

## STANDARD FORMATS FOR ARCHIVING

### PREFERRED DEPOSIT FORMATS

#### Media What media format we will accept\*

Audio	.wav, .aiff, .mp3** (96khz, 24bit is our archival target; however 48khz, 24bit, or as close to the archival as possible will be accepted)
Video	.mts (AVCHD), .avi, .mov, .mpg*
Images	.tif, .jpg
Text	.txt, .xml, .pdf (.rtf and .docs should be converted to .txt or .pdf prior to submission)
Annotations	.eaf, .xml, .txt
Lexicons	.xml, .txt

#### Media What media formats are available in the archive

Audio	ARCHIVAL FILE: .wav 96khz, 24bit ACCESS COPY FOR ONLINE STREAMING: .mp3 AVAILABLE FOR DOWNLOAD: .wav, .mp3
Video	ARCHIVAL FILE: .mkv: JPEG2000 ( <a href="https://en.wikipedia.org/wiki/JPEG_2000#Motion_JPEG_2000">https://en.wikipedia.org/wiki/JPEG_2000#Motion_JPEG_2000</a> ) ACCESS FILE FOR ONLINE STREAMING: .webm: WebM ( <a href="https://en.wikipedia.org/wiki/WebM">https://en.wikipedia.org/wiki/WebM</a> ) AVAILABLE FOR DOWNLOAD: .mp4
Images	ARCHIVAL FILE: .tif ACCESS FILE: .jpg AVAILABLE FOR DOWNLOAD: .tif, .jpg
Text	ARCHIVAL AND ACCESS FILES: .txt, .xml, .pdf
Annotations	ARCHIVAL AND ACCESS FILES: .eaf, .xml, .txt
Lexicons	ARCHIVAL AND ACCESS FILES: .xml, .txt

\*If you have a format other than those listed above, please contact me so I can advise you on what we can be done.  
\*\*If you have files such as .mp3 (audio) or .mpg (video), we will certainly accept them; however, if you are collecting new recordings, please avoid these formats, as they are lossy, compressed formats.

### POST-PROCESSING OF MEDIA FILES

#### *Resampling audio files*

It may be necessary for you to do a sample rate conversion for the audio files you have recorded. This is a process that alters the size and quality of your media. Reasons for such a process are:

1. If you have not recorded at the 96 kHz, 24 bit rates, you, or PARADISEC, will need to create archival versions of your audio. This process does not add quality to your audio file; it is our way of futureproofing PARADISEC. If all archived audio files have the same structure, we can enact batch processes for migrating the materials as the technology changes.
2. You may want to work with your audio files in the field, but the 96 kHz, 24 bit file size is too large and is slowing down your elan/clan/transcriber/F4 or F5 programs. Resampling a smaller working copy would decrease the size (and quality) of your audio for easier manipulation. Remember to retain your archival version, do not write over it. Please see [RESAMPLING AUDIO USING AUDACITY](#) for more information on this.



### *Transcoding videos*

Most digital video cameras record in .avi, .mov, .mp4, or .mts (AVCHD: <https://en.wikipedia.org/wiki/AVCHD>) formats. Ideally, you should choose the highest quality format possible when recording. To further utilise your videos, you will likely need to change them (transcode) to a more usable format. As mentioned above, create a working copy when you transcode, retaining your original file to send to PARADISEC.

1. You may not be able to use formats such as .mov or .mts with your transcribing software because you do not have the correct codec installed on your computer to play the files, your transcribing software cannot open that format, or they are just too large to be useful.
2. PARADISEC will transcode all video files that are archived. This is another way we are futureproofing the archive, as well as adhering to the standards of peak archiving bodies such as the International Association of Sound and Audiovisual Archives (IASA: <http://www.iasa-web.org/>). When you send in .mov, .mp4, .mts, etc. files, we will create an industry-standard archival copy (JPEG2000), a streaming copy (currently this is a WebM file), and an access copy .mp4.
3. You may want to do some editing of one of your video files, perhaps adding subtitles, or simply create a few clips. Your editing software may not allow all formats, so you would need to transcode from one format to another.

For more information on video post-production and video workflows, see [EXTRACTING AUDIO FROM VIDEO FILE USING VLC](#), [WORKFLOW FOR VIDEO FILES](#), and [FILE TRANSFER FOR ARCHIVING AND POST-PRODUCTION PROCESSING](#).

### *Processing image files*

This is especially relevant if you are taking high resolution images of field notebooks or other textual materials that you wish to archive. If you use a camera that allows you to take RAW images or large tif files, then this is the best option for high quality photos. Canon cameras, for example, generate a .cr2 file which would then be converted to a .tif, our archival format. We automatically generate .jpgs from the .tifs you send to us. The .jpg file become the access copies (though you can also download the larger .tif files). For more information on processing image files, see [DIGITISING FIELD NOTEBOOKS](#).

*After reading through this guide, if you still have questions, or you wish to request a service, feel free to email me ([julia.miller@anu.edu.au](mailto:julia.miller@anu.edu.au)), or better, visit the **CoEDL Service Request Form**. CoEDL members use the Member login at the bottom of the CoEDL webpage. Then click the General Members tab, the link to the request form in in the left-hand panel. (<http://members.dynamicsoflanguage.edu.au/general-members/job-request/>).*