

WP2 Note on a Workflow for the Digitization¹ and Archiving of Audio/Video Recordings

Peter Wittenburg
March 13, 2003

This note describes a possible workflow to come from audio/video recordings in DV format to media files fully integrated into an organized archive. It should be added that

- this method was applied successfully to establish an online available, browsable and searchable archive with about 3000 h of digitized audio/video material covering about 15.000 sessions (units of analysis),
- the IMDI metadata standard was used to describe the resulting resources,
- other methods may be used and
- other methods may come up with new type of technology such as for example MPEG2 camcorders or DV camcorders with MiniDisc (for such cases new workflows will be described).

We describe three diagrams that stand for different detailed workflows:

- In [diagram 1](#) the case is discussed that a researcher sends his/her audio/video for digitization to a central site.
- In [diagram 2](#) the case is discussed that a researcher does the digitization him/herself. It is assumed that the tape has a continuous time code.
- In [diagram 3](#) the same case as in 2 is described, however, there is no continuous time code on the tape, i.e. the DV camcorder is used in a non-continuous mode.

In all these cases it is assumed that the camcorder is not used for still pictures and that no mode switches (for example switch to long play) occur on one tape. The Adobe Premiere software cannot handle such exceptions and stops with the transfer. So all batch mode operations mentioned in the diagrams will not be carried out. In such cases different procedures have to be taken..

In this note we will not describe the IMDI metadata set in detail. One can find suitable descriptions for such resources on the ISLE IMDI website (www.mpi.nl/ISLE) or by looking at the MPI or DOBES corpora for concrete examples (www.mpi.nl/corpus, www.mpi.nl/DOBES). We also do not make statements about the necessity of good tape and file labels which should be entered into the metadata descriptions. The metadata descriptions can serve as glue for all sorts of related resources that is important from a management perspective. On request the MPI team can certainly give advice with respect to further details.

¹ The term “digitization” is not used fully in the correct sense, since for example Digital Video (the proprietary Sony DV format) is already a digital format. With “digitization” in this note we mean “transferring to computer”.

For older formats such as reel-to-reel audio, Betamax, Hi-8 etc special workflows have to be discussed. Some archivists such as the MPI can handle such formats. Mostly a central digitization scheme is appropriate (see workflow 1).

It should be noted here also that with respect to video encoding formats the authors of this note have the following views:

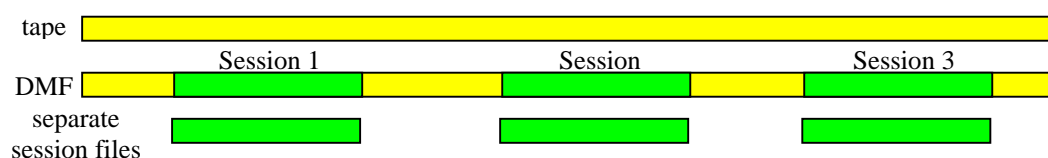
- The **DV** format from Sony is proprietary format that needs much storage capacity (30 MB/s). It can be exchanged in simple ways between camcorders and notebooks for example via the i-link interface. It is not an archiving format.
- The **MPEG2** format is accepted as an archiving format by many large institutions, i.e. it is a openly documented format, offers sufficient video quality and will be used for a number of years. The necessary bandwidth is about 6 Mbps. It is used to generate different usage formats such as MPEG1 and MPEG4.
- The **MPEG1** format is seen as a presentation format that can be transmitted for example via the current and emerging Internet channels such as XSDL and UMTS. Its bandwidth requirements are in general about 1 Mbps.
- The **MPEG4**² encoding format will increasingly often replace the MPEG1 format. It allows to select different qualities (that correspond with different bandwidth requirements) and offers a higher quality by the same bandwidth compared to MPEG1.

Therefore, all workflow schemes presented assume that the user has DV recordings and the archivist wants to have an MPEG2 copy in his archive.

Workflow 1: central transfer to computer

For the corresponding workflow diagram as a PDF file [click here](#). The steps in words are

1. The researcher sends a DV tape with a unique label to the archivist.
2. The archivist creates a Digital Master File that is a digital copy of the whole tape.
3. The DMF is sent back for further processing by the researcher.
4. The researcher analyzes the DMF to identify useful sessions, i.e. scientifically relevant fragments on the whole recording
5. The researcher creates a metadata description for each session including the exact times when each session starts and stops. At later phases of course further sessions can be identified. The procedure remains the same.
6. The researcher sends the metadata descriptions to the archivist who can create the real session movie files by an automatic script.
7. The session resources are integrated into the archive corpus in accordance with the metadata descriptions that yields an organized corpus.
8. The archivist sends back the individual sessions to the researcher to allow him/her to create annotations etc.
9. At a given moment the researcher sends the annotations with updated metadata descriptions. They include all the information that is necessary for the archivist to identify where these additional files belong. In general the archivist also sends back the original tape, except that it has special archival facilities.



This method is not efficient from the viewpoint of the researcher, since it includes a delay for working on the digitized material. However, it turned out to be very robust in terms of manageability.

² MPEG4 is a standard that describes a media object scenario, but it also includes a new encoding proposal.

Workflow 2: local transfer to computer with continuous time code

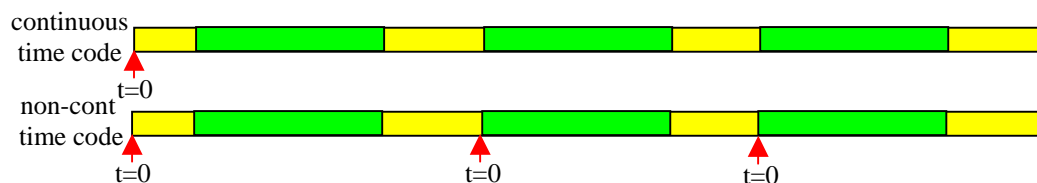
The Sony camcorders for example allow you to create a continuous time code on one tape³. This simplifies the procedures, since you always can refer back to these times.

For the corresponding workflow diagram as a PDF file [click here](#). The steps in words are

1. The researcher labels the DV tape with a unique name
2. He or she identifies the sessions, i.e. scientifically relevant fragments of the whole tape, and determines the start and stop times.
3. He or she creates metadata descriptions including these exact times. The metadata description has to be preserved for later use.
4. The researcher uses Adobe Premiere for capturing the session media segments, i.e. the program creates AVI-files that are direct copies of the defined tape segments. Premiere has an option for creating of several files from one DV tape in a batch mode. The batch file has also to be stored for later redigitizing of the tape to MPEG2 format by the archivist.
5. The researcher creates condensed mpeg1 versions of the AVI files for immediate annotation. The TMPGEnc.exe program can be used for transcoding AVI into MPEG1 files. To view the transcoding settings [click here](#).
6. The researcher can start to annotate the sessions and where necessary updates the metadata descriptions files.
7. At a given moment the researcher sends the annotation files, updated metadata descriptions, batch files and DV tapes for redigitizing the data to MPEG2 format.
8. In general the archivist sends back the DV tape, except that he has special archival facilities.

Workflow 3: local transfer to computer without continuous time code

The user of camcorders may be unaware of the digitization problems and create a not continuous time code on one tape. This complicates the procedures, since then there are more than one start point on one tape.



For the corresponding workflow diagram as a PDF file [click here](#). The steps in words are:

1. The researcher labels the DV tape with a unique name
2. He or she creates a Digital Master File that is a digital copy of the whole tape. The researcher uses Adobe Premiere for capturing, i.e. the program creates an AVI-file that is a direct copy of the whole tape. The only possibility to process tapes without continuous time code is to capture them manually, i.e. start playing and capturing at once by manually pressing buttons. The start point of capturing that is only available during the capturing phase has to be documented. It is necessary for later redigitization of the tape to MPEG2 format by the archivist and to assure that all time indexes set will be correct.
3. The researcher creates a condensed MPEG1 version of the DMF-AVI file. This can be achieved by using the TMPGEnc.exe program. To view the correct transcoding settings [click here](#).
4. The researcher looks through the MPEG1 DMF and define time points for sessions, i.e. scientifically relevant fragments of the whole recording.

³ To produce a continuous time code: every new shot follow the one before without a break; use end-search-button, functions only, if tape is not taken out of the camera yet, or press edit-search button to play back until shortly before time code-end.

5. He or she creates metadata descriptions including the exact times when each session starts and stops. Please keep the metadata description file for later use.
6. The researcher uses the M1-Edit-CL program for cutting session files from the DMF. The program M1-Edit-CL needs a batch file in a special format ([click here](#)) to carry out the cutting operation. The batch file has to be stored for later redigitization of the tape to MPEG2 format.
7. The researcher can start to annotate the created files and updates the metadata descriptions.
8. At a given moment the researcher sends the annotations, updated metadata descriptions, batch files, start points of capturing the DMF and media tapes for redigitizing the data to MPEG2 format.
9. In general the archivist sends back the original tape, except that it has special archival facilities.

Workflow 4: Digitizing of audio

Audio handling is much more standardized and simpler than video handling. So only few remarks have to be made. It is widely accepted as a de-facto standard that the WAVE format (16 bit PCM) is used. New devices supporting memory sticks or flash cards are widely used, but it should be mentioned that these devices often use compressed sound such as encoded via the MP3 or ATRAC methods. They cannot be used as archiving formats and cut away parts of sound signals that will be necessary for later usage. ATRAC (used by Sony in its MiniDisc recorders for example) is a proprietary format. So the authors strongly recommend to use recording devices that create at least 16 bit linear PCM⁴.

There are several devices (USB supported etc) that allow the user to directly add a good microphone and a high-quality analog-digital-converter and to use the notebook as the recording device. This is the simplest system. Files will be created and they have to be associated with metadata descriptions. For the sampling programs such as CoolEdit can be used for the control.

Different types of recorders support the SPDIF protocol (such as DAT recorders or the new flash memory recorders). They can be connected via the iLink interface to a notebook, i.e. stepwise sessions can be transferred to the notebook. Programs such as CoolEdit can be used to control the transfer. Again, files will be created and they have to be associated with metadata descriptions.

In the case of using analogue input channels, CoolEdit allows one to sample material session by session. Again, files will be created and they have to be associated with metadata descriptions.

MPI Services

For individual researchers and institutions that adhere to the workflow described, the MPI for Psycholinguistics can act as an archivist, i.e. house the data, register the metadata descriptions and make everything available according to the access rights policy statements. To a limited extent this includes all phases such as the digitization steps.

⁴ Many recorders such as DAT recorders create 44 kHz stereo 16 bit linear PCM signals that offer even better quality.