CARLI Consortium of Academic and Research Libraries in Illinois

GUIDELINES FOR THE CREATION OF DIGITAL COLLECTIONS

Digitization Best Practices for Audio

This document sets forth guidelines for digitizing audio materials for CARLI Digital Collections. The issues described concern sample rates, bit-depths, file formats, and equipment for the analog-to-digital conversion of audio materials. Background information on digital audio and a bibliography of selected additional readings are also provided.

This document was created by the CARLI Digital Collections Users' Group (DCUG), standards subcommittee.

For questions about this document, please contact CARLI at support@carli.illinois.edu

Best Practices: Sample Rates, Bit-depths, File Formats

There are several standards associated with digital audio. For commercial audio CDs, manufacturers established the standard of 44.1 kHz sample rate with a 16-bit file, based on the fact that the highest frequency recordable is one half of the sample rate. This means a 44.1 kHz recording will support a top frequency of 22 kHz. Since most humans cannot hear sounds above 20 kHz, the music industry adopted the 44.1 kHz sampling rate for digital audio recordings knowing that it would capture all human-audible sound.

Digital audio for use in libraries and cultural institutions—especially for preservation-quality audio and audio digitized from an analog source—should use a higher sample rate and file bit-depth. Several specific reasons for richer files include: the accurate capture of noise like clicks, pop, and other inaudible information that resides in frequencies higher than 44.1 kHz; desire to communicate inaudible harmonic information that impact perception of sound; ability to record and provide content that, although not necessarily heard, helps listeners understand and hear better space, depth, and instrument location in stereo and surround sound recordings; and to accommodate future user applications. A higher bit-depth allows more audio information to be stored within the digital file, including support for a greater dynamic range. Based on these reasons, the current standard for archival digitization of analog recordings requires the sample rate to be 96 kHz with a 24-bit depth.

Master audio files should be saved uncompressed and in a widely-used file format, with maximum likelihood of continued support. The WAV PCM file type, developed by Microsoft, is readable by most audio software programs. Broadcast WAV (BWF) is an extension of the WAV

PCM format that includes metadata in the file header. Apple Computer's AIFF format is also widely used, although less so than WAV and BWF. Of these formats, WAV PCM or WAV BWF are preferred for long-term storage and maximum flexibility. AIFF is also acceptable.

Access copies for audio files may be saved in compressed formats that allow for quicker transfer or streaming via the Internet. The MP3 format is widely-supported, is playable on nearly all handheld devices, and is commonly used for web delivery. MP3 is the preferred format. MP3 files saved at 192 Kbps (the bit-rate) are recommended for good quality compressed audio. Windows Media, Real Audio, and Apple's Quicktime are also widely used formats for streaming audio, and format selection should based on local preference and the institution's ability to provide support for users of that format.

In summation, the following table should serve to identify the proper file format, sample rate, and bit-depth of captured audio **for the preservation master**.

Recommendation Level	**File Format**	**Sample Rate**	**File Bit-depth**	**Usage Example**
Minimum	WAV / AIFF	44,000	16	Audio from commercial CD
Special Consideration	WAV / AIFF	48,000	24	Human voice only, no instrumental music
Recommended	*WAV / AIFF*	*96,000*	*24*	*Oral History recording, especially if there is music. Natural sounds or sounds from nature.*

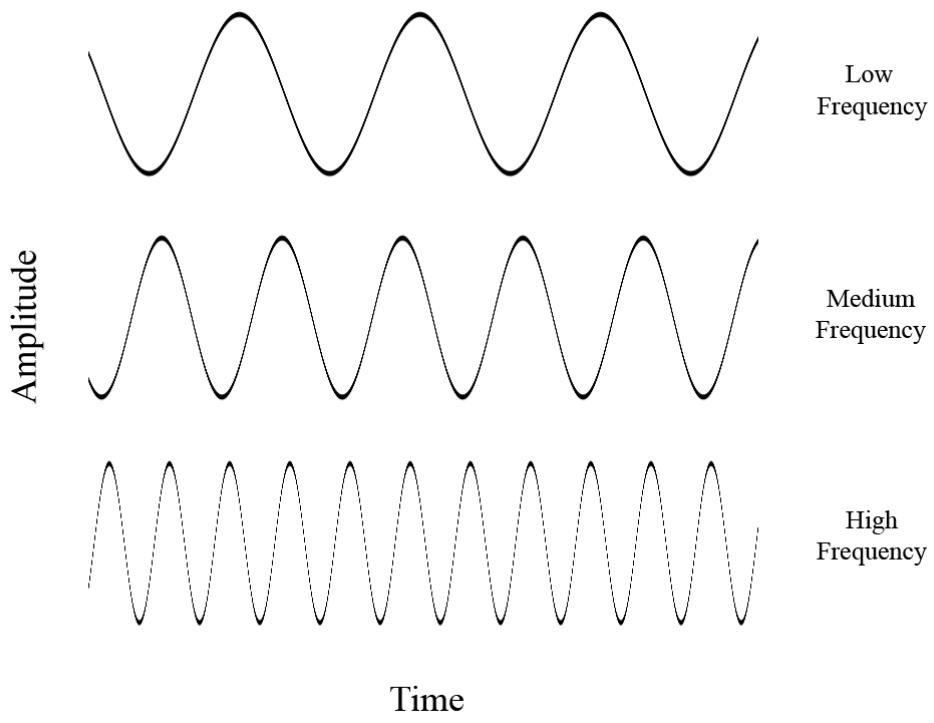
The following table should serve to identify the proper file format, sample rate, and bit-depth of captured audio **for the presentation format**.

Recommendation Level	**File Format**	**Bit rate**	**Usage Example**
Minimum	MP3	128 Kbps	Oral history, without music
Recommended	*MP3*	*192 Kbps*	*Most audio*
Special	MP3	320 Kbps	Very high fidelity required.

Background Information

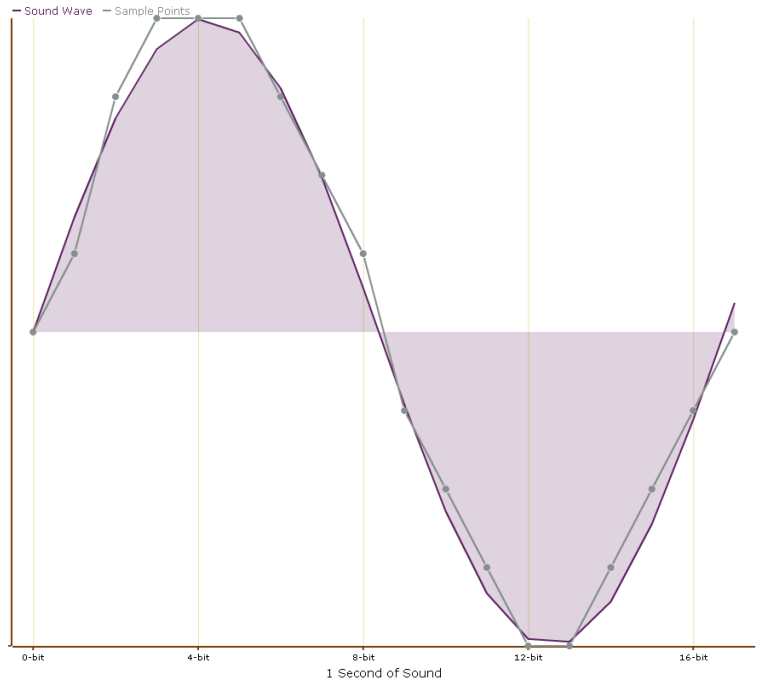
The quality and resolution of digitized audio is determined by two factors: the number of times per second the amplitude of the wave is measured (the *sample rate*) and the bit-depth used to record each measurement.

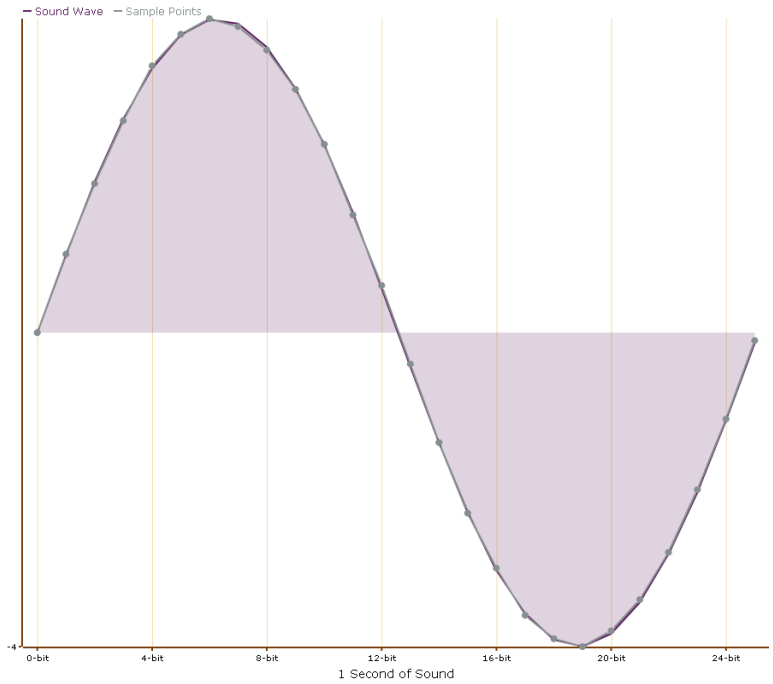
Amplitude is measured or described in kilohertz (kHz). Using the commercial audio CD as an example, the music is recorded at 44.1 kHz, meaning that each second of audio is made up of 44,100 distinct amplitude measurements in a single point of the WAVE file, be it a trough or a peak. The illustration below demonstrates how lower frequencies contain fewer waves in a specific amount of time while higher frequencies include more waves in the same period of time.



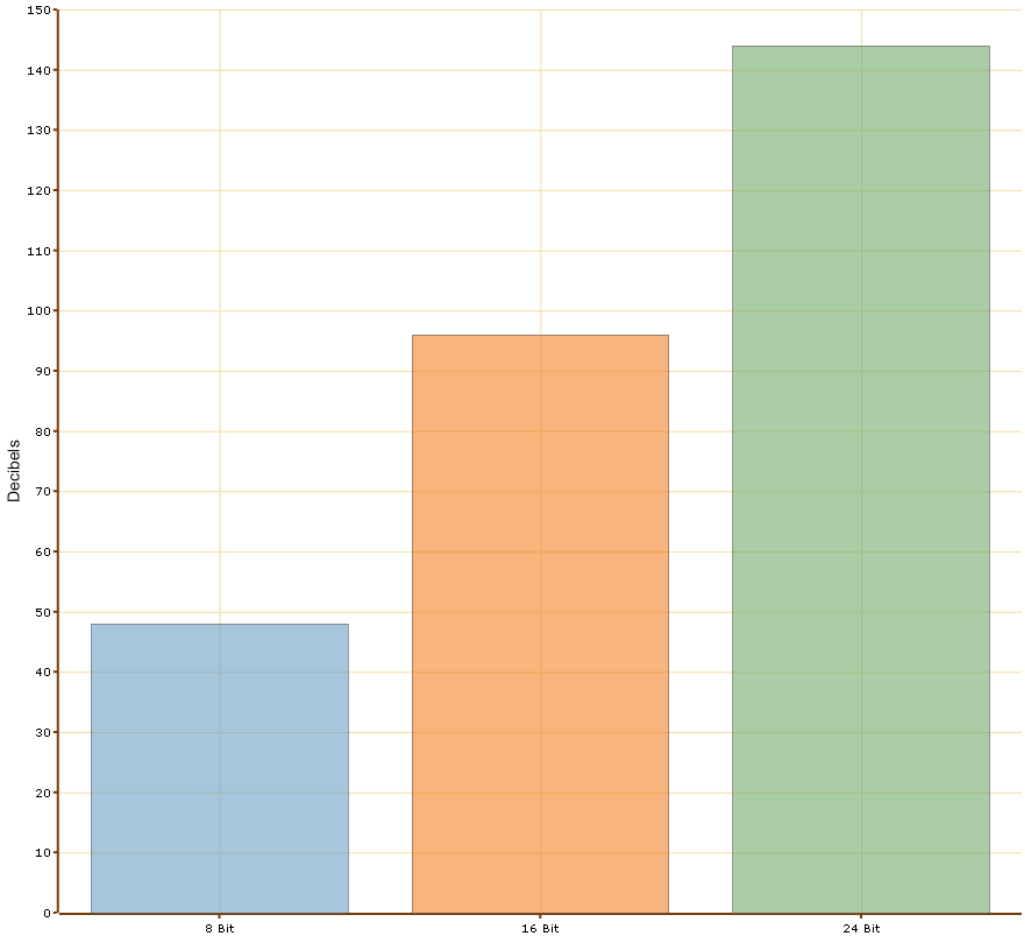
The second factor is bit depth, which refers to the quality or richness of the measurement at each of those 44,100 distinct points. If one can imagine, from left to right, margin to margin on this page, a series of waves with crests and valleys, a digital converter chooses various points along those waves to capture sound, at which point it will convert that audio information into a digital reproduction. The more points captured along each wave, the higher the bit-depth will be. Bit-depth is the number of points captured per sample, or amplitude wave. The following image illustrates how the digital capture points might be distributed when digitizing audio, or saving a

digital audio recording, as a 16-bit file and a 24-bit file. Note how there are a third more points of the sound wave captured in the 24-bit example.





Additionally, the bit-depth is mathematically proportional to the dynamic range of sound, meaning that higher bit-depths capture a greater range of sound. A 16-bit recording is only capable of capturing audio up to 96 decibels; a 24-bit recording can capture up to 144 decibels.



Higher bit-depths can represent a wider volume range, possessing a greater "dynamic range." [1]

[1] CDP Digital Audio Working Group, Digital Audio Best Practices Version 2.1, October 2006, pg. 8, http://www.mndigital.org/digitizing/standards/audio.pdf

As noted above, the quality of the digital audio on commercial audio CDs has a sample rate of 44.1 kHz and the file is saved as a 16-bit file. For preservation purposes, however, Mike Casey and Bruce Gordon state in *Best Practices for Audio Preservation* that "this [commercial audio CD sampling rate and bit depth] combination is now almost universally considered inadequate for audio preservation of analog recordings."[2] The International Association of Sound and Audiovisual Archives (IASA), the National Digital Information Infrastructure and Preservation Program (NDIIPP), and the Council for Library and Information Resources (CLIR) also recommend higher sampling rates and bit-depths because they create a reproduction that retains more information. Caroline R. Arms and Carl Fleischhauer at the Library of Congress write, "creating these 'high resolution' digital audio files is analogous to practices employed in the still image preservation world, where the term 'rich' is sometimes applied to high quality preservation masters."[3]

There are certain considerations one must take into account when creating these richer files. One, the equipment required to do so is not standard and may require additional pieces.[4] Two, the resulting files will be larger and require more digital storage space. A stereo recording that meets the CD industry standard is only one quarter the size of the same recording at the levels suggested by the archival community. The following chart shows the approximate differences in files size of a 60-minute recording at different sample rates.

Sample Rate kHz	Bit Resolution	No. of Channels	File Size Megabytes	File Size Gigabytes
44.1 kHz	16	2 (stereo)	591 mb	0.59 gb
44.1 kHz	16	1 (mono)	296 mb	0.30 gb
44.1 kHz	24	2 (stereo)	887 mb	0.87 gb
44.1 kHz	24	1 (mono)	444 mb	0.44 gb
96 kHz	24	2 (stereo)	1931 mb	1.93 gb
96 kHz	24	1 (mono)	966 mb	0.97 gb

[5]

[2] Casey, Mike and Bruce Gordon, "Best Practices for Audio Preservation," http://www.dlib.indiana.edu/projects/sounddirections/papersPresent/index.shtml

[3] Library of Congress, National Digital Information Infrastructure and Preservation Program, "Sustainability of Digital Formats Planning for Library of Congress Collections" http://www.digitalpreservation.gov/formats/index.shtml

[4]

[5] An external converter is recommended by IASA, Technical Committee, *IASA-TC 04*, 6

Audio Capture and Encoding Software

Capture and Encoding

Sony Sound Forge
Form the website: Sound Forge is a professional digital audio production suite. It is capable of capturing raw audio and converting it to web-ready formats. Sound Forge also provides advanced functionality for recording, editing, and saving audio files. Not available for Macs.
http://www.sonycreativesoftware.com/soundforge

Audacity
Audacity is a free, open source digital audio editor and recorder. It allows the user to modify audio files, apply effects, cut/delete/add segments, repeat parts, amplify or decrease the volume, import from MIDI, record live audio, and digitize recordings from cassette tapes, vinyl records, or minidiscs. Audacity is available for a Mac, Windows, or Linux platform.
http://audacity.sourceforge.net/

Adobe Audition
Like Sony Sound Forge, this is a professional, feature-rich audio capture and editing solution. It is only available for the Windows operating system.
http://www.adobe.com/products/audition/

SoundTap
SoundTap lets you record audio to either wav or mp3 format. For Windows-based machines. Formats that can be converted include streaming radio, VoIP calls and Instant Messaging conversations. It is MAC and PC compatible.
http://www.nch.com.au/soundtap/

Transcoding/Encoding Software

Note: In addition to those listed below, Sound Forge, Audacity, and Adobe Audition have the ability to save (convert/export) files in specific formats.

MediaCoder Audio Edition
Audio Edition is an open source audio transcoding tool based on MediaCoder, which provides broader support for media encoding such as video. All the audio codecs are integrated into a single piece of software. The software supports most audio formats and can convert audio streams in video files. It can also convert audio files from DVD/VCDs as well as CD formats.

Network streaming audio can be directly converted using Media Coder. It is native to Windows but can run on Mac and Linux with WINE.
http://www.mediacoderhq.com/index.htm

FFmpeg
FFmpeg is an open source, command line program that can record, convert and stream digital audio and video in numerous formats. FFmpeg is a command line tool that is composed of a collection of free software / open source libraries. It includes *libavcodec*, an audio/video codec library used by several other projects, and libavformat, an audio/video container mux and demux library.
http://ffmpeg.mplayerhq.hu/

Audio Playback Software

Note: These are in addition to popular playback software such as Windows Media Player, Apple iTunes, and Quicktime that come preloaded with most operating systems.

VLC Media Player
VLC Media Player is an open source media player with wide support for all media (both audio and video) formats. It is available for Windows, Mac, and Linux.
http://www.videolan.org/vlc/

MPlayer
MPlayer is a free and open source media player. The program is available for all major operating systems, including Linux and other Unix-like systems, Microsoft Windows, and Mac OS X. Versions for OS/2, Syllable, AmigaOS and MorphOS are also available.
http://www.mplayerhq.hu/

Bibliography of Selected Readings

CLIR (Council on Library and Information Resources) Reports

Publication 128: Survey of the State of Audio Collections in Academic Libraries
http://clir.org/pubs/abstract/pub128abst.html

Publication 135: Copyright Issues Relevant to Digital Preservation and Dissemination of Pre-1972 Commercial Sound Recordings by Libraries and Archives
http://clir.org/pubs/abstract/pub135abst.html

Publication 137: Capturing Analog Sound for Digital Preservation: Report of a Roundtable Discussion of Best Practices for Transferring Analog Discs and Tapes
http://clir.org/pubs/abstract/pub137abst.html

E-MELD (Electronic Metastructure for Endangered Language Data) Classroom, Digitization of Audio Files
http://emeld.org/school/classroom/audio/index.html

Federal Agencies, Digitization Guidelines Initiative, Audio-Visual Working Group
http://www.digitizationguidelines.gov/audio-visual/

Harvard Library, Weissman Preservation Center, Guidance for Digitizing Audio
http://preserve-new.harvard.edu:8883/guidelines/audiodig.html

Historical Voices, Digitizing Speech Recordings for Archival Purposes
http://www.historicalvoices.org/papers/audio_digitization.pdf

Indiana University Digital Library Program, Sound Directions, Digital Preservation and Access for Global Audio Heritage
http://www.dlib.indiana.edu/projects/sounddirections/

Indiana University Digital Library Program, Variations3, Administrator's Guide – Audio Digitization
https://wiki.dlib.indiana.edu/confluence/display/V3/Administrator%27s+Guide+-+Audio+Digitization

iasa (International Association of Sound and Audiovisual Archives), The Safeguarding of Audio Heritage: Ethics, Principles and Preservation Strategy
http://www.iasa-web.org/de/safeguarding-audio-heritage-ethics-principles-and-preservation-strategy

JISC Digital Media, Audio Advice
http://www.jiscdigitalmedia.ac.uk/audio/

Library of Congress, NDIIPP (National Digital Information Infrastructure & Preservation Program) Sustainability of Digital Formats, Sound
http://www.digitalpreservation.gov/formats/content/sound.shtml

Multimedia Data information from Cardiff School of Computer Science course CM0340 by David Marshall
http://www.cs.cf.ac.uk/Dave/Multimedia/node141.html

NINCH (National Initiative for a Networked Cultural Heritage) Guide to Good Practices, Audio/Video Capture and Management
http://www.nyu.edu/its/humanities/ninchguide/VII/

NISO, A Framework of Guidance for Building Good Digital Collections
http://www.niso.org/publications/rp/framework3.pdf

University of Texas, College of Education, Digitizing Sound Files
http://www.edb.utexas.edu/multimedia/PDFfolder/DigitizingSound.pdf

Wellesley College, Library & Technology, Audio Editing
http://new.wellesley.edu/lts/techsupport/macs/itunesaudio

Appendix

Sample Workflow

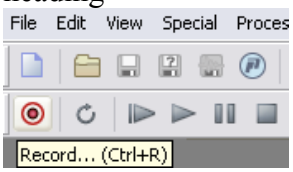
This sample workflow will demonstrate how to apply the audio digitization standards and best practices outlined in this document. For example purposes, the audio to be digitized is an oral history recorded on cassette tape in 1992. The interviewee is a jazz musician discussing his experience working with Miles Davis in the 1970s. There are musical excerpts.

This section uses Sony Sound Forge as the audio capture and export software.
http://www.sonycreativesoftware.com/soundforge

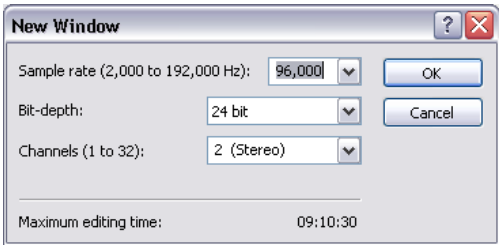
It is possible to divide the workflow into two sections:

Creating the Digital Audio Master

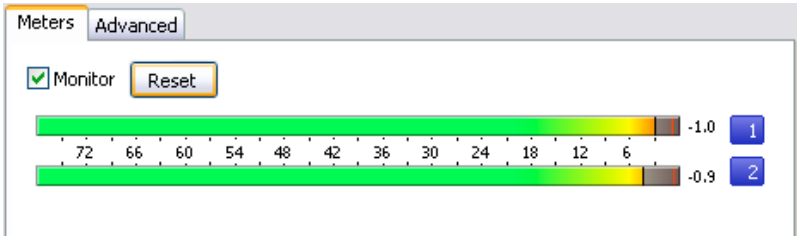
1) Connect cassette player to computer.
    a. Software utilized: Sony Sound Forge.
    b. Hardware utilized: Cassette player and stereo audio cable to connect from the cassette player's headphone jack to the computer's microphone jack. Ideally, the computer used to capture the audio would have a separate sound card, which would increase fidelity to the original. A high quality cassette player would also be preferable.
2) Open Sound Forge.
3) To record sound, click the button that looks like a red bulls-eye underneath the File heading



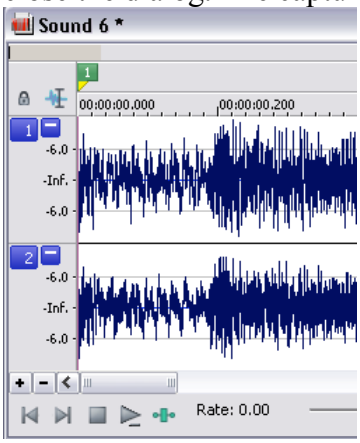
4) A window will appear that will permit you to select the recording setting ("recording attributes") and check the audio levels.
    a. To alter the recording attributes, click "New…" at the top of the "Record – Sound" dialog.
    b. Set the "Sample Format" to 96,000.
    c. Set the bit-depth to "24 bit."
    d. Set Channels as appropriate (probably "2" – stereo).
    e. Hit OK.



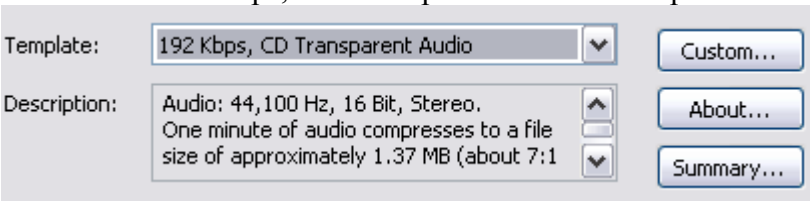
5) Check the recording levels.
   a. Set volume on the playback device (the cassette player) to its highest setting and playback the audio.
   b. Adjust the playback volume until the playback meters for each channel peak just before the right side (0). Be sure to evaluate the loudest section of the audio original by this method.



6) Once the playback levels are set properly, you can begin recording by clicking the record button—the red bulls-eye—on the open "Record – Sound" dialog, above the playback meters pictured above.
7) Once recording is complete, hit stop on the still-open "Record – Sound" dialog and then close the dialog. The captured audio will appear below.



8) Trim any dead air from the beginning and end of the recording.
9) "Save as…" and select "Wave (Microsoft)." This will save the file at 96 kHz and 24-bit, per the "recording" settings.

Creating the portable audio file

10) Open the WAV file in Sound Forge if the captured audio project is no longer open.
11) "Save as…" and select "MP3 Audio." File -> Save As…
12) Select the "192 Kbps, CD Transparent Audio" Template. Name the file and hit save.