

Audio Procedures

Version 1.52

Created date:	26 January 2012
Last updated:	24 August 2017
Review Due:	24 August 2018
Authors:	Jen Mitcham, Kieron Niven, Ray Moore and Jenny O'Brien
Maintained by:	Digital Archivists
Previous version:	Live

1 Purpose of this document

This page is designed to document current ADS procedures for production of dissemination and preservation copies of digital audio files. It contains a list of current dissemination/preservation formats and how to migrate files to required formats. More information on this data type, can be found in the *Guides to Good Practice* for Digital Audio.¹

- Casey, M. and Gordon, B. (2007) *Sound Directions. Best Practices for Audio Preservation*. Trustees of Indiana University: Indiana.²
- JISC (2010a) *Uncompressed Audio File Formats*.³
- JISC (2010b) *Choosing a Digital Audio File Format*.⁴
- Knight, G. (2010) *Significant Properties Testing Report: Audio Recordings*. JISC.
- Library of Congress (2011) *Sustainability of Digital Formats*. Planning for Library of Congress Collections.⁵

¹ http://guides.archaeologydataservice.ac.uk/g2gp/Audio_Toc

² <http://www.dlib.indiana.edu/projects/sounddirections/papersPresent/index.shtml>

³ <http://www.jiscdigitalmedia.ac.uk/audio/advice/uncompressed-audio-file-formats>

⁴ <http://www.jiscdigitalmedia.ac.uk/audio/advice/choosing-a-digital-audio-file-format>

⁵ http://www.digitalpreservation.gov/formats/content/sound_preferences.shtml

2 Formats

Offered format	Accepted	Preservation	Presentation	Notes
Waveform Audio .wav	YES	Waveform Audio .wav	Compressed audio formats (preferably supplied by depositor) that are also dependent on content and requirements (see below) :typically MPEG-1 Audio Layer III .mp3	Embedded metadata should be preserved, and where none has been created we will require software used to create/encode, bitrate (kbps) and sampling frequency range (KHz). May be supplied in a compressed form using some form of codec, so be sure to check on delivery.
Broadcast Wave Format .bwf Preferred format	YES	Broadcast Wave Format .bwf	a/a	Embedded BEXT-chunk, iXML-chunk or other metadata should be preserved. May be supplied in a compressed form using some form of codec, so be sure to check on delivery.
Audio Interchange File .aif (likely from Mac users)	YES	Waveform Audio .wav , or Broadcast Wave Format .bwf	a/a	Embedded ID3 or other metadata should be preserved. May be supplied in a compressed form using some form of codec, so be sure to check on delivery.
SUN au.au	YES	Waveform Audio .wav , or Broadcast Wave Format .bwf	a/a	Where possible embedded metadata should be preserved.

FLAC .flac	YES	Waveform Audio .wav	a/a	.flac also holds metadata ⁶ , make sure this is preserved if needed.
-------------------	-----	----------------------------	-----	--

Uncompressed audio files are the most accurate digital representation of a soundwave, but can also be the most resource-intensive method of recording and storing digital audio, both in terms of storage and management. Their accuracy makes them suitable for archiving and delivering audio at high resolution, and working with audio at a professional level, and they are the 'master' audio format of choice (JISC 2010).

Uncompressed audio types are simply wrapper formats that utilise forms of bitstream recording, specifically pulse-code modulation (PCM), with the addition compatibility and system data that allows them to be used. Obviously the offset of uncompressed audio is much large file size; a 60 minute interview recorded in stereo at 24-bits and a sample rate of 96kHz, for example, will produce an uncompressed file size of 2GB. Significantly the Library of Congress advocates that “fidelity characteristics (bitstream encoding) should be used as the primary consideration; choice of file formats as secondary” (Library of Congress 2011), but realistically both should be given equal consideration. Some information on fidelity is included in the Library of Congress guidelines:

- High sampling rate (expressed as kHz), typically 96kHz or higher.
- 24-bit sampling, or higher.
- Uncompressed linear PCM codec preferred over compressed
- High data rate (e.g. 128 kilobits per second) preferred.
- Surround sound (5.1 or 7.1) encoding should only be included if it essential to content integrity. In most case uncompressed encoding in stereo should be preferred

The G2GP suggests three uncompressed preservation formats:- Waveform Audio (WAV), Broadcast Wave Format (BWF) and Audio Interchange File Format (AIF, AIFF). More recently Multichannel Broadcast Wave Format (MBWF), which “combines RF64 audio with a BEXT chunk”, has been touted as a “long-term solution suitable for archiving uncompressed 5.1 surround sound and other X.1 multi-channel formats” (JISC 2010a), but is unlikely to affect us in the near future. It may be worth noting that both WAV and BWF have a size limitation of 4GiB (in practice it is more like 2GiB).

Both WAV and BWF are the most commonly used preservation formats, although neither is essentially an ‘open’ format; whilst AIF, the lossless format used by Apple, is variously regarded as ‘open’ and proprietary. If uncompressed preservation is not possible or desired (not sure when this would be) then Free Lossless Audio Codec (FLAC) and Apple Lossless Audio Codec (ALAC) are open alternatives that provide lossless compression.

In instances where the depositor supplies only compressed versions of their audio files, they should in the first instance be strongly encouraged to send uncompressed originals. When this is not possible because the original recordings were made in a compressed form, or the original recordings have not been retained, uncompressed versions can be created from the compressed versions (usually WAV/BWF format). Obviously, such conversions are not ideal due to a loss in quality and information when compared to an uncompressed original.

⁶ <http://xiph.org/flac/format.html>

Associated metadata

It is important that any copyright information/permissions are stored alongside the requisite file in a suitable preservation format. Similarly transcriptions of audio files should be stored in a suitable preservation format.⁷

Embedded metadata

During conversion care should be taken to preserve any existing metadata found within the audio file, in an ideal world we should ask for metadata to be supplied separately in an XML or TXT form but this is not always possible. IF we receive separate metadata it should be preserved and disseminated alongside the relevant files (see below for notes on storage).

Like other file types metadata is often automatically generated on creation/editing, but the quality and nature of this metadata varies between file types. The metadata embedded in WAV files is often fairly limited, restricted to information necessary to play/use the file. In contrast BWF (an extension of the WAV format) was specifically designed to contain more extensive metadata, which is often embedded in 'chunks', typically BEXT, and more recently iXML. It is also worth noting that sound files may also be associated with project files (PTS, PTF, AUP) produced by software packages used to create/edit the file. These may also contain important metadata should be preserved.

When disseminating audio files it is generally good practice to supply any metadata found in the original file, although in some instances this may not be possible (e.g. BWF => MP3). Where this metadata is not embedded within the converted file it should be supplied separately. This action allows for more advanced reuse of the audio file.

3 Documentation / Metadata

The current *Guides to Good Practice* recommends that the following documentation is supplied with each file.

Element	Description
Software	The software (or device) used to create the file.
Bit Depth	e.g. 16 or 24 bit.
Bit Rate	Optional, often recorded as kbps
Sample Rate (KHz)	e.g. 44.1kHz
Codec Used	The codec used in creating the file e.g. FLAC or ACC.
Length of Recording	Preferably recorded in hours, minutes, seconds (hh:mm:ss).
Copyright Clearances	These are very important for audio files, especially oral histories/interviews.
Transcriptions of interviews	Transcripts of interviews can be important documentation particularly in clarifying those involved in recordings and allowing specific individuals to be identified.

⁷ see ADS Data Procedures: Binary and Plain Text.

This table is derived from the *Guides to Good Practice*.⁸

4 Accessioning checks

- Do we have the necessary documentation (see below)
- Depositors should be encouraged to supply uncompressed and compressed version of audio files. The compressed versions are to be used for web delivery.
- The necessary rights and clearances have been attained from contributors with regard to copyright
- Embedded metadata (BEXT, iXML chunks)
- Associated project files (PTS, PTF, AUP)
- Transcriptions (where appropriate)

Significant properties

The *Guides to Good Practice* recommends that the following significant properties should be maintained during preservation (after Knight 2010: 8-10)

- Duration - i.e. the length of the audio file in Timecode character format (TCF). Checks should be made to ensure that the file matches its intended length.
- Bit Depth - indicates the number of bits of information stored per sample and is an indicator of audio quality e.g. 16 or 24 bit.
- Sample Rate - an indicator of the number of samples per second, sample rate is usually expressed in hertz e.g. 44.1 kHz (a common sample rate) and is, like bit depth, an indicator of the quality of the file.
- Channels - A descriptive or numeric value that indicates the number of distinct streams within an audio object or a description of their configuration e.g. "2 (stereo)".
- Channel Assignment: Channel Number and Sound Map Location - Two values that allow specific channels to be mapped to specific output locations to configure the output sound.

When choosing a file format for delivery, the main factor is choosing a format and compression (if any) with the widest accessibility to your audience. You may find that you need to compromise on file size and quality (through compression) to effectively deliver your files. (JISC 2010b)

MP3 files provide an adequate balance between file size and quality, but such conversions are invariably 'lossy' (although ACC and OCG are also worth considering). In circumstances where lossless conversion and dissemination are required other formats may be more appropriate, for example FLAC, ALAC, etc. Other formats may be more useful when the file has specific content; Speex (OCG) and WMA (depending on codec), for example, have been specifically designed for voice/speech audio files. Dissemination format should be judged on a case-by-case basis.

5 How to convert files

Conversion of audio files can be a complicated process and this is why the best case scenario would be for the depositor to supply both compressed and uncompressed versions of audio files for both preservation and dissemination. Any file conversions should be carried out according to the strategy identified by the Digital Archivist. There are a wide range of

⁸ http://guides.archaeologydataservice.ac.uk/g2gp/Audio_3

tools available that can convert audio files; however, some care should be taken to check the settings used in the conversion process, and that any embedded metadata is preserved.

Software

The ADS has little/no software which can deal with audio files (e.g. Adobe Audition), largely due to the rarity of their deposition, however there are a number of freeware programmes available which can be utilised to create preservation and dissemination versions. SoX is an easy to use, command-line programme that has been tested by InSPECT and the National Archives⁹ and seems to provide reliable conversions (much more so than Audacity or similar). The latest versions are available through sourceforge.¹⁰ It supports most audio formats and automatically recognises the quality level of the source file and retains this in the destination format.

Unfortunately SoX does not directly support the MP3 format, consequently you will need to install an encoder/decoder .dll files that allows conversion. For encoding/writing try a web search for lame-enc.dll, libmp3lame-0.dll, libmp3lame.dll, or cygmp3lame-0.dll, whilst for decoding/reading try libmad-0.dll, libmad.dll or cygmad-0.dll. These .dll files should be placed in the same directory as SOX.EXE file. SoX is simple to use, a typical conversion looking something like this:

1. From the DOS prompt navigate to the folder where SoX is installed
2. Enter the command 'sox [file location]/audio.wav file [file location]/audio.mp3'
3. Jobs a good one!

Audacity can also be used to create dissemination versions of files, although be aware that it does not support all audio file types (e.g. BWF) and seems to take little account of embedded metadata during conversions (if you files contain no or separate metadata this shouldn't be a problem).

1. Decide on dissemination format you wish to create. Generally OCG or MP3 (although see discussion above). If you are creating .mp3 files you will need to install an encoder to export as MP3 or use the Lame option in Mediacoder¹¹
2. Open Audacity. Use File => Preferences then the File Formats tab to adjust the settings for the export format.
3. File => Open to open the file
4. File => Export as .mp3 OR File => Export as OCG to export files to desktop
5. Listen to file, check length of file and carry out other checks as documented in the AHDS Audio Preservation Handbook

Alternatively, you can use the command-line programmes outlined above, e.g. SoX.

Extracting Embedded Metadata

If the Digital Archivist has decided that embedded metadata is a significant property of a file, then this should be extracted and saved in a suitable preservation format. Many audio software packages contain metadata editors/extractors that allow the export of metadata, but there are also a host of freeware standalone tools that provide similar functionality (e.g. Meta-Extractor¹²; BWF MetaEdit¹³; JHOVE).

⁹ http://www.significantproperties.org.uk/audio-testingreport.html#_Toc226278971

¹⁰ <http://sox.sourceforge.net/>

¹¹ <http://lame1.buanzo.com.ar/#lamewindl>

¹² <http://meta-extractor.sourceforge.net/>

¹³ <http://sourceforge.net/projects/bwfmetaedit/>

- Download and unzip the Meta-Extractor files in an appropriate place (e.g. C:\Program Files\metadata-extractor). There is useful installation guide if you require further guidance¹⁴
- Open the metadata.bat file and the graphical interface will appear.
- Set the 'destination' to a meaningful location, or leave it in its default setting.
- Click 'Tools' => 'New Object' => 'Simple Object' or 'Complex Object'. You can then add files to the list of files from which the metadata is to be extracted.
- Once the list is complete click 'Tools' => 'Process', and the .xml files will magically appear in your destination folder. These can then be moved to the required location (see below).

There is useful User's guide in the documentation section of the website.¹⁵

File naming

Where possible files should retain the same name as the original. On occasion (and normally for dissemination), it may be necessary to create different versions of the same file. In these cases a logical naming strategy should be used, and should be accompanied by explanation in the Processes section of the CMS.

Extracted metadata should also be named consistently, for example//

myaudio_meta.xml

Then placed in prescribed location (see below).

6 Post-migration checking

We should ensure data consistency by undertaking a check of files post migration. The number of files to be checked is at the discretion of the Digital Archivist. Specific attention should be paid to:

- bit depth and file duration are retained and the same as the original (NB: There may be some slight discrepancy of the former when converting dissemination formats due to the inherent characteristics of lossy encoding formats that make use of variable bit depth for each sample. Whilst slight discrepancies in file duration may be consequential of the handling of milliseconds in some software packages).
- Original fidelity has been retained.
- Embedded metadata, where possible, has been preserved.

Compare the metadata for both preservation and dissemination formats to make sure the conversion has preserved the significant properties of the original file.

Storage

Data should be stored in appropriately named folders, as described in the ADS Repository Operations manual.¹⁶ Any directory structure from the SIP should be retained in the AIP. In some cases editing/restructuring may be necessary, but such restructuring should be recorded in the Processes section of the CMS.

Otherwise, store data in one of the following directory structure:

/preservation

¹⁴ <http://meta-extractor.sourceforge.net/documentation.htm>

¹⁵ <http://meta-extractor.sourceforge.net/documentation.htm>

¹⁶ <http://archaeologydataservice.ac.uk/advice/RepositoryOperations.xhtml>

```
/{original_structure}
  myaudiofile.wav
```

```
/dissemination
  /{original_structure}
    myaudiofile.mp3
```

Storing metadata

File and embedded metadata (copyrights, documentation, etc) should be stored in an appropriate archival format with the preservation/dissemination files in a "documentation" folder within the requisite file type, for example:

```
/preservation
  /{original_structure}/
    /documentation
      mytunes_metadata.docx
      mytunes_embedded_metadata.xml
```

```
/dissemination
  /{original_structure}
    /documentation
      mytunes_metadata.pdf
```

For dissemination, any embedded metadata can be left within the file, although care should be taken to make sure this metadata remains unchanged, particularly following conversion. If this metadata is supplied by depositor separately, then this should be presented with the dissemination data. For example:

```
/dissemination
  /{original_structure}
    /documentation
      mytunes_embedded_metadata.xml
```

If metadata has been extracted from an existing file the Digital Archivist should make sure that these files are named logically (i.e. relate them to the file(s) and what they contain). Any metadata extraction should be recorded in the Processes section of the CMS.