the adoption by archives of media with greater permanence, and presumably the establishment of recording formats appropriate to these media. But the risk here is that this approach commits the archival community to maintaining formats and equipment that may be just as obsolescent as the industry formats being avoided.

Fourth, there is the role that the type of preservation copy plays in facilitating access by researchers. The production of digital masters makes it relatively efficient to produce service copies, e.g., streaming copies that could be put on the Web. In the case of audio, we make service copies in a post process that does not require a trained engineer to perform. (Any teenager can do it.) Making copies from analog masters is a bit more time consuming and troublesome. On the video side at this time, it's a horse race. If you stick to mastering to conventional tape--analog or digital--at the time of copying you can send a second signal stream to a second device--analog or digital--to make your viewing copy.

Sidebar: At the Library, copyright considerations mean that we must limit access to much of our recorded sound and video collections, we can't put them on the web. But since the Motion Picture, Broadcasting, and Recorded Sound Division is scheduled to move to a new building in Culpeper, Virginia, in 2005, and since we want our reformatted content to be continue to be accessible in reading rooms on Capitol Hill, we expect the digital service copies that we place in the Library's secure storage systems to help us accomplish that goal.

Finally, I think many of us have the feeling that it's time wake up and smell the coffee: the "media independent" digital era is here. We need to figure out how to take advantage of it. This feeling is reinforced by another: the next generation of content to reach us from outside our institutions will be digital and intangible to begin with, and its preservation will no doubt depend upon techniques similar or identical to those we must establish for digitally reformatted content.

2. A Tangent on Reformatting as Transformation

On the face of it, the statement "copy to tape" sounds like the statement we might use for conventional paper-collection reformatting: "copy to microfilm." But the outcome is different: reformatted sound and video (and here I am *not* referring to theatrical motion

pictures) generally recreate the original experience for the ears and eyes, the copy sounds or looks like the original. In contrast, as Nicholson Baker reminds us, no one would mistake the images of book pages in a microfilm reader (or for that matter on a computer screen) for the paper pages that we used to turn.

There are, however, transformative elements in the realm of reformatted sound and video, analog or digital. Recorded sound items are often multi-part or complex: phonograph records have sleeves or jackets with culturally significant pictures or writings, tapes may have a folk music collector's handwritten notes, etc. These visual elements are typically omitted in analog reformatting, although most archives retain the originals for consultation by researchers. Digital approaches do permit putting scanned images together with the audio in digital objects that reproduce the whole recorded-sound item, but this entails deconstruction identical to that of the microfilmed or scanned book.

The reformatting of video programs occasions a different kind of transformation, rather more subtle. Most of our historical items have what is called a *composite* signal--the type of signal used in broadcasting, which mixes color and brightness information. When these are reformatted--and this is true of both analog and digital copies--the signal is converted from composite to what is called *component* by a device with the delightful name of *comb filter,* which sorts out luminance elements from chrominance elements as the copy is made, changing the look of the image in small ways.

3. High Resolution Reproduction and Related Puzzles

In 1999, we at the Library began prototyping the digital reformatting of audio and we hope, during 2003 and 2004, to do some prototyping with video as well. Roughly speaking, our approach is to create high resolution digital files that reproduce the content elements: audio, still images, and (in the future) video streams. But determining what high resolution means and why we seek has proven a bit less self-evident than one would imagine. Our audio discussions have reminded me of the discussions of image resolution that several of us have engaged during the last decade.

One tends to start with considerations of spatial resolution for images and the corollary, sampling frequency, for sound. For the uninitiated, think of this as "how many pixels are there in every inch of your picture?" or "how many sound samples did you make every second." In both cases, the higher the number, the higher the resolution.

The starter question for imaging is: "What are the relevant features in the original and how big are they?" Identifying the smallest relevant feature in an object to be imaged gets a little tricky once we leave behind the relatively simple question of "can you read the fine print in this document?" Although you could, I suppose, move to the other end of the spectrum and try to resolve detail to the level of, say, the grain in the original photo negative. The CIA has an easier go of it: if their analysts are looking for one-meter-long bomb, their satellite images had better resolve to one meter on the ground.

The starter question for audio is: "What is the range of sound frequencies that we might expect in this original item?" What is captured by a 78 rpm disc from the acoustic era? From 8-10,000 cycles per second? On paper, you might say that if we digitally sample at a bit more than twice that frequency--let's say 25,000 cycles per second, we would capture the full range of frequencies. Or: if the folk music collector used a Nagra tape recorder, recorded at 7.5 ips with a Neuman condenser microphone, what is the highest frequency tone that we might expect to hear? That system is not likely to capture frequencies above from 12-15,000 cycles per second. So--on paper, again--if digitally sample at 44 or 48,000 cycles per second, we ought to capture the full range of frequencies.

So we make some test copies and move to the next question: "can you see or hear the difference?" A-B comparisons for imaging can be tricky and the outcome will depend in part on what device you use to "see" the image. But we do it, looking at color or gray scale images at 300 ppi, 400 ppi, and even higher. Shall we stop when you cannot see any difference? Likewise, people make and play audio examples back and forth at each other, asking "Can you hear the difference between copy A and copy B on these super-studio loudspeakers?" Once we got up at the high end, most of us could not; some could, or said they could.

In the end, our answers to these questions failed to provide the steering effect we wished for. The engineers did not want to work at 44 or 48,000 cycles but rather at 96,000, with some people eyeing 192. Their desire did not turn on the inherent fidelity of the original, nor because golden ears could hear the difference, but rather reflected ideas like the following:

- "Just in case."
- "Suppose your operator makes mistakes, won't you want a extra-data cushion to let you fix it later?"
- "There may be hard-to-hear harmonics that you won't want to lose."
- "In the future we'll have better enhancement tools and post-processing, so save as much information as you can."

For comparison, this exchange from a visit I once paid to another organization's imaging program, where they were scanning rare books in color at 400 dpi. "Can you see the improvement over 300," I asked. "No," was the reply, "but we wanted to play it safe and give ourselves a margin for error for future possibilities."

A similar refrain has come from the video expert Jim Lindner. Jim was inspired by future possibilities for indexing, that is, extracting information that would support discovery. For those familiar with MPEG-7, some of this extraction concerns what are called "low-level" features: data about colors, shapes, and sounds that might be used in the famous query "find me more like this one." Jim urged us to consider capturing high frequency information, wanting to get the pine needles on the pine trees, even when the apparent resolution on your display monitor didn't show you the needles very well.

In our prototyping, this kind of thinking has prevailed for now. The reasons for working at high levels of resolution pertain to factors that are not objectively measurable, even if you

had a measuring tool. The result is that many people tend to work at the upper limit of available technology. Digital reformatting is still an emerging practice and we won't have clarity until we have more experience.

That was sampling frequency--now, what about bit depth? My sense is that our engineers are convinced that it is worth working at 24 bits per sample--three bytes--instead of the 16 bits--two bytes--used in CD-audio and digital audio tapes (DAT). (Audio engineers sometimes call bit depth "word length.") Assuming your equipment does its job, the additional byte gives you greater precision in locating the sample point on the original soundwave, permitting the wave to be recreated more accurately. The imaging analogy is that 24 bits per pixel can represent more colors than 8 or 16 bit sampling and thus offers the possibility of greater color fidelity, assuming you do everything else right. When you talk to practitioners about this, you may also hear them express ideas like this: "Greater bit depth permits later manipulations that are less damaging to the bitstream--you will not develop gaps in your histogram [or whatever the audio equivalent of a histogram is]."

What is the role of objective measurement? In imaging, this is related to the use of targets and, in audio, the equivalent of a target--more on that below. The outputs produced by targets 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. They don't measure actual "content" images or sounds directly.

Steve Puglia of the National Archives helped us with some digital imaging projects a few years ago and joined us in assessing the state of the art. 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, imaging experts like the Eastman Kodak scientist Don Williams have wrestled with what are called *performance measures* for digital imaging systems. You can't believe your scanner when it says 300 ppi, Williams warns us. Instead, he recommends measuring what actually comes through an imaging 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.

I wish I had a better grasp of the state of the art regarding audio "targets." 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, and I am sure that those of us working on the problem in the archive and library community will get smarter with time.

I have not studied this but I have a hunch that there is a lot more engineering and science already in place for video, thanks to the fact that the video signal is inherently very complex (it needs more engineering help) and thanks to the broadcasting industry, who rely on a very wide-ranging set of standards.

We try to take performance into account in our current prototyping by using professional workers and professional equipment. For example, professional analog-to-digital convertors (the devices that actually sample the analog waveform and spit out the bits) are generally external to the computer workstation (or digital audio workstation) and are superior to "pro-sumer" a-to-d devices, often installed as a card in the desktop computer.

4. What Do We Do While We Are Waiting for Better Answers?

So what do we do in the meantime? Our prototyping project has proceeded to make files, ready to adjust our specifications as time and discoveries indicate. For now, we create pulse code modulated (PCM) files, saving them in the WAVE format. Like TIFF for images, WAVE is an open, well documented industry "standard" that is widely implemented and used. Note that it is the "PCM-ness" of the file that is important, not the "WAVE-ness." This idea can be compared to our use of TIFF files for image masters, where it is the uncompressed bit-mapping that is more important than the "TIFF-ness" of the file.

PCM sampling is fairly straightforward: take the audio waveform and sample it on a periodic basis. My friend Richard Wright at the BBC, has written, "PCM data, irrespective of sampling rate, word length, method of packing data into bytes and left-to-right or right-to-left arrangement of bits and bytes, can be decoded by relatively simple trial-and-error, and we can expect this to be the case indefinitely. PCM is in this sense a 'natural' representation for audio, and has very good long-term prospects regardless of the remaining problems of format migration."

There has been some talk about an alternate scheme for representing sound in a digital bitstream, most often associated with the Sony corporation and called DSD. It is a very high frequency one-bit stream and, to tell the truth, I don't understand how it works. DSD one-bit-deep sampling is not widely implemented, so we are taking a wait and see attitude.

As mentioned a moment ago, we produce masters at 96 kHz (kiloherz, or thousands of cycles per second) and 24-bit word length. At this time, we make two service copies: first, a down-sampled WAVE files at compact-disc specifications: 44.1 kHz and 16-bit words, and second, an MP3 file that is very handy in our local area network. And we make images of accompanying matter, like disc labels, tape boxes, and documents.

We avoid or minimize cleanup tools when making masters. And for mono discs in our collections, we copy in stereo to allow for a future process to "find the best groove wall." In principle, there is no objection to cleaning up the listening copies but we have preferred the

idea (not yet put into practice) of supplying end users in our reading rooms with software clean up tools that they can apply as they listen, to suit their own preferences. We look forward to the development of expert systems, automated tools to help us judge quality or at least spot anomalies for us to inspect later. Some are emerging from the PRESTO project organized by broadcasters in Europe.

And a word about metadata (first of three): the preceding remarks highlight two kinds of administrative information we will want to record for the historical record:

- What equipment and copying approach did we use?
- What are the technical characteristics of the digital file we created?

5. For Audio and Video it Is Digital *and* Analog, Not Digital *Versus* Analog

Audio and video reformatting demand skills and tools from both the analog and digital realms. I know that it is a little too simplistic to say that to digitize a photo, you just lay it under your digital camera or place it on your digital scanner. Items like contrasty glass-plate negatives (or any negatives for that matter) require skill, judgement, and professional equipment, as do printed halftones, still the Achilles heel of book and newspaper reformatting. But to our fevered minds, these imaging problems seem relatively manageable compared to the challenge of extracting audio from deteriorated discs or tapes. Take one of Alan Lomax's 1930s field recordings of folk music. These are instantaneous discs (meaning cut and then playable in the field), typically acetate on an aluminum base. Our team starts by cleaning the item, tricky if it is moldy or happens to be exuding palmitic acid (but of course, photos need to be cleaned too). Then--by actually playing the disc--the engineer confirms the rotational speed (generally in the absence of a recorded reference tone), determine how to set the tone arm; and uses trial and error to identify the best stylus, which varies according to the level of and type of wear on the groove. How do you know when you have your best setup, how do you know when you have extracted all the sound you can?

On the video side, 2-inch quadruplex videotapes are the poster children for playback problems: 2-inch videotape players are no longer manufactured, parts are difficult to obtain and sometimes have to be hand made, the tapes are at risk of shedding oxide as they are played, and the engineer's vigilance is needed at every moment to minimize what is called "banding," where the multiple heads render segments of the picture in slightly different ways.

Thus playing back originals to the best effect is an art and a science, and the act of digitizing inhabits the analog realm as well as the digital. Organizations like ours need skilled workers who possess both digital and analog skills. The truth is that many young people today arrive with a good familiarity with digital technology. It isn't quite yet the case that we have to explain to them what a 12-inch vinyl lp is but, for some, this antique format is not part of their personal experience. The 2-inch videotapes are of course quite beyond the realm of most people's knowledge.

And a word about metadata (word number 2): the preceding remarks highlight two additional kinds of administrative information we will want to record for the historical record:

- What are the technical characteristics of the source item from which we made our copy?
- What treatment did we apply to the original in association with the act of copying?

6. Reformatting and Preservation

Sometimes our rhetoric makes reformatting sound like it is coterminous with preservation. "I preserved that film," an audio-visual archivist might say, or "I preserved that brittle book," from a microfilming specialist. What they mean is that a copy has been made. Now these speakers know very well that the copy must be properly stored in a well-designed vault at a prescribed temperature and humidity, with attentive monitoring to be sure nothing goes wrong. But the vault is kind of in the background, its importance goes without saying.

Most of my acquaintances in the digital library community are less likely to use the word *preservation* for production or reformatting processes. For them, it is the *vault* that is front and center. We associate the term *preservation* with what we call the digital *repository.* Our focus is as much concerned with *keeping* as with *making.* NARA has an excellent, forward-looking program with a strong *keeping* element and Ken Thibodeau will discuss this later. Meanwhile, at the Library, the new National Digital Information Infrastructure Preservation Program (NDIIPP) is doing its share of pathfinding toward a repository. And these repositories will contain not only digital content that results from reformatting but also born digital content newly arrived in the institution.

What do we do while we are waiting for the repository? We use UNIX filesystems established in the Library's storage area network. Although not as sophisticated as a future repository will be, our storage area network has an active backup system in place, a system that has sustained the 7 million or so files from our American Memory program for six or seven years now. We keep trying to make 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. But we know that we still have a way to go with all of this.

And a word about metadata (word number 3): the preceding remarks suggest additional kinds of administrative information we will want to associate with our digital object:

- What are the hardware and software environments in which these files may be played or "rendered?"
- What is their history as digital entities?
- Is there data that may be used to check for things like file integrity? ·
- And more . . . .

OK, why the metadata subtext? I want to highlight the extent of metadata we wish to compile, and suggest how its capture can represent a challenge of its own. My three

comments have been limited to administrative metadata. But there is also a need for the familiar descriptive information that we used to put on a catalog card, rights-related information, and more. We have explored this need for extensive metadata as early adopters of the emerging XML metadata structure called METS--Metadata Encoding and Transmission Standard. There is information about METS at the Library of Congress website, along with other standards information provided by our Network Development and MARC Standards Office.

7. Worrying about the Need for Infrastructure and Other Policy Questions

Let me close with a musing that harks back to the repository and even to our interim use of UNIX filesystem storage. It is clear that *keeping* digital content in the manner outlined by the digital library community requires a significant information technology infrastructure, meaning both gear and people. Well, that may be fine for larger organizations like NARA and the Library of Congress, but what about smaller or independent libraries and archives? We talk to many small sound and video archives and they clearly are not in a position to mount this level of IT infrastructure. What are they to do? Is there an approach that is reasonable without requiring the full panoply of servers, backup systems, and intermittent data archiving? Is it wise at this time to work in a hybrid manner, digital and analog, in spite of the extra cost?

The provision of a technical infrastructure and the need for economies of scale is is one of the many thorny policy questions under discussion in the digital library community today, one that clearly relates to the national digital information infrastructure. Should there be *many* libraries and archives--thought of as those who organize, catalog, and provide access to content--served by *few* repositories--the keepers of the bits? Could the many archives bring digital content to certain state of readiness and then depend on the smaller number of full-service repositories for long-term preservation? How might such a many-few structure be established? Who would pay for what? My familiarity with digital library conversations suggest that there are far fewer proposed answers to questions like these than there are to the problems in technology.