Discussion Paper on Video File Formats and Wrappers
Carl Fleischhauer – Library of Congress
March, 2005

Introduction

Two of my Library of Congress colleagues--Greg Lukow and Barbara Humphrys--have passed along Sara Fuchs's AMIA-L query about film and video preservation. Although no one in our circle has a definitive response to her questions, there are a few things to say.

As you will read, my response here is "format-centric." This reflects my current work assignment contributing to this Web site: http://www.digitalpreservation.gov/formats/index.shtml. I hope that some of this commentary will be of interest to members of the listserve, but my real hope is that the AMIA-L readership will provide additional information and correct errors in my understanding. For that I will be very grateful!

Incidentally, Sara asked about the NDIIPP partnership with the public television group (WNET/Channel 13, WGBH, PBS, and NYU). This three-year project is just getting rolling and we anticipate that useful reports and presentations will begin to emerge later this year and next.

Scope for these comments

This is about intangible file and wrapper formats (and the encodings they contain), not tangible media like DVDs. By wrapper, I am thinking of archetypal examples like WAVE or QuickTime and not of "bundling" or "packaging" (OAIS sense) formats like MPEG-21 or METS. My comments are more or less at the essence level rather than at the package level, but it is the case that some wrappers are rather package-like (QuickTime, even, and AAF and MXF). There is not a hard line between wrappers and packages.

Why focus at the essence level when we know that packages will be needed in order to carry out long-term preservation in an OAIS-oriented repository? The essence emphasis today reflects the need to develop implementations at that level first--you gotta have something to put in the packages--and there is still work to be done at the essence level. In addition, there is a "meanwhile" that moving image archivists can take advantage of: various non-moving image digital library programs are in the process of establishing package formats and we can benefit from their experiences.

Discussion

Even at the essence level, there are two senses to the term format: (1) file format, e.g., QuickTime or MXF, and (2) encoding format, e.g., Sorenson 3, MPEG-4 Advanced Video Coding, or lossless JPEG 2000. In both categories today, we confront moving targets.

The Library's recent planning has included explorations of reformatting, i.e., copying an older item (usually analog) for the sake of preservation, and also developing a response to the arriving wave of born digital content. The target formats for reformatting will often (not always) also be good "preferred" formats for born digital.

Format in this context means "intangible." Physical media are a very secondary concern. You will read nothing here about tangible DVDs, hard drives, and SACDs. These are the same limits that Caroline Arms and I adopted for our Web site, cited above.

Why pick one format over another? Here are the considerations that occur to me:

1. Sustainability. Is this a format that will fit into a preservation-oriented program? Will it be viable, renderable for the long term, or does it have characteristics that lend themselves to migration or system emulation? Does it score well on the sustainability factors identified at the LC formats Web site cited above: disclosure, adoption, licensing and patent claims, transparency, self-documentation, external dependencies, and technical protection considerations?

2. Fitness to purpose, more or less what we call quality and functionality factors on the LC formats Web site.

3. Considerations pertaining to available tools. Sometimes you can't make the object you wish to because no one has provided a tool for the purpose. This finding, of course, can be a good trigger for a lobbying or tool-making activity.

4. Considerations associated with a repository implementation, especially a repository that expresses parts of the OAIS reference model. For example, if we were to use Herbert van de Sompel's repository system as developed at the Los Alamos National Laboratory, we would package content in MPEG-21 structures. There may be things about repository implementations that influence decisions about formats at the essence level.

Back to essences. For sound recordings, the tools and their outputs are reasonably well established in actual practice. By general consensus, today's preferred bitstream encoding is LPCM (linear pulse code modulated sampling), and the preferred wrappers are variants on the WAVE file, including the BWF (broadcast WAVE file, which adds space for some additional metadata). [At the formats Web site, see WAVE_BWF_LPCM: http://www.digitalpreservation.gov/formats/fdd/fdd000003.shtml (this page and the others referenced here include links to related formats)]

For video recordings and film scanning, matters are still in an exploratory phase. Regarding video, Jim Lindner's helpful report for the dance heritage project (http://www.danceheritage.org/preservation/DigitalVideoPreservation1.pdf)advocates that use of JPEG 2000 Core Coding ("part 1" of the specification), in a lossless mode, as the preferred encoding for video reformatting. As a careful reader of the literature on JPEG 2000 encodings, and with one ear to the ground regarding practices in the realm of still images, Jim's preference makes sense to me.

Some may worry because the production of JPEG 2000 video frames (about 54,000 every half hour) means accepting costs in terms of data creation and handling, transmission, and storage. But this approach is absolutely worth exploring. [At the formats Web site, see JPEG 2000: http://www.digitalpreservation.gov/formats/fdd/fdd000138.shtml]

In the video context, it is the wrapper that needs more definition.

The Dance Heritage report nominated SMPTE's Material eXchange Format (MXF) as the preferred wrapper. Specified in SMPTE 377M, MXF is an object-based wrapper (like QuickTime and others) that can embrace multiple elements, not the least of which is picture data and sound data. Interestingly, there has been some recent support for MXF from the Digital Cinema Initiative

(DCI), whose specification-still-in-the-making also highlights JPEG 2000 picture encoding (albeit presumably in a lossy mode). [At the formats Web site, see MXF: http://www.digitalpreservation.gov/formats/fdd/fdd000013.shtml And Digital Cinema Distribution Master: http://www.digitalpreservation.gov/formats/fdd/fdd000177.shtml]

The issues with MXF--insofar as I have been able to sort them out—are indicative of the cutting-edge nature of the standard. Not all of the MXF specification documents have been published; those that have been are dated 2004. In order to be placed within an MXF file, essences (bitstreams) must be "mapped to" what is called the MXF Generic Container (SMPTE 379M). The published set of MXF specifications includes Generic Container mappings for MPEG Streams (381M), DV-DIF Data (383M), SDTI-CP Essence and Metadata (385M), Type D-10 Essence Data (386M), Type D-11 Essence Data (387M), and A-law Coded Audio (388M). Waiting in the wings, still in the hands of working groups, are mappings for uncompressed video and LPCM audio. Nowhere--and I plead with readers of this post for better information--have I found mention of a work group mapping JPEG 2000 to the MXF Generic Container. Meanwhile, it is certainly the case that the implementations of MXF that one reads about, e.g., SONY's MSW-2000 series of video recorders (the video recorder with an ethernet port on the back, cheek by jowl with the co-ax), are oriented to MPEG-2.

The JPEG 2000 developers at ISO/IEC have already published their own wrapper specification for moving image frames encoded as JPEG 2000. On the face of it, this looks like it ought to be perfect for JPEG 2000 frame images. This wrapper is specification part 3 in the JPEG 2000 family, titled Motion JPEG 2000. It also object-based, a subtype of what ISO/IEC calls the Base Media File Format (BMFF), a format with a very large debt to QuickTime. The BMFF is also the basis for the jp2 file format for JPEG 2000 still images and the latest spin on the MPEG-4 file format. But I do not know to what degree the Motion JPEG 2000 specification is being implemented in tools and systems. (Comments welcome!!!) [At the formats Web site, see Motion JPEG 2000: http://www.digitalpreservation.gov/formats/fdd/fdd000127.shtml]

One interesting thing about MXF is that it is a subtype of the Advanced Authoring Format (AAF). AAF is promulgated by an independent organization, the AAF Association, and it and MXF both seem to be "on the uptake" in term of adoption. Since there is a supertype-subtype relationship, at least some of the new equipment/systems coming on the market, e.g., from AVID (I think), can produce both AAF and MXF files.

The difference in auspices (AAF Association as compared to SMPTE) seems to weigh in favor of MXF. Some have suggested that the SMPTE auspices may make it more likely that "source code" relevant to MXF file-making will remain stable for the long term in an SMPTE context, perhaps more so than the AAF Association context. [At the formats Web site, see AAF: http://www.digitalpreservation.gov/formats/fdd/fdd000004.shtml]

There is another factor that makes the MXF subtype more appealing in a preservation context than its AAF parent: the approach to essence management. MXF employs the MXF Generic Container, which is relatively device- and operating-system-independent. In contrast, AAF uses what might be called a little filesystem for the persistent object store, based on Microsoft's Component Object Model (COM) interface and Structured Storage. The AAF Association offers an open-source Software Development Kit (SDK), which bridges the COM gap to non-Microsoft operating systems like Linux, Irix, and Macintosh OS X. The AAF SDK Web site states: "On non-Windows platforms COM support is provided by [the SDK's] own portable library." But this may

mean that the COM-support elements within the SDK software are themselves an external dependency, fine in the short term but worrisome for the long term.

Film scanning is in an exploratory moment of its own. The M/B/RS film lab in Dayton, Ohio, has been working with a scanner on a getting-acquainted basis. Their explorations have brought several interesting matters to the fore. First, film scanning and digital-shooting-in-lieu-of-film raises in striking fashion the issue of dynamic range, i.e., the range of brightness in a scene or in film being scanned. Unlike video reformatting, where the dynamic range in the source is more or less a settled matter, the digital film literature is full of agonies about the desirability of 16-bit-per-channel representations, and comparisons of linear and logarithmic handling of the tonal scale.

Clearly a desirable encoding and formatting outcome would be to "master" to a format that accommodates an extended range and probably a linear rendering. Some commentators have suggested that a shortcoming of the DPX specification and DPX tools is a shortfall in support for extended dynamic range. Second, the Dayton team's initial forays encountered difficulties in finding tools that offered capabilities for certain desired formats (partly issues with color space possibilities) or that were able to manipulate files with 10-bit-per-channel data. [At the formats Web site, see DPX:
http://www.digitalpreservation.gov/formats/fdd/fdd000178.shtml]

The types of problems reported by the Dayton team are ones that can be addressed by additional experimentation and dialog with hardware and software companies. But they remind us in a vivid way of the cutting edge nature of work in this realm. Onward now, on our voyage of discovery!

* * * * *

Carl Fleischhauer
Office of Strategic Initiatives
Library of Congress
March 2005