

*Penn Libraries*

*Scholarship at Penn Libraries*

---

*University of Pennsylvania*

*Year 2006*

---

Digital Audio in the Library

Richard Griscom

University of Pennsylvania, [griscom@pobox.upenn.edu](mailto:griscom@pobox.upenn.edu)

Unpublished, incomplete book manuscript. Version 1.11 (19 July 2006)

<http://richardgriscom.com/da> HTML version is also available.

This work is licensed under the [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License](http://creativecommons.org/licenses/by-nc-sa/3.0/).

To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/> or send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.

This paper is posted at ScholarlyCommons.

[http://repository.upenn.edu/library\\_papers/5](http://repository.upenn.edu/library_papers/5)

# Digital Audio in the Library

Richard Griscom

June 2006  
version 1.11 [19 July 2006]

# Contents

<b>I</b>	<b>Digital Audio Technology</b>	<b>8</b>
<b>1</b>	<b>Fundamentals of Recorded Sound</b>	<b>10</b>
1.1	Analog and Digital Sound . . . . .	10
1.2	Capturing Sound . . . . .	11
1.3	Reproducing Sound . . . . .	12
1.3.1	Speakers . . . . .	12
<b>2</b>	<b>Fundamentals of Digital Audio</b>	<b>13</b>
2.1	Concepts . . . . .	13
2.1.1	Sampling . . . . .	13
<b>3</b>	<b>Formats for Digital Audio</b>	<b>16</b>
3.1	Terminology . . . . .	16
3.2	Structure of an audio file . . . . .	17
3.3	Open and proprietary formats . . . . .	17
3.4	Uncompressed formats . . . . .	18
3.4.1	PCM . . . . .	18
3.4.2	Sun/NeXT (extension .au or .snd) . . . . .	19
3.4.3	AIFF (extension .aiff or .aif) . . . . .	19
3.4.4	WAVE (extension .wav) . . . . .	19
3.5	Compressed formats . . . . .	20
3.5.1	“Lossless” Compression . . . . .	20
3.5.2	Open formats . . . . .	21
3.5.3	Proprietary Formats . . . . .	21
3.6	“Lossy” Compression . . . . .	22
3.6.1	MPEG . . . . .	22
3.6.2	Proprietary Formats . . . . .	24
3.6.3	Open Format . . . . .	26
<b>4</b>	<b>Delivery to the Listener</b>	<b>27</b>
4.1	Accessing Digital Audio Files . . . . .	27
4.1.1	Downloading (by John Anderies) . . . . .	27
4.1.2	Streaming . . . . .	28
4.1.3	Peer-to-Peer (P2P) networks . . . . .	29

4.1.4	Podcasting (by John Anderies)	30
4.2	Software for Playback	31
4.2.1	Common Audio Players	32
<b>5</b>	<b>Hardware and Software to Support Digital Audio</b>	<b>36</b>
5.1	Equipment	36
5.1.1	Encoding workstation	37
5.1.2	Audio components	39
5.1.3	Servers	41
5.2	Software (by John Anderies)	43
<b>II</b>	<b>Digital Audio in the Library</b>	<b>45</b>
<b>6</b>	<b>Sources for Digital Audio Recordings</b>	<b>46</b>
6.1	Digitizing Library Local Collections	46
<b>7</b>	<b>Curricular Listening</b>	<b>48</b>
7.1	Streaming Service	49
7.1.1	Creating and Managing the Sound Files	49
7.1.2	Encoding	49
7.1.3	User interface	55
7.1.4	Staffing	55
<b>III</b>	<b>Appendix</b>	<b>57</b>
<b>A</b>	<b>Resources for Questions on Copyright</b>	<b>58</b>
<b>B</b>	<b>Glossary</b>	<b>62</b>
<b>C</b>	<b>Bibliography</b>	<b>63</b>

# Preface

In his 1966 article presenting the material—the “stuff,” as he called it—of music libraries, James B. Coover described the trio of books, scores, and sound recordings as the “meat and potatoes” of music library collections. While the book and score had been around for centuries, the “record” was the new kid on the block, and by the mid-1960s—several years after the introduction of the long-playing record—this kid had proved to be a handful, demanding of time and resources. Coover longed for the “halcyon days” when librarians had to deal “only with 10- and 12-inch 78 rpm records, single- and double-sided, inside-out or outside-in, made of acetate or shellac.... But they are gone, and even though the variety of records available then presented some difficulties, the problems were in no way comparable in breadth or depth to those encountered today.”<sup>1</sup>

And that was over forty years ago, long before the onset of the cassette, eight-track tape, compact disc, minidisc, and digital audio tape. Throughout the relatively brief life of the sound recording, this hyperactive and needy member of the music-library family has been showered with attention by its weary guardians. While the format and content of books and scores have remained relatively unchanged over the course of centuries, near constant innovation in sound-recording technology has caused upheavals in collection development and facility planning at least every few decades. Large collections and costly equipment have been rendered obsolete as one format has succeeded another.

Until recently, these technological innovations have been realized through the introduction of new physical media—discs with grooves, magnetically charged tape, microscopically pitted aluminum discs. As each new format arrived on the scene, librarians met it with a mix of excitement and wariness. The excitement came from considering what the new technology offered—in most cases, enhanced fidelity, ease of use, and storage. The wariness came from calculating the expense of adopting the new format. Librarians approach new technology with deliberate caution, since the potential impact on collection budgets, shelving space, and facilities is great. It makes little sense to adopt a new format until it is clear the format has staying power. For this reason, libraries have usually been slow to embrace new technology, and a healthy skepticism has allowed them to avoid being stuck today with legacy collections of 8-track tapes and minidiscs.

During recent years, the development of digital sound technology has taken the sound recording down a new path. Advances in computer and networking technology have allowed the sound recording to take on a virtual existence, and the sound recording is no longer confined to a physical object—something to be purchased, stored on a shelf, and circulated. It is now also a file of data—something to stream over the internet, to download onto an iPod.

The benefits have been great. The listener is now off the leash, able to listen anywhere there is a network connection, and with the growth of wireless access points, the options increase daily. Librarians are also great beneficiaries, since we can now offer substantial collections of commercial sound recordings

---

<sup>1</sup>James B. Coover, “Materials of a Music Library,” in *Manual of Music Librarianship* (Geneva, N.Y.: Music Library Association, 1966), 3.

without the inconvenience of finding space to store them, and we can provide reserve listening services that make it possible for dozens of class members to listen to the same Bach fugue simultaneously—and scattered across campus—just hours before an exam.

The challenges and frustrations continue, however. Lying behind these digital audio services is technology that can be confounding. When working with a collection of physical sound recordings, we could meet listeners' needs simply by providing equipment for the various media in our collections and making sure our collections were properly cataloged and shelved. With digital audio, our role in bringing the music to the listener can be far more complex. For small-scale digital audio installations supporting curricular listening, a librarian might be expected to encode sound recordings for curricular listening, to maintain the server that stores and delivers the audio files, and to create a user interface for listeners to locate and select files for listening.

The technological knowledge and skills needed to manage digital audio services can be daunting. Sound can be captured in a number of digital formats and then compressed for network delivery in an even larger number of other formats. Librarians who are new to digital audio may find themselves grappling with technical concepts that are foreign to them—bitrates, codecs, streaming—and feel unequipped to provide the services that digital audio technology makes possible.

## Scope and Organization

In summer 2005, John Anderies approached me about working with him on writing a book for librarians on digital audio and digital audio services. Because of other commitments, John wasn't able to see the project to its end, but the scope and organization of this text is the product of the planning we did together in fall 2005. John also drafted a few sections of this document, and for these his byline is indicated in the heading.

This draft is an incomplete realization of the book John and I had planned. We had worked out a tight deadline for writing the book, because we knew that the rapid changes in digital audio would mean that the parts of the book would become obsolete quickly. Once John realized he would not be able to finish the project, I decided that if the sections I had written were to be of any use at all to readers, the text needed dissemination quickly. There would be no time for me to complete the book on my own or to bring another author into the project. For these reasons, I decided to deposit the book in *ScholarlyCommons@Penn* (<http://repository.upenn.edu>) under a Creative Commons license, so that others could make use of my work and perhaps even build upon it.

The book is intended to serve as a guide and reference for librarians who are responsible for implementing digital audio services in their libraries. In the treatment of technological matters, I have tried to keep the novice in mind as the primary reader, but I also hope that a librarian approaching the book with some knowledge of digital sound technology will find the content valuable as a

reference.

The book was written to serve several purposes:

- to introduce the concepts of digital sound recording technology
- to compare various common digital audio formats and offer guidance on their strengths and weaknesses
- to describe the process of encoding and compressing sound
- to explain the various ways sound can be delivered and listened to over networks
- to survey current practices in digital audio services in libraries

This book is divided into two parts. Part 1, “Digital Audio Technology,” covers the fundamentals of recorded sound and digital audio, including a description of digital audio formats, how digital audio is delivered to the listener, and how digital audio is created. Part 2, “Digital Audio in the Library,” covers digitizing local collections, providing streaming audio reserves, and using digital audio to preserve analog recordings.

The appendix offers resources on copyright issues affecting digital audio services, and the book concludes with a glossary, bibliography, and index.

## Acknowledgments

I thank the following music librarians who responded to a call for volunteers to complete a questionnaire on digital library services, announced on MLA-L in November 2005: Leslie Anderson (California State University, Long Beach), Jane Beebe (Amherst College), Leslie Bennett (University of Oregon), Sara J. Beutter (Vanderbilt University), Linda Blotner (University of Hartford), Anita Breckbill (University of Nebraska–Lincoln), Pamela Bristah (Wellesley College), Ken Calkins (University of California, San Diego), Sarah Canino (Vassar College), Kathy Carbone (California Institute of the Arts), Alexander Cari (Texas Christian University), Paul Cauthen (University of Cincinnati), Keith D. Eiten (Wheaton College), Linda Fairtile (University of Richmond), John Gibbs (University of Washington), Jon Haupt (Iowa State University), Raymond Heigemeir (Stanford University), Barbara Hirsch (University of California, Santa Barbara), David Hunter (University of Texas at Austin), Amber C. Johnson (Mansfield University), Carolyn A. Johnson (Connecticut College), Pam Juengling (University of Massachusetts, Amherst), Karen Jung (Southeastern Louisiana University), Rebecca Littman (University of Wisconsin–Milwaukee), Amanda Maple (Pennsylvania State University), Paula Matthews (Princeton University), Erin Mayhood (Boston University), Nancy Nuzzo (University at Buffalo), Jennifer Otervik (University of South Carolina), Antoinette Powell (Lawrence University), Mary Prendergast (University of Virginia), Alisa Rata (Southern Methodist University), Tracey Rudnick (University of Connecticut), Darwin Scott (Brandeis

University), Kristina L. Shanton (Ithaca College), Bradley Short (Washington University), Gerald A. Szymanski (Eastman School of Music), Robert D. Terrio (Westminster Choir College), Christia R. Thomason (North Carolina School of the Arts), Sha Towers (Baylor University), Kent Underwood (New York University), and Marlene Wong (Smith College). I also thank Jim Farrington (Eastman School of Music) for sharing chapters of his *Audio and Video Equipment Basics for Libraries* (2006) in advance of publication.

This draft was written mostly on the SEPTA R2 train en route from Wilmington to Philadelphia, where I work at the University of Pennsylvania. Most of the research, thinking, and reflection about digital audio that informed the text occurred at home in Wilmington, and throughout my year's work on this draft, my wife, Lisa, was a constant source of support, encouragement, and patience.

Richard Griscom  
Wilmington, Delaware  
June 2006

**Part I**

**Digital Audio Technology**

A librarian does not necessarily need a thorough knowledge of the technology of digital audio in order to plan and manage digital audio projects and services. In larger institutions, library systems support staff or the institution's information technology department will have sufficient expertise with the technical side of digital audio to make informed decisions. Even with excellent technical support, however, a librarian managing a digital audio service will benefit from a basic understanding of the underlying technology. A librarian equipped with knowledge of key concepts will be prepared to work more effectively with technical staff, who are not always aware the issues that must be considered in providing service to library users. In a smaller institution with limited technical support, the librarian may have little support and be expected to manage every aspect of a digital audio project, from staff supervision right down to selecting sampling rates and streaming speeds. In these cases, technical expertise is essential.

This part provides an overview of the technology of recorded sound and digital audio, including a description of digital audio formats, how digital audio is delivered to the listener, and how digital audio is created.

# Chapter 1

## Fundamentals of Recorded Sound

The sound waves that occur in nature—the clang of a bell, for example, or the roar of a passing train—are continuous, without interruption and without measure. These sound waves are variations in air pressure, generated by something vibrating, such a violin string set into rapid motion by a bow, or the halves of a bassoon reed that quickly beat against each other whenever a stream of wind passes through them.

### 1.1 Analog and Digital Sound

Beginning with Edison’s invention of the phonograph in 1877, a number of technologies have been used to capture and reproduce sound waves, and they fall into two broad categories: analog and digital. To understand the difference between analog and digital, let’s think for a minute about a clock, which is probably the most common example of a technology with both analog and digital equivalents.

Imagine an analog wall clock with the three traditional hands, designating hours, minutes, and seconds. The hands move smoothly; as the second hand sweeps around the dial, the minute hand makes its almost imperceptible progress from one-minute mark to the next. The motion of the clock is like time itself, smooth and without interruption.

Digital clocks measure time in precise increments. They convey no sense of the smooth flow of time; the digits representing the seconds change instantly, one after another, and after sixty seconds have passed, the minute digit increments immediately. A digital clock chops up the uninterrupted flow of time into precise units—hours, minutes, seconds, tenths of seconds, and even hundredths or thousandths of seconds.

The contrast between the analog clock and the digital clock has a parallel in analog and digital sound technologies. The motion of sound waves, like the

passing of time, is continuous; sound waves, by nature, are analog. Analog equipment, such as Edison's tinfoil-wrapped cylinder, are able to record and reproduce sound as continuous waveforms. Like the motion of the analog clock, the representation of sound on the cylinder is smooth and uninterrupted. Digital audio equipment, on the other hand, divides the continuous sound waves into discrete samples, captured at precise intervals, just as a digital clock divides time into precise intervals. When the sound is played back, the samples are reproduced in sequence at the same precise intervals, creating the illusion of continuous sound waves.

## 1.2 Capturing Sound

Any sound intended to be stored as a sound recording— analog or digital— must first be captured. Sounds that are produced naturally—a child's voice, a piano, the rustling of leaves, a brass band—are captured using a microphone. The variations in air pressure that make up the sound hit a diaphragm in the microphone and cause it to vibrate sympathetically. The vibrating diaphragm creates a weak series of voltage pulses that are transmitted through the wires of the microphone to its plug.

The sound captured using the microphone can be stored in a number of ways. The microphone input can be routed to a recording component, such as a tape deck, where the voltage pulses are stored as magnetic patterns on tape, or to an analog-to-digital converter, which converts the voltage pulses into a series of binary digits that can be stored on a compact disc, DVD, DAT, or hard drive.

Most digital-audio projects do not involve capturing natural sound.<sup>1</sup> Instead, they take existing recordings—commercial recordings, recordings of local concerts and lectures, field recordings captured by researchers—and convert them into digital audio files, which can then be stored, duplicated to media for playback, or delivered over networks.

For audio that is already in a digital format, no capture or conversion is needed. The process involves simply reading the binary digits on the source recording and copying them to another medium.

The sound on an analog recording, however, must be converted to digital audio. A microphone is not used to capture the audio, because the sound waves are already represented as a series of voltage pulses on the analog recording. The analog sound is reproduced on a traditional audio component (typically a turntable or tape deck), and the resulting analog signal is routed directly to a computer or digital audio recorder for conversion and storage as digital audio.

---

<sup>1</sup>An example of an exception would be an audio reserves project for which the instructor provides a spoken introduction or announcement for each audio example.

## 1.3 Reproducing Sound

The process of reproducing sound reverses the process of capturing it. Analog recordings contain physical representations of voltage pulses—as grooves on a disc or magnetized patterns on tape—and these are read by an audio component and converted into voltage pulses. For a digital recording, the binary digits are read by a digital component—a CD player, DAT deck, computer, or personal digital player—and translated into voltage pulses by a digital-to-analog converter.

### 1.3.1 Speakers

The voltage pulses produced by an analog or digital component are sent to an amplifier and then to headphones or speakers. A speaker consists of a coil of wire attached to one or more cone-shaped diaphragms. When the voltage pulses are received by the coil, they are converted to magnetic pulses that attract and repel the back of the speaker cone. The motion of the speaker cone creates variations in air pressure that reproduces the sound captured and stored on the sound recording. Low-frequency sounds require that the speaker cone move slowly, and high-frequency sounds require rapid movement. Because no single cone size can move both slowly and quickly enough to cover the full spectrum of perceivable sound, speaker systems often consist multiple cones that are optimized for specific frequency ranges. The typical categories are the woofer (25 Hz-300 Hz), midrange speaker (800 Hz-16 kHz), and the tweeter (6 kHz-30kHz). Computer sound systems often consist of two small midrange speakers and a “subwoofer” to reproduce lower frequencies.

## Chapter 2

# Fundamentals of Digital Audio

### 2.1 Concepts

The basic process of digital recording can be described simply: sound is captured thousands of times each second, the captured samples are converted to digital data, and the data is stored to a device. To play back the digital sound, the process is reversed.

Underlying this basic process are a number of more complex concepts, and understanding them will help you make informed decisions that will improve the quality, usability, and longevity of the digital audio files you create.

#### 2.1.1 Sampling

When sound is recorded digitally, “snapshots” are taken of the analog sound at precise intervals. The snapshots are then processed by an analog-to-digital converter (ADC), which translates the analog sound into binary bits that can be stored on a disc, memory chip, or some other digital medium. The snapshots are called “samples,” and the process of capturing and converting the samples is called “sampling.”

Here is another way to look at the concept of sampling. Think of a bouncing ball. As we watch the ball, its motion is continuous and smooth. If we want to capture the image of the bouncing ball to view later or to share with others, we might film it with a movie camera. As we film the ball, the camera takes a rapid sequence of still photographs. In fact, it takes twenty-four photographs each second. Once the film is developed, we can see by looking at the strip of film that each photograph shows the ball frozen at a specific point in its continuous motion, up and down. When the film is run through a projector, these still images are reproduced in their original sequence, twenty-four images

per second, recreating, through illusion, the continuous and smooth motion of the bouncing ball.

When we sample sound, we take thousands of “snapshots” of the sound each second, and when these samples are played back in sequence, the resulting sound creates an illusion of smoothness and continuity, like the illusion of smooth motion in the film of the bouncing ball.

The two parameters of sampled sound that most directly affect its quality are its sampling rate and its resolution (or bitdepth).

### **Sampling Rate**

The sampling rate is simply the number of samples captured each second, usually measured in kilohertz (kHz), or thousands of samples per second. The audio on a compact disc, for example, is sampled at a rate of 44.1 kHz, so 44,100 samples are taken each second.

Thinking again of our motion picture of the bouncing ball, imagine two movie cameras: one that takes forty-eight photographs each second and another that takes only twelve photographs each second. The first camera, because it samples at a higher frequency, will more faithfully reproduce the continuous motion of the ball’s movement than the second camera. The same is true of the sampling rates for sound. The more samples per second, the higher the fidelity of the sound.

When selecting a sampling rate for audio, something known as the Nyquist Theorem comes into play. In 1927, an AT&T physicist named Harry Nyquist determined that the sampling rate must be at least two times the highest frequency to be reproduced.

Because human hearing can perceive frequencies no higher than 20 KHz, according to the Nyquist Theorem, sampling at just over 40 KHz will capture the full spectrum of perceivable sound. Similarly, since the frequencies of human speech typically lie under 3 KHz, spoken word can be sampled effectively at a far lower rate. A sampling rate of 8 KHz is more than sufficient for recorded speech.

### **Resolution (Bitdepth)**

When filming our bouncing ball, we want the camera to take a sufficient number of images per second to reproduce the ball’s smooth motion, but we also want each of the images to be distinct and clear so that the ball itself is reproduced as realistically as possible. This is resolution.

Imagine we are using a digital video camera to capture the series of images of the bouncing ball. The larger the number of pixels used for each image, the clearer the ball will be. If we use a small number of pixels—a low resolution—the ball will be fuzzy and indistinct. The higher the resolution, the more realistic the resulting images, and the more convincing the representation of the ball as it bounces.

Resolution in uncompressed digital audio is similar to resolution in a digital photograph. The larger the number of bits used to capture each sample of the sound, the higher the fidelity of the reproduced sound. Digital audio samples are stored in strings of bits called words. The number of bits in each word determines the resolution, or bitdepth, of the digital audio. Typically, the lowest resolution used for digital audio is 8 bits. Compact-disc audio has a resolution of 16 bits, but professional recording studios capture sound using 20-, 24-, and even 32-bit resolutions.

As in digital photography, the trade-off with resolution is storage space. The higher the resolution, the more storage required. Also, for a stereo recording, each channel is sampled separately, so a single sample requires two words. You can use this formula to calculate the size of an uncompressed audio file: *size in bytes = sampling rate × resolution × number of channels × number of seconds / 8* (the number of bits in a byte). For example, to calculate the size of a file containing one minute of stereo compact disc audio:  $44,100$  (sampling rate of 44.1 kHz)  $\times 16$  (resolution, in bits)  $\times 2$  (channels)  $\times 60$  (seconds)  $/ 8$  (bits in a byte) = 10,584,000 bytes, or 10.1 MB.<sup>1</sup>

### Bitrate

The concept of resolution cannot be applied to compressed audio, since the software doing the compression may adjust the resolution to suit the content of the sound.<sup>2</sup> A passage played by a full symphony orchestra, for example, would require more bits than a passage played by a solo oboe, simply because the sound is more complex. For compressed files, instead of measuring bitdepth we measure the bitrate, or the number of bits used to store a second of sound. Bitrates are measured in thousands of bits per second (kbps), and as with resolution, the higher the bitrate, the better the sound. The bitrate for a typical MP3 file is 128 kbps.<sup>3</sup>

Here is a useful formula to calculate the size of a compressed audio file: *size in megabytes = number of seconds x bitrate in kbps / 8,388.608* (the number of kilobits in a megabyte). Using this formula, one minute of audio compressed at 128 kbps would have a size of .9155 MB (60 seconds  $\times$  128 kbps  $/$  8,388.608).

---

<sup>1</sup>One megabyte is 1,048,576 rather than 1,000,000 bytes. For more information on the confusing math of bits and bytes, see MB, Mbps, and MHz in the glossary.

<sup>2</sup>Bruce and Marty Fries, *Digital Audio Essentials*, (Sebastopol, Calif.: O'Reilly, 2005), 175.

<sup>3</sup>It is also possible to calculate the bitrate of uncompressed audio. The compact disc audio example covered in the preceding section has a bitrate of 1,411 kbps (44,100 samples per second  $\times$  2 channels  $\times$  16 bits per sample), or about eleven times the bitrate of typical 128 Kbps compressed audio.

## Chapter 3

# Formats for Digital Audio

Once audio has been sampled and converted to digital data, it can be processed and stored in a number of different formats. During the early development of digital audio, sound engineers devised formats for sampling and storing audio data that met the particular requirements of whatever operating system they happened to be using, and as a result, multiple formats emerged for the storage of digital audio.

As time passed, certain formats gained enough of a following to become de facto standards for certain applications. During the 1980s, with the advent of the personal computer, microprocessors increased in speed and capacity, and then during the 1990s, network access became commonplace, and new formats were developed to make the most of these technological advances as well as to meet emerging needs for compressed streaming audio and streaming media. Instead of the technology settling down to one or two established formats—as has happened with audio and video media in the past—the number of formats has increased rather than decreased.

A regular user of the internet confronts dozens of media formats for audio and video, and at this point, it seems doubtful that any one format will prevail. Fortunately, today's software can play files in most of the standard formats and files can be easily converted from one format to another.

In this section, we will review the formats used to capture, encode, and store digital audio. Of the dozens of audio formats that have been developed through the years, many are now used only infrequently, so we will look only at those that are likely to have some application in a library setting.

### 3.1 Terminology

Terminology sometimes becomes blurred in discussions of digital audio. Often the same name is applied to the software that creates the audio, the computer algorithm that compresses and decompresses the audio, the file format that is used to store the compressed audio data, and the player that plays back the

resulting audio file. For example, the Windows Media Encoder can be used to create compressed Windows Media Audio data stored in a Windows Media Audio (.wma) file, which can be played back using a number of different players, including the Windows Media Player. As I discuss digital audio, I will try to maintain distinctions in the terminology used for file formats, compression/decompression algorithms, software, and players. Here is a summary of the terminology used in this book:

*digital audio data*: the binary data that represents the audio

*digital audio format*: the format of the digital audio data

*codec*: a computer algorithm used to compress and decompress digital audio data in a particular audio format

*digital audio file*: a file containing digital audio data

*digital audio file format*: the format of a digital audio file

The most common area of confusion lies in the term “format.” A distinction should be made between the format of the digital audio file and the format of the digital audio that the file contains. Think of a pitcher containing a beverage: a pitcher is similar to an audio file. Instead of a beverage, an audio file contains audio data. Similarly, the type of pitcher (round or octagonal; plastic or glass) would correspond to the file format, and the type of beverage (lemonade, iced tea, margaritas) would correspond to the audio format. The file format and the audio format are different concepts, and they exist independently of each other.<sup>1</sup>

## 3.2 Structure of an audio file

An audio file consists of several parts: a header, the audio data, and, optionally, metadata and a wrapper. The header provides information about the data in the file—the sampling rate, number of channels, bit depth, and similar technical specifications. The audio data—the bits representing the samples taken of the audio—make up the bulk of the file. Audio files may also include metadata—text describing the content of the audio file (performer, copyright information, track name, source album, etc.)—and a wrapper, which controls use of the file. Digital rights management and streaming capability, for example, are usually provided by a wrapper.

## 3.3 Open and proprietary formats

Some digital audio formats are open, which means that the specifications of the format—how the data is structured, the algorithms used to encode the data—

---

<sup>1</sup>A good example is the QuickTime file format, which can contain data in a number of different formats—for example, MP3 data or AAC data.

are freely available, and use of the format is free of legal restrictions. Usually open formats are maintained by a national or international standards organization. Advocates argue that use of open formats will help guarantee long-term access to data and encourage cooperative development of the formats.

Other formats are proprietary; for these, a private concern—usually a commercial enterprise—maintains control over the format and the release of details on its structure, encoding, and decoding. In many cases, the owner of the format will release information on the structure of file and how it is encoded but retain rights over the decoding algorithm. Owners of proprietary formats are interested in promoting use of their format and often take actions to discourage the use of competing formats.

Some proprietary formats are actually based on open formats. Apple, for example, sells tracks on its iTunes Music Store in a proprietary format that uses AAC-encoded audio (an open format) with a proprietary digital rights management wrapper that restricts use of the file.

Many popular, well-established formats are proprietary, and librarians often choose to base their digital audio services on proprietary formats because they are familiar to patrons, and software to play back the files is readily available—sometimes even packaged with the computer’s operating system. There are some risks, however, in basing audio services on proprietary formats. Support can be very good until the sponsoring company abandons or alters the format. Companies often promote their own proprietary audio formats to the detriment of others with the hope of securing a greater market share, and they make adoption of their format attractive by offering convenient tools for encoding sound in the formats. Often proprietary formats are developed for specific hardware and software, which will place limits on the playback options for listeners. For these reasons, a proprietary format that works well on one operating system may present problems for another.

## **3.4 Uncompressed formats**

There are two broad classifications of audio formats: uncompressed and compressed. For uncompressed formats, the audio data consists of the digital audio samples as they were originally captured at their original bitdepth. Uncompressed formats do the best job of capturing and reproducing sound, and for that reason they are used extensively in the recording, mastering, and storage of digital audio. In this section, we will review the most common formats for uncompressed digital audio.

### **3.4.1 PCM**

Pulse Code Modulation (PCM) is the process most often used to transmit and store uncompressed digital audio data. Most uncompressed digital audio file formats—including WAV, AIFF, and CDDA—use PCM as the format for the audio

data. PCM is not new technology; it was developed in 1937 by British engineer Alec Reeves while working for International Telephone and Telegraph.<sup>2</sup>

When an analog-to-digital converter translates analog audio samples into binary “words,” it uses PCM to transmit the individual bits of the words as voltages (“1” as a positive voltage; “0” as the absence of voltage), which can then be reconstituted as binary data for storage in a computer file or on a compact disc. PCM is the audio equivalent of ASCII text; because of its simplicity, most audio programs can play PCM.<sup>3</sup> It can accommodate a number of different resolutions (8-, 16-, and 24-bit depths are common), sampling rates (usually between 22 kHz and 96 kHz), and channel configurations (for example, mono, stereo, and 5.1 surround sound).

### 3.4.2 Sun/NeXT (extension `.au` or `.snd`)

The `.au` file format—“au” is short for “audio”—was developed by Sun for use with telephone transmissions processed by Unix computers, and it became one of the earliest formats commonly used for audio files on personal computers. It is now primarily only of historical interest. The extension `.snd` is used for files in this format on Sun, NeXT, and Silicon Graphics computers. Although `.au` files usually contain PCM audio, the format can also handle several compressed formats.<sup>4</sup>

### 3.4.3 AIFF (extension `.aiff` or `.aif`)

The Audio Interchange File Format (AIFF) was developed by Apple for use with the Macintosh, but it is recognized by a number of Windows and Linux audio editing programs as well. AIFF accommodates uncompressed PCM audio with a variety of channels, sampling rates, and resolutions.

### 3.4.4 WAVE (extension `.wav`)

The WAVE (Waveform Audio) is a proprietary file format developed by Microsoft for use in Windows 3.1. It is actually a variant of the RIFF bitstream format and is a “wrapper” format capable of containing audio data of various types, including compressed audio data. The default (and most common) type of data contained in a WAVE file is PCM data, which can be accommodated in a variety of channels at various sampling rates and resolutions.

WAVE is the format most frequently used in Windows operating systems for uncompressed audio. Many compact disc “ripping” applications store the resulting raw data in WAVE format, so it is often used as an intermediate format when preparing compressed audio for streaming.

---

<sup>2</sup>Ken C. Pohlmann, *Principles of Digital Audio*, 5th ed. (New York: McGraw-Hill, 2005), 52-53.

<sup>3</sup>Fries & Fries, 159.

<sup>4</sup>Pohlmann, 569. For details on the structure of the header of an `.au` file and the audio formats it can accommodate, see [http://en.wikipedia.org/wiki/Au\\_file\\_format](http://en.wikipedia.org/wiki/Au_file_format) (Accessed 28 November 2005).

## 3.5 Compressed formats

Because uncompressed audio files are so large—about 10 MB of storage for every minute of CD-quality audio—they are impractical for streaming and downloading over the internet.<sup>5</sup> For network use, audio files are “compressed” to reduce their size, allowing for quicker downloads and real-time streaming.

Computers compress and decompress audio data by using software called a codec (COmpress/DECompress). Sometimes the term “codec” is used interchangeably with “audio format,” but there is an important difference: a codec is software that is used to interpret an audio format. In fact, in some cases several different codecs exist to compress and decompress a single audio format.

With compression, there is a tradeoff between file size and sound quality. Codecs that provide high levels of compression discard parts of the original audio to reduce the amount of data. The more data that is discarded, the smaller the audio file, but the loss of data also results in a degradation in sound quality.

Audio compression formats fall into three groups: formats defined by international standards (such as MPEG), proprietary formats (such as Windows Media and RealAudio), and open-source formats (such as Ogg Vorbis).

It is important to select the compression format that best meets your particular needs, and those needs often concern more than audio quality. The projected longevity of the format, its market share, its technical support, the requirements and limitations it imposes on hardware and software—all of these can be just as important as sound quality.

### 3.5.1 “Lossless” Compression

Some compression formats are able to reduce the size of an audio file without discarding any data. This is lossless compression.<sup>6</sup> When the resulting compressed file is decompressed, it is identical to the original uncompressed audio file. Lossless compression can be used to distribute and archive digital audio, and digital players can decode the most common audio formats for playback. The rate of reduction varies generally between 25 percent and 50 percent, depending on the content of the source file.

Because they reduce the size of an average compact-disc audio file by no more than 50 percent, however, lossless compression formats are generally impractical for use in streaming—at least over networks slower than 600 kbps. Their primary application is in the archiving of master recordings, where it is essential both to preserve content and to save storage space. With lossless com-

---

<sup>5</sup>In order to stream CD-quality audio (two channels sampled at 44.1 kHz with 16 bytes per sample) a computer would need to receive a stream at 1.4 Mbps. Broadband networks will undoubtedly eventually achieve this speed, but in 2006 downloading a minute of CD-quality audio over a relatively fast 768 kbps DSL connection takes nearly two minutes (110 seconds).

<sup>6</sup>General compression algorithms for computer data—such as PKZIP and the Unix command `gzip`—can be applied to audio data files, but they typically reduce an audio file’s size by only 10 to 20 percent. Lossless audio compression formats use algorithms specific to audio data that produce higher compression rates, generally reducing the size of a file by between 30 and 50 percent.

pression, if the original media is lost or damaged, an exact duplicate of the original can be recovered at any time.

Most lossless encoders offer various levels of compression. The tradeoff is between file size and the amount of time required to encode a file; higher compression comes at the cost of speed. Often the encoding software will offer guidance in selecting an appropriate compression level to suit your needs.

### 3.5.2 Open formats

#### FLAC (extension `.flac`)

The Free Lossless Audio Codec (FLAC) was developed by the Xiph.org Foundation and is a free, open-source format that has no restrictions on use and no licensing fees. There is also a metadata component: a “cue sheet” metadata block can be used to store a compact disc’s track listing and index points. FLAC can be used with any PCM data with bitdepths from 4 to 32, sampling rates from 1 Hz to 1 MHz, and one to eight channels.<sup>7</sup> Typical compression rates run between 30 and 50 percent. A technical strength of FLAC is its ability to be decoded quickly, which makes it suitable for streaming over fast networks. FLAC data is often contained in Xiph.org’s Ogg file format.<sup>8</sup>

#### WavPack (extension `.wv`)

As its name suggests, WavPack is used to compress WAV files, and it can accommodate files with multiple channels, at sampling rates from 6 to 192 KHz, and at 8-, 16-, 24-, and 32-bit resolution. The compression rate ranges from 30 percent to 70 percent, depending on the source file.

The WavPack encoder offers options for both lossless and lossy compression as well as a “hybrid” mode, which creates a lossy compressed file and a second “correction” file that can be used to restore the compressed file to its original lossless state. The encoder is available in versions for Windows, Linux, and Mac OS X. All versions are run from a command line, but an optional Windows interface is available.<sup>9</sup>

### 3.5.3 Proprietary Formats

The three major proprietary formats—Windows Media, RealAudio, and QuickTime—now offer lossless codecs for use with archiving as well as streaming. The usual caveats related to proprietary formats apply here as well: the sponsoring companies may offer encoding software, and their popular media players (Windows Media Player, RealPlayer, Quicktime Player) may be able to

---

<sup>7</sup>Pohlmann, 472.

<sup>8</sup>See also “Ogg Vorbis” on p. 26 below.

<sup>9</sup>Most of this information is taken from the WavPack site: David Bryant, “WavPack Audio Compression,” <http://www.wavpack.com> (Accessed 14 December 2005).

play back the files, but if the sponsoring company were to discontinue support for the format, users would likely be left to rely on legacy software.

### **Monkey's Audio (extension .ape)**

Monkey's audio was developed by Matthew T. Ashland "for fun to keep myself busy during the cold Minnesota winter."<sup>10</sup> The Monkey's Audio encoder can encode mono and stereo WAV files at resolutions of 8, 16, or 24 bits and at any sampling rate. According to Ashland, the program has been optimized for use with compact-disc audio (stereo, 16-bit resolution, 44.1 KHz sampling). It achieves compression rates of about 40 to 50 percent. ID3 tags are supported, but third-party tagging programs are required to use the more expansive ID3v2 tags.

The current Monkey's Audio (version 3.99) is available only for the Windows platform, although the official website mentions that versions for Apple and Linux are in development.<sup>11</sup> Although Monkey's Audio is not an open-source project, use of the encoder and the format are free for personal or educational purposes; permission must be granted by the author for commercial use.

## **3.6 "Lossy" Compression**

The most common compressed audio formats use "psychoacoustic models" to discard audio data that cannot be heard or that is typically ignored by the human ear. By eliminating this data, a file can be reduced in size while minimizing the effect on the sound. These formats that selectively discard data are known as lossy formats, and they can produce files that are anywhere from one-fourth to one-thirtieth the size of the original uncompressed audio, with a corresponding degradation in fidelity.

### **3.6.1 MPEG**

Among the lossy formats, those based on MPEG standards are the most popular. MPEG is a suite of open standards for compressed audio and video developed by the Motion Picture Experts Group, a working group established in 1988 under the direction of the International Standards Organization.

MPEG's standards have been released in families, each designated by number. MPEG-1 (approved in 1992), supports video encoding as well as mono and stereo audio encoding at three sampling rates; MPEG-2 (1994) increases the number of sampling rates and provides for broadcast-quality video and surround sound; MPEG-4 (1998) supports a broad range of multimedia and is

---

<sup>10</sup>Response to the question "Could you tell me about the future of Monkey's Audio?" in the FAQ contained in version 3.99 of the Monkey's Audio software (Accessed 13 December 2005).

<sup>11</sup>See <http://www.monkeysaudio.com>.

able to integrate synthetic audio systems (such as MIDI and text-to-speech programs); MPEG-7 (2001) provides tools for managing metadata.<sup>12</sup>

Of the many formats provided by the MPEG standards, the most common are MP3 and AAC.

### **MP3 (extension .mp3)**

MP3, officially known as MPEG-1 Audio Layer III, is an audio subset of the 1992 MPEG-1 standard. (Layer III also received some enhancements in the MPEG-2 standard.) MP3 files were front and center in the digital music revolution of the 1990s and gained notoriety through their open sharing on peer-to-peer networks.

The MP3 format has been particularly popular because it can produce “near CD” quality<sup>13</sup> audio at a compression rate of 11 to 1. In other words, one minute of compact-disc audio, which requires about 10 MB of storage, can be compressed to an MP3 file smaller than 1 MB. Despite the development of compression formats that produce better sound quality at identical bitrates—such as AAC and Ogg Vorbis—MP3 remains the most popular audio format on the internet, and it has become the lingua franca of personal digital audio players.

Although the specifications of the MPEG standards are open and freely available, the Fraunhofer Institute and Thomson Multimedia—the companies that helped finance the development of the standards—hold patents on many of the algorithms used to code and decode MPEG files.<sup>14</sup> In 1998, when the Fraunhofer Institute issued a letter stating that it would begin charging royalties to developers of MP3 encoders, some distributors removed MP3 codecs from players, and some developers decided to begin work on truly open formats, such as Ogg Vorbis (see below).

### **Advanced Audio Coding (AAC) (extensions .aac, .mp4, .m4a, etc.)**

Advanced Audio Coding (AAC) was developed under MPEG-2 and enhanced under MPEG-4. In the MPEG family of standards, AAC is the heir apparent to MP3. Until the introduction of AAC, MPEG audio formats were “backward compatible,” which means that files created with earlier standards could be played with decoders for the newer standards. With the introduction of AAC, MPEG abandoned backward compatibility in order to take advantage of newer coding algorithms and took the practical precaution of assigning it a name that

---

<sup>12</sup>The numbering scheme for MPEG is not strictly sequential; there is no MPEG-3, -5, or 6.

<sup>13</sup>“Near CD” quality is entirely subjective and can vary depending on the listener. For some, compressed audio encoded at 128 kbps might be nearly indistinguishable from a CD; for more sophisticated listeners, the rate needs to be pushed up to 196 kbps.

<sup>14</sup>For information the Fraunhofer Institute’s and Thomson’s patent claims, see Gabriel Bouvigne, “Patents and MP3,” <http://www.mp3-tech.org/patents.html> (Accessed 10 December 2005). Thomson’s licensing information is available at <http://www.mp3licensing.com> (Accessed 10 December 2005).

would distinguish it from its “MP” predecessors.<sup>15</sup>

AAC provides better sound quality than MP3—particularly at lower bit rates—and it supports sampling rates from 8 kHz to 96 kHz, compared to MP3’s 16 kHz to 48 kHz. One claim to fame for the MPEG-4 AAC format was its adoption by Apple as the basis for the audio format used by its iTunes music store.<sup>16</sup> In fact, because of the close association of AAC with the iPod, it is often mistakenly assumed that AAC stands for “Apple Audio Codec.”

Files with the extension `.aac` are MPEG-2 AAC files; only a few audio players are able to support these files. AAC audio data is more frequently contained in an MPEG-4 file (similar in structure to a QuickTime file), which is supported by most popular audio players. A number of confusing file extensions are applied to MPEG-4 files, and their interpretation can be challenging. Although the official MPEG-4 file extension is `.mp4`, this extension is not found as frequently as the ones applied by Apple for use with the iPod and iTunes: `.m4a` (“MPEG-4 audio”) is used for files ripped using iTunes, `.m4p` (“MPEG-4 protected”) is used for files purchased on the iTunes Music Store (the “protected” refers to embedded digital rights management), `.m4b` (“MPEG-4 bookmarkable”) is used for audio book files that can be “bookmarked,” and `.m4v` (“MPEG-4 video”) is used for audio/video files.<sup>17</sup>

Although AAC is a part of the open MPEG standards, the situation with licensing is similar to the one with MP3: the patent rights to the codecs used with AAC are held privately—in this case by AT&T, Dolby, the Fraunhofer Institute, and Sony—and developers who incorporate AAC codecs into their software must pay royalties to the patent holders.<sup>18</sup>

### 3.6.2 Proprietary Formats

#### Windows Media Audio (extensions `.wma` and `.asf`)

Windows Media Audio (WMA) was introduced by Microsoft in 1999 as a competitor to MP3, and while it was slow to catch on at first, its popularity has increased in recent years. Several online music stores—including Napster—use WMA (with Digital Rights Management) as the basis of their service, and a growing number of portable digital players support the format. WMA files are usually wrapped in an Advanced Systems Format (ASF) file, a fully documented format that provides streaming capability.

<sup>15</sup>David Austerberry, *The Technology of Video and Audio Streaming*, 2nd ed. (Burlington, MA: Focal Press, 2005), 122. The MPEG-4 variety of AAC has been confusingly referred to as MP4, which reinforces the perception of AAC as the successor to MP3 while mixing apples with oranges: the “3” in “MP3” refers to Layer 3 in the MPEG family of standards, but there is no Layer 4 that would correspond to “MP4”; in this case, the “4” refers to the MPEG-4 group of standards, which provides for much more than AAC.

<sup>16</sup>The AAC audio purchased on Apple’s iTunes store is contained in files in Apple’s proprietary `.m4p` format, which includes a digital rights management wrapper that controls use of the file.

<sup>17</sup>“MP4,” <http://www.audiocoding.com/modules/wiki/?page=mp4> (Accessed 3 February 2006)

<sup>18</sup>Fries & Fries, 170.

Microsoft's Windows Media offerings include an encoder (Windows Media Encoder), various software development kits, and a player (Windows Media Player). There are both lossy and lossless codecs available for WMA.

### **RealAudio (extensions .ram and .ra)**

The earliest live audio offerings on the web were radio broadcasts streamed using RealAudio, introduced by Progressive Networks (now RealNetworks) in 1995. This new format—and technology—led to the rapid growth of streaming audio and video webcasts during the late 1990s. With the subsequent development of competing formats, RealAudio's market share has deteriorated, but it is still a popular choice for streaming radio broadcasts, and it is still the format of choice for streaming digital audio reserves in music libraries. One advantage of RealAudio is the support of SMIL files, which allow a series of audio files to be played consecutively without prompting from the user. This feature is particularly useful with longer works, such as operas and multi-movement works, which are typically divided into multiple tracks on compact disc recordings.

RealNetworks applications in support of RealAudio include a player (RealPlayer), encoder (RealProducer), and streaming server (RealServer). In July 2002, RealNetworks launched Helix, an open-source initiative that builds on programming code released by the company. The Helix Community currently offers a player (Helix Player), encoder (Helix Producer), and server (Helix Server), all of which are developed to support RealAudio.

### **QuickTime (extension .mov)**

QuickTime, developed by Apple Computer, is a popular format for streaming video and multimedia presentations encoded in various formats. The first version was released in December 1991, and Apple initially used QuickTime to provide video, graphics, and audio content on CD-ROMs. It remains the most popular format for CD-ROM video. In fact, Apple was fairly late in offering streaming capability for QuickTime, which was not made available until the release of Version 4 in June 1999—four years after the introduction of RealAudio.

The QuickTime is a “container” format that is particularly useful for synchronizing the content of numerous multimedia files, which may be stored in different locations. The process of editing a multimedia presentation in QuickTime is much simpler than in other formats, and because of this facility, MPEG adopted the QuickTime .mov format as the basis for MPEG-4 in 1998. In an odd twist, Apple held off on incorporating the resulting MPEG-4 standard into QuickTime following a dispute with the MPEG-4 license holders over licensing fees. The two parties reached a compromise, and QuickTime 6 was released in July 2002. As of this writing, the latest release is Version 7.0.3.

### 3.6.3 Open Format

#### Ogg Vorbis (extension .ogg)

Ogg Vorbis is a free and open audio format developed and maintained by Xiph.org.<sup>19</sup> Chris Montgomery began the Ogg Vorbis project at the Massachusetts Institute of Technology soon after the Fraunhofer Institute announced in September 1998 that it would begin charging licensing fees for use of the MP3 format.

It is a fairly new format; the specifications were established in May 2000. Strictly speaking, Ogg is a file format and Vorbis is an audio format. Ogg can be used as a container for audio in other Xiph.org formats (such as FLAC), and Vorbis can exist as raw data without the Ogg container. Nonetheless, the audio format is commonly referred to as simply Ogg Vorbis.

The quality of Vorbis audio at a given bitrate is comparable to AAC and superior to MP3 and Windows Media Audio. At this point, though, the format is not in wide use, perhaps because the patent owners MP3 and AAC codecs have not been aggressive in collecting royalty payments for their use.

#### AIFC (extension .aifc)

AIFC or AIFC (Audio Interchange File Format Extension for Compression) is a version of the AIFF format that accommodates compressed data. The codec can achieve compression rates as high as 83 percent.

---

<sup>19</sup>Xiph.org was originally named Xiphophorus, which explains the choice of a fish as the official logo for Ogg Vorbis. <http://www.streamingmediaworld.com/audio/tools/vorbis> (Accessed 13 December 2005).

## Chapter 4

# Delivery to the Listener

### 4.1 Accessing Digital Audio Files

Throughout most of the history of recorded sound, the traditional method for distributing and accessing recordings has been a physical object—a cylinder, a disc, a reel of tape, a cassette. During the first two decades following the introduction of digital audio technology, this tradition continued with the development of several physical formats for digital audio data: the compact disc, digital audio tape, and the minidisc.

Digital audio technology is quickly moving away from traditional physical distribution to network distribution. Although network distribution has been available for well over a decade, it was only with the introduction of digital audio players in 1998 and Napster in 1999 that digital audio files were commonly distributed over networks without the use of physical media, other than the hard drive of the destination computer. Most industry observers predict that physical distribution of sound recordings will eventually be abandoned altogether.

In this section, we will look at the four most common ways that digital audio files are distributed from one computer to another over networks: making them available on servers for downloading, streaming them in real time, sharing them over peer-to-peer networks, and syndicating them as a podcasts.

#### 4.1.1 Downloading (by John Anderies)

To download is to transfer the content of a digital file from a remote computer and store a copy on a local computer. The remote computer is usually called a server, and the destination computer is called a client workstation. The file is usually transferred via protocols such as HTTP (Hyper Text Transfer Protocol) or FTP (File Transfer Protocol). The amount of time required to download a sound file depends on several factors: the size of the file, the amount of bandwidth available for the transfer, and to some degree, the performance of the client workstation itself.

In the simplest instance, a sound file must be completely transferred and saved to the client workstation before playback may begin. A newer method, called “progressive downloading,” allows playback to start before the sound file is completely transferred to the client workstation. Under the right conditions, this can mean almost instantaneous playback, and for that reason the end result of progressive downloading is not unlike the next delivery method we’ll consider: streaming.

### 4.1.2 Streaming

Unlike the process of downloading, which transfers whole files to the client workstation, streaming divides the file into small packets of data, which are sent in a continuous stream to the client workstation, which discards the packets after playing them.

Streaming technology evolved in the mid-1990s when network and dial-up speeds increased, computer audio technology became commonplace on desktop computers, and audio compression formats made it possible to reduce the data content of audio files while maintaining an acceptable level of audio quality. Initially, streaming technology was most commonly used to deliver radio broadcasts in real time over the internet. In the mid-1990s, academic libraries recognized the potential of streaming technology to improve the delivery of listening assignments to students, and by 1996, several libraries were digitizing listening assignments and making them available through streaming servers.

Streaming technology offers several advantages over downloading for distributing audio over a network. Digital audio files can be very large (roughly one megabyte of data for every minute of compressed audio encoded at 128 kbps),<sup>1</sup> and depending on the speed of the network connection, downloading one minute of compressed digital audio can take anywhere from a few seconds to several minutes. Streaming audio allows the user to listen without having to download the entire file. When a user requests an audio stream, the streaming server begins sending packets of data to a buffer—a digital holding tank—on the client workstation, and a second or two later, once the small buffer is filled, the player can start playing. Also, because the player works with only a small part of the file at a time and discards it once it has been played, streaming audio technology discourages illegal copying and distribution of copyrighted material, since the entire sound file is never stored on the listener’s computer.

Using streaming audio for listening assignments offers several advantages over traditional delivery methods. Streaming audio technology allows library users to listen to recordings on any computer connected to the internet, and with wireless access points in libraries, in airports, in coffee shops, in hotels, listening can be done practically anywhere. Also, in the past, because only a few people could listen to a sound recording at a time, students often had to wait in line for an assigned recording to become available, even when the library made

---

<sup>1</sup>The exact figure is .9155 MB. For a formula to calculate the size of a compressed audio file, see p. 15

multiple copies. Streaming audio allows an entire class of students to listen to the same selection at the same time.

### 4.1.3 Peer-to-Peer (P2P) networks

Downloading and streaming are client–server technologies. Both involve a central “server” that stores the audio data and delivers it to a software “client” running on a user’s computer. With a client–server system, the data is under the complete control of the central server; the server controls what is distributed and who is able to access it. This central control in the client–server model is a strength as well as a weakness. The clients are dependent on the single server’s ability to handle all of the requests it receives for data. If the server is overwhelmed—or is down completely—no content can be delivered.

Peer-to-peer networks (often abbreviated “P2P”) avoid dependence on a central server by spreading the responsibility for delivering content across the individual computers on the network. Each computer acts as a server, and the data moves laterally from peer to peer. Once an audio file has been downloaded to a computer, that computer can then make the file available for sharing with others. A peer-to-peer network draws on the computing power, bandwidth, and content of the individual computers participating in the network. If one computer is down, there is no noticeable degradation of service because the remaining computers on the network are available to fill the void.

Peer-to-peer networks are informal. All that is needed to participate is the appropriate software application, and computers join and leave the network on a whim. The number of computers sharing content changes minute by minute, so content available one day might not be available the next. The participants on the network are usually not known to each other, and they are traceable only by their network address.

The informal, amorphous nature of peer-to-peer networks make them a convenient vehicle for anonymous filesharing, and they have gained notoriety for enabling the illegal sharing of copyprotected sound recordings, videos, and films. Most of the files shared on peer-to-peer networks are copies of commercial content being distributed without the permission of the owner, and for that reason these networks have become targets for litigation by the recording and film industries. Librarians are advised not to tap into the commercial audio resources available on these networks. The networks, however, are also a rich source for nonprofessional bootleg recordings and unreleased outtakes. By no means is all of the content on P2P networks being shared illegally, but caution is advised.

The first popular peer-to-peer network was Napster, established in June 1999 by Shawn Fanning, who wrote the underlying program while a student at Northeastern University in Boston.<sup>2</sup> Napster became especially popular with

---

<sup>2</sup>Strictly speaking, Napster was not a pure peer-to-peer network, since it relied on central servers to maintain lists of connected systems and the files they made available for sharing. This also opened it up to successful litigation. More recent peer-to-peer networks have abandoned this central-server model.

college students, who were able to take advantage of fast broadband campus networks for speedy, efficient filesharing. In December 1999, the recording industry, alarmed that tens of thousands of copyprotected files were being shared free of charge through Napster, pursued legal action against the company, and a March 2001 injunction shut down Napster as a free file-sharing service.<sup>3</sup>

The recording industry's ongoing attempts to shut down companies that develop and distribute filesharing programs has led to a game of cat and mouse; as soon as Napster was gone, filesharers moved to other platforms, and despite aggressive legal action by the recording industry—not only against the developers of peer-to-peer software but against individual filesharers—their popularity continues unabated. As of the end of 2005, the four most popular peer-to-peer filesharing networks are eDonkey, BitTorrent, FastTrack, and Gnutella, and each network can be accessed through a number of software clients.

Filesharing accounts for the majority of data traffic on the internet, and since 2005, movies and videos have surpassed music as the most popular content shared. A 2005 study of global internet traffic by CacheLogic showed that 60 percent of all internet traffic is the product of filesharing, and of the files shared on the four major platforms, 62 percent were video and only 11 were audio.<sup>4</sup>

Although peer-to-peer networks have little—if any—application in a library setting, they have been of great interest to the library community for the intellectual property challenges they have raised. In fact, most of the litigation involving digital audio and video has centered on attempts to curb filesharing. In 2005, the Supreme Court ruled unanimously that Grokster (the developer of software used with the FastTrack network) could be held liable for users' copyright infringements.<sup>5</sup> The future of music distribution will undoubtedly be shaped by such legal decisions as they are handed down.

#### **4.1.4 Podcasting (by John Anderies)**

Podcasting is essentially the same as downloading, but with the added element of syndication technology, which delivers the sound file to the client workstation automatically as part of a subscription.

Syndication protocols (such as RSS and Atom) allow users to subscribe to weblogs and other online content. The subscriptions are known as “feeds,” and they are read by using a news- or feed-reader. At particular intervals, the reader makes a call to the server that provides the feed to see if there is new content that can be retrieved. If so, it is automatically downloaded to the client workstation.

In the case of weblogs, the content of a feed is usually text-based; podcasts are the audio equivalent—an audio feed. Because these audio files are often

---

<sup>3</sup>The Napster brand and logo were acquired by Roxio in 2002 for use with a commercial service released in October 2003, but beyond the name, the commercial Napster bears no resemblance to the original file-sharing platform.

<sup>4</sup>Jason Meserve, “P2P Traffic Still Dominates the 'Net,” *Networkworld*, 29 August 2005, 16.

<sup>5</sup>Case no. 04-0480, 545 U.S. \_\_\_\_ (2005).

loaded onto iPods and other MP3 players for listening once they have been downloaded, the process has been termed podcasting.

Podcasting technology allows inexpensive, quick, and easy distribution of audio content. Because the audio is downloaded to the client workstation, however, most podcasts consist of noncommercial content—spoken commentary, movie reviews, travelogues, idle observations on life—and steer clear of musical content because of intellectual property issues. Several universities are now using podcast technology as a means of distributing lectures and other noncommercial course-related audio content through an Apple project called iTunes U.

## 4.2 Software for Playback

A key component to any digital audio service is the audio player—a software program used to play back digital audio, either as a stream or as a downloaded file. Depending on the type of audio service you provide, the decisions you will need to make about players will range from the simple to the fairly complex.

The developers of the common proprietary audio formats (Microsoft, Real, and Apple) provide players designed specifically to play back those audio files. While these players—like the audio formats—are proprietary, they are distributed free of charge to users in order to promote both the audio format and the player. In some cases, a basic, stripped-down version of the player is offered at no cost, and a “premium” version with added features is available for a charge. These features may include, for example, the ability to rip and burn compact discs.

Often the choice of player will depend on what type of audio files you plan to stream. Many of the proprietary players will not play sound files in formats supported by competitors. For example, Windows Media Player will not currently play protected AAC (iTunes) files, and the iTunes player will not currently play Windows Media Audio. If all of your digital audio is encoded in the same format, then it makes some sense to use the proprietary player associated with that format. On the other hand, if you offer audio in multiple formats, you must either provide multiple players or identify a single player that can handle all of the relevant formats.

For most libraries, there are two broad categories of listeners to digital audio services: those in-house who rely on library workstations for listening and those who use their own computers—either in-house or remotely—to access the service. The software requirements for the workstations used by these two categories of listener are different: an in-house listening station simply needs players that are compatible with whatever operating system and browser are used on public workstations. Often systems staff will include audio players as part of a standard public workstation “disc image,” and the only ongoing responsibility of the music librarian is to make sure the players continue to function whenever the browser software and the operating system are upgraded.

The situation is not so simple when providing service to remote users, who

	WAVE	FLAC	MP3	RealAudio	WMA	AAC	Vorbis
WMP	Yes	No	Yes	No	Yes	No	No
RealPlayer	Yes	No	Yes	Yes	Yes	Yes	No
QuickTime	Yes	No	Yes	No	No	Yes	No
iTunes	No	No	Yes	No	No	Yes	No
VLC	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Figure 4.1: Capabilities of five common audio players

will prefer to access the service with their chosen operating system, browser, and audio player. In order to provide the best service to the largest number of users, you should run tests on computers running both Windows and MacOS to determine which software configurations are compatible with your service. (Linux users are used to being neglected and often derive some satisfaction from discovering workarounds on their own. And they'll tell you about them.) Of course, it would be impossible to test all permutations of operating systems, browsers, and players, but you should be able to recommend at least one successful combination of browser and player for both Windows and MacOS. Be sure to provide your users with a detailed list of systems requirements for your service—complete with web links to the pages where players and browsers can be downloaded—and update it regularly.

### 4.2.1 Common Audio Players

There are dozens of computer applications that play back digital audio—far more than could be covered sufficiently in this book—so I will highlight four freely distributed proprietary players that are often used with digital audio projects as well as one open-source player that is a good alternative choice for libraries that offer digital sound in multiple formats.

To test the capabilities of the players, I took sample files in five formats—MP3, RealAudio, Windows Media Audio, AAC (in both an MPEG-2 and an MPEG-4 format file), and Ogg Vorbis—and attempted to play them on the latest version of each player, as it is, “out of the box,” without importing additional codecs. The table in figure 4.1 summarizes the results.<sup>6</sup>

From this table, it would be easy to conclude that MP3 is the ideal format (since it can be played by all the players) and VLC the ideal player (since it can play all the formats), but the situation is a bit more complicated than that. Other formats provide better fidelity than MP3 at similar bitrates,<sup>7</sup> and the support community for VLC is far smaller than for the proprietary players.

Many of the popular proprietary players—such as Windows Media Player, RealPlayer, and QuickTime Player—are also designed to serve as media content browsers to access news reports, videos, shopping sites, and so forth. Like

<sup>6</sup>The WAVE, FLAC, MP3, AAC, WMA, and Ogg files were created using Easy CD-DA Extractor, and the RealAudio file was created using RealProducer.

<sup>7</sup>See, for example, “Advanced Audio Coding” and “Ogg Vorbis” on p. 23 and p. 26 above.

iTunes, they can also synchronize content with portable digital players and connect directly to digital music stores, where use licenses can be purchased for individual tracks. The latest version of Windows Media Player and RealPlayer can also play DVDs. These extra features add to the size of the application and the demands placed on computer resources. Unless these added features are needed by your users, they will be best served by a stripped down, freely distributed “basic” version of the player, which will load more quickly and perform more reliably than the full-featured versions.

In this section, we will look only at the functionality of the players for playing back audio. The use of other features of the players—such as ripping CDs and encoding files—will be covered elsewhere.

### **Windows Media Player**

Windows Media Player (WMP) is a proprietary player developed by Microsoft that has been bundled with the Windows operating system since the release of Windows 98 Second Edition in 1998. The player is also available for independent download, but the most recent version, WMP 10 (released in September 2004), requires Windows XP and is not compatible with earlier versions of Windows.

Although earlier versions of WMP were released for MacOS and Solaris, there are no plans for future development of WMP for non-Windows platforms. WMP 9 for Mac OS, released in 2003, was a disappointing product that performed poorly. Microsoft now distributes a QuickTime Player plugin, WMV Player, to support Windows Media on the Mac platform.

*Audio formats supported:* for encoding: Windows Media Audio (48 to 192 kbps; variable bitrates also available), Windows Media Lossless, MP3 (128 to 320 kbps; requires installation of plugin); for playback: MP3, Windows Media Audio (WMA) and other native Windows formats, such as Windows Media Video (WMV) and Advanced Streaming Format (ASF).

*URL for download:* <http://www.microsoft.com/windows/windowsmedia>

### **RealPlayer**

RealPlayer is a proprietary multimedia player developed by RealNetworks to support its various RealMedia formats. The first version, released in April 1995 (under the name “RealAudio Player”), was one of the earliest players to support streaming audio, and RealPlayer remains one of the oldest audio players with an ongoing history of development. There are full-featured versions of the player for Windows that play back DVDs, download tracks to portable digital players, rip and burn CDs, and provide an iTunes-like catalog of audio tracks as well as an integrated web browser. RealNetworks charges a fee for the full-featured versions of the player. For simple playback of audio and video, however, there are free “basic” versions that are better suited for a library setting. This basic player is available in versions for Windows, MacOS, Linux (and other versions

of Unix), as well as several handheld and mobile devices. The Helix community also offers a basic player, known as Helix Player.

RealNetworks maintains an archive of “legacy” versions of the player, which is valuable to users searching for an audio player that is compatible with older versions of Windows and Mac OS.

*Audio formats supported:* for encoding: WAV, MP3 (32 to 320 kbps; variable bitrate also available), RealAudio (32 to 320 kbps), RealAudio lossless, MPEG-4 AAC (96 to 320 kbps), and WMA (64 to 192 kbps); for playback: MP3, RealAudio (including RealAudio lossless), WMA, AAC.

*URL for download:* <http://www.real.com>; Real Legacy Software Archive: <http://forms.real.com/real/player/blackjack.html>; Helix Player: <http://player.helixcommunity.org>

### **QuickTime Player**

Apple’s QuickTime Player is designed to play QuickTime audio and video files. For many years, the player has also been licensed to developers for use with QuickTime files in third-party software applications. Apple itself uses the QuickTime Player as the playback engine for its iTunes software.

*Audio formats supported:* MP3, AAC in an MPEG-4 (or QuickTime) file, AAC encoded with FairPlay (.m4p files from the iTunes Music Store), Apple Lossless.

*URL for download:* <http://www.apple.com/quicktime/download/standalone.html>. (Caveat: the default download site for QuickTime Player (<http://www.apple.com/quicktime/download>) offers the player bundled with iTunes.)

### **iTunes**

When iTunes for MacOS was released in January 2001 (a Windows version would follow in October 2003), it set new standards for convenience and usability in personal music management software. The program was developed by Apple to serve many purposes. For a typical user, the chief application of iTunes is to manage the content of an iPod, but iTunes is also an independent media player, rich with features. The program can be used to organize playlists, rip compact discs, encode audio, burn CD-R copies, download podcasts, listen to internet radio stations, and purchase music through the integrated iTunes Music Store. A distinct disadvantage of iTunes as a player, however, is that files must be incorporated into an iTunes playlist before they can be played back; it is not possible simply to open a file and play it. The player itself is based on the QuickTime Player, and the two are most frequently bundled together for download.

*Audio formats supported:* for encoding: WAV, AIFF, MP3 (16 to 320 kbps; variable bitrates also available), AAC (16 to 320 kbps; variable bitrates also available), Apple Lossless; for playback, MP3, MPEG-4 AAC (.m4a), AAC encoded with FairPlay (.m4p files from the iTunes Music Store), Apple Lossless. Will play back WMA files only after converting them to MPEG-4 AAC.

*URL for download:* <http://www.apple.com/itunes/download> (includes QuickTime Player)

## **VLC**

VLC is an open-source media player with versions for Windows, Mac, and various Linux distributions, and it is the only player that could play back all seven of the sample audio files (see fig. 4.1 on p. 32). It is a relatively small program and requires little CPU power, unlike some of the large, feature-rich media players. Although it is not nearly as popular as the four players just covered, I include it to show that there are alternatives to the proprietary players, which often include features that are not needed for simply playback of audio content.

*Audio formats supported:* WAV, FLAC, MP3, RealAudio (using the Cook codec), Windows Media Audio, AAC, Vorbis.

*URL for download:* <http://www.videolan.org/vlc>

## Chapter 5

# Hardware and Software to Support Digital Audio

This section covers the computer hardware, audio components, and software typically used in a digital audio projects.

### 5.1 Equipment

The choice of equipment for a digital audio project will vary depending on the recording formats you are working with and the amount of modification you want to make to the sound signal. A modest course-reserve encoding project based only on compact disc recordings can be put in place with nothing more than a laptop, while a full-blown audio preservation project that includes 78s, LPs, and tapes will require professional-quality audio components to play back the sound as well as additional components to process the audio signal.

Because of the fundamental difference between digital and analog recordings, different processes are used to convert each to digital audio files.<sup>1</sup> A digital recording—such as a compact disc—already consists of digital audio data, so the process involves simply reading the digital data on the disc and storing it as a file on a computer. This conversion can be done by any computer with a CD-ROM drive that is running CD extraction (or “ripping”) software. The time needed to extract the audio from a CD will vary according to the speed and throughput of your computer’s CD-ROM drive and microprocessor, but because the computer is processing data and not sound, extracting CD audio always takes considerably less time than playing back the audio in real time.

Analog recordings must first be converted into digital data before being stored as a file on a computer. This work is done by a digital audio converter (DAC), which is a component part of the computer’s sound card or, optionally,

---

<sup>1</sup>For a description of the differences between analog and digital recordings, see “Analog and Digital Sound” on p. 10 above.

its external digital audio interface.<sup>2</sup> The analog recording is played back on a traditional audio component—typically a turntable or tape deck—whose output is patched into an amplifier or preamplifier, which in turn is patched into the computer’s sound card or audio interface.

The source audio can be modified by additional components and software either during the creation of the digital audio file or afterwards. Some common types of manipulation include eliminating stretches of silence at either end of the recording, adjusting the equalization to boost or suppress certain frequency ranges, or filtering out tape hiss, surface pops on a recording, and other extraneous noises.

### 5.1.1 Encoding workstation

An appropriately equipped computer workstation is the single essential piece of equipment for a digital audio project. The quality of the uncompressed audio created from analog sources and the efficiency of its encoding will be dependent on the capabilities and performance of this central component. When selecting a computer for encoding, the most important features to consider are the microprocessor (speed and bit capacity), memory (size and speed), hard disk (capacity and performance), CD-ROM/DVD drive (speed), and sound card or audio interface (bitdepth and sampling rate).

The specifications of computer workstations are constantly improving. Over increasingly shorter intervals of time, microprocessors double in speed, memory doubles in size, and hard drives double in capacity, while workstations decrease in price. The technology for external storage is also rapidly changing; over the past few years we have moved from recordable CD-ROMs to recordable DVDs and external hard drives to small USB flash-memory drives.

Because the technology is developing so quickly, it would be misleading to suggest specifications for a digital audio workstation, since any recommendation would be outdated as soon as it is made. You can feel confident, though, that when selecting a workstation, you will never regret investing extra money to purchase a faster processor, more memory, or a larger hard drive.

Other specifications for the workstation are ultimately less important, but they also happen to be hotly contested. The operating system, for example: Apple or Microsoft—or even Linux? About one quarter of the survey respondents report that they are using Macs for their digital audio project, with the remainder presumably using Windows.<sup>3</sup> Each has its ardent partisans, and although historically the Mac has had an edge over the Windows platform in the development of hardware and software for multimedia, at this point the platforms perform equally. There is no convincing technological reason to choose one over the other, so your decision should be guided by your choice of software,

---

<sup>2</sup>For more detailed information on the digitization of sound, see “Fundamentals of Digital Audio” on page 13 above.

<sup>3</sup>No one reported using the Linux operating system on their workstation, but software does exist to support a digital audio project using Linux. In fact, in 2001 the author launched a successful digital audio reserves project at the University of Illinois using a Linux workstation for encoding.

since all audio programs are not available in versions for both platforms. If you have no preference for software, then stick with whichever platform is already established in your library, or the one you're most familiar with.

For Windows computers, which microprocessor is best, Intel or AMD? Both perform acceptably well. The factors to consider are compatibility and price. Intel has set the standard for PC microprocessors since the introduction of the IBM PC in the early 1980s. Because Intel chips are an industry benchmark, developers of operating systems, software, and computer peripherals make sure their products are compatible with Intel microprocessors. AMD, on the other hand, manufactures microprocessors that perform as well as Intel's and meet their specifications but at much lower prices. The decision between Intel and AMD can mean a difference of several hundred dollars in the cost of a computer. Although Intel is perceived to be the safer choice, there's no good reason not to purchase a computer with a cheaper AMD chip.

If your digital audio project is small in scale, and your budget will not allow the purchase of a dedicated workstation, you can easily mount a digital audio project using a workstation that is shared with other applications. There are advantages, however, to devoting a workstation exclusively to digital audio. When a computer is shared with other applications, it can easily become cluttered with programs and multiple processes running in the background, which will degrade the general performance of the machine. You will find that even a computer fresh out of the box will have several programs running in the background by default.<sup>4</sup>

Thirty-one of the respondents to our survey reported on the equipment they use for their digital audio project. Of these, five (16%) are digitizing only compact discs, and two of these are doing the work on a laptop. Most libraries devote a single computer to encoding; only three respondents (about 10%) reported that their encoding workstation is shared with other applications.

### **Sound card or audio interface**

Every computer manufactured today is equipped to record and play back sound. Sound capabilities are provided by a soundcard, either integrated into the computer's circuitry or installed as a separate component plugged into a slot on the motherboard. Typically, this preinstalled sound card, however, is of mediocre quality, intended simply to play back system sounds and music for recreational listening.

If you are digitizing analog recordings (LPs and cassette or reel-to-reel tapes), the quality of audio your workstation can produce will be greatly improved through an upgrade to the sound card supplied with the computer. Until a few years ago, this meant installing or replacing a card inside the computer. With the introduction of Firewall and USB2, the most convenient way to upgrade your computer's sound system is to purchase an audio interface, a stand-

---

<sup>4</sup>If you don't know how to find these and disable them yourself, ask someone in your systems office to review these background processes with you to make sure that only essential ones are running.

alone component that plugs into the computer's Firewall or USB2 port. Also, an independent, external audio interface can easily be moved from one computer to another, which simplifies upgrading and replacing the encoding workstation.

Audio interfaces range widely in features, functionality, and cost. Some of the more popular manufacturers are M-Audio, Edirol, Tascam, and Mark of the Unicorn (MOTU). A few years ago, a typical stock sound card installed in a computer offered 16-bit resolution and a 48 kHz sampling rate. As use of digital audio technology has grown, audio interface technology has become more sophisticated, offering deeper resolution (24-bit is now standard) and faster sampling rates (up to 192 kHz). Three survey respondents offered information on the specific brand and model of the audio interface used for their project, and in all cases, the interface offered 24-bit resolution and a sampling rate of up to 96 kHz.<sup>5</sup> Some studio-quality audio interfaces offer 32-bit resolution, and inevitably the technology will continue to advance over time.

### 5.1.2 Audio components

#### Playback

A digital audio project drawing exclusively on a compact disc collection requires no equipment other than a computer with a CD-ROM drive. Most libraries, however, are digitizing analog recordings as well as compact discs, and these formats require additional components to play back the recordings. The components are usually patched into a mixer or amplifier, which in turn is patched directly into the computer's sound card or audio interface. To work with tape recordings, components are added that are appropriate for playback of tape through a traditional sound system—for example, a cassette, reel-to-reel, or DAT deck.

The quality of sound that can be reproduced from older analog recordings will depend on the condition of the recording and the quality of the component playing it. Libraries that have been providing listening services for several decades will have turntables and tape decks that are probably, at this point, underused and can be co-opted for the digital audio project.

The high-quality playback that is essential to preservation work can be provided only by high-end components, and some are designed with preservation in mind. A laser turntable, for example, allows playback of an LP or 78 without contact with the disc, so repeated playback results in no wear whatsoever.

The equipment configuration used for most digital audio projects consists of a computer with a CD-ROM drive, a turntable, a cassette deck, and an amplifier.<sup>6</sup> The analog recording format most frequently digitized by libraries is the

---

<sup>5</sup>An argument can be made that there is no reason for the sampling rate ever to exceed the 44.1 kHz rate of CD audio because of the 20 kHz limit to human hearing and the Nyquist Theorem's formula for sampling at twice the rate of the highest frequency to be reproduced. Sampling above 44.1 kHz certainly produces higher-quality audio, but the enhanced quality cannot be perceived by the human ear.

<sup>6</sup>For more information on audio components, see Jim Farrington, *Audio and Video Equipment Basics for Libraries*, Music Library Association Basic Manual Series, no. 5 (Lanham, MD: Scarecrow Press, 2006).

LP. Eighteen (58 percent) of the thirty-one survey respondents have a turntable devoted to their digital audio project. Next in frequency, with fifteen (48 percent) is the cassette deck. After that, the numbers fall off sharply. Three libraries use DAT decks, and two use reel-to-reel decks.

Some compact discs cannot be read by a computer's CD-ROM drive. This is most often occurs with older compact discs. These discs must be played on a traditional compact-disc player, which is patched into the system as if it were a turntable or tape deck. Two of the survey respondents have independent CD players devoted to their digital audio projects.

### **Mixer**

If your digital audio project draws on a number of audio components for sound audio, then you will benefit from adding a small mixer, a scaled-down version of the large mixing boards seen in recording studios. A mixer allows you to switch smoothly between multiple input sources without removing and replacing cables. An amplifier can provide the same functionality, but typically an amplifier can accommodate no more than three or four input sources. A eight-input or sixteen-input mixer should fill the needs of most digital audio projects.

Each input of a basic mixer has jacks to receive the input signal and a dial or slider called a potentiometer (or "pot"). The pot is used to control the volume of each input and allows the user to fade a source in or out. For example, if you typically include a spoken announcement or description at the beginning of an audio-reserve selection, the mixer's potentiometers will allow you to make a seamless transition from the announcement to the music by turning down the volume of the microphone after the announcement while turning up the volume of the turntable or other audio component. More sophisticated mixers have input sensitivity controls, filters, equalizers, and other features.

A mixer also provides multiple outputs. One output will be routed to your computer's soundcard or audio interface. You may choose to have another output routed to an amplifier equipped with speakers so that you can cue up sound recordings and play back the audio you have created on something other than the small speakers supplied with most computers. Mixers also have a headphone jack so that you can monitor exactly what is being output by the mixer.

### **Other components**

Libraries that are digitizing for preservation may use disc-cleaning machines, filters, noise-reduction modules, and equalizers.

### **Input types**

A sound card or audio interface usually has two types of input jacks: low signal and high signal. The low-signal input is used for microphones and other devices that produce a weak signal requiring amplification by the sound card. The jack is usually marked "mic in" and is often colored red.

The high-signal input is used for electronic components, such as tape decks, DVD players, VCRs, amplifiers, receivers, and mixers. A number of different labels are used for the high-signal input. Some of the most common are “aux in”, “line in,” or “audio in.”

A turntable is one of the few audio components that uses a low-signal input, but turntables are usually routed into into a preamplifier or amplifier rather than directly into the sound card. The preamplifier or amplifier, in turn, would be plugged into a mixer or directly into the high-signal input jack.

It is important not to confuse the low- and high-signal inputs. An electronic component plugged into the sensitive low-signal input will overpower it and produce distorted sound. On the other hand, a microphone plugged into the high-signal input will be barely audible.

### **5.1.3 Servers**

A server is a computer that provides services on request to computers on a network. The services are often related to shared resources—such as printers, data files, media files, web pages—and the server is expected to provide them at any time, day or night.

In most library settings, the installation and maintenance of servers is under the purview of an information technology department, either in the library or at a broader institutional level, and rare is the case where a librarian is expected to maintain a server.

All network-based digital audio projects, however, rely on services provided by servers, so anyone managing digital audio project should have a basic understanding of what servers are, how they differ from desktop computers, and what kind of software they run to provide streaming audio services.

#### **Servers compared to desktop workstations**

A computer used as a server contains many of the same components as a desktop workstation—a microprocessor, memory, a hard drive, a network connection. In fact, a standard-issue desktop computer is equipped to provide the functionality of a server, and for small applications that deliver modest services to a few users, a desktop computer can simultaneously fill the dual roles of desktop workstation and server. A computer intended to act as a server works most efficiently, however, when it is dedicated solely to server tasks and its hardware is optimized for performance as a server.

Since most of the work of a server is related to delivering data, computers designed for use as servers are equipped with high-performance hard drives and network interfaces that deliver data far more quickly than a standard desktop computer. Also, very little processing power is required to deliver data, so servers have no need for the powerful microprocessors required by desktop computers to manipulate and display data. Because commands are issued to servers remotely from other computers, there is no need for a video card, monitor, key-

board, or other external peripherals. A typical server consists of a standalone box or a component mounted on a rack.

### **Operating systems**

Like desktop computers, servers can run a variety of operating systems, and systems specialists enjoy debating the relative merits of each. The most common are Unix (in its many flavors, most notably FreeBSD, Solaris, and Linux—which itself comes in many flavors), Microsoft’s Windows Server System, and Apple’s OS X Server. Unlike the world of desktop computers, where Microsoft dominates, among server operating systems, Unix traditionally has been the major player. In fact, Unix is most common operating system for web servers by a wide margin, although Microsoft recently been making gains.<sup>7</sup>

The choice of operating system is driven partly by the server hardware. Microsoft and Unix operating systems run on Intel-based processors (Unix can also run on several other classes of processor), and Apple’s operating system runs on Apple servers.

The choice of operating system is usually transparent to the user of a server, and one should function as well as another. Some information technology departments, however, have rigid preferences for certain operating systems and hardware, and when planning a digital audio project, you might be expected to work with whatever server operating system happens to be supported. Smaller information technology departments, in particular, will be resistant to adopting a new server platform to support a single project.

This becomes an issue when planning a digital audio project, because the choice of operating system will, in turn, define the possible choices for the software that will drive the digital audio service.

### **Software for delivering audio**

For digital audio projects, servers provide two basic functions: storing audio files and delivering them to users. There are several ways a server may deliver digital audio to a client. The two most common are (1) using a web server to make a digital audio file available for downloading and (2) using a streaming server to deliver a stream of digital audio data to a digital audio player.

The first option is the easiest, since web servers are ubiquitous, and the question becomes simply one of storage space. When a user keys the appropriate URL into a browser or clicks a link to the URL on a web page, the web server delivers the file to the browser, and once the file is transferred completely, it can either be played back or stored on the local computer. Some audio formats—such as Microsoft’s Advanced Streaming Format (ASF)—provide

---

<sup>7</sup>According to a February 2006 survey of more than 76 million websites at [http://news.netcraft.com/archives/web\\_server\\_survey.html](http://news.netcraft.com/archives/web_server_survey.html), the Apache web server (a Unix application) was used by 67 percent of the sites, while Microsoft’s server was used by 21 percent.

“progressive playback,” which allows playback of a file to begin before it is fully downloaded.<sup>8</sup>

Delivering audio as a stream requires a streaming server, which is simply a server running streaming software. If your institution already maintains a streaming server, you might investigate the possibility of adding your digital audio service to the existing server. In fact, most digital audio projects end up sharing a streaming server with other streaming media projects. Seventy-nine percent of the respondents to our survey reported that their service shares a server with other applications.

As mentioned earlier, the choice of server operating system can dictate the choice of streaming server software, which will in turn dictate the type of audio files you can deliver through your service. According to our survey, the two most popular streaming servers are RealNetwork’s Helix Server (along with its predecessor, RealServer) and Apple’s QuickTime Streaming Server. The Helix Server software is available for the Windows 2003 Server, Linux, and Solaris operating systems, and the server can deliver files in RealAudio, Windows Media, QuickTime, MP3, and AAC formats. The QuickTime Streaming Server software runs on Apple’s OS X and can stream audio contained in QuickTime (.mov) files as well as AAC (.mp4) files.

Some smaller libraries have even set up a digital audio service by ripping CD tracks in iTunes and making the resulting library available on the library’s network.

## 5.2 Software (by John Anderies)

### Audacity

Audacity is a free, open-source digital audio editor that runs on the Mac, Windows, or Linux operating systems. It allows the user to play, record, and edit sound files including the WAV and Ogg Vorbis formats. With the addition of a plugin called the LAME MP3 Encoder, Audacity will handle the MP3 format as well. Audacity does not have the built in capacity to rip compact discs but is an excellent tool for converting analog audio to digital audio, as well as for manipulating the digital file once it is created.

### iTunes

In addition to being an online audio player, storage system, and music store, iTunes may be used as a digital-audio editor too. Developed by Apple Computer, iTunes is available free for download and is compatible with both Mac and Windows operating systems. iTunes can play back a variety of file formats

---

<sup>8</sup>As we mentioned above, though (see “Streaming,” p. 28), using download technology to deliver copy-protected audio exposes you to the risk of legal action. Be sure to check with your institution’s legal department before setting up a listening service that allows users to download and save copies of copy-protected audio.

including AAC, AIFF, MP3, MPEG-4, and WAV. However, it is limited to ripping compact discs and converting digital audio from one format to another.

### **QuickTime Pro**

Like iTunes, QuickTime Pro was developed by Apple Computer and is compatible with both Mac and Windows operating systems. QuickTime Pro is available for purchase and includes a number of features that its free relation (QuickTime) does not include. QuickTime Pro may be used to convert and compress digital audio in a variety of formats including AIFF, MPEG-4, and WAV.

### **RealProducer**

RealProducer Basic and RealProducer Plus are proprietary products of RealNetworks that convert digital audio in a range of file formats (including AIFF, AVI, MP3, MPEG-4, and QuickTime) into compressed RealAudio files (.ra, .ram, .rm) for distribution through RealServer or Helix Server. RealProducer Basic is free, while RealProducer Plus is available for purchase. Both run on the Windows and Linux operating systems only. RealProducer Plus includes features such as unlimited target audiences/bitrate streams and batch processing.

### **Sound Forge**

Sound Forge is a digital audio editing and creation tool produced by Sony Media Software (formerly produced by Sonic Foundry) and available for purchase. It is compatible with the Windows operating system only. Sound Forge will rip compact discs as well as convert analog audio to digital audio. It offers a full suite of audio editing tools and can export files in AIFF, MP3, Ogg Vorbis, RealAudio, WAV, and WMA formats.

## **Part II**

# **Digital Audio in the Library**

## Chapter 6

# Sources for Digital Audio Recordings

### 6.1 Digitizing Library Local Collections

Early in the days of digital library projects, we read quite a bit about the prospect of entire research libraries being digitized; some writers were so bold as to predict a date. In these early days, it was assumed that libraries—either singly or cooperatively—would take on the responsibility for the digitization of library collections. During the past decade, however, commercial enterprises have taken the lead in the digitization of print content, and recording companies have been entering into licensing agreements with various online services for the delivery of commercial sound recordings over the internet.

As library digitization projects have become more numerous and more sophisticated, libraries have moved beyond the idealistic (and ultimately impractical) goal of digitizing “everything” to the more realistic goal of digitizing only content that is rare and in some cases unique—content outside the scope of commercial digitization enterprises, content that highlights the materials that distinguish one library different from another. Focusing on special collections not only brings unique content to the public; it also provides a promotional tool that can showcase a library and its host institution.

Most music libraries own such collections of unique, noncommercial sound recordings. For a college or university library, these might be recordings of concerts, recitals, and lectures that have taken place on campus, or field recordings made by a researcher and donated to the library. For a public library, it might be recordings of local community musical groups or guest lectures.

Because these recordings are unique, they are also irreplaceable, so most libraries have imposed restrictions on their use in order to protect them. Digitizing these special recordings and making them available digitally accomplishes two goals: the recordings are more easily accessible—available to all listeners, both inside and outside the library—and they are preserved. Once the sound

recording has been digitized, there is no further need to use the original recording, so the original sound recording can be stored permanently and is protected from any damage it might receive through use.

## Chapter 7

# Curricular Listening

In academic libraries, the sound recordings that are used most heavily are usually those assigned by instructors for course-related listening by students. In larger universities, music appreciation and music history survey courses often have enrollments of hundreds of students spread over several sections. The puzzle of how best to deliver listening assignments to large groups of students simultaneously has been dogging music librarians for decades.

Some instructors will require students to purchase a set of recordings that accompany the assigned text for the course, in which case the library is relieved of the responsibility. In most cases, though, the instructor will prepare a listening list that is tailored to suit his or her preferences for repertory and performances.

When supporting a customized listening list, the library cannot accommodate the needs of a large class—particularly on the night before an exam—by simply placing the library’s copy of the various recordings on the reserve shelf. Through the years, music libraries have turned to state-of-the-art audio technologies for solutions to the problem of providing an effective reserve listening service for heavily enrolled classes.

In the 1970s and 1980s, reserve listening was provided by copying LP recordings to reel-to-reel or cassette tapes in multiple copies for students to borrow. (Some libraries piped the recordings from a central tape player to multiple listening carrels in a listening center.) In the 1990s, with the advent of recordable compact discs, this same technique was transferred to the new technology; libraries burned circulating copies of the listening assignments on CD-Rs.

In the mid-1990s, many libraries quickly adopted new streaming-audio technology to provide reserve listening over the internet, and today, a growing number of libraries are making use of commercial subscription services (such as Classical Music Library and Naxos Music Library) and digital music players (such as iPods) to provide reserve listening assignments.

In order to learn more about current practices in music libraries, I posted

a note on MLA-L,<sup>1</sup> the email discussion list of the Music Library Association, calling for volunteers to fill out a survey that asked questions about software, hardware, access, and staffing for the digital audio services offered in their libraries. Forty-two librarians responded to the survey, and while the survey is by no means scientific, it was the best method I had to identify current practices in the field. So before proceeding, I offer this caveat: when I refer to “most libraries,” “few libraries,” or “no libraries,” I am drawing conclusions based on practices in the forty-two libraries represented in the survey results.

## 7.1 Streaming Service

The earliest widespread application of digital audio technology was in the mid-1990s, when several music libraries began providing streaming audio for recordings placed on reserve for course assignments. The focus on curricular listening assignments was a natural choice, since reserve recordings are heavily used and usually constitute a comparatively small, well-defined collection.

Streaming technology is still the most popular means of providing network delivery of reserve listening. It allows libraries to provide around-the-clock reserve listening both on campus and off, and students appreciate the convenience of being able to whenever and wherever they please.

### 7.1.1 Creating and Managing the Sound Files

The most time-intensive activities in providing a streaming-audio reserves service are encoding, describing, organizing, and preserving the digital sound files. There are a number of ways to approach these tasks, and careful thought and planning will pay off once your project is in production.

### 7.1.2 Encoding

A crucial step in planning a streaming audio project is selecting the format and bitrate for the compressed audio files that will be streamed to users. The quality of the audio produced for your project will be affected by your choices, since some formats produce higher-quality sound than others at identical bitrates. The bitrate will affect the performance of your service, since streaming audio at higher bitrates require greater bandwidth and faster internet connections.

When selecting a format for your streaming digital audio project, the factors that are likely to have the greatest impact on your decision are the ones that, in the end, will probably be invisible to you: the streaming server and the staff in charge of maintaining it. If your project is to be hosted on an existing media server maintained by personnel outside your departmental library, you will probably be expected to work within the limitations of that server, in which case certain decisions will already have been made for you.

---

<sup>1</sup>mla-l@listserv.indiana.edu; archives available at <http://listserv.indiana.edu>.

A factor of less concern in selecting a streaming audio format is sound quality; all compressed formats, at a sufficient bitrate, will deliver acceptable audio to users. Keep in mind, though, that newer formats—such as AAC and Ogg Vorbis—produce audio of higher quality than older formats—MP3, for example—at an identical bitrate.

The choice of bitrate for the audio files will determine not only the quality of sound but the minimal connection speed for your service. Most users now access the internet using fast broadband connections—ethernet, cable, or DSL—and these can easily handle streams of 128 kbps. If you know that a significant number of your users connect to the internet using slower modem connections, you might consider encoding at two different rates (perhaps 48 kbps and 128 kbps) or use a format (such as RealAudio's Surestream) that can accommodate streams at multiple rates. It is clear, however, that the days of slow, dial-up modems are numbered, so it is better to err on the side of higher bitrates if you want to extend the usability of your compressed files.

Another factor to keep in mind when selecting a streaming audio format is your users' preferences for operating systems. Media players exist for most streaming formats in versions for Windows, MacOS, and Linux, so most streaming services will be compatible with all three operating systems. RealAudio and the MPEG-based formats, because they have a long history and are well established, are safe choices if you want to be sure that your service is accessible on all platforms. On the other hand, if your library uses Macs for public workstations and most of your users own Macs, then selecting format tailored for Windows (such as Windows Media Audio) would be a poor choice. If a strong preference for a specific operating system exists at your institution, then it will probably influence the choice of the streaming server, and the capabilities of the server will, in turn, influence your choice of a streaming audio format.

In the end, a successful streaming audio service can be based on any common compression format. Your decision should be made in reverse. Start by talking with the staff who will manage the streaming server that will host the service. They will tell you which formats are compatible with the server. Once you have a short list of possible formats, try playing audio streams in each format off the web.<sup>2</sup> If you want to support multiple players or operating systems, be sure to try the streams using every possible combination and take note of whether software or plugins need to be installed to playback the stream. This should give you some idea of what will be required of your users to configure their systems to play back streams in each format. Finally, consider the quality of the audio.

The survey revealed that most libraries use one of two formats for their streaming audio: MP3 (47 percent of the respondents) and RealAudio (39 percent of the respondents). Use of Quicktime (presumably AAC files in a QuickTime wrapper) was reported by 11 percent of the respondents. Also, some of the libraries that use the MP3 format specified that the files are streamed in a

---

<sup>2</sup>Radio stations are an excellent source for testing the compatibility of various audio streams. The Public Radio Fan site (<http://www.publicradiofan.com>), for example, indexes stations according to a number of criteria, including streaming format.

QuickTime wrapper. One library (3 percent) bases its audio service on Windows Media Audio.

### **Naming and Organizing Files**

Over the course of a few semesters, even the smallest digital audio reserve project can generate hundreds of audio files. Labeling and organizing a large group of files presents special challenges that can be approached in a number of ways. The two fundamental decisions to be made are (1) how to assign names to the audio files and (2) how to organize them into folders (or directories). It is best to settle on a system for naming and organizing files before you start encoding. Try a number of methods and test them on a small scale to predict how they might work with your service. Fixing problems later, after you have processed hundreds of files, will be difficult and costly, so this will be time well spent.

From the survey of forty-two libraries providing digital audio services, I learned that there are nearly as many ways to name and organize digital audio files as there are libraries providing digital audio services. Although few libraries use identical methods, all use a combination of data elements drawn from the same five categories of data. Each category has certain advantages and disadvantages.

#### **Musical content**

*Typical data elements:* composer's last name, title of composition, title of component part(s) of composition, performer

*Advantages:* Content of the audio files can be easily identified

*Disadvantages:* Often requires a long string of characters in order to produce a unique file or folder name; names and titles must be standardized (authority control); the source recording can be identified only by searching the catalog

#### **Container (the physical sound recording)**

*Typical data elements:* label name, label number, disc number, track number

*Advantages:* Unique file and folder names can be easily constructed; source recording can be quickly identified

*Disadvantages:* Requires a search of file metadata or the library catalog to determine the content

### **Shelving number (used for filenames only)**

*Typical data elements:* call number or accession number used for shelving in the local collection

*Advantages:* Produces short filenames that are easily constructed; the source recording can be retrieved quickly

*Disadvantages:* Requires a search of file metadata, a catalog search, or retrieval of the item to determine content of the audio file

### **Bibliographic record number (used for filenames only)**

*Typical data elements:* local bibliographic record number, bibliographic utility (OCLC, RLIN) record number

*Advantages:* Short, unique filenames that are easily constructed and efficiently searched in the local catalog

*Disadvantages:* Requires a search file metadata or a catalog search to determine content

### **Curricular function**

*Typical data elements:* year/semester, course name, course number, instructor

*Advantages:* Easy management of files and folders by course and instructor; file and folder names can be directly related to the instructor's course syllabus

*Disadvantages:* Data elements change from semester to semester; multiple copies of file must be created when more than one course uses the same track; requires a search of file metadata to determine content

A filename should fill two functions: to provide a unique, concise identifier for the file that will distinguish it from all other files, and to provide information that will allow staff to determine the content of the file and its source. In some ways, these two functions are at cross purposes: data that does the best job of providing short, unique filenames (bibliographic record numbers, for example) require database searches to reveal a file's content; data that does the best job of describing the file's content (composer, title, label, label number) often produces long, unwieldy filenames.

Looking over the current practices in the libraries represented by the survey, files are most often named by using data either based on the musical content, shelving number, or bibliographic record number. Using these elements allows files to be used for different courses and reused from semester to semester. Few libraries construct filenames based on the curricular function of the file. Here are some examples of how librarians have approached the naming and organization of sound files, taken from responses to our survey:

MUS101/Puccini-Tosca-Sabata-EMI-1953-01-Ah\_Finalmente [In folders by course number; individual tracks named using composer last name, work title, conductor, label, label number, disc number, track title]

2005FA/MEN3335/dahl\_concerto\_mvt1 [In folders by year and semester, then course number; files named using composer last name, title, movement]

Music240/00000000-02-11 [In folders by course number; individual tracks named using OCLC number, disc number, track number]

muen/275/bach\_cpe\_magnificat\_magnificat\_anima.4211482 [In folders by course, number; files named using composer last name and first name, work title, component part, label number]

MozartSym41i [Composer, short title, movement]

cd1234.2.3 [File name using accession number, disc number, and track number]

Blue-Ridge-Ramblers.Jug-Rag.CD12134-2-1 [File name using performer name, track title, accession number, disc number, track number]

Schumann-Dichterliebe-07-Ich grolle nicht [File name using composer last name, work title, track number, title of component part]

cd-26504.05 [File name using accession number, track number]

When naming a file based on its musical content, you can select the data using a number of sources. The most reliable source would be the appropriate MARC fields in the recording's bibliographic record—the 100, 240/245, 700 \$a \$t, etc.—so that the form of the composer's name and the title of a work will be consistent for all files. The process of looking up authorized headings can be time consuming, however.

Some libraries rely instead on metadata supplied by large databases of compact-disc data, such as Gracenote's CDDB and Freedb.<sup>3</sup> Because most encoding programs automatically query these databases for metadata, information on composer, performer, album title, and track titles can be imported automatically, eliminating the need for data entry for all but the most obscure compact discs. Also, encoding programs can often be configured to construct file and folder names automatically, based on the metadata elements retrieved from CDDB or Freedb.

The disadvantage to these compact disc metadatabases is that the data is contributed by users, and no standards of consistency are applied to the data. Also, CDDB metadata can include "extended characters" (letters with umlauts and accents as well as other special characters). If you are relying on CDDB data to name your files, be sure to test filenames that include extended characters and punctuation before you move your project into production. Some servers and audio players can process these filenames without a problem, but others choke completely on the extended characters and are unable to play the file. Often the encoding program or player offers an option to strip out extended characters from filenames and metadata. At least one librarian uses an external program to convert the characters.

---

<sup>3</sup>See <http://www.gracenote.com> and <http://www.freedb.org>.

## Backup protection of content

Any digital audio project, regardless how small, represents an investment of time, and because the creation of digital audio from analog sources like LPs and cassettes must be done in real time—a half hour of music requires at least a half hour for encoding—a project that relies heavily on analog sources requires a significant investment in time. Protecting this investment, on the other hand, requires little time and expense, especially when compared to the work that would be required to recreate the project.

The most time-intensive step in the encoding process is the creation of the uncompressed audio file. If this file is preserved, the time spent creating it will not need to be lost in the event of a natural disaster, the malicious vandalism of a hacker, or something more innocuous—like the migration to a different format for audio streaming.

Compressed audio files, on the other hand, are often not backed up, for two good reasons. First, the time required to recreate a compressed file is relatively insignificant (so long as you have retained the source CD or uncompressed audio file) and as microprocessors grow faster, that time becomes increasingly negligible. Second, as network speeds increase and compression technologies become more sophisticated, it is inevitable that an audio project will eventually migrate to a new compression format for delivery, in which case the compressed files will eventually need to be recreated.

Judging from the responses to the survey, the majority of libraries do not create and maintain independent archival copies of audio files—compressed or uncompressed—for their digital audio-reserve services. Less than one third of the libraries routinely archive the source files. This does not mean that most libraries routinely destroy source files once they have been used. Often they will reside on the encoding computer, but no backup is kept beyond this initial copy, and if a disaster were to strike the encoding computer, the work would be lost.

Even libraries who do routinely maintain archival copies of uncompressed files choose not to back up uncompressed files ripped from compact discs. The compact discs themselves can serve as a backup, and if the compressed files created from a particular CD are lost—or need to be recreated in at a different bitrate or in a different format—the collection's copy of the source CD can be put into service. Because replacement files can be digitized from a compact disc at a rate much faster than real time, backing up uncompressed source files for compact discs becomes a practice that consumes time and storage space when very little risk is being assumed.

When backing up data, regardless of the medium, it is important to keep the backup copies separate—physically and virtually—from the data being protected. There are three back up methods reported in the survey:

*Remote server.* Files are copied to a remote server, which itself is routinely backed up by IT personnel. This method is simple and quick, since it places the bulk of responsibility on the administrator of the server, and

the library simply needs to make sure that files are copied to the server as soon as they are created.

*Recordable discs.* Files are copied to recordable compact discs (CD-Rs) or recordable DVDs. This method has the advantage of keeping the archival copies easily accessible and under the full control of the library, but it also requires the library to maintain a regular back up schedule and devise a routine for organizing and storing the CD-Rs.

*External hard drive.* Files are copied to an external hard drive, typically a USB drive that can attached easily to the computer for backup and then removed for storage in a safe place. The advantages are ease and speed of the backup; the disadvantage is relying on a storage device that itself might be subject to failure.

With backups, redundancy is important. Better to set up two redundant backup methods than to rely on one whose failure would spell disaster.

### **7.1.3 User interface**

Of the thirty-four librarians who offered information on the interface used for their digital audio reserves service, nineteen (56 percent) report that links to the audio selections are provided through a course management system, such as BlackBoard or WebCT. Seven (21 percent) use independent web pages, six (18 percent) use their online catalog and/or its reserve module.

Although none of the survey respondents report using Apple's iTunes software, we know from distribution-list postings and reports at conferences that some libraries provide listening assignments by maintaining a network-accessible shared "library" on iTunes.

### **7.1.4 Staffing**

The survey revealed that libraries have used a number of successful staffing models for their audio reserves service. Most services are launched with little or no increased staffing. Roughly three quarters of the respondents were able to launch a digital audio reserves project by redeploying existing staff out of necessity; it is often a decision between adjusting job assignments of existing positions or not taking on the project at all.

For those libraries that did receive an increase in support for the project, the added staff ranged from one 20-hour per week student assistant to 1 FTE support staff and five student assistants. In several cases, the work for the project initially was absorbed by existing staff, but once the project was established and grew in scope, a case could be made with the library administration for adding staff.

### **Library's responsibilities**

For a digital audio reserves project, the library is almost always responsible for encoding the audio, uploading it to the server, and providing an interface for access.

Once a digital audio reserves project is up and running, most of the ongoing work involves digitizing and encoding the source recordings. Among the thirty-three respondents to the survey who answered questions about staffing, twenty-five (76 percent) report support staff are involved in digitization and encoding. Twenty-two (67 percent) use student assistants, usually in combination with support staff, although four libraries (12 percent) have only students working on recordings. Thirteen (39 percent) reported that a librarian is involved, and in four cases librarians do all of the digitization and encoding work.

For preservation projects and large digital audio reserves projects, the library may employ a full-time audio specialist to digitize and encode the audio.

Back in the days of tape-based reserve services, the responsibility for creating listening tapes rested with the instructor—or more commonly, with the instructor's teaching assistant—and the library simply handled the reproduction and circulation of the tapes. This model continues at several institutions who replied to the survey. Teaching faculty and staff create audio files for their listening lists and sometimes also mount them on a central server or on the course-management software site for their course. This is the exception to the norm, though, and although some faculty and staff prefer to have control over their curricular listening assignments, it is in the best interest of the students and the institution to have curricular listening services centralized to provide a uniform interface and uniform quality and to make sure intellectual property laws are observed.

### **Systems staff responsibilities**

Systems staff are typically responsible for maintaining the server hardware and the streaming software. Depending on how technology support is shared in an institution, these staff might be employed by the departmental library, the general library, or the campus IT department.

**Part III**

**Appendix**

## Appendix A

# Resources for Questions on Copyright

When setting up a digital audio reserves service or any other service that involves the delivery of copy-protected recordings over a network, you should work with your institution's legal department to insure that the proposed service is acceptable within their interpretation of the Copyright Law. You may find that your institution's legal services department will play a large role in determining the content of your digital audio service and its access. Institutional legal departments can vary greatly in how much risk they are willing to allow their institution to assume. Some will prohibit any services involving copy-protected recordings—even if access is restricted. At the other extreme, there are institutions that will allow instructors to rip CDs and upload MP3 files to a courseware site for students to download to their iPods. You are best advised to clear your service with your legal department rather than to see it shut down a few weeks into production.

As with any law, the copyright law can be read a number of ways, depending on one's point of view and personal interests. In February 1996, the Music Library Association's Legislation Committee issued the statement below, which supports the digitization of reserve materials and their delivery over networks. The statement can be useful in explaining to apprehensive library administrators how the law makes provisions for digital audio reserves.

### **MLA's "Statement on the Digital Transmission of Electronic Reserves"<sup>1</sup>**

Music educators cannot effectively teach the structure of a musical work without providing aural access to the complete work. Attempting to comprehend an entire musical composition through

---

<sup>1</sup>Music Library Association, "Statement on the Digital Transmission of Electronic Reserves," c1996-2002, <http://www.lib.jmu.edu/org/mla/guidelines/accepted\%20guidelines/Digital\%20Reserves.asp> (Accessed 19 November 2005).

excerpts, or even sections, is no more effective than attempting to comprehend a novel, architectural plan, poem, or painting in the same manner. At best, only a sense of style is conveyed, not compositional structure. Additionally, educators who teach the history, culture, theory, composition, or performance of music require the flexibility to select the compositions they teach based on educational relevance and instructional objectives. Recognition of the appropriateness of providing such flexibility in instruction is expressed within Section 110 of the copyright law, which states:

Notwithstanding the provisions of section 106, the following are not infringements of copyright:

(1) performance or display of a work by instructors or pupils in the course of face-to-face teaching activities of a nonprofit educational institution, in a classroom or similar place devoted to instruction, unless, in the case of a motion picture or other audiovisual work, the performance, or the display of individual images, is given by means of a copy that was not lawfully made under this title, and that the person responsible for the performance knew or had reason to believe was not lawfully made;

The American Library Association's "Model Policy Concerning College and University Photocopying for Classroom, Research and Library Reserve Use" (*C&RL News* [April 1982]: 127–131), as drafted by Mary Hutchins, states the view that the library reserve room may be considered an extension of the classroom. The Music Library Association fully supports this view as well as the consequent view that students enrolled in a class have the educational right to aurally access its assigned musical works both in the classroom and through class reserves. The MLA also believes that the dubbing or digital copying of musical works for class reserves falls within the spirit of the fair use provision of the copyright law.

In light of the above, the Music Library Association supports the creation and transmission of digital audio file copies of copyrighted recordings of musical works for course reserves purposes, under the following conditions:

Access to such digital copies must be through library-controlled equipment and campus-restricted networks.

Access to digital copies from outside of the campus should be limited to individuals who have been authenticated: namely, students enrolled either in a course or in formal independent study with an instructor in the institution.

Digital copies should be made only of works that are being taught in the course or study.

Digital copies may be made of whole movements or whole works.

Either the institution or the course instructor should own the original that is used to make the digital file. The Library should make a good faith effort to purchase a commercially available copy

of anything that is provided by the instructor.

The library should remove access to the files at the completion of the course.

The library may store course files for future re-use. This includes the digital copy made from an instructor's original if the library has made a good faith effort to purchase its own copy commercially.

## For Further Reading on Copyright

The following citations are offered for reference when grappling with questions of copyright in the management of digital audio services.

### U.S. Copyright Law

Title 17 of the U.S. Code <http://www.copyright.gov/title17> (Accessed 18 November 2005).

Digital Millennium Copyright Act (H.R. 2281)  
<http://lcweb.loc.gov/copyright/legislation/hr2281.pdf> (Accessed 18 November 2005).

### Overviews

Fries, Bruce, and Marty Fries. *Digital Audio Essentials*. Chapter 17, "Digital Audio and Copyright Laws." Sebastopol, Calif.: O'Reilly, 2005. ISBN 0596008562. pp. 317-30.

Offers ten hypothetical test cases to illustrate what practices are and are not acceptable under existing laws.

Frith, Simon, and Lee Marshall, eds. *Music and Copyright*. 2nd ed. New York: Routledge, 2004. ISBN 0415972523.

Jeweler, Robin. "Copyright Issues in Online Music Delivery." In John V. Martin, ed., *Copyright: Current Issues and Laws*, 97-107.

Martin, John V., ed. *Copyright: Current Issues and Laws*. New York: Nova Science, 2002. ISBN 1590332687.

Schrader, Dorothy. "Digital Millennium Copyright Act, P.L. 105-304: Summary and Analysis." In John V. Martin, ed., *Copyright: Current Issues and Laws*, 131-52.

Vaidhyanathan, Siva. *Copyrights and Copywrongs: The Rise of Intellectual Property and How It Threatens Creativity*. New York: New York University Press, 2001. ISBN 0814788068.

Weimer, Douglas Reid. "The Copyright Doctrine of Fair Use and the Internet: Case Law." In John V. Martin, ed., *Copyright: Current Issues and Laws*, 109-15.

## **Electronic Reserves**

American Library Association. Association of College & Research Libraries. "Statement on Fair Use and Electronic Reserves." November 2003. <http://www.ala.org/ala/acrl/acrlpubs/whitepapers/statementfair.htm> (Accessed 19 July 2006).

Music Library Association, "Statement on the Digital Transmission of Electronic Reserves," c1996-2002, <http://www.lib.jmu.edu/org/mla/guidelines/accepted%20guidelines/Digital%20Reserves.asp> (Accessed 19 November 2005).

## **Preservation**

Besek, June M. *Copyright Issues Relevant to Digital Preservation and Dissemination of Pre-1972 Commercial Sound Recordings by Libraries and Archives*. CLIR Publication, no. 135. Washington, DC: Council on Library and Information Resources and Library of Congress, 2005. ISBN 1932326235. <http://www.clir.org/pubs/reports/pub135/contents.html> (Accessed 9 December 2005).

OCLC, Inc. "Digitization & Preservation Online Resource Center: Copyright Online Copyright Resource Kit," <http://digitalarchive.oclc.org/da/ViewObject.jsp?fileid=0000016179:000000676940&reqid=1269> (Accessed 26 March 2006).

## Appendix B

# Glossary

**b** Bit

**B** Byte. 8 bits.

**Kb** Kilobit. 1,000 bits.

**KB** Kilobyte. Despite the “kilo” prefix, a kilobyte is not one thousand bytes. One kilobyte is 1,024 ( $2^{10}$ ) 8-bit bytes, or 8,192 bits.

**kbps** Kilobits per second. A kilobit is 1,000 bits.

**kHz** Kilohertz. 1,000 cycles per second.

**Mb** Megabit. 1,000,000 bits.

**MB** Megabyte. 1,048,576 (2<sup>20</sup>) bytes, or 1,024 kilobytes. It is also 8,388,608 (1,048,576 \* 8) bits, or 8,388.608 kilobits. Despite the “mega” prefix, a megabyte is not one million bytes.

**Mbps** Megabits per second.

**MHz** Megahertz. 1,000,000 cycles per second.

## Appendix C

# Bibliography

### General

- Austerberry, David. *The Technology of Video and Audio Streaming*. 2nd ed. Burlington, Mass.: Focal Press, 2005. ISBN 0240805801.
- Bailey, Andy. *Network Technology for Digital Audio*. Boston: Focal Press, 2001. ISBN 0240515889.
- Ballora, Mark. *Essentials of Music Technology*. Upper Saddle River, N.J.: Prentice Hall, 2003. ISBN 0130937479.
- Edstrom, Brent. *Musicianship in the Digital Age*. Boston: Thomson Course Technology, 2006. ISBN 1592009832.
- Farrington, Jim. *Audio and Video Equipment Basics for Libraries*. Music Library Association Basic Manual Series, no. 5. Lanham, MD: Scarecrow Press, 2006. ISBN 0810857162.
- Fries, Bruce, and Marty Fries. *Digital Audio Essentials*. Sebastopol, Calif.: O'Reilly, 2005. ISBN 0596008562.
- Middleton, Chris. *The Complete Guide to Digital Audio: A Comprehensive Introduction to Digital Sound and Music-Making*. Boston, MA: Muska & Lipman, 2003. ISBN 1592001025.
- Pohlmann, Ken C. *Principles of Digital Audio*. 5th ed. New York: McGraw-Hill, 2005. ISBN 0071441565.
- White, Glenn D., and Gary J. Louie. *The Audio Dictionary*. 3rd ed. Seattle: University of Washington Press, 2005. ISBN 0295984988.

## Digital audio formats

### General

“Introduction to Digital Formats for Library of Congress Collections.”

<http://www.digitalpreservation.gov/formats/intro/intro.shtml>.

Bosi, Marina, and Richard E. Goldberg. *Introduction to Digital Audio Coding and Standards*. Boston: Kluwer Academic, 2003. ISBN 1402073577.

Thorough, but highly technical, information on digital audio coding algorithms and the MPEG family of standards. Treatment of psychoacoustic modeling is particularly good. Provides no coverage, however, for popular non-MPEG standards like WMA, Real, and Ogg Vorbis.

### FLAC

Coalson, Josh. “FLAC – Free Lossless Audio Codec”

<http://flac.sourceforge.net> (Accessed 10 December 2005)

### Monkey’s Audio

Ashland, Matthew T. “Monkey’s Audio: A Fast and Powerful Lossless Audio Compressor.” <http://www.monkeysaudio.com> (Accessed 12 December 2005).

### MPEG

Fraunhofer IIS <http://www.iis.fraunhofer.de/amm/techinf> (Accessed 18 November 2005).

Moving Picture Experts Group. “The MPEG Home Page.”

<http://www.chiariglione.org/mpeg> (Accessed 18 November 2005).

### AAC

Fraunhofer IIS. “Audio & Multimedia: MPEG-2 AAC.”

<http://www.iis.fraunhofer.de/amm/techinf/aac> (Accessed 18 November 2005).

Apple, Inc. “QuickTime Technologies AAC Audio.”

<http://www.apple.com/quicktime/technologies/aac> (Accessed 9 December 2005).

## **MP3**

Bouvine, Gabriel. "MP3-Tech" <http://www.mp3-tech.org> (Accessed 10 December 2005).

———. "Patents and MP3" <http://www.mp3licensing.com> (Accessed 10 December 2005).

Fraunhofer IIS. "Audio & Multimedia: MPEG Audio Layer-3." <http://www.iis.fraunhofer.de/amm/techinf/layer3> (Accessed 18 November 2005). Thomson, Inc. "Mp3licensing.com Home" <http://www.mp3licensing.com> (Accessed 10 December 2005).

## **Ogg Vorbis**

Xiph.org. "Ogg Vorbis Documentation" <http://www.xiph.org/vorbis/doc> (Accessed 10 December 2005).

Xiph.org. "Vorbis Audio Compression" <http://xiph.org/vorbis> (Accessed 13 December 2005).

## **Quicktime**

Apple, Inc. "Quicktime." <http://developer.apple.com/quicktime> (Accessed 6 December 2005).

## **RealAudio**

RealNetworks, Inc. "RealNetworks Documentation Library" <http://service.real.com/help.library> (Accessed 10 December 2005).

## **Windows Media**

Microsoft, Inc. "Microsoft Windows Media." <http://www.microsoft.com/windows/windowsmedia> (Accessed 6 December 2005).

## **Servers**

### **Operating Systems**

#### **Apple**

Apple, Inc. "Mac OS X Server" <http://www.apple.com/server/macosx> (Accessed 23 March 2006).

## **Linux**

FreeBSD Project. "The FreeBSD Project." <http://www.freebsd.org> (Accessed 23 March 2006).

Red Hat, Inc. "Red Hat: The Open Source Leader." <http://www.redhat.com> (Accessed 23 March 2006).

Sun Microsystems, Inc. "Solaris Enterprise System." <http://www.sun.com/software/solaris> (Accessed 23 March 2006).

## **Windows**

Microsoft, Inc. "Microsoft Windows Server System: Home." <http://www.microsoft.com/windowsserversystem> (Accessed 19 March 2006).

## **Streaming Server Software**

### **Helix**

RealNetworks, Inc. "Products and Services Media Servers." [http://www.realnetworks.com/products/media/\\_delivery.html](http://www.realnetworks.com/products/media/_delivery.html) (Accessed 19 March 2006).

### **QuickTime**

Apple, Inc. "QuickTime – Streaming Server." <http://www.apple.com/quicktime/streamingserver> (Accessed 19 March 2006).

### **Windows**

Microsoft, Inc. "Windows Media Services 9 Series." <http://www.microsoft.com/windows/windowsmedia/9series/server.aspx> (Accessed 19 March 2006).

## **Preservation**

Council on Library and Information Resources. *Capturing Analog Sound for Digital Preservation: Report of a Roundtable Discussion of Best Practices for Transferring Analog Discs and Tapes*. CLIR Publication, no. 137. Washington, DC: Council on Library and Information Resources; Library of Congress, 2006. ISBN 1932326251.

<http://www.clir.org/pubs/reports/pub137/pub137.pdf> (Accessed 13 April 2006).

Stanford University Libraries, Preservation Department, Conservation Online. "Audio Preservation." <http://palimpsest.stanford.edu/bytopic/audio> (Accessed 2 April 2006).

Rosen, Jody. "How Pop Sounded before It Popped." *New York Times*, "Arts and Leisure" section, 19 March 2006. On the University of California, Santa Barbara, Cylinder Preservation and Digitization Project.

## Metadata

### ID3

"ID3v2." <http://www.id3.org> (Accessed 18 November 2005).

## Peer-to-Peer Networks

Katz, Mark. "Living in Cyberspace." Chapter 8 of *Capturing Sound*. (see under "Digital Audio and Culture" below).

Merriden, Trevor. *Irresistible Forces: The Business Legacy of Napster & the Growth of the Underground Internet*. Oxford: Capstone, 2001. ISBN 1841121703.

Oram, Andrew, ed. *Peer-to-Peer: Harnessing the Benefits of a Disruptive Technology*. Sebastopol, CA: O'Reilly, 2001. ISBN 059600110X.

United States Congress. House Committee on the Judiciary. Subcommittee on Courts, the Internet, and Intellectual Property. "Reducing Peer-to-Peer (P2P) Piracy on University Campuses: A Progress Update." One Hundred Ninth Congress, First Session, 22 September 2005. Serial no. 109-56 <http://purl.access.gpo.gov/GPO/LPS66466> (Accessed 17 April 2006).

## Subscription Services

Quist, Ned., Darwin F. Scott, and Alec McLane. "Naxos Music Library." *Notes* 61, no. 2 (December 2004): 512–16.

## Portable Digital Audio Players

### Digital Audio in Libraries

Honan, Mathew. "Libraries Turning to iPods and iTunes," *Playlist*, 13 February 2006 <http://playlistmag.com/features/2006/02/library> (Accessed 26 March 2006).

Lutz, Marilyn. "The Maine Music Box: A Pilot Project to Create a Digital Music Library." *Library Hi Tech* 22, no. 3 (2004): 283–94.

Maple, Amanda, and Tona Henderson. "Prelude to a Digital Music Library at the Pennsylvania State University: Networking Audio for Academic Library Users." *Library Resources and Technical Services* 44, no. 4 (October 2000): 190–95.

Stewart, M. Claire, and H. Frank Cervone. "Building a New Infrastructure for Digital Media: Northwestern University Library." *Information Technology and Libraries* 22, no. 2 (June 2003): 69–74.

Sullivan, Kathryn, John J. Stafford, and Cindy Badilla-Melendez. "Digital Music Project at Winona State University." *Information Technology and Libraries* 23, no. 2 (June 2004): 70–73.

Walker, Diane Parr. "Music in the Academic Library of Tomorrow." *Notes* 59, no. 4 (June 2003): 817–27.

### Digital Audio and Culture

Katz, Mark. *Capturing Sound: How Technology Has Changed Music*. Berkeley: University of California Press, 2004. ISBN 0520241967. See in particular chapter 8, "Listening in Cyberspace," which focuses on the MP3/P2P phenomenon and intellectual property issues.

Lysloff, Rene T. A., and Leslie C. Gay. *Music and Technoculture*. Middletown, CT: Wesleyan University Press, 2003. ISBN 081956513X.

### Digital Audio and the Recording Industry

Barfe, Louis. *Where Have All the Good Times Gone?: The Rise and Fall of the Recording Industry*. London: Atlantic, 2004. ISBN 1843540657.

Burkart, Patrick, and Tom McCourt. *Digital Music Wars: Ownership and Control of the Celestial Jukebox*. Critical Media Studies. Lanham: Rowman & Littlefield Publishers, 2006. ISBN 0742536688.

Coleman, Mark. *Playback: From the Victrola to MP3, 100 Years of Music, Machines, and Money*. New York: Da Capo, 2003. ISBN 0306809842.

Hull, Geoffrey P. *The Recording Industry*. 2nd ed. New York: Routledge, 2004. ISBN 041596802X. See in particular chapter 11, "The Recording Industry and the Internet."

Katz, Mark. *Capturing Sound: How Technology Has Changed Music*. Berkeley: University of California Press, 2004. ISBN 0520241967.

Kusek, David, and Gerd Leonhard. *The Future of Music: Manifesto for the Digital Music Revolution*. Edited by Susan Gedutis Lindsay. Boston: Berklee Press, 2005. ISBN 0876390599.

### **Digital Rights Management (DRM)**

Howe, J. "Licensed to Bill." *Wired* 9, no. 10 (October 2001): 140–49.

## COLOPHON

The typeface used for this document is Bitstream Charter, designed by Matthew Carter in 1987. The text was composed using  $\text{\LaTeX}$  on a computer running Fedora Core 5, a distribution of the Linux operating system.