Iowa Heritage Digital Collections
Preservation
University of Iowa Libraries
Iowa City IA 45540

IHDC Participants

Iowa Heritage Digital Collections

Iowa Heritage Digital Collections

A Guide For IHDC Participants

Beginnings

Digital Audio Collections

2005-6 Edition

DIGITIZING AUDIO

- Why should I digitize?
- How do I digitize?

What is the IHDC?

The Iowa Heritage Digital Collections (IHDC) is envisioned as an online collection of Iowa history and culture created by bringing together in digital form documents, images, maps, finding aids, interpretive and educational materials, and other media from collections held by a wide range of organizations throughout Iowa. To view the present collection, visit our website: *http://iowaheritage.org*. For information about joining the project, contact *nancy-e-kraft@uiowa.edu*.

Why should I digitize?

Carl Fleischhauer's Reasons (Library of Congress)—

- Analog media has a more limited life expectancy
- Quality is lost with each generation of analog-to-analog preservation
- Few analog machines (e.g. record players, reel-to-reel) are being produced/serviced ("hardware and media obsolescence")
- Easy access (Digital media is easily distributed, e.g. over the internet, and thus more available to researchers. If it's more accessible, it will be more useful.
- Future collections will be digital.

This readable and thorough article explains how discs are constructed, the formats available, and the advantages/disadvantages of each. There is also a helpful glossary and bibliography at the end. Even if new formats are introduced, this is still an essential article.

Bibliography:

- Association for Recorded Sound Collections (ARSC), *Education and Training in Audiovisual Archiving and Preservation, http://www.arsc-audio.org/ETresources.html Contains a list of resources, including listserves.*
- Colorado Digitization Program, *Digital Audio Best Practices and an Introduction to Digital Audio, http://www.cdpheritage.org/resource/audio/std_audio.htm Here are two pdf documents that include useful technical guidelines.*
- Conservation Online (CoOl), Stanford University Libraries, *Audio Preservation. http://palimpsest.stanford.edu/bytopic/audio/ Contains reference for preservation history and practice.*
- *Free Online Dictionary of Computing (FOLDOC) http:// wombat.doc.ic.ac.uk/foldoc. This is a place to look up those terms you've forgotten or have never heard of before you began to digitize your library's collection.*
- The National Initiative for a Networked Cultural Heritage [NINCH], *The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials (2002) http://www.nyu.edu/its/humanities/ninchguide/index.html See especially Chapter VII on "Audio/Video Capture and Management."*
- Proceedings from the Symposium, (7-2003), *Sound Saving: Preserving Audio Collections, http://www.arl.org/preserv/sound_savings_proceedings/ Contains many full text presentations (on copyright law, reasons for preservation, a history of recording formats et al.) as well as a list of resources.*
- *Wikipedia http://www.webo pedia.com This is another source for finding out what a word means. Generally this encyclopedia is accurate, readable and extremely useful. Keep in mind, however, that anyone may contribute to its information base (even you !).*

Tape is made of an oxide layer emulsion (iron oxide + binder; in other words rust + glue) PLUS a base of one of the following:

Acetate *(translucent) Common Problems: grows/shrinks with humidity; brittle; vinegar syndrome (vinegar odor as tape breaks down, along with distorted sound and "stickiness." Note that it can be 'treated' if caught early also note that it will 'spread' to surrounding tape areas); lubricant failure.*

Polyester = "mylar" *(opaque) Appearing in the mid 1960s, this solved the problems with acetate tape—it flows nicely, is easy to cut and splice, stable in humidity, lubricant failure rare, and doesn't break (it "stretches—but note that this damages the tape as well). It also has much higher fidelity recordings than acetate. Common Problems: "Sticky Shed Syndrome" Caused by moisture (humidity) interacting with the glue that holds the oxide to the polyester base and causes tape to stick to machine on playback. If You hear a "screeching" sound, stop immediately for permanent damaging is occurring. Contact us for advice on this problem.*

Paper *The earliest tapes, very rare, very delicate.*

Philips Compact Audio Cassettes: *Cassettes use Polyester tape. Common Problems: most are mechanical (loose pressure pad, roller failed, shell malformed, etc). Note that "shells" can be changed to solve these problems.*

Digital Mediums

The place to start when considering all aspects of digital production—migration, preservation, etc. is Fred R. Byers' *Care and Handling of CDs and DVDs A Guide for Librarians and Archivists* (October 2003) which is available online in pdf form at the IHDC website: *http://iowaheritage.org/site-templates/digitization.html.*

How do I digitize?

Getting Started—Plan and Consider:

- **The first collection:** Choose a small, coherent collection to begin. Consider your local research demands, the historical significance of the material and the condition of the material. You may want to arrest the decline of some materials by digitizing them first.
- **Media and Hardware**: Choose how you want to deliver the new digitized records—via CD, DVD, or streamed from a website. Then consider what sort of storage media you'll need in order to accomplish your goals. IHDC recommends gold single sided CDs (and be sure to write only on the inner circler lest you introduce inks which will cause media degeneration); and QuickTime streaming. You will need good analog players too—perhaps a turntable or reel-to-reel tape player capable of various speeds. Contact us with your specific questions.
- **Metadata:** Consider what sort of information you want to keep about your items. IHDC has created data dictionaries to guide you [see *http://iowahertage.org*]. Certain information is required in order to facilitate cross database sharing and searching. Some metadata, however, is intended for 'in house' use only. You will have to decide, by following the guidelines in the data dictionaries which fields you will need and whether they should be hidden fields, searchable fields, or fields governed by a controlled vocabulary.
- **Software:** The software selected for the Iowa Heritage Digital Collections is CONTENTdm which is 100% Web compatible. Servers and collections can be administered from remote desktops; while item and metadata entry can occur at as many as 50 remote locations using the Acquisition Station. Collections are readily accessible to end users through standard Web browsers. You may need specific software

to turn your analog collection into a digital collection before it can be shared online with CONTENTdm. There are many freeware or inexpensive shareware applications available. Contact IHDC if you're uncertain about how to proceed.

Common Digital Formats:

—*Uncompressed*

Note — IHDC recommends that you create and keep an archival copy in the uncompressed WAVE format selecting one of the following specifications:

- *96kHz sample rate, 24-bit resolution is archival*
- *44.1kHz sample rate 16-bit resolution is CD quality*

In order to share your digitized recording over the internet, you will create a compressed copy (see the next page for format choices). Keep the WAVE file though. When you create a "record" in the Contentdm work station, link your metadata to the compressed (and/or streamed) file that you have generated from the WAVE file.

- **AIFF/RIFF (.aif, .rif) File** (Audio Interchange File Format/ Resource Interchange File Format): Uncompressed audio format developed by Apple (AIFF) and Microsoft (RIFF). These are large binary files built from multiple data structures called "chunks." The quality is high.
- **WAVE (.wav) File** (Waveform Audio File Format): A high quality file format developed by Microsoft. It can be used on most platforms (including Macs). File sizes are very large.
- There are other, less often used, file types (e.g. Liquid Audio; .au; .paf; .sf, .voc), and we can help you find information about their utility. A Google search, though, will probably bring up an excess of information on these more esoteric types, or try the *Wikipedia http://www.webopedia.com or the Free Online Dictionary of Computing (FOLDOC) http://wombat.doc.ic.ac.uk/foldoc.*

—*Compressed* *[used for web delivery]*

- **MP3 (.mp3):** Originally part of the MPEG (Motion Picture Experts Group) movie format in which the third tract was the audio track. MPEG audio is an open standard family for compressed audio that

Typical Analog Formats

Analog Discs— *breakable but otherwise pretty stable, so from a preservation stand point perhaps not your first priority (although access may be enhanced by digitization).*

- **78 rpm**

Shellac (commercial discs) **,** various speeds (70.29 to 80 rpm), various groove/stylus sizes (1.0-3.0 mil or larger), center hole not centered, usually 10", usually less than 3 minutes playing time. *Common Problems: Speed & sizes not standardized; centering off (and hence the familiar wobbily sound), easily broken*

Acetate (recording discs)**,** various speeds (70.29 to 80 rpm), various groove/stylus sizes (1.0-3.0 mil or larger), center hole is centered (there may be extra holes and the label may be hand written), the base is aluminum / glass with a recording layer (acetate laminate), common sizes 10", 16"- *Common Problems: delamination, palmitic acid (caused by separation of the plasticizer (castor oil) in the acetate coating—contact us for cleaning information), glass based discs break*

- **LP or long playing records** ("MicroGroove"—this very small groove (0.8 mil or thousandths of an inch) is what defines an LP
- **PVC** (flexible); usually 12" (or 7" for "45s"), either 33 1/3/ or 45 rpm; 15-20 minutes playing time *Common Problems: a "less robust" medium.*

Analog Tape *—comes in*

many speeds –30 ips, 15 ips, 7 1/2 ips, 3 3/4 ips, 1 7/8 ips, 15/16 ips (note how higher speed is derived by doubling)

many sizes: 14", 12", 10 1/2", 7", 5", 4", 3", 2" —play length derived from speed, reel size, tape thickness (1.5 mil, 1.0 mil, 0.5 mil) and track format (full track, half track, quarter track, multi-track).

Rights Issues

"Rights complicate access, but rarely impact preservation."

Keep in mind what George Blood said in his presentation at the 2005 ALA Annual Meeting:

> *"...let me touch briefly on the issue of Intellectual Property and Copyright. Ambiguities on rights should not be a primary factor in preservation. If you have endangered audio recordings (or video, or paper), if you wait 'till the rights are cleared up, you may not have anything to preserve, or the cost of preservation may have increased significantly. Rights issues may influence access, but if you have a mandate to preserve the recordings, rights on use & access do not eviscerate your mandate on preservation. If you have a highly restricted object (say letters of a famous author), a user may have to jump through hoops to be permitted to view it, but you want that information to be available to them."*

IHDC has several sample rights statements on its website at *http://iowaheritage.org/site-templates/digitization.html*. You may use one of them, or modify one to fit your own situation, in the "required" rights field on your metadata template. Determination of copyright is the responsibility of the institution holding the collection, not of IHDC.

includes MP2, MP3 and AAC. These files are highly compressed (10 times smaller than WAV files) but, when encoded at levels above 128 kbps, play back "at near-CD quality." (Note that MPEG-1 is slower loading.) These files, due to their size, can easily be streamed over the internet (though some argue that RealAudio and Windows media are better). Most software programs (RealPlayer, Winamp, etc.) and an increasing number of CD players recognize MP3 files. There are several proprietary formats based on MPEG audio—such as MP4 (version 2, 1999) which is favored by some because information (such as album cover graphics) can be embedded.

- **Ogg Vorbis (.ogg):** This open source file format begun in 2000 supports various compression rates. The fact that it is a non-proprietary format and, unlike most formats, includes an extensible (XML) metadata schema (so potentially good for libraries) and supports multichannel compression. Thus, it's growing in utility and popular support.

- **QuickTime (.qt, .mov):** Developed by Apple, the QuickTime format (like Windows Media and RealPlayer) is often used for streaming content over the Internet. It utilizes a proprietary compression scheme and requires the free and widely available QuickTime player. The quality is excellent and it may be used for large high resolution productions. Like the others it handles mpeg content.

- **RealAudio:** (.rm) RealAudio is an often seen file format developed primarily for streaming media over the internet. It utilizes a proprietary (and lossy) compression algorithm (10:1) and requires free RealMedia software for playback. The sound quality is adequate.

- **Windows Media:** (.wma) Similar to RealAudio and QuickTime, Windows Media was developed by the Microsoft Corporation to provide a method of streaming content (both audio and video) online. It uses a proprietary compression scheme and requires Windows Media software to be played back. High quality.

A summary of the main contenders for web delivery [2005]

Format	*Optimized for Streaming?*	*Compression Codec*	*Pros and Cons*	*Format*
Mp3 (.mp3)	Primarily for download, can be streamed	MPEG-2, Level 3 compression	• Widely used format - supported by many media players • Proprietary codec, low bit rates result in poor audio quality	*Mp3 (.mp3)*
Real Audio (.ram, .rm)	Yes	Proprietary - RealAudio, Inc. developed codec	• Primarily used for streaming audio content • Proprietary codec, requires a free RealAudio player, poorer quality	*Real Audio (.ram, .rm)*
Windows Media	Yes	Proprietary – Microsoft develope codec	• Primarily used for streaming audio and video content • Proprietary codec, requires a free Windows Media player	*Windows Media*
Quick Time	Yes	Proprietary – Apple developed co	• Used for streaming (and high quality) audio and video content • Proprietary codec, requires free QuickTime player	*Quick Time*
Ogg Vorbis	Either download or streaming	Open Source – free codec	• An alternative to proprietary media formats with potential to develop features useful to libraries • Accepted by few media players	*Ogg Vorbis*