RUcore and NJDH Standards Analysis for Audio Objects

Recommended minimum requirements for preservation sampling of audio

Introduction

This document will set forth two standard requirements for audio. One will establish a minimum and recommended sampling rate – the quality level at which the audio is digitized – for the digital audio masters and presentation copies. The second standard will recommend specific file formats for the preservation master and derivatives, for implementation into the Workflow Management System (WMS).

Although the standards will be different, the philosophy behind preservation and presentation will be same as for all other object types. It will be mandatory to archive an uncompressed archival master, to ensure an object of the highest quality is preserved. Additionally, a small but diverse number of presentation copies will be archived as well. These presentation copies are to be stored and accessible in formats that the end user will find easy to play back, and will be "low-bandwidth friendly" whenever possible, allowing users with slower internet connections to have access to these objects as well.

Sampling and Digitization Rationale

As with all other objects, obtaining a high quality sample of the original for preservation in RUcore will assure the best chance of long term preservation without having to go back to the original source for a resample in the future. This will also allow us to ensure that the presentation copies provide a comparatively high fidelity that sacrifices little in quality. In the digital realm, audio is represented by a digital sampling at a set frequency, to obtain a granular but reasonably accurate representation of the analog original. Sampling is the process of converting a signal (e.g., a function of continuous time or space) into a numeric sequence (a function of discrete time or space). The higher the sampling rate – it is assumed – the more accurate the digital representation will be.

For audio, there has been a wide practice of following the *Nyquist-Shannon Sampling Theorem*, a doctrine which is used to assert that 44.1kHz is an acceptable minimum sampling rate for all audio. This belief is based on the established fact that most human ears perceive sound up to an upper frequency threshold of 20,000Hz, and sampling must occur at twice the upper limit to achieve an acceptable digital copy. Consequently, a number of digital recordings, including CDs, adhere to this standard sampling rate (thus the term "CD Quality" is attributed to this sampling rate).

This 44.1kHz sampling rate is not without its detractors. Over time, audiophiles have consistently complained that they perceive a loss of fidelity when analog recordings are digital remastered to CD Audio. While some audio experts have insisted that these complaints are based on purely psychological factors, there is some support for a need for a higher sampling rate. There are inherent risks in losing quality to the sampling process, causing a degradation that is not accounted for in Nyquist. However, a higher sampling rate may be able to compensate for these sampling losses.

As a result, the standard set forth accounts for the CD-Audio minimum sampling rate and accepts it as a minimum, while recommending a higher level whenever the opportunity to sample at a better rate presents itself.

Recommended Standards for NJDH and RUcore Audio Sampling

- **Minimum sampling rate: 44.1kHz 16-bit (CD Audio)**
  This is the minimum acceptable rate to ensure a good preservation master. Most Compact Discs (CDs) are mastered at this rate. As such, all audio obtained from CDs will be archived at this rate. Additionally, 44.1kHz is a suitable sampling rate for RUcore partners when mastering recordings of spoken-word speech (i.e. interviews, speeches, press conferences and lectures), that are not accompanied by high-fidelity sound or music.

- **Recommended Sampling rate: 96kHz, 24-bit audio**
  This is widely considered an ideal rate for high quality audio recordings, including DVD-Audio. For most audio formats, this sampling rate is the maximum sampling rate that also supports Quad (Dolby 4.0) and Surround (5.1) audio. When repository content partners are making a first generation sample of musical or high-fidelity recordings from an analog master, it is recommended that this sampling rate be used whenever technically possible.

- **High Level (Maximum) Sampling rate: 192kHz, 24-bit audio**
  This sampling rate is often touted by audiophiles as one of the best sampling rates to work with in the editing of audio recordings and creating master samples. However, this format is generally not supported in current mass-produced formats for Quad or Surround sound. As such, recordings sampled at this rate should be limited to Mono or Stereo recordings. In general, this sampling rate, and higher rates, are recommended if there is a reasonable justification for using such a high sampling rate, and it is believed that the 96kHz rate will not be sufficient for accurate reproduction of the original sound.

Recommended File formats for preservation and presentation of audio objects

The following formats are recommended for the preservation and presentation of audio.

- **For Preservation: Standard WAV or Broadcast WAV Format (BWF)**
  BWF is an extension of the popular WAV audio format. It was first specified by the European Broadcasting Union in 1997, and updated in 2001. WAV records audio using Pulse Code Modulation (PCM), the industy standard method for digitizing audio and is used in CDs and DVDs.
  The stated purpose of these two file formats is the seamless exchange of digitized audio between different computer platforms. BWF also specifies additional metadata, allowing audio processing elements to identify themselves, document their activities, and permit synchronization with other recordings. This metadata is stored as an extension chunk in an otherwise standard digital audio WAV file.

- **No compression of archival master is recommended**
  As of this writing, the Audio and Video Standards Working Group recommends that no compression of the preservation master occur. While there are some lossless compression formats available (e.g. Shorten and FLAC), the open source formats that are currently available are not mature, nor do they have a large enough user base to justify their use. Doing so may expose the repository to the risk of being unable to later decompress and access these masters if at some point in the future, support and development for the chosen compression scheme is abandoned. However, the working group does recommend that the issue of lossless compression for archival masters be re-assessed at a later date, to determine whether an open standard is more widely accepted, likely to be readily available and supported for the foreseeable future, and suits our needs.

- **For presentation Audio: MP3 (required) and Ogg Vorbis, (optional) using Variable Bitrate (VBR) encoding**
  Both file formats are widely used by computer end users and supported by most popular audio playback hardware and software.

  MP3 enjoys wider acceptance, but is a format that is encumbered by proprietary compression algorithms. However, current licensing restrictions indicate that we would not be required to pay royalties for non-commercial, non-profit-generating use. Ogg Vorbis, while not quite as widely accepted, still enjoys support from the audiophile community and is an open source format, without any proprietary encumbrances. Given the wide acceptance of MP3 (MPEG-1, level 3), which is developed from the International Standard, MPEG-1, is the current recommended presentation format. An archive may also elect to support Ogg Vorbis, which will be reconsidered as the primary presentation format, if use of this format expands.

Evaluating collection objects that do not meet standards

The working group recognizes that there has been a period of at least two decades where digital audio has been recorded and exists prior to the establishment of these guidelines. It is important to acknowledge that there is a prevalence of digital audio objects that may be of immense value to repository partners, but for which there is no analog master available and the best digital master may not meet our established digitization standards.

In light of this, it is important to stress that the standards we have established are recommendations, and must not be the only criteria for accepting or dismissing a potential audio object. While we believe it is of the utmost importance that collection partners strive to meet the standards in order to ensure longevity of their collections, the advisory committee should consider the overall content and value of the collection before making a decision as to its inclusion. In particular, the committee may want to evaluate:

- The playback quality of the objects, and whether the audio quality can subjectively be deemed acceptable in spite of not meeting standards.
- The importance, prominence, and significance of the content
- Whether further degradation of the content can be inhibited by storing the object as an archival master, or converting an object with lossy compression into a lossless format.

If the advisory committee decides that the benefits of storing an object or collection into the repository outweigh its lack of standards compliance, then the standards can be waived for that object or collection. However, in doing so, the point should be stressed to the collection partner that long term preservation of the object *cannot* be guaranteed. While the repository and the team supporting it will put forth its best efforts to sustain the collection, the collection partner should be made aware that the chances of losing the object to format obsolescence or degradation of integrity are greatly increased because the object has not been digitized to our specifications.