

Preserving Audio

Data Types Series

Artefactual Systems and the Digital
Preservation Coalition



DPC Technology Watch Guidance Note

July 2021



Digital Preservation Coalition

The Data Type Guidance Note Series

Each Guidance Note in the Data Types series is designed to provide a primer on the current state of community knowledge about data types commonly encountered by those seeking to preserve digital holdings. Digital preservation is about keeping information findable, usable, and trustworthy over the long-term. The best approach for any repository will vary according to the scope and content of its holdings, available resources, and the expectations of its funders and users. There are however, broadly applicable good practices that have been established as a result of many years of research, practical implementation, and consensus building. These are presented here as a starting point, along with additional resources for further exploration.

This series of Data Type Guidance Notes has been authored by staff at Artefactual Systems in collaboration with the Digital Preservation Coalition. These notes have been developed in conjunction with the UK Nuclear Decommissioning Authority.

Digital preservation is an evolving field and continues to change and develop in response to external drivers and fresh challenges. New formats, standards, and examples of good practice will emerge over time and the information contained within this report will need to be updated. We welcome comments and feedback to: info@dpconline.org.

1 Overview

Audio file formats represent sound waves that have been captured and stored in a digital format. They include digitized versions of analogue formats, such as vinyl, wax cylinder, or magnetic tape, and audio data that was recorded originally as a digital file. Audio formats are composed of a container that holds one or more streams of raw or encoded audio data. The digitization and preservation of audio data is a well-established field, with agreed good practice on how to create, digitize, and preserve audio objects.

Audio files comprise a number of technical features that can influence the quality and long-term preservation or access of the file itself:

- **Bit depth** is the number of [bits](#) contained in each audio sample. The bit depth fixes the dynamic range of an encoded audio stream ([IASA, 2009](#); [Library of Congress \(LC\), 2019](#)). The higher the bit depth, the more information that is being conveyed within each audio sample. More information being conveyed results in a higher quality and better sounding audio stream. The size of a sample is determined by the sample rate (see the section on sampling rate). The bit depth should be at least 24 in order to ensure that the digitization process captures the full dynamic range ([IASA, 2009](#)).
- **Channels** (also called **Tracks**) are independently processed or recorded audio signals organized in a sound field ([FADGI, 2020a](#); White, et al, 2011). Audio data can be recorded as or compressed down into a single stream, but it is common for audio files to contain two, eight, sixteen, or more channels that create the entire audio experience.
 - Mono audio files contain one channel, or two channels with the same content.
 - Stereo files contain two separate channels, for example one for the left side and one for the right side of headphones.
 - Audio channels can also represent surround sound. This attempts to create an aural experience with any number of additional audio channels intended for individual speakers. These speakers are then arranged in specific placements around a physical environment, or send audio streams that play back in a specific way.

Channels can influence spatial properties and determine how channels sound together when played back.

- **Compression** algorithms can be used to reduce the storage size of an audio file. They compress the information within the file through a [lossless](#) or [lossy](#) method. Lossy compression reduces quality to achieve smaller file sizes. This leads to an irreversible degradation of audio quality. Lossless compression reduces size while retaining all original information, thereby preserving audio quality.
- **Sampling rate** (or sampling frequency) is the number of samples (values or a set of values that correspond to a specific point in time ([Wikipedia, 2020a](#))) per second taken from a continuous signal to make a discrete or digital signal ([LC, 2019](#)). Frequencies are measured in hertz (Hz), or cycles per second ([FADGI, 2020a](#)).

2 Preservation challenges

Audio preservation often requires specialized knowledge about formats, carriers, equipment, and software. Broader knowledge on these topics will allow creators and preservation generalists to understand good practice for working with audio materials.

2.1 Timeliness of digitization from analogue carriers

Large volumes of audio recordings on analogue carriers are at risk of loss due to media degradation and obsolescence as hardware necessary to playback and/or digitize is no longer available or becomes prohibitively expensive ([Casey, 2015](#)). Some organizations reference 2025 as ...”the date where it will become significantly harder to digitize these collections because equipment and skilled staff will no longer be available” ([NAA, 2020](#)).

2.2 Analogue material assessment

Analogue audio material requires access to increasingly rare equipment to render the content. Even with appropriate and often expensive equipment, playback can damage the physical material and degrade content, which means that transfers must be done with care and precise monitoring. Only minimal appraisal can be carried out on physical carriers without playback.

2.3 Technical features

Audio data types have a number of unique technical features that impact their preservation (as described in the Overview above). These technical characteristics play a crucial role in authentic representation of aural experience, and therefore are significant for preservation decisions. Understanding the technical characteristics properly ensures appropriate preservation actions can be taken without losing channels, quality, or other data.

3 File formats

There is no single perfect format for the preservation and future use of audio. Decisions made on file formats should be dependent on the features and functionality to be preserved and the future use cases to be supported. Note that the tables below do not provide an exhaustive list of formats suitable for preservation and access. The most suitable format for preserving the important features and functionality of a file may be the original format that it was created in. It is recommended that careful research and analysis is carried out before migrating files to a new format.

| File format | Extensions | Brief summary |
|----------------------------------|------------|--|
| Broadcast Wave Format (BWF) | .wav | Broadcast Wave Format (BWF) is an extension of the WAVE format, standardized as EBU-TECH 3285 (EBU, 2011). BWF allows for additional metadata to be embedded into the file, making the file self-describing. This includes metadata about the original broadcast extension, quality, the MPEG audio extension, descriptive XML, and other data. BWF files are limited to 4GB of audio data. Numerous organizations consider BWF to be a preferred preservation format (LC, 2020-2021 ; NEDCC, 2020 ; NARA, 2020 ; IASA, 2009 ; ARSC, 2015). |
| Free Lossless Audio Codec (FLAC) | .flac | Free Lossless Audio Codec (FLAC) is an open lossless audio encoding standard. FLAC represents both the container and encoding. FLAC produces an audio file that is lossless in integrity but a smaller file size than uncompressed audio, without size limitation, and supports |

| | | |
|-----------------------------------|-------|--|
| | | <p>metadata tagging, embedded images (e.g. for an album's cover art), and CRC-32 checksums per frame (Rice, 2013).</p> <p>FLAC is not as widely used as WAV due to the fact that it is a newer format, is less established in a preservation context, and has less explicit tool support. However, it is considered an acceptable preservation format (LC, 2020-2021; NARA, 2020; ARSC, 2015).</p> |
| IRENE | .tiff | IRENE is a digital imaging technology that produces high-quality digital image files representing the grooved structures of analogue media (NEDCC , 2020). This is appropriate for unique wax cylinders, shellac discs, tin foils, lacquer discs, and other rare formats. |
| RF64 | .rf64 | RF64 is an extension of the WAVE format that allows file sizes to exceed 4GB (EBU , 2007). While structurally similar to WAV, RF64 is not as ubiquitous and may not be as widely supported. |
| Waveform Audio File Format (WAVE) | .wav | <p>WAVE is a wrapper format that can contain audio streams and other data, up to a maximum of 4 GB. The general recommendation is to store uncompressed audio sampled at 96 kHz with 24-bit audio sampling in a WAVE container.</p> <p>WAVE lacks the additional file header metadata supported by Broadcast Wave, but is still considered an acceptable preservation format (LC, 2020-2021; NARA, 2020).</p> |
| AAC | .aac | AAC is a lossy audio compression algorithm. It was designed to be the successor of the popular MP3, having higher sound quality while using the same bit rate (Wikipedia , 2020). |
| MP3 (MPEG-1 Audio Layer III) | .mp3 | MP3 (Standards ISO/IEC 11172-3 ; ISO/IEC 13818-3) is a popular format for storing audio data that has had huge commercial success. MP3 describes both the container and the encoding. MP3 can store data up to 16-bit depth only. |
| Ogg Vorbis | .ogg | Ogg Vorbis is an open-source audio codec (xiph , 2016). The format is considered to have better compression ratios than MP3, creating a better-sounding access derivative while taking up less space. However, it is not as commonly used or as well-known as the older MP3 format. |

4 Metadata standards

There are structural, administrative and descriptive metadata standards for both analogue and digital audio formats.

- **AES57** is an Audio Engineering Society standard for describing structural and administrative metadata for both digital and analogue audio formats ([AES](#), 2020).
- The **Broadcast Wave Format** specification EBU-TECH 3285 ([EBU](#), 2011) records metadata for: Description, Originator, Originator Reference, Origination Date, Origination Time, Coding History, Quality Report, and Cue Sheet ([Lucker et al](#), 2019). FADGI provides additional guidelines on the BWF specification with recommendations pertaining to the following metadata sets: BEXT, LIST-INFO, and XML chunks ([FADGI](#), 2019).
- **ID3** is an open metadata tagging system established in 1998 for describing audio, particularly music ([ID3](#), 1998-2020). ID3 includes tags such as title, artist, album, and musical genre. It is the de facto standard for describing MP3 files ([Wikipedia](#), 2020b).

5 Tips for creators

Whether creating born-digital audio or digitizing analogue audio collections, consider consulting an archive for advice and guidance on current good practice for creation.

5.1 Born-digital

- Keep copies of original source material and any associated files, cuts, cues, and metadata that are significant to the entire piece.
- Metadata should be embedded into the files whenever possible. This allows the objects to be self-describing rather than requiring the use of an external, proprietary system to give context and meaning to the work.

5.2 Digitized

- The Association of Recorded Sound Collections *Guide to Audio Preservation*, Section 6.1 Conversion to Digital Files ([ARSC](#), 2015) provides extended guidance on how to digitize and preserve audio files. ARSC recommends producing a primary preservation copy, a primary high-quality access copy, and dissemination derivatives created for immediate delivery and end-user access.
- Some appraisal of content prior to digitization can be performed using physical attributes, documents or metadata. Resources are available to help with initial assessment:
 - [meemoo, Flemish Institute for Archives](#) (2020), provides a questionnaire for determining specific formats based on their physical attributes.
 - AVCompass provides an educational resource on giving “direction in organizing, preserving, seeing and appreciating” audiovisual collections, from a physical inventory standpoint ([Bay Area Video Coalition](#), 2015).
- Document the physical attributes of audio materials ([ARSC](#), 2015). Understanding the original source material can provide a better understanding of when, where, and how the audio was recorded, and provide insight into the original condition when digitized.
- The recommended sampling rate for preservation audio is 96kHz, with 48kHz being the recommended minimum ([IASA](#), 2009; [Plichta and Kornbluh](#), 2002). A higher sampling rate can result in better audio quality, with more opportunities to create a larger range of audio frequencies. There are situations in which the recommended sampling rate may be lower.

For example, CDs are sampled at 44.1 kHz, and there is no value in increasing the sampling rate beyond 44.1 kHz.

- To migrate digital originals on physical media, for example optical discs, DAT, or MiniDiscs, keep files at their native sampling rate and bit depth ([ARSC](#), 2015). There is no value in trying to get a higher quality sound out of what has already been digitally created.
- If performing digitization work in-house, acquiring dedicated professional equipment to support digitization and quality assurance is necessary to create high quality digital surrogates of analogue audio media ([Casey et al.](#), 2007). Necessary equipment includes an analogue-to-digital converter, professional-level sound card, converter cables, and professional digitization software supporting 96kHz and 24-bit capability ([Plichta and Kornbluh](#), 2002). The United States Federal Agencies Digital Guidelines Initiative provides guidelines on performance metrics, cost, and impairments in its *Audio Digitization System Performance Guidelines* ([FADGI](#), 2019).
- FADGI provides guidance on embedding metadata in Broadcast WAVE files ([FADGI](#), 2021).

6 Tips for archivists

6.1 General guidance

The following resources provide guidance on preserving and providing access to audio files:

- DPC's Technology Watch Reports, *Pragmatic Audiovisual Preservation* ([Blewer](#), 2019) and *Preserving Moving Pictures and Sound* ([Wright](#), 2012).
- British Library's How to care for sound recording collections ([British Library](#), 2021).
- Association of Recorded Sound Collections *Guide to Audio Preservation*, Section 6.1 Conversion to Digital Files ([ARSC](#), 2015).
- *IASA-TC 04 Guidelines on the Production and Preservation of Digital Audio Objects* (IASA, 2009).
- UW Preservation: Audio LibGuide ([University of Washington Libraries](#), 2021).

6.2 Community assistance

Engage with the community for advice and support. Professional communities provide guidance and recommendations through published recommendations, working groups, conferences, and online groups.

- The Association for Recorded Sound Collections ([ARSC](#)) offers an annual conference, academic journal, listserv, newsletter, and preservation recommendations.
- Audio Engineering Society ([AES](#)) offers standards and technical documents, publications, conventions and educational opportunities, and committees.
- The International Association of Sound and Audiovisual Archives ([IASA](#)) offers an annual conference, academic journal, technical publications, listserv, and preservation recommendations.

6.3 Equalization

- Perform equalization (also referred to as "EQ") on audio files. Equalization balances audio frequency components and makes adjustments to those levels. This can involve adjusting overall volume, or emphasizing specific sounds that would otherwise not be as clear in the overall audio mix. When working with audio data, be it for quality control, transcoding and producing preservation or access copies, or removing "hiss" from old analogue recordings, it

is important to consider the equalization levels and adjust them when necessary. Document these changes in supporting metadata.

6.4 Multiple channels

- Understand the multiple channels in an audio file and what they represent. The total number and purpose of each channel, and how the channels relate to each other, must be considered on a per-file basis. Audio files can contain multiple channels meant to be played simultaneously (e.g. an orchestral composition), or sequentially (e.g. multiple language tracks that are meant to be played one at a time).
- Think about how multiple channels should be mapped when creating access derivatives. Single or dual channel derivatives may be preferable, depending on the file structure and content.

6.5 Quality assurance

- Plan extra time for performing quality assurance. Full quality assurance of the entire content takes longer than other non-time-based data due to the materials having a playback duration. Often, spot-checking is used instead of complete review.
- Quality assurance can be performed using the creating software or other tools available as part of a Digital Audio Workstation (DAW). If reviewing materials digitized by a vendor or coming from a donor, the free and open-source tool, QCTools ([MediaArea, 2021](#)), is another option for waveform visualization. Possible approaches for quality assurance include:
 - Visually inspect the waveform to identify signal peaks or other irregularities.
 - Random audio sampling.
 - Error detection that can be applied to each channel ([Casey, et al, 2007](#)).

6.6 Characterization

Characterization can be useful to identify file formats, extract metadata, identify broken or encrypted content, or check conformance to profiles or standards. Tool support and effectiveness can vary considerably for different file formats.

- Identify file formats with a tool such as [DROID](#), [FIDO](#), or [Siegfried](#) that uses the [PRONOM file format registry](#).
- MediaInfo ([MediaInfo, 2021](#)) can be used to assess and extract metadata from Audio files.

7 References

Association for Recorded Sound Collections [ARSC] (2015) *ARSC Guide to Audio Preservation*.

Available at: <https://web.archive.org/web/20201112003800/https://www.clir.org/wp-content/uploads/sites/6/pub164.pdf>

Audio Engineering Society [AES] (2020) *AES Standard - AES57-2011 (r2017): AES standard for audio metadata - Audio object structures for preservation and restoration*. Available at:

<https://web.archive.org/web/20200924083349/https://www.aes.org/publications/standards/search.cfm?docID=84>

Bay Area Video Coalition (2015) *AVCompass*. Available at:

<https://web.archive.org/web/20201001083015/http://www.avcompass.bavc.org/>

Blewer, A. (2020) *Pragmatic Audiovisual Preservation Technology Watch Report: Digital Preservation Coalition*. Available at: <http://doi.org/10.7207/twr20-10>

British Library (2021) *How to care for sound recording collections*. Available at:

<https://www.bl.uk/help/how-to-care-for-sound-recording-collections>

Casey, M. and Gordon, B. (2007) *Sound Directions: Best Practices for Audio Preservation*. Available at:

https://web.archive.org/web/20201127013950/http://www.dlib.indiana.edu/projects/sounddirections/papersPresent/sd_bp_07.pdf

EBU (2007) *RF64: An extended File Format for Audio*. Available at:

https://web.archive.org/web/20210112155946/https://tech.ebu.ch/docs/tech/tech3306v1_0.pdf

EBU (2011) *Specification of the Broadcast Wave Format (BWF)*. Available at:

<https://web.archive.org/web/20201112021545/https://tech.ebu.ch/docs/tech/tech3285.pdf>

Federal Agencies Digitization Initiative [FADGI] (2019) *Guidelines: Embedding Metadata in Broadcast WAVE Files - BWF MetaEdit Help*. Available at:

https://web.archive.org/web/20201016165644/http://www.digitizationguidelines.gov/audio-visual/documents/help_home.html

Federal Agencies Digitization Initiative [FADGI] (2020a) *Glossary*. Available at:

<https://web.archive.org/web/20201202215552/http://www.digitizationguidelines.gov/glossary.php>

Federal Agencies Digitization Initiative [FADGI] (2021) *Guidelines: Embedding Metadata in Broadcast WAVE Files – Audio-Visual Working Group*. Available at:

<https://web.archive.org/web/20210712135529/http://www.digitizationguidelines.gov/guidelines/digitize-embedding.html>

ID3 (1998-2020) *Home - ID3.org*. Available at:

<https://web.archive.org/web/20191005010913/http://www.id3.org/>

International Association of Sound and Audiovisual Archives [IASA] (2009) *IASA-TC 04 Guidelines on the Production and Preservation of Digital Audio Objects, Second Edition*. Available at:

<https://web.archive.org/web/20201128104430/https://www.iasa-web.org/tc04/audio-preservation>

Library of Congress [LC] (2020-2021) *Recommended Format Statements: Audio*. Available at:

<https://web.archive.org/web/20201117163339/https://www.loc.gov/preservation/resources/rfs/audio.html>

Library of Congress [LC] (2019) *Sustainability of Digital Formats: Planning for Library of Congress Collections: Sound*. Available at:
<https://web.archive.org/web/20201017182325/https://www.loc.gov/preservation/digital/formats/content/sound.shtml>

Lucker, P., Sijtsma, C., and van Veenendaal, R. (2018) *Significant Significant Properties*. Available at:
<https://osf.io/mc2ea/>

MediaArea (2021) *QCTools*. Available at:
<https://web.archive.org/web/20201207080406/https://mediaarea.net/QCTools>

MediaInfo (2021). Available at:
<https://web.archive.org/web/20210712050906/https://mediaarea.net/en/MediaInfo>

meemoo, Flemish Institute for Archives (2020) *What image and sound do you have stored in your (personal) archive?* Available at:
<https://web.archive.org/web/20201123192918/https://knowyourcarrier.com/>

NARA (2021) *Records Management Regulations, Policy, and Guidance, Appendix A: Tables of File Formats*. Available at: <https://www.archives.gov/records-mgmt/policy/transfer-guidance-tables.html#digitalaudio>

Northeast Document Conservation Center [NEDCC] (2020) *Audio Preservation with IRENE*. Available at: <https://web.archive.org/web/20200921072431/https://www.nedcc.org/audio-preservation/irene>

Plichta, B. and Kornbluh, M. (2002) *Digitizing Speech Recordings for Archival Purposes*. Matrix, Michigan State University. Available at:
https://web.archive.org/web/20201125181528/http://www.historicalvoices.org/papers/audio_digitization.pdf

Rice, D. (2013) *Flac in the Archives*. Available at:
<https://web.archive.org/web/20200809232557/http://dericed.com/2013/flac-in-the-archives/>

University of Washington Libraries (2021) *University of Washington Preservation: Audio LibGuide*. Available at:
<https://web.archive.org/web/20201204211237/https://guides.lib.uw.edu/c.php?g=342245&p=2305095>

White, G. and Louie, G. (2011) *The Audio Dictionary (REV - Revised, 3 ed.)*. Seattle: University of Washington Press.

Wikipedia (2020a) *Sampling (Signal processing) - Wikipedia*. Available at:
[https://web.archive.org/web/20210107143058/https://en.wikipedia.org/wiki/Sampling_\(signal_processing\)](https://web.archive.org/web/20210107143058/https://en.wikipedia.org/wiki/Sampling_(signal_processing))

Wikipedia (2020b) *ID3 - Wikipedia*. Available at:
<https://web.archive.org/web/20201204071027/https://en.wikipedia.org/wiki/ID3>

Wikipedia (2020c) *Advanced Audio Codec - Wikipedia*. Available at:
https://web.archive.org/web/20201220190706/https://en.wikipedia.org/wiki/Advanced_Audio_Coding

Wright, R. (2012) *Preserving Moving Pictures and Sound*. Available at:
<http://dx.doi.org/10.7207/twr12-01>. DOI: 10.7207/twr12-01

xiph (2016) *Vorbis audio compression*. Available at:
<https://web.archive.org/web/20200417172825/https://xiph.org/vorbis/>