

Out of the Darkness:
Preserving the Digital Output of Independent Punk and Metal Labels

by
Daniel Finn

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Arts
Moving Image Archiving and Preservation Program
Department of Cinema Studies
New York University
May 2014

ACKNOWLEDGMENTS

I would like to thank the following people, without whose help this project would not have been possible; Mona Jimenez, Peter Oleksik, Kara Van Malssen, Howard Besser, Dan Streible, Tim Harwich, Greg Anderson, Clay Marshall, Andy Low, Scotty Heath, Nicholas Sansano, Brad Boatright, Mike Josephson, Max Burke, Sean Patrick Rhorer, and Cory Kempton. My deepest gratitude goes out to all of you for your various and innumerable contributions. \m/

INTRODUCTION

Digital technology has fundamentally altered the way music is written, recorded, and distributed. The distribution side has undergone particularly transformative shifts in the last 15 years since the advent of the MP3 and the ease with which compressed audio files can be exchanged online through any number of licit and illicit pathways. Major labels and the RIAA have been very noisy about how these traits of digital audio technology are hurting the music industry. On the other hand, the same traits have opened up pathways for independent artists and labels to get music out to more potential listeners than ever before.

However, for all parties involved there are still growing pains as distribution avenues and workflows continue to change. Unless they are very new, many independent labels will have had experience with analog tapes being the master elements for their releases. But in the last decade the vast majority of independent artists have chosen to record digitally as the quality of affordable digital recording has been able to match professional recordings. Now labels and artists deal with file-based workflows until physical copies are pressed. In the case of digital-only releases the entire workflow is file-based.

These issues have not caused immediate problems for releasing material. In the area of long-term storage and preservation of digital content, though, independent practices have not caught up to the available technology. Digital preservation can require extensive up-front investment and dedicated staff to manage. Major labels have the capital and resources to ensure the long-term accessibility of their digital assets, but independent labels are often unable to afford the time or money to plan for the survival of their digital materials in the long-term. Some labels are not concerned with the archival perspective at all. Their focus and resources are directed only towards what they will release next.

I want this thesis to get independent artists and labels thinking about what it means to safeguard their work in a digital context. There is overarching concern across many disciplines that lack of awareness about the fragility of digital media may lead to a digital dark age in nearly every sphere of digital culture.¹ One of the prime examples of concern is email. Will the historians and scholars of the future be able to access our generations' emails the way letters have been researched in the past? How will that material get accessed? It's not as if you can expect to find a box with a bundle of emails stored away in a person's attic. There is no reason to consider independent music any less at risk. This is an incredible moment in the history of music. The Internet has allowed musicians and listeners to eschew the traditional gatekeepers to publishing and distributing music. Yet the technology that allows such ease of accessibility and transmission is not built for indefinite storage. If consideration is not taken as how to keep this material available, it is very possible that vast quantities of great music may vanish in a matter of decades.

As such, the thesis seeks to serve a number of purposes. The first is to raise independent producers' awareness of digital preservation risks and the latest preservation practices. Second, to provide a snapshot of independent digital audio preservation practices and their accompanying problems to the preservation community. Finally, it will provide labels and artists with strategies to cost-effectively reduce the risk of losing important and irreplaceable material.

The scope of the survey includes 5 labels. These are Forcefield Records, Prosthetic Records, Robotic Empire Records, Southern Lord Records, and Tankcrimes Records. Interviews

¹ Kurt D. Bollacker, "Avoiding a Digital Dark Age," *American Scientist* (March-April 2010), accessed May 3, 2014, <http://www.americanscientist.org/issues/pub/2010/2/avoiding-a-digital-dark-age>.

were also conducted with other professionals in music production² to further contextualize the production process through the lens of digital preservation.

It is important here to address that all the surveyed labels operate in the genres of punk and metal. Besides the obvious reason that punk and metal are totally awesome, some other factors influenced the choice. I focus on punk and metal in part because my personal background and contacts are in this area of music. I imagined this would make it easier for me to elicit substantive responses and new points of contact, as opposed to cold calling individuals or organizations with which I had no prior experience.

Furthermore, I do not believe that different genres of music create substantially different digital preservation issues. Audio file formats and digital storage media have the same physical characteristics regardless of the genre of music they contain. That is not to deny that some differences may exist. Vinyl is one of the central ways in which punk and metal albums continue to be distributed. As such, preservation options I address involving this stage of production may not apply to styles of music for which vinyl are not commonly created.

For the most part though I expect the production workflow to be similar to other styles of music. Especially in the context of independent distribution, digital technologies have allowed independent artists to handle the entire production chain themselves if they want. A site like Bandcamp, which is a hybrid social media and digital distribution website, features artists from every conceivable genre, and allows artists to digitally distribute downloads or do online mail order from their profile pages. Labels can link various artist pages to a label profile and generate income this way. Several of the labels interviewed have active Bandcamp profiles. Regardless of the genre many independent artists are making use of many of the same recording, mastering,

² Such as mastering engineers, a pressing plant employee, and digital distributors.

and distribution technologies. As such I believe this research will be beneficial to any independent labels or artists interested in preservation, no matter their genre.

Another notable feature of the labels I have chosen is their age. Most of the labels were founded in 1998 or 1999. I chose labels that started around this time or later since from this moment onwards the majority of the releases they put out were digitally recorded and mastered. For the sake of simplicity I did not pick older labels whose catalogs bridge the gap between analog and digital masters, as this would bring complications into the discussion that are not germane to the main focus of preserving specifically digital material.

Although independent music rarely attains levels of distribution and exposure that major label releases receive, independent releases afford artists greater control of their work. Such heartfelt work can often spur new musical trends and modes of sonic expression. It is my sincere hope that the following pages will help content producers protect their work and allow for greater access to vital musical expressions.

BACKGROUND ON LABELS

The following provides a brief description of the history and background of each of the labels I interviewed, as well as the details of our interviews pertaining to their preservation activities.

Forcefield Records

Tim Harwich and two of his friends started Forcefield Records in 2007 in Richmond, VA. As of 2009 Harwich became the sole operator of Forcefield. The label began with a focus on local punk and metal, but has since focused primarily on metal and put out releases from

bands from all over the country, as well as some international acts. They have put out incredible debut records from bands like Windhand, Inter Arma, Cough, and Bastard Sapling. At the time of writing Forcefield have 43 releases.

I interviewed Harwich in January 2014 at the Forcefield HQ.³ Harwich stores all masters he receives on an external hard drive. Unfortunately, a few of the releases are absent, and there is no backup or secondary storage. He said at any given time there may be overlap between the external hard drive and the internal hard drive of his PC, but the external hard drive is considered to be the primary storage location.

At this point the absence of some of the masters has not created a catastrophic problem. One example he provided was a split release between Cough and Wounded Kings. When a repress of the vinyl was going to be made, he realized his hard drive did not contain the master he needed. Fortunately, the mastering engineer had retained a copy, and Harwich was able to get the project done with that copy.

One cause for some of the absences results from the fact that sometimes he does not touch the digital master during production. This happens when the mastering engineer knows the CD or vinyl pressing plant handling the release, and sends it directly to them in the interest of time.

The other way in which Harwich tries to preserve material is by keeping the master lacquers for albums that are pressed to vinyl.⁴ However, this option is often not feasible for him due to the added expense.

³ AKA his home.

⁴ Lacquers are the first element in the generation of vinyl records. Lacquers begin as plain discs with no grooves. The master audio from the mastering engineer is used to etch the grooves of the record into the lacquer. One lacquer is used for each side of the album. These lacquers are then used to create plates, or metal mothers, which are metal negatives of the grooves. These plates are used to press each individual vinyl, which then gets sleeved, shipped, and ultimately spun on listeners' turntables.

That is the extent of Harwich's preservation activity; having one copy for a majority of the releases on one hard drive, and maintaining lacquers for a minority of the releases.

Prosthetic Records

Founded in 1998 in Los Angeles, Prosthetic started when E.J. Johantgen and Dan Fitzgerald left Columbia Records to start their own label specializing in heavy metal. They have since put out albums by killer bands like Skeletonwitch, Mutilation Rites, Dragged into Sunlight, Kylesa, and Trap Them. As of this writing Prosthetic have released 157 albums.⁵

I spoke with Clay Marshall, current General Manager of Prosthetic, via several e-mails in February 2014. His e-mails do a great job of capturing what I have found to be the prevailing attitudes towards preservation for the labels I have interacted with. As a result I am quoting heavily from his correspondence. Here is what Marshall said about the digital storage of Prosthetic's digital material:

Nowadays we keep everything we're sent... We store everything on our server on a shared drive. It's kind of like virtual nesting cubes: Each band has a folder, and within that folder, there are various subfolders that ideally house everything we'd ever need -- artwork, music, scanned press clips, etc. We don't really have time to work backwards, though, so I wouldn't be surprised if some of our older band folders (that is, bands who are inactive or have broken up) aren't well stocked, so to speak. For example, we used to keep press folders with physical copies of reviews, so I'd be shocked if all of that has been scanned.

⁵ "Releases," Prosthetic Records, accessed April 24, 2014, <http://prostheticrecords.com/?cat=4>.

Marshall on the subject of keeping an inventory:

No such inventory exists. At the risk of making us seem unorganized, we know where to look for something, so creating a list that tells us whether we have something or not seems kind of superfluous. (Rather than checking the list, we might as well check the band folder.)

When asked whether Prosthetic has backup storage for the digital material, Marshall responded:

Not exactly. If we were the "master keeper" of these files, we'd do a far better job making backups, but most of the time, we'd be able to ask to receive another copy from the mastering plant or the band, or just re-create a digital master by ripping a CD. Further, most of our manufacturers keep copies of our masters, so even if we were hit by a heavy metal terrorist attack, I imagine we'd be able to rebuild easily enough. That said, we keep an archive of physical copies of everything, so I guess you could consider that a de facto digital archive as well, in that those physical copies could replenish our digital files if need be.

The idea of restoring digital masters from physical copies of the release is potentially feasible. Especially if the album was recorded at CD quality and the CD used for the restoration is in good shape. Digitizing with vinyl may result in hearing some artifacts that would not have been present in the digital master, such as the iconic clicks and pops from dust. A better option would be digitizing from the master lacquer of the vinyl release. Another problem could arise

from using a CD to restore the master when the original master was recorded and mastered at a resolution higher than CD quality audio. While there are some issues with the strategy, keeping the physical archive is an effective method to ensure that the material can be accessed.

The organization of Prosthetic's digital storage is great. As a result, using the directory tree as a de facto inventory is an effective practice. The important characteristic is that there is a standard of storing, describing, and locating material. However, there is still risk associated with having a single copy of all that information.

Robotic Empire

Andy Low started Robotic Empire in Richmond, VA in 1999. According to the Robotic Empire webpage, "a wide amount of genres have been covered over the years as our love for many things left-of-center have kept the variety levels high."⁶ This breadth includes releases from bands like Pageninety-nine, The Red Chord, Daughters, Torche, Grails, and Thou. At the time of writing Robotic Empire have 116 releases.

I spoke with Andy Low over several e-mails in February 2014. He has master materials on DAT, CD-R, and an external hard drive. The external hard drive stores not only digital audio masters but also artwork files and other files related to "Robo"⁷ business. There are also data CD-Rs containing artwork files. No inventory exists for this material, which Low admitted was fairly disorganized and "an archivist's nightmare."⁸

Similarly to Forcefield, Low mentioned that some masters are absent, also the result of instances when the master traveled directly from the mastering facility to the pressing plant. This has only happened a few times. Low is confident he has masters for 95% of his releases.

⁶ Robotic Empire, "About Us," accessed April 24, 2014, <https://www.roboticempire.com/about-us/>.

⁷ Robo is the label's vernacular nickname Low used in correspondence.

⁸ Andy Low, email to author, February 4, 2014.

Low also stated that Robotic Empire has never had an issue with accessing something he needed in his “loose” organizational system, but admitted problems could easily arise if one or more items needed to be accessed urgently.

Southern Lord Records

Greg Anderson and Stephen O’Malley founded Southern Lord in 1998 as a way to get their musical projects released under their own terms. Early releases on the label included albums from Burning Witch, Khanate, Sunn O))), Electric Wizard, The Obsessed and Boris. With these Southern Lord quickly gained quite a reputation for releasing high quality albums from bands playing doom, black metal, sludge, ambience, and other forms of experimental metal. At this point Southern Lord has issued 255 releases.

I met with Greg Anderson on August 6, 2013 at the Southern Lord HQ in Los Angeles.⁹ The conversation regarding preservation was brief. All of their digital material is stored on two hard drives. There is only the primary copy for all the material. There is also no inventory for what is on the drives. When I asked him what types of drives he used he replied, “Whatever’s cheap.” I asked, “Aren’t you worried you might lose something and not be able to get it back?” To this he replied that between Southern Lord, the band, the pressing plant, and the mastering house, someone ought to have a working copy.

At one point I asked whether he ever thought of the record label as an archive, or as a body of work representing his own hard work along with the hard work of friends making music they all have passion for. His answer was simply no; he doesn’t think about it like that. Once

⁹ HQ here essentially means three rooms in a small office off Vine Street minutes away from the Walk of Fame and Hollywood, an interesting location for a metal label.

he's released something, in a lot of ways he's done with it. He would rather focus on new material.

Instances in which problems accessing masters have arisen were again very rare. The example Anderson cited for me resulted in them using a slightly lower quality version of a master for a repress than might have been available if everything had been preserved. The loss of quality didn't seem to ruin the project or leave a bad taste in the mouths of anyone involved.

In this case, the lack of preservation activity is not due primarily to lack of time or resources, but because the label does not see preservation as part of its objectives.

Tankcrimes Records

Scotty Heath founded Tankcrimes in Oakland, CA. The first release for the imprint came out in 2002. It was then and continues to be a one-man operation, all out of Heath's garage.¹⁰ At the time of writing Tankcrimes has put out 74 releases from bands like Ghoul, Cannabis Corpse, Agoraphobic Nosebleed, Direct Control, Toxic Holocaust, and Inepsey.

I spoke with Scotty Heath via e-mail in February 2014. Heath preserves digital masters and other digital material via an external hard drive, and considers the mastering houses to be the backup location. Earlier masters are on CD-R, and there are a couple of DAT tapes from his earliest releases.

Similar to the other interviewees, there have not been any major issues with the corruption or inaccessibility of a master or other digital object, despite the lack of backups or an inventory administered by the label itself.

¹⁰ "Tankcrimes Founder Runs Mini Punk Empire Out of His Garage," Noisecreep, March 3, 2011, accessed April 13, 2014, <http://noisecreep.com/tankcrimes-founder-runs-mini-punk-empire-out-of-his-garage/>.

There are a number of commonalities in all these cases. The labels themselves do not expend a lot of effort on preserving their digital materials. In all cases the focus is primarily on organizing upcoming projects and releases. As a result a level of disorganization is present for the materials they maintain of past releases. In each case here there is a bit of what Andy Low called the “archivist’s nightmare.” And perhaps even more conducive to night sweats for preservationists is that Low, admitting his own shortcomings in personal archiving for the label, noted, “It’s rough, but I’ve seen worse.”

The following section seeks to elaborate on what causes archivists’ uneasiness in scenarios like the ones described above. By framing the discussion of independent audio production within a larger conversation about digital preservation, I hope to highlight the ways current practices fall short of what they could be.

INDEPENDENT MUSIC AND DIGITAL PRESERVATION

Audio recording is now a predominantly digital process. According to Mike Josephson of Pirate’s Press¹¹ as of 2012 “99% of the masters [Pirate’s Press] receive[s] are in digital form.”¹² He adds that “it is rare to get a recording that is exclusively analog as most people, even if they record to a 2" tape through an analog board for example, will quite often dump those tracks into a computer for the mixing/mastering process.”¹³ In *Perfecting Sound Forever*, Greg Milner writes that as of 2000 recording in ProTools was becoming standard practice and that “by 2007, between 70 and 80 percent of all pop music (and probably nearly 100 percent of all hip-hop,

¹¹ Pirate’s Press is a vinyl record pressing plant that has handled many of the releases from Forcefield, Robotic Empire, and Southern Lord.

¹² Mike Josephson, email to author, December 10, 2012.

¹³ Ibid.

R&B, and dance music) was mixed in the box.”¹⁴ Recording in the box here refers to using a digital audio workstation (DAW) such as ProTools. Reliance on the millions of dollars of equipment in traditional recording studios became unnecessary. Big record labels saw this technology as a way to cut costs, but it also served to lower the ceiling of access for many independent artists who could now more readily seek out freelance engineers to get professional quality recordings out of ProTools or other DAW suites.

Even if the final destination for a release is a vinyl record or cassette tape, in the vast majority of cases the master elements will be digital files. Digital storage technology like optical media, hard drives, and data tape do not have the same kind of “shelf life” that their analog forebears do. As a result it is important to understand their life cycles and the associated risks to the long-term survival of digital audio masters.

DIGITAL PRESERVATION BASICS

In order to move forward a basic understanding of what digital preservation entails is necessary.¹⁵ The preservation of any object must be considered an ongoing process. One thing that distinguishes digital preservation is how evident this principle becomes. Due to the particularities of digital media, the lifecycle of a digital object is remarkably short when compared to past methods of storing information. In general information storage systems have historically become more complex as humans generate and desire to retain increasing amounts of information.

¹⁴ Greg Milner, *Perfecting Sound Forever* (New York: Faber and Faber, Inc., 2009), 338-339.

¹⁵ Much of the basis for the following information is indebted to a lecture from Kara Van Malssen. Kara Van Malssen, “Preservation of Digital Collections,” Lecture from New York University, New York, NY, September 3, 2013.

The first information storage method humans used is language. Information (ideas/thoughts) is encoded into words that can be transmitted (spoken), decoded (heard and understood), and saved for later use (remembered). Writing followed, and is an example of what can be termed static media. Static media can be read by the naked eye. Once created these media do not change, hence the term static. At this stage there are very few external dependencies necessary to understand the information. By external dependencies I mean requirements not included in the information itself necessary to appreciate and understand that information. The only outside dependencies in the example of writing would be knowledge of the language in question and literacy in its written form. Books, magazines, and newspapers are common examples of static media. Static media have a very long useful life if they are stored in environments where the climate is agreeable.

Magnetic media such as VHS and audiotape present new complexities. Information stored magnetically is not human readable. Content is stored as voltages that have a larger number of external dependencies in order for the content to be seen and used. These include the hardware required for playback, such as monitors and playback decks. The media items themselves are fragile. Industry changes led to many formats of varying popularity. Once a format became unpopular, production of the hardware necessary to view those tapes stopped, and soon that hardware became rare. Even tapes in good condition are unusable when there is no working playback deck available.

Digital media compounds these issues. There are an even greater number of external dependencies, and the information itself is effectively intangible. The cycle of obsolescence for the dependent components is more rapid, and information indicating exactly what these dependencies are is not always present.

Whereas for older types of physical media preservation efforts can often be boiled down to maintaining proper storage conditions, digital media require a much more active approach. As an example, take two versions of the same movie. One copy is on film and one is a digital file. If the film copy is printed on preservation grade film stock, under the correct storage conditions that object can be expected to sit on a shelf without degradation for as long as a century or more. To contrast, after putting the digital file onto a new hard drive, preservationists will be nervous about leaving that drive on a shelf for more than five years without additional intervention. The cause of this anxiety is the short lifespan of digital storage technology. Another cause is the fact that a hard drive has no way to alert anyone to its deterioration the way film, books, or tapes do.

Hard drives don't yellow, become brittle, emit odd odors, or crack apart. They probably just look a little dusty. You can't know if the media survives until you try and use it. There is no way to discover a problem without performing systematic monitoring. For these reasons digital preservation puts so sharply into focus the principle that preservation is a constantly ongoing process, and how active digital preservation needs to be compared to former preservation methods. In the same hundred years that one piece of film could have survived without hassle, the digital version will need to have undergone dozens of refreshes, migrations, and various other interventions.

The following section aims to generally outline the basic risks facing digital media.

GENERAL DIGITAL RISK FACTORS

Bit Rot

Bit rot can describe a number of different phenomena in which the bits, the 1s and 0s, stored on particular storage media can decay, as though radioactive. Most commonly the term refers to the probability of any bit randomly “flipping;” a 1 becomes a 0 or vice versa.¹⁶

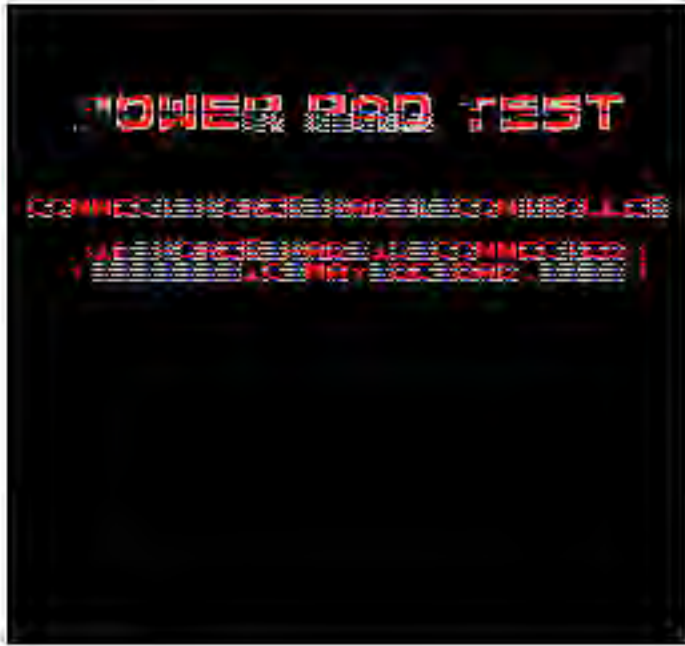
Depending on the location within a file, one flipped bit can cause minor or catastrophic damage.

If you have ever seen a bizarre color block on a video file, or heard strange chirps in an audio file that weren’t there previously, bit rot is a likely culprit. Below are two images, one from a game cartridge that is working well, and another from the same game cartridge suffering from bit rot¹⁷:



¹⁶ “Bitrot and Atomic COWs: Inside ‘Next-Gen’ Filesystems,” ArsTechnica, accessed March 4, 2014, <http://arstechnica.com/information-technology/2014/01/bitrot-and-atomic-cows-inside-next-gen-filesystems/>.

¹⁷ “A Small Lesson in Bit Rot” accessed May 3, 2014, <http://my.ais.net/~xtreme/SF/Bit-Rot/>.



Software Obsolescence and Upgrades

Many audiovisual and other software suites feature built-in obsolescence schedules to increase sales of future releases. Files from earlier versions will become unsupported after a certain version is released. If you have updated your software but haven't also updated a file created with an earlier version, you may no longer be able to access the file.

One example that arose during my interviews comes from using ProTools. Nicholas Sansano, a recording, mixing, and mastering engineer as well as professor at the Clive Davis Institute for Recorded Sound at the Tisch School of the Arts at New York University, described the issue to me.¹⁸ His career in recording began prior to digital recording, and so he has experience with most iterations of the ProTools software. He has stored all of his ProTools session files from every project that used the software. These go back to the first versions of the

¹⁸ Nicholas Sansano, interviewed by author, digital recording, New York City, October 12, 2012.

software from the mid-1990s. His earliest ProTools recordings cannot be accessed on the current version of the software. In order to get those files working, one would have to find a computer with the appropriate version of ProTools. One solution is trying to break out a dusty computer from that time period with the appropriate version of ProTools and hoping everything on that still works. Emulation is the other solution for recreating legacy computer environments.

Emulators “run on modern hardware, but mimic an obsolete software environment, allowing old software to run.”¹⁹ In this example, one would find an emulator for an older Windows or Mac operating system, acquire the appropriate version of ProTools, install that version of ProTools within the emulator, and then open the old session file. The only other option is to continuously update files as new versions of the software emerge. Every time a new version of ProTools is released, this would require opening the files from the previous generation, saving them as session files in the new generation, and repeating this process for all files with successive iterations of the software.

Hardware Obsolescence

Various components can become obsolete. As connection types (SASI for example) go out of fashion, or if a peripheral fails commercially and drivers are no longer made for newer computing environments, the piece of hardware will no longer be usable. One reason for rapid obsolescence is that “a one or two order of magnitude improvement in power, speed, efficiency, or cost per value has occurred every several years in areas such as CPU speed, memory chip density, storage device capacity, video processing rate, and data transmission rate.”²⁰ Newer

¹⁹ “Obsolescence: Hardware and Media,” Digital Preservation Management: Implementing Short-Term Strategies for Long-Term Problems, accessed March 4, 2014, <http://www.dpworkshop.org/dpm-eng/oldmedia/obsolescence2.html>.

²⁰ Ibid.

software that makes use of increased power will not be able to run on older hardware. The computing power required to make digital audio production viable simply didn't exist prior to the 1990s and 2000s. By now many of the earliest hardware devices used in early digital production such as interfaces, MIDI controllers, and any number of other devices likely don't have the connection types necessary to interface with modern computers. Even if they could connect to a computer, modern recording software suites likely do not support them.

Storage Media Obsolescence

Many types of removable digital storage media have come and gone; punch-cards, a host of floppy disk sizes, ZIP and JAZ disks, various hard drive configurations, to name just a few. If a drive is unavailable, or if the drive cannot be connected to a computer, then the information on the storage device might as well not exist.

The simultaneous decrease in size of storage media and increase in storage capacity is a trend that has been continuing for some time. The journey from 8" floppy disks to USB drives that fit in your pocket has been relatively short.²¹ As demand for portable, high quality audio and video files continues to rise there is no sign of this trend slowing down.

Storage Media Failure

In addition to becoming obsolete due to a lack of the components storage media depend on, the media themselves can and will simply fail. With hard drives for instance, predicting failure is "not always possible, and sometimes a hard drive will just die."²² Some predictive signs of imminent hard drive failure do exist, including strange noises, isolated instances of data loss

²¹ Ibid.

²² Joel Falconer, "How to Tell When Your Hard Drive Is Going to Fail," Lifehack, accessed May 6, 2014, <http://www.lifehack.org/articles/technology/how-to-tell-when-your-hard-drive-is-going-to-fail.html>.

or disk error, computer crashes, and very slow access times.²³ Depending on the storage medium causes of failure and lifespan vary, but inevitably all fail. Ask not for whom the bell tolls.

Lack of Metadata

Metadata literally means data about data. As an example, metadata for an album includes information such as album title, artist name, track listing, and so on. In files metadata refers to a litany of possible things. One common example of lacking metadata is lacking information about what a particular file is called or where it is stored. The absence of this information is another way that a file that technically exists essentially disappears because no one can retrieve it.

Maintaining metadata about their digital material is not something many independent artists or labels concern themselves with. Their digital distributors on the other hand need to do this. Metadata is imperative for digital distribution because it is the foundation of functions like searching and distribution.²⁴

Bandcamp, the social media digital distribution hybrid mentioned above, provides automatic metadata tagging for artists.²⁵ Downloading files uploaded by a band either from their website or other services has often led to confusing situations. If a band uploads a file named “Master 2(final).wav,” and a user downloaded it, the user now has a file with that name on her computer with no way for her to know what the song title was supposed to be, which album it belongs to, or what artist made the song. To solve this problem, Bandcamp siphons data from the band’s profile page. In order to allow streaming, the band has to enter content like the name of an album, song titles, and artwork in addition to the audio itself. Whenever an artist enters

²³ Ibid.

²⁴ Troy Allen, “The Importance of Metadata in Content Management,” CMSWire, January 10, 2011, accessed April 24, 2014, <http://www.cmswire.com/cms/enterprise-cms/the-importance-of-metadata-in-content-management-009746.php>.

²⁵ “Audio Basics,” Bandcamp, accessed April 24, 2014, http://bandcamp.com/help/audio_basics.

information such as song titles, album names, song credits, or anything else, Bandcamp automatically stores this data. When it generates the files for users to download, Bandcamp automatically includes this data with the derivative song files. The derivative file names include the artist and song title, and the derivative file formats allow embedding of other metadata Bandcamp harvests from the artist's profile page.

Lack of Organizational Support and Resources

Knowing about all these potential risks is not enough to abate them without commitment to action. However, action requires resources, time, money, and people. Long-term preservation can easily become a secondary or nonexistent concern in the face of tight schedules, constrained budgets, and limited staff. Though this is not a risk factor that is unique to digital objects, it is arguably one of the largest causes of information loss, and especially pertinent in the context of this project.

SPECIFIC DIGITAL STORAGE FORMAT RISKS

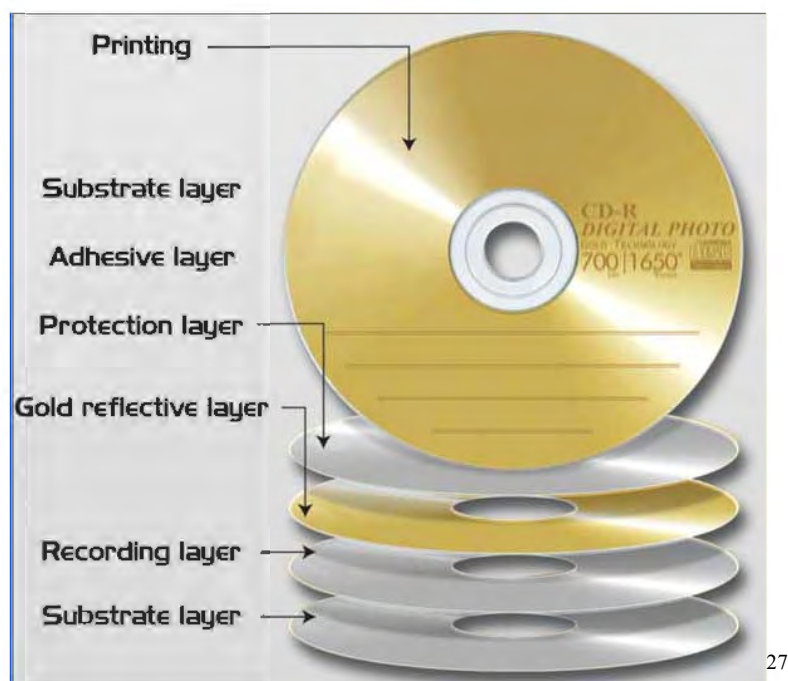
Building off of the previous section, the following is a selection of risk factors to specific digital audio storage formats. The list is not intended to be comprehensive, but addresses the most common formats that the surveyed labels discussed using for digital storage.

*CD-R*²⁶

²⁶ Fred Byers, *Care and Handling of CDs and DVDs: A Guide for Librarians and Archivists*, Copublished by Council on Library and Information Resources and National Institute of Standards and Technology, October 2003, accessed March 5, 2014, <http://www.clir.org/pubs/reports/pub121/reports/pub121/contents.html>.

Many factors affect the life expectancy of a CD-R, including manufacturer, environmental conditions, handling, condition of the disc at the time of recording, and quality of the recording process. Even certain batches from the same manufacturer of the same product can have varying useful lifetimes. Due to the variance in makeup of optical media and the lack of a lot of independent testing, estimating the longevity of CD-R can be difficult.

Life expectancy for a disc means the length of time data can be accessed successfully. However because many standards are used in the creation of discs and drives, a disc that is unusable in one drive could be usable in another. Therefore it is important to try a disc that won't read in a number of drives.



27

²⁷ "Everything You Wanted to Know About CDs," The Electronic Web Matrix of Digital Technology, accessed April 24, 2014, http://www.co-bw.com/Audio_everything_you_wanted_to_know_about_CDs.htm.

The main components of the CD-R are the plastic base, or substrate, a reflective layer of gold, silver, or silver alloy, and an organic dye layer on which the data is written. The data and reflective layers are much more susceptible to degradation than the plastic base. Generally only severe mishandling injures the base. It can also bend if stored in a non-vertical position over long periods of time. Metal and dye layers are susceptible to environmental factors like high heat, high humidity, and exposure to sunlight. As stated given the lack of testing and variety of disc compositions, accounts for life expectancy vary from 20 - 100 years.

The resolution of CD audio is 16-bit, 44.1kHz. Many reports on the best practices of importing audio recommend the target resolution for complex audio such as music should be 24-bit, 96 kHz. This is unnecessary in the case of CD audio, as the tracks are natively 16-bit, 44.1 kHz, and the master files that produced the CD audio would have to have been at this resolution as well.

*DAT*²⁸

Digital Audio Tape is comprised of an uncompressed digital signal stored on magnetic tape. As such DAT is subject to the same physical degradations as magnetic tape. High temperatures cause the tension of the tape to rise and tighten the pack of the tape, which can deform the tape. This can also lead to print-through, where magnetic information is passed between nearby layers of tape causing unwanted pre- and post-echo. Exposure to humidity causes a chemical reaction that leads to the composition of the tape breaking down and shedding of magnetic material, which causes loss of information as well as problems for the audio heads

²⁸ Susan Eldridge, "Digital Audio Tapes: Their Preservation and Conversion," Smithsonian Institution Archives, Summer 2010, accessed March 5, 2014, http://siarchives.si.edu/sites/default/files/pdfs/digitalAudioTapesPreservation2010_0.pdf.

reading the tape. Lubricants used in the tape as well as surface dust and dirt can also clog playback heads.

As production of DAT decks ceased in 2005, the biggest risk to DAT is obsolescence. Information on DAT should be migrated to a more secure format as soon as possible.

DAT can store audio in a number of different resolutions. These include 12-bit, 32kHz, as well as 16-bit at either 32, 44.1, or 48kHz. After inserting a tape into a DAT deck, the deck should display the bit-depth and sampling rate used to record the tape.

*Flash and USB Storage*²⁹

A number of limitations make using this type of storage unwise for long-term stability. Due to their small size and portability, they are readily susceptible to physical damage. Spilt beverages, friction in bags and pockets, exposure to dust and other particulate matter, or any number of things can physically damage the media. There is a limited amount of rewrites before data integrity suffers. Disconnecting the media without first safely ejecting it from the computer can cause a corruption of the file system and loss of files that are being written or edited.

*Hard Disk Drives (HDD)*³⁰

Hard drives fail due to three types of reasons, mechanical, electronic, and logical. Mechanical failure results from the breakdown of the complicated moving parts. The head that reads information from the disk hovers mere nanometers above the disk itself, and when the head

²⁹ Tessa Miller, "How Can I Reduce the Risk of Loss and Corruption on a USB Drive?" Liferhacker, July 26, 2013, accessed March 5, 2014, <http://liferhacker.com/how-can-i-reduce-the-risk-of-data-corruption-and-loss-o-923595543>.

³⁰ "Why Does the Hard Disk Drive Fail?" MicroCom Worldwide Data Recovery, accessed March 5, 2014, <http://www.data-master.com/why-harddrives-fail-Q11.html>.

touches the surface this causes a head crash. The moving head breaks and scratches the disk. The only way to recover data in this event is by going to a specialized vendor.

Electronic failure affects the drive's circuit board, which in turn affects a variety of processes the drive needs to operate correctly; orchestrating the movements of the internal parts, read and write functions, and countless other technical details. The primary cause is overheating, so it is best to avoid storing a drive near heaters or in areas frequently exposed to sunlight.

Electronic failure can occur as easily in new or old drives.

Logical failure essentially describes a corruption of the file system. Improperly formatting the drive or deleting key system files can cause this type of failure. Aging can also cause logical failure.

The following section details preservation strategies used to mitigate the above risks.

ESSENTIAL REQUIREMENTS FOR DIGITAL PRESERVATION

Successfully preserving digital information can be distilled to 3 requirements.³¹

1. Bit Preservation: The stream of 1s and 0s retains its integrity. Every 1 remains a 1, and every 0 remains a 0.

In its most essential form, every digital file is composed of a string of 1s and 0s. How these numbers are structured depends on the file format. Some of the numbers refer to the content itself, while others are used as directions to the software decoding the information. In order to access the file as intended, every number in the chain has to remain the same. This requirement means taking care to avoid bit rot.

³¹ Kara Van Malssen lecture.

2. Accessibility and Usability: The data object can be found, retrieved, interpreted, and used.

In addition to the file itself, preservation of digital media requires preserving data about the file. What use is saving an object if no one can find and access it? In order to allow researchers to access a digital object, they have to be able to find out that it exists, retrieve a copy, use contextual information surrounding the material to interpret what it is and what it is supposed to do, and last but not least, be able to open it.

3. Sustainability: An entity has to exist that claims responsibility over points 1 and 2.

Maintaining files and information surrounding them is a lot of work, and so the most important requirement is the agent(s) behind that activity. The responsible entity should be capable of the tasks at hand, and be able to document how it handles these tasks according to established practices and policies.

There are many ways to achieve these three goals. The data being preserved and the people or organization preserving it informs the paths taken. The following section will detail digital preservation best practices generally. A case study will follow illustrating how major labels deal with long-term storage of their digital materials more specifically.

STRATEGIES AND BEST PRACTICES

As the development of digital technology is rapidly changing, digital preservation standards and practices have to evolve as well. This section will provide some insight into the basic principles behind digital preservation best practices. Given the breadth and proliferation of

information on digital preservation, it is not the intent of this section to break significant ground, but rather to try and make the general ideas more accessible to people who do not have a background in preservation and are not familiar with its terminologies.

Before getting started, here is a brief introduction to the publications to which I am most indebted for the subsequent information. The International Association of Sound and Audiovisual Archives (IASA) released their *Guidelines on the Production and Preservation of Digital Audio Objects*³² in 2009, and it remains an excellent resource. *Sound Directions: Best Practices for Audio Preservation*³³ from 2007 is a thorough guide to many aspects of audio preservation. For a more general overview of digital preservation concepts and strategies, an online tutorial called “Digital Preservation Management: Implementing Short-Term Strategies for Long-Term Problems” is very useful.³⁴

In this context the one shortcoming with the above resources is that all are aimed at traditional collecting institutions, such as libraries, archives, and museums. As such a lot of their content and suggestions are not applicable or not possible in a setting like an independent record label. The aims of the label, unlike the missions of libraries, archives, and museums, do not necessarily have anything to do with ensuring long-term availability of substantial amounts of materials, and so certain assumptions inherent in the language do not hold true, and may only add to confuse parties interested in preserving smaller amounts of material. Yet the essential concepts therein are very useful and need to be considered.

³² IASA Technical Committee, *Guidelines on the Production and Preservation of Digital Audio Objects*, ed. by Kevin Bradley. Second edition 2009. (= Standards, Recommended Practices and Strategies, IASA-TC 04). www.iasa-web.org/tc04/audio-preservation.

³³ Mike Casey and Bruce Gordon, *Sound Directions: Best Practices for Audio Preservation*, Indiana University and Harvard University, 2007, accessed March 4, 2014, http://www.dlib.indiana.edu/projects/sounddirections/papersPresent/sd_bp_07.pdf.

³⁴ “English Tutorial,” Digital Preservation Management: Implementing Short-Term Strategies for Long-Term Problems, accessed April 24, 2014, <http://www.dpworkshop.org/dpm-eng/oldmedia/index.html>.

A number of different strategies exist to preserve digital objects, but no single one of them is a complete solution. Additionally, each one needs to be repeated over time to ensure continual access.

Bitstream Copying

A bitstream copy denotes making a bit for bit copy of a digital object. It can also be called creating a backup copy. This is an integral part of all digital repositories and all digital preservation endeavors. However, alone it merely tackles short-term issues such as corruption of the individual file, hardware failure, or storage media failure. In addition to having a second copy, the media on which that copy is stored (CD-R, HDD, etc.) should be kept in a separate location from the primary copy. This is in order to avoid both copies being lost in the event of a flood, fire, or other disaster. Bitstream copying, or backing up your data, is the ‘minimum maintenance strategy for even the most lightly valued, ephemeral data.’³⁵ This is considered because a bitstream by itself is not necessarily capable of communicating anything. The bitstream at its foundation is simply the arrangement of ones and zeroes. To continue to be accessible, the bitstream must be usable. An important standard in digital preservation known as the Trusted Digital Repository standard separated digital preservation activity into two categories; those that preserve a bitstream and “those that provide continued accessibility of its contents.”³⁶ Saving a string of numbers is one thing, but it is “quite another to preserve the content, form, style, appearance, and functionality” of that string.³⁷

³⁵ Digital Preservation Management: Implementing Short-Term Strategies for Long-Term Problems, “English Tutorial, accessed March 4, 2014, http://www.dpworkshop.org/dpm-eng/eng_index.html.

³⁶ Ibid.

³⁷ Ibid.

Refreshing

Refreshing can be described as backing up the backup. Specifically, it entails copying the digital object onto the same type of medium. Copying a master from a an older CD-R onto a brand new CD-R is an example of refreshing.

Migration

Migration is similar to refreshing, except it usually denotes a change in either the storage medium or the file format. The change will usually be due to the approaching obsolescence of the current storage medium or format. In this process “preserving the essential characteristics” of the file is paramount.³⁸ Migration is about keeping a digital object usable as new generations of computing technology appear.

Replication

Depending on the source, replication can refer to bitstream copying or migration. Replication can also have a less formal connotation. For example, peer-to-peer file sharing networks are examples of replication. It is worth noting that despite controversy, this method is already in place for all kinds of music. However, the most commonly shared versions of music recordings will be in MP3 or other lossy, compressed formats. As such we cannot rely on this version of copying to ensure the survival of master quality versions of recordings. Perhaps this may change if audiophile formats like FLAC gain popularity.

³⁸ Ibid.

Fixity

Fixity in digital preservation refers to the certainty that a copy is a completely accurate copy. “Since digital objects are easily modified, a mechanism is necessary to maintain fixity over time, or to consciously document when a digital object has been altered.”³⁹ One of the most popular mechanisms of ensuring the fixity of a given copy is a checksum. Checksums are alphanumeric strings generated by an algorithm that scans the entire bitstream of a file. By creating a checksum of a file before making a bitstream copy, the archivist can verify the copy’s checksum against the original checksum to ensure they are true copies. The algorithms are designed to generate significantly different strings even if a single bit is different. Any time a preservation entity generates a bitstream copy, refreshes a file, or migrates a file to a new storage medium, checksums are used to ensure fixity. In the event that migration includes a change of file format, the resulting checksums would not match given these changes. In such cases other quality control strategies need to be adopted to ensure that the significant characteristics of the original file are carried along into the migrated copy.

Analog Backup

Analog backup refers to keeping an analog copy of a digital object. Printing out emails is an example. This strategy is especially pertinent for the labels addressed, and Clay Marshall explicitly mentioned it in the case of Prosthetic. The strategy is implicitly in place for the other labels, as all of them have copies on hand for most of their releases available for mail order. Since many releases in these genres feature a vinyl component, it is advisable whenever possible to try and save a copy of the lacquers from the vinyl creation process. As vinyl continues to be

³⁹ Ibid.

popular and vital to many music genres, it is likely that such an element in good condition will be accessible for decades to come.

Digital Archaeology

A last resort, digital archaeology refers to taking damaged media or storage devices to a professional to see what can be salvaged. Vendors specializing in data recovery from disasters or mechanical/electronic failure of drives can be expensive, and almost certainly cannot retrieve all of the information that was once present.

As stated above, no single one of these strategies is sufficient alone. A functioning preservation strategy must be prepared to incorporate several or all of them. Which strategies are implemented and how this is done depends on the requirements of the material and of the entities involved in preservation. To illustrate how a digital preservation workflow is implemented in the context of audio production, I interviewed a vendor that provides long-term digital storage solutions primarily to major label clients.

BEST PRACTICE CASE: BMS/CHACE

BMS/Chace is a vendor for long-term storage of digital assets.⁴⁰ They provide these services mostly to major record labels, but also to individual artists and archival institutions. Clients include Sony, Columbia, Epic, Geffen, Arista, Garth Brooks, and the Library of Congress, to name only a few.⁴¹

⁴⁰ BMS/Chace Home Page, accessed April 24, 2014, <http://bmschace.com/>.

⁴¹ "Clients," BMS/Chace, accessed April 24, 2014, <http://bmschace.com/clients/>.

I interviewed Cory Kempton, Director of Label Services at their Los Angeles office in June 2013. Their basic workflow for a label's release can be broken down in the following way.

1. Verify masters, mixes, stems, and digital audio workstation (DAW) session files. The amount of material that comes standard here is important to note. Not only are masters delivered, but also unmastered mixes, "submixes" known as stems, and the multitrack DAW session files. DAWs contain software that allows the recording, mixing, and mastering of audio files. DAW session files contain all the raw recorded data and editing decisions that form the sound of mix and master files. They are also called multitrack files because instrumentation is broken up by track. A mix file takes all of the recorded tracks and editing decisions and bounces these into an uncompressed stereo audio file. Stems are a subset of mixes. Stems only contain partial amounts of the recorded information, not all of the tracks. Many productions (audio, video, film or otherwise) group audio production into stems. In a feature film stems might include mixes that contain only dialogue, only music, or only special effects. For an audio production there may be different stems for similar groups of instruments. Breaking mixes into stems facilitates a number of major market activities like foreign language dubbing, remixing, and creating radio friendly edits of songs. Maintaining the session files, or the DAW multitrack files, entails saving the project in the ProTools file format, or the native file format from whatever DAW is used, as opposed to a WAV or AIFF file. BMS/Chace receives session files from DAWs like ProTools, Cubase, Logic, and Frooty Loops. Keeping these files is useful because it allows a lot of potential freedom for future remixing or remastering efforts. It is analogous to having access to all of an analog

recording session's original tapes, not just a single tape with all of the instruments mixed down to two stereo channels. This can be helpful if in the future multichannel formats become popular. Super Audio CD (SA-CD) supports multichannel recordings but has not really penetrated the consumer market.⁴² When multichannel SA-CDs are made the engineer uses multitrack files or original tapes, as opposed to a two-channel stereo master. The engineer is allowed greater freedom to decide where she wants information to go in each channel. All files are delivered to the BMS/Chace office on hard drive, or are picked up from the recording studio by a BMS/Chace employee. All of the various directories containing the masters, mixes and DAW files are then verified to be functional and authentic. They are then properly organized (studios often do not do a great job of organizing these directories themselves).

2. Further technical and preservation metadata is inserted via Widget X, proprietary software designed by BMS/Chace. This metadata can include things like the archival engineer performing the preservation actions, information about what has been performed, and anything else useful to maintaining the integrity of the files, their relation to one another, and their usability.
3. The verified directories with inserted metadata are uploaded to the BMS/Chace cloud storage service. From here they can be transmitted to clients via FTP, or file transfer protocol, a very stable method of transferring data online. This serves as a more accessible backup for clients. Since the hard drives ultimately go back to the client that

⁴² SA-CD.net, "FAQ," accessed May 5, 2014, <http://www.sa-cd.net/faq>.

funded the recording, if they find the hard drive has failed the easiest method to regain the data would be via FTP.

4. These data packages are also archived onto LTO tape. At the time I was there they were using LTO 4. LTO stands for linear tape open, and is a form of data tape that has become popular for what is known as deep storage. LTO tapes are designed to be stored for long periods without use and are considered the last resort fail safe. When dealing with a large enough business and amount of data, they are an attractive option due to a relatively small cost per gigabyte of storage. This option requires periodic reinvestment in addition to the costs of starting up. LTO 4 refers to the fourth generation of the technology. With each new generation tape capacity increases, lowering the cost per unit of storage. Changes to the file system have also been made during the course of the development of various generations. Each generation requires a new drive to write to the tape cartridges and to subsequently access the material. Reinvestment is necessary because any given generation of drive can only support two generations backwards. At the moment LTO 6 is the most recent available version.⁴³ With an LTO 6 drive one can read LTO 5 and LTO 4, but not LTO 3. Therefore adopting an LTO workflow presupposes periodic reinvestment and migration every two generations in order to keep the material accessible.

Setting up this sort of infrastructure would be overkill in the context of the labels I interviewed. It's even outside of the scope of many major labels, which is why they hire a vendor. Paying for these services from a vendor like BMS/Chace is likely outside of most

⁴³ Ultrium LTO, "About," accessed May 4, 2014, <http://www.lto.org/About/faq.html>.

independent labels' budgets also. Their only independent client is Epitaph, one of the largest independents. Yet this workflow provides some of the best possible solutions to digital preservation issues for the type of digital objects BMS/Chace manages. Multiple copies of each object are kept on different formats and in separate geographic locations. This helps prevent loss due to disasters like flood and fire, and protects against the obsolescence of media. Metadata is meticulously generated and travels with the items to which they pertain, making the items findable, renderable, understandable, and capable of being traced to an authentic origin. Multiple generations of the audio production process are stored to allow for future technological developments or commercial opportunities.

While all of this is fantastic and top of the line preservation activity, this degree of intervention is usually only appropriate for companies operating at the scale of major labels or other enterprises with large volumes of data. It is way above what is needed for the vast majority of independent labels and the media they generate. For instance, saving DAW sessions is not something that is normally helpful to the labels I spoke with because they have no practical reason for them to revisit this version of the recordings after an album's release. This is not to say it is a fruitless practice. Many remixes and mashups featuring bands from the 1990s have been very popular recently. One example is Swiss Andy and Cecil Otter's project Wugazi. Wugazi combines vocal tracks from hip-hop's Wu-Tang Clan with instrumentals from the punk band Fugazi.⁴⁴

What is important to note is that the BMS/Chace workflow provides a horizon towards which independent labels should look to see how seriously digital preservation is taken when significant resources are available. If not these specific methods, many of the preservation

⁴⁴ Dan Solomon, "How the Wugazi Mashup Came to Be," MTV Hive, July 13, 2011, accessed May 6, 2014, <http://www.mtvhive.com/2011/07/13/how-the-wugazi-mashup-album-came-to-be/>.

principles behind BMS/Chace's workflow can be incorporated into independent audio production workflows.

With this in mind, I will return to the context of independent audio production to further analyze its preservation issues.

LOOKING BACK TO INDEPENDENT

Though best case, enterprise level solutions are not called for with independent labels, there are still issues that need to be addressed. Given the vast array of risks that can beset a digital object, keeping a single copy of something is certainly inadvisable. The other large-scale issue is the lack of organization of labels' digital media. As most of the interviews showed, many labels seem on the surface to be susceptible to problematic data loss.

Yet despite a lack of backups and formal organization⁴⁵, another commonality is that there have not been any instances of irreplaceable loss. Between all five labels there have only been a handful of instances that forced a label to resort to a lower quality element to replace a master. There are no instances where a project could not be undertaken due to corrupt media. There are furthermore no instances where some form of a release could not be accessed. Is this simply due to the relatively small amounts of data that each label has under their care? This cannot be the case since there are already missing elements in some of the labels' catalogs. The collections are not the largest around, but they are large enough to have a few blind spots.

Still, why is this situation not worse? Why haven't the archivists' nightmares of digital ephemerality sprung out to haunt these labels as well? The answer is due to something inherent in the production of digital audio. Another commonality echoed by all interviewees is their reliance on other members within the production chain to serve as de facto backup sites. Clay

⁴⁵ I.e. some form of inventory, database, or other form of cataloging materials in their possession.

Marshall even went so far as to suggest that the label is not the “master keeper” for these elements. Greg Anderson echoed the sentiment. He identified other places he could obtain a master if Southern Lord lacked it such as the band, the mastering facility, the pressing plant, and the digital distributors. Heath also espoused this idea for Tankcrimes, saying the mastering houses back up his files. The way digital audio is produced inherently creates an environment of multiple stakeholders.

The idea of identifying and leveraging multiple stakeholders for digital preservation initiatives has been an integral strategy in the field for many years. Just as digital technology easily transcends the borders of systems, organizations, and nations, so too can digital preservation efforts. This statement from Brian Lavoie and Lorcan Dempsey, made in 2004, outlines the concept very well:

For the most part, digital preservation systems have been designed "holistically", combining raw storage capacity, ingest functions, metadata collection and management, preservation strategies, and dissemination of archived content into a physically integrated, centrally administered system. But other organizational structures are also possible: for example, digital preservation activities might adopt a "disaggregated" approach, where the various components of the preservation process are broken apart into separate services distributed over multiple organizations, each specializing in a focused segment of the overall process.⁴⁶

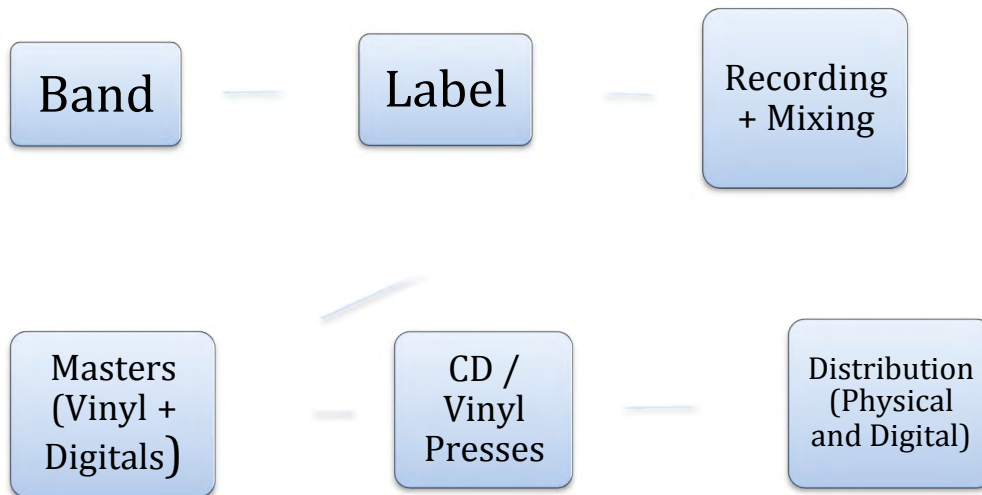
⁴⁶ Brian Lavoie and Lorcan Dempsey, “Thirteen Ways of Looking at... Digital Preservation,” *D-Lib Magazine* 10, no. 7/8 (July/August 2004), accessed April 24, 2014, <http://www.dlib.org/dlib/july04/lavoie/07lavoie.html?pagewanted=all>.

In order to more effectively carry out preservation activities, major projects frequently feature multiple organizations dissecting and allotting the various tasks involved in the project. Similarly, the very nature of digital audio production features and leverages multiple stakeholders. Different agents along the production chain maintain a copy of master files in order to more efficiently perform jobs on that project as needed. A common example is a pressing plant maintaining a copy in case represses are ordered. Mastering engineers will keep a copy of a digital master in case the label later wants to get a vinyl master to produce a vinyl version of a release.

A description of digital audio production for independent music will help to identify the potential stakeholders that exist for any given release.

GENERALIZED PRODUCTION WORKFLOW

Due to the varying priorities bands, the needs of a label, and other considerations, none of the labels has the exact same process for each release. A lot of variables can affect how a release comes out. A major benefit to working independently is having room for flexibility. However, for the purposes of this discussion, I will stick to a very basic sketch of the production process.



First, a band exists and makes music. Later a label finds out about the band, digs their work, and wants to help them release music. The label and band come to an agreement. This can be for one or multiple releases, or for one format of a single release. It is not uncommon for one label to handle digital and CD distribution, while a second label releases a vinyl version. In any event, agreements are made. The form of the agreement also varies. It can range from a complex document filled with legalese to a simple oral agreement.

The recording will either have happened already at the band's expense, or will be an expense that the label pays part or all of. The result of a recording session is a file called the mix, produced by the recording engineer. Next, a mastering engineer takes the mix file and masters it to make the master file. This is a good place to further differentiate between mixes, masters, and digital audio workstation (DAW) session files.



The above image is a screenshot from a session in ProTools.⁴⁷ Recording in a DAW emulates the recording process in a multitrack console. Tracks are separated based on instrumentation. Each vocalist has a track, bass guitar has a track, rhythm guitar has a track, drums have several tracks, and so on. In the screenshot above these tracks are visualized as the individual rectangular bars. Each bar displays the waveform of the specified instrument's recording. Once every instrument is recorded, the recording engineer mixes them together. Each track can be treated individually at this stage. The engineer can add a lot of reverb to one guitar but no other track. She can pan one vocal track and leave everything else as it is. These edits are made using built-in effects or plugins. The two windows on the right of the screenshot depict two plugins. All of these editing decisions are visible and can be altered to taste in the DAW session

⁴⁷ Avid, "ProTools 11," accessed May 4, 2014, <http://www.avid.com/US/products/pro-tools-software/features>.

using that session file. The session file saves all of the data relating to each track and every plugin or effect used.

Once the band is happy with how all of the instruments sit together in the mix, the engineer bounces all the tracks down to a stereo mix file from the session file. Mixes are usually in WAV or AIFF format. The mastering engineer receives the mix and then applies additional processing in order to get the quality of the recording as high as possible. Often this means global equalization adjustments, getting the volume to an adequate and stable level, and making sure that the different songs of the same record are sonically cohesive with one another. At this stage every change is global. Instead of looking at all the waveforms for each instrument, the engineer is manipulating two waveforms representing the left and right channels of the stereo recording. Adding a little reverb would affect every element in the mix; the engineer cannot target a specific instrument. While master files are the end result and should be the best sounding version of the three, the earlier files allow for flexibility. Saving the DAW session file can be helpful if you want more than one mix. This is often desirable when you want to mix to different channel setups. If you wanted a song to have stereo and 5.1 surround sound versions, this would require two different mixes, as opposed to different masters from the same mix. This is why major label workflows save all session files, mixes, and masters.

Saving a mix file would be necessary if you wanted more than one master. In the event that vinyl will be produced, a vinyl master may be created. There can be as many as three masters, a high quality digital master, a CD-quality digital master, and a vinyl master. To get a better understanding of how this process works, I spoke with Brad Boatright, a mastering engineer who handles many releases for Southern Lord. He receives a mix file at the resolution in which it was recorded. In this context resolution refers to the combination of two digital audio

characteristics, bit-depth and sampling rate. As he said to me, “99% of all mixes I receive are at 24 bits these days... I’d say 40% of mixes are submitted at 44.1 kHz, 40% at 48 kHz, and then an even distribution of 88.2 and 96 kHz after that.”⁴⁸ This means the vast majority of mixes are recorded at a resolution greater than CD quality audio, which is 16-bit, 44.1 kHz. For most digital masters he usually dithers the bit-depth to 16-bit and truncates the sampling rate to 44.1 kHz. Vinyl masters are generated at the resolution provided. He will also print digital masters at their native resolution for labels and artists that “request high resolution digital masters for archival purposes.”⁴⁹ His masters are always printed as WAV files.⁵⁰

Next, the relevant masters are sent to pressing plants and digital distributors. The vinyl house gets the vinyl master, the CD house gets the CD master, and the digital distributors get the digital master at CD resolution or higher. Vinyl and CDs are pressed and sent to the label’s physical distribution partner(s).

Finally, the label, the artist, the recording engineer, the mastering engineer, the pressing plants, and the digital distributors hold on to copies of the master elements in the manner and duration that suits them.⁵¹

INHERENT PRESERVATION

The number of different people involved in the production process provides a hint as to why none of the labels reported a case where a working copy of a release could not be found.

Although labels themselves tend to neglect backing up their files, this is partly due to the fact

⁴⁸ Brad Boatright, email to author, March 5, 2014.

⁴⁹ Ibid.

⁵⁰ Boatright went on to note that he does not use the Broadcast WAV (BWF) format yet since for his purposes the metadata applications are not stable. Additionally, he often creates MP3 copies with embedded metadata and album art as a courtesy to the bands.

⁵¹ And, of course, ideally lots of people purchase and listen to the resulting release.

that they are aware they do not necessarily need to. Bands, mastering engineers, pressing plants, and digital distributors are all potential sources for a copy of a master should the label's copy become lost or unusable.

The digital preservation strategy of leveraging multiple stakeholders is therefore inherent to the independent production process. However, problem areas could still arise over the long-term. Most of those stakeholders only maintain copies for short-term purposes. Mastering engineers and pressing plants tend to keep copies around 1 to 3 years. Brad Boatright spoke to me about how he maintains his files:

I archive everything on external hard drives as uncompressed folders containing the masters. I also save the DAW sessions in these folders. About twice a year I'll do some cleaning up and trim things down a bit. For example, digital masters that were already released on Bandcamp and have aged might get purged, although I'll save the session file in case vinyl masters are ever needed.⁵²

For establishments with stricter policies, if further work is required, the timing generally restarts. If a release gets a second pressing on vinyl, for instance, the pressing plant restarts their clock on maintaining the master. Pirate's Press keeps their masters for 2 years, or shorter if storage space is low.⁵³

It is only on a short-term basis that the amount of stakeholders is high. If a record didn't need repressing or remastering for five years, the number of stakeholders with potential copies would effectively revert to two, the label and the artist. Another issue is that all the stakeholders

⁵² Boatright, email.

⁵³ Josephson, email.

know about each other. Therefore all know that everyone else ‘will probably have a usable copy.’ So while multiple copies exist in different places, some of the copyholders may not place a lot of value on the copy they retain, and as such may not do very much to ensure its survival. And if this is the case at the point where the label and the band are the only stakeholders, then the copies face a high risk of becoming inaccessible.

At this point something else Boatright mentioned about his data retention practice requires discussion. He mentioned he feels safe purging masters that have been uploaded to Bandcamp.⁵⁴ Presumably, this is the case because the site has to maintain uploaded master files in order to generate the derivatives purchased by customers visiting the site. To keep the service operating at an acceptable level, it is a fair assumption that they would have to safeguard the digital materials they source for derivative audio and image files. Otherwise users would be dissatisfied either with downloads being unavailable upon request or with corruptions in the downloaded files. The same should hold true for most digital distributors, as they need to ensure delivery of satisfactory derivatives for the labels they serve.

As such digital distributors may be one of the most effective stakeholders. In order for that to be the case, they would have to store the master audio and artwork files in such a way that they are retrievable by either the label or artist, and that these files are safeguarded against decay. The effectiveness of the digital distributor also relies on the length of time they store these files. It should be the case that the digital distributor maintains a well-preserved copy of all master files for the duration the service is employed by the label. If this is the case, the digital distributor copy should be available for a longer period of time than the copies maintained by either engineers or pressing plants.

⁵⁴ Boatright, email.

A more in-depth discussion of how digital distribution companies operate is warranted to more accurately depict their potential value as stakeholders for the preservation of independent labels' digital output.

DIGITAL DISTRIBUTORS AS STAKEHOLDERS

Since digital distribution is still in its adolescence, there are a plethora of choices. The methods of distribution can vary as well. One form of service is to provide artists and labels with the means to generate income through many popular online marketplaces and streaming services.⁵⁵ In such instances the company handles all of the heavy lifting required to submit music, artwork, and relevant metadata to each service. In other cases distribution can take the form of a single portal. Bandcamp is such an example. Bandcamp does not offer services to release music through other online marketplaces; it only handles transactions made through the Bandcamp domain. Some companies work with major labels, some with independent labels, and others offer services to unaffiliated artists looking to self-release their material. In each case the workflow and services rendered vary.

Another source of complexity is that many labels make use of more than one option at the same time. Southern Lord for instance uses Sony's RED service for domestic digital distribution, The Orchard⁵⁶ for European digital distribution, and also has a Bandcamp page highlighting its artists' profiles. To simplify the discussion I will focus on three examples, Sony RED, Bandcamp, and Archival Digital.

Sony RED

⁵⁵ Such as iTunes, Amazon, Spotify, Pandora, etc.

⁵⁶ The Orchard, "Homepage," accessed May 4, 2014, <http://www.theorchard.com/music-distribution/>.

RED Distribution LLC is a sales and marketing division of Sony Music Entertainment that deals primarily with independent labels.⁵⁷ They provide many services apart from distribution, but for now I will focus solely on their digital distribution service since that is the only way anyone I talked to uses them. At the moment they represent 114 labels.⁵⁸ RED provides the first type of digital distribution service described above. They act as an intermediary between the label and a large number of active online marketplaces.

Unfortunately, I have not been able to speak with someone who works at RED about the company's process. I was however able to interview Max Burke, who works with Sony DADC.⁵⁹ Sony DADC offers optical disc replication and digital distribution services to a wider group of clients, including "entertainment, education, and information industries."⁶⁰ In his role in Sony DADC, Burke often works closely with RED in the capacity of generating derivative files and delivering those files according to the requirements of each marketplace.

Burke was able to describe RED's digital workflow to me.⁶¹ The label interacts with a label representative from RED. The RED label rep acquires the master audio and artwork files along with basic metadata. At this stage they generate other metadata themselves, such as UPC and ISRC numbers.⁶² Once they receive the files from the label, the files are given to another branch within Sony, media production. Media production ingests the master files into Sony's proprietary content management system. It is called the Asset Offering and Management Application, or AOMA.⁶³ This content management system is a piece of software that allows various branches in Sony to perform necessary operations on various types of media files, such

⁵⁷ RED Music, "About," accessed April 24, 2014, http://www.redmusic.com/?page_id=275.

⁵⁸ RED Music, "Labels," accessed May 4, 2014, <http://www.redmusic.com/labels/>.

⁵⁹ Sony DADC, "About," accessed May 4, 2014, <http://www.sonydadc.com/en/about-us/>.

⁶⁰ Ibid.

⁶¹ Max Burke, interviewed by author, phone interview, April 28, 2014.

⁶² UPC numbers are used to track sales of entire albums, ISRC numbers are used for individual songs.

⁶³ Max Burke, email to author, May 5, 2014.

as transmitting and storing the files in-house, transcoding files into derivative versions, delivering derivatives to partners or clients, and managing metadata. Burke takes the master files ingested by media production and performs all of the transcoding to create packages specific to each partner.⁶⁴

None of these derivative packages are retained after they are delivered to their respective online marketplace services. This is understandable, as it would require a great deal of storage space to save everything generated. All of the master files are retained within AOMA. According to Burke, “AOMA stores geographically redundant copies of all masters... The master resides in the system indefinitely until a label takes action to purge or delete it.”⁶⁵

This situation is very helpful to the preservation of master elements. In addition to digital distribution services, RED effectively serves as a de facto robust digital backup service. Not only do they keep a copy of all of the master audio, artwork, and metadata, but more than one copy is maintained and these are geographically separated. Another advantage is that the label has control over the deletion of the files.

Bandcamp

Bandcamp takes a different approach to distribution.⁶⁶ They only distribute through their site. Labels and artists alike use Bandcamp, and signing up for the basic service is free. They require you to upload lossless audio (WAV, AIFF, or FLAC), and from there create a number of derivatives, among these the highest possible quality MP3s, but also formats like AAC, FLAC, and OGG VORBIS. Streaming copies are low quality MP3s to ensure effective service.

⁶⁴ Partner is Burke’s word for what I am referring to as an online marketplace, e.g. iTunes.

⁶⁵ Burke, email.

⁶⁶ Bandcamp, accessed April 24, 2014, <http://bandcamp.com/>.

Bandcamp also acts a way for artists to sell their music via their profile. Labels can have profiles as well that serve as a hub for the various artist profiles they represent. Artists can also sell physical merchandise and treat the site like an online mail order portal. Bandcamp generates income by taking a cut of all sales. Bands are able to offer downloads for free if they wish.

As described above in the general risk factor section, Bandcamp automatically updates metadata to the master and derivative files. This is a service that other social-network-digital-market-place hybrid sites have not provided in the past. When a band enters metadata about its music, Bandcamp automatically adds information to the derivative files it generates for users.

I have been unable to get a response from anyone at Bandcamp regarding the storage and potential retrieval of the uploaded lossless files. To test if it is possible to retrieve the uploaded files I started a free account and uploaded a WAV file. From the site's online interface there is no option for the owner of the profile to download the version of the file they uploaded. It may be possible to retrieve these by contacting Bandcamp administration, but I have not been able to confirm this or achieve contact, and the point is not mentioned in their help pages.⁶⁷ The only options for download available to the artist are the same derivatives available to any user of the service. Although WAV or AIFF options are not available, there are high quality lossless options such as FLAC. In a scenario where the original WAVs or AIFFs are lost, these high quality derivatives would be valuable versions of the audio files. But I had hoped that retrieving the original versions would be an option for the artist controlling the profile.

Though there appear to be difficulties in retrieving a master in its original form from digital distributors, access to high quality derivative files could be satisfactory. Metadata tagging is another beneficial service they provide. It is still uncertain though how they store these files,

⁶⁷ Bandcamp, "Help," accessed May 4, 2014, <https://bandcamp.com/help>.

how derivatives are generated, and whether artists could access material they uploaded in the event they lose their own copies

Archival Digital

Archival Digital, LLC is a one-man digital distributor run by Sean Patrick Rhorer. He handles digital distribution for Forcefield Records, and in total services 60 labels.⁶⁸ It is a relatively new company. Rhorer started it after a physical and digital distributor he worked for went out of business, but he realized that many of his clients through that job still wanted to work with him.

He offers services similar to RED's, acting as an intermediary between the labels and online market places. For every release the label fills out a metadata form. The forms are Excel spreadsheets that track metadata such as artist, album name, track listing, UPC and ISRC numbers, genre, and others. He provides whatever the label cannot fill in.

After derivative packages are created the master audio, artwork files, and metadata are stored locally on Rhorer's hard drive. He stores only one copy of the material. According to Rhorer, "File errors and corruptions are certainly a possibility, but not something I've ever encountered. I also assume the labels don't consider my data to be the primary point of storage for their content, so if I did have an issue, I could replace needed content fairly easily."⁶⁹

Depending on the digital distributor, labels appear to get varying levels of de facto preservation service gratis. This should be seen as a great tool for preservation, and may help

⁶⁸ Sean Patrick Rhorer, email to author, May 5, 2014.

⁶⁹ Rhorer, email.

take some stress off of label owners concerned with preservation but lacking in time and resources to seek out solutions themselves. However, since different distributors care for the master material they receive differently, labels should familiarize themselves with those policies before assuming that the distributor will always have the master material after the initial delivery.

After investigating all of the various stakeholders, what I have described as the inherent preservation of the digital audio production chain provides a more robust preservation system than I had imagined going into this project. Although the specific care individual sites pay to their copy of master elements along the production chain can vary, the net result is a host of entities collaboratively preserving their digital output. Even though it is unintentional, it is an excellent short-term solution. It explains why there have been no horror stories from any of the labels regarding the disappearance of a particular master or being unable to pursue a project. The large number of places where labels can reasonably expect to find usable digital master material explains why Clay Marshall does not even consider Prosthetic as the “master keeper” of its digital output. Labels have not pursued preservation not only due to a lack of resources, but also because there has yet been any real motivation to do so.

I would go so far as to argue that the characteristics of this inherent preservation system bear a resemblance to one of digital preservation’s most respected standards, the Open Archival Information System (OAIS). The following section describes what OAIS is and draws comparisons between how that model defines the functions of the digital archive, and how digital audio production practices perform similar functions. In this way I hope to highlight the unexpected strength of current practice, which has to date allowed few instances of irrevocable loss. In the end, though, I hope also to highlight what OAIS has and what is missing from the

audio production system in the context of long-term preservation, and in this way cast light on areas that can be improved to ensure access to independent material in the future.

OAIS

The Open Archival Information System, or OAIS for short, is a reference model that has become an accepted standard in the field of digital preservation.⁷⁰ Beginning in 1982 a group called the Consultative Committee for Space Data Systems, a large group of international space agencies, started developing this model in order to preserve their digital data.⁷¹ Since then it has been accepted by myriad organizations that want to effectively preserve digital information.

The OAIS model also takes into account the idea of multiple stakeholders. In his introductory guide to the model, Brian Lavoie makes a statement about the challenges surrounding the preservation of digital information that prefigures the necessity and effectiveness of leveraging a system with multiple stakeholders, “Just as the benefits of digital information environments transcend people, systems, and domains, so do the challenges which accompany them.”⁷²

OAIS is a conceptual model that defines what an archive of digital information must do to properly care for digital material. The archive is an entity comprised of people and systems that have taken it upon themselves to do two things. The first is to preserve information, and the second is to make that information available to a Designated Community. The term Designated

⁷⁰ The Consultative Committee for Space Data Systems, “Reference Model for an Open Archival Information System (OAIS)” Magenta Book, June 2012, accessed April 24, 2014, <http://public.ccsds.org/publications/archive/650x0m2.pdf>.

⁷¹ Brian Lavoie, *The Open Archival Information System Reference Model: Introductory Guide*, DPC Technology Watch Series Report 04-01, January 2004, accessed April 24, 2014, www.dpconline.org/docs/lavoie_OAIS.pdf.

⁷² Ibid.

Community refers to the intended users of preserved material. They possess or have access to all of the necessary external dependencies to acceptably use the information.⁷³

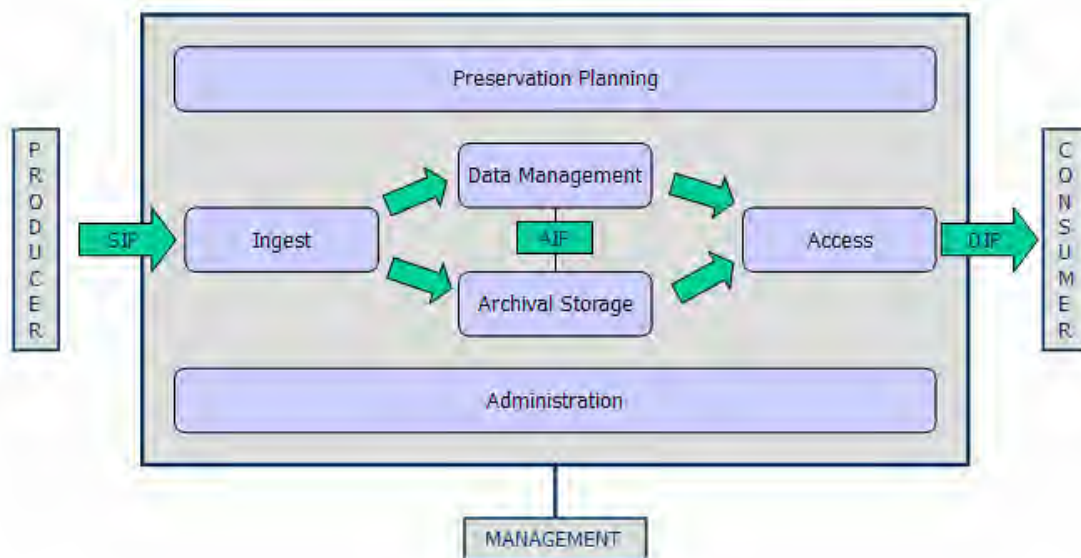
In order to accomplish these two objectives, the model demands the archive meet six requirements:⁷⁴

1. Negotiate for and accept information in an appropriate state from the producers of that information.
2. Obtain sufficient control over the producers' information to be able to preserve the material in the long term
3. Determine on its own or with other people/organizations who the designated community or users of the information are, hence who should be able to understand the material
4. Ensure that the information can be understood by those users; that they will not need the expertise of the people who produced the information to use it
5. Follow documented policies and procedures to protect the information from reasonable risks, and ensure the information provided to users is authentic or can be traced to an authentic origin
6. Provide access to users

⁷³ Ibid.

⁷⁴ Ibid.

The Open Archival Information System (OAIS) Reference Model



75

Above is a simple functional diagram for the OAIS model. I will refer to this in the following discussion. I will start with the least identifiable parts first, the SIP, the AIP, and the DIP.⁷⁶ These comprise OAIS' description of how information travels from the producer, through the archive, and to the consumer.⁷⁷

From the producer and to the archive the information takes the form of the Submission Information Packet, or SIP. As a way to parallel to what I have spoken to earlier, a similar exchange would be when a label gives their master to the digital distributor. They provide the audio master files, along with associated artwork and metadata such as album info and ISRC numbers.

⁷⁵ "Digital Archives," Archivemata, November 8, 2005, accessed April 24, 2014, <http://archivemati.ca/2005/11/08/digital-archives/>.

⁷⁶ Those dudes in the green boxes.

⁷⁷ The term consumer here is interchangeable with user or Designated Community.

The Archival Information Package, or AIP, is created when the archive transforms the SIP during ingest into a form suitable for preservation. To continue the digital distribution analogy, after taking the master audio files for an album, the distribution company would package together the provided metadata with the album and artwork. It will also generate metadata it needs. All of this will be packaged together in such a way that allows all of this information to be kept intact in a logical manner, and allow the distribution company to eventually provide access to it. All of this sits in the archive as one logical package.

The AIP allows the archive to create the Dissemination Information Package, or DIP. DIPs are given to the user/consumer. To complete the digital distribution analogy, the entire package of master audio and artwork files with associated information, its AIP, is used to create derivative files that are delivered to users via online marketplaces. Derivative files include MP3s of audio files and smaller image files from master artwork image files. These derivatives can be used along with the appropriate album info, artwork, and commercial information for iTunes to sell to individuals, or for people streaming on Spotify to hear the track, look up the artist and album, see the track length, view a thumbnail of the album cover, or any of the other browsing activities listeners in which listeners can engage.

The SIP, AIP, and DIP all stem from the same concept, that of the information package. The information package is defined as “the content to be preserved, along with its associated metadata.”⁷⁸ This definition comprises the digital object itself along with all the necessary information to render and understand the object. In the music example, the songs, the artwork, all song names, track order, genre tags, etc.

The AIP is the strictest version of the information package concept. It includes important components like the object and its metadata, as well as additional metadata created by the

⁷⁸ Lavoie, OAIS Introductory Guide.

archive. This can include fixity information, or information that ensures the integrity of the digital object, that its 1s and 0s have not altered. Also included in the AIP is information regarding the context and origin of the packaged information. Keeping track of this information ensures that the versions accessed by users will be authentic versions of what the archive took in from the producer.

The central gray box in the above diagram delimits the archive itself; that which the OAIS model defines as the entity that preserves and provides access. The role of management is to determine the policies and procedures the archive follows in turning the SIP into the AIP and the AIP into the DIP, as well as the policies and procedures surrounding how it manages the AIP over the long-term.

While this description is admittedly blunt, I hope it is sufficient for me to continue my thoughts from earlier.

As I quoted Lavoie above, he said that the challenges and benefits of digital objects transcend people, place, organization and systems. My conception of how the independent music production chain and how this OAIS model relate involves altering and multiplying the way these boxes are defined in the diagram. By which I mean many of these functional roles (producer, archive, consumer, management, etc.) have de facto analogues. These roles can be performed by multiple organizations simultaneously.

The producer can be imagined as the label, the artist, or both. The consumer is most easily compared to a listener, but could also be the label or artist in the event their assets become corrupt and they need to access a copy from another storage location.

As I did in my examples explaining information packages above, we might place the digital distribution companies in the position of the archive, the central gray box. Therefore this box would have to be multiplied, with several layers on top of each other, as distributors like RED, Orchard, and Bandcamp operate in parallel but importantly different ways often times for the same labels at the same time.

For the way Bandcamp is set up, they operate in a way that overlaps the archive and management roles. As the archive they create the package downloaded by users, and as management they determine how those services are carried out. Bandcamp only allows three types of lossless file submissions, and have a set number of access file types.

In the case of RED, RED and the label might occupy the role of Management. RED and other Sony departments determine how they manage data storage, but the label can make decisions as well. For instance the label defines which online market places will be used.

The role of archival storage is essentially split up between many entities. At first copies are held by the band, the label, the mastering engineer, the pressing plant(s), and the digital distributors. Each also acts as the management in terms of determining how and for what duration the material is stored.

To be fair, because digital distribution has grown in popularity and quality it makes sense that there would be structural consonance with a preservation model whose primary goal is access. Both systems are concerned with serving users with digital information that they can understand and enjoy properly. My main objective at trying to draw various parallels here is to emphasize that many of the behaviors described in preservation texts are behaviors already being performed in the production of digital audio, only under different names and with different levels of rigor.

At this point, then, it is important to move on to how OAIS is different from the audio production process, and thereby showcase why although labels have not yet had issues, there is reason to invest in more robust preservation practices if they want their output to survive into the future.

As stated the archive as defined in OAIS agrees to do two things; to preserve and provide access. It is a long-term commitment to preserve and provide access. Not one stakeholder in the music production chain as it stands explicitly agrees either to preserve or provide long-term access. All they have agreed to do is provide short-term access. Many labels consider the job done after a record is released. Pressing plants and mastering houses only keep material for a short period, and this is primarily to assist their jobs should more work be done on the same record, not as a preservation action. Digital distributors can resemble something like a digital repository only due to the fact that effective online distribution relies on stable information. They need to maintain the integrity of the files, but this maintenance is not built on ideals of preservation. Although distributors may keep files for longer periods than pressing plants or mastering houses, this maintenance is not a long-term commitment to preservation. The time frame for which distributors could be considered effective stakeholders can more appropriately be defined as just as long as sufficient commercial demand for the service exists, and no longer.

Therefore the above diagram's square for preservation planning has no true analogue in the independent music production chain, and neither do many functions taken on by the various functional entities in this diagram for the sake of long-term preservation and long-term access.

As such I return to the point I have made earlier about who the stakeholders end up being over a long-term period; the label and the artists. Should preservation be of no concern to either,

it is a real possibility that their material will disappear in the coming decades. If, on the other hand, these parties would like to take an active role in trying to see the music survive in coming decades and new generations, then there are a number of options that can be taken up. The below section enumerates steps that can be taken to help preserve material that might otherwise be neglected.

RECOMMENDATIONS

Due to the long-term risks to independent media, I have assembled a list of recommendations that should be manageable for most independent labels and artists that have the desire to ensure the longevity of their productions. My hope is that these suggestions are low-impact on the labels' resources, financial or otherwise.

DAT Transfers

Though in my interviews there were very few DAT tapes with master material on labels' shelves, I still want to stress the importance of migrating material off of DAT due to its prevalence in the beginning of digital recording. DAT decks stopped being manufactured in 2005, and are not physically stable.⁷⁹ If the masters exist nowhere else, it is imperative to migrate information off of DAT as soon as possible. Susan Eldridge has written a very complete guide to the issue of converting DAT tapes.⁸⁰ As Eldridge notes, "there are many facets of this process, including the selection of hardware, software, file formats, and metadata."⁸¹

If there are only a handful of tapes, a vendor may be best. Finding a deck can be time consuming and potentially expensive to ensure that it works.

⁷⁹ Eldridge.

⁸⁰ C.f. Note 17

⁸¹ Ibid.

If you have enough tapes that purchasing a deck and doing the transfers yourself is preferable financially, the best way to do the transfers is through a digital-to-digital conversion. This provides a lossless transfer through a less demanding quality control process than doing analog-to-digital. Digital-to-digital transfers require using a DAT deck with either an AES/EBU or S/PDIF output. You then need to find an audio interface with an AES/EBU or S/PDIF input. AES/EBU and S/PDIF are cable specifications that were specifically designed to handle digital audio streams. AES/EBU cables look like XLR, but normal XLR should not be used, since its guts were not designed for digital audio.

The interface should also be compatible with 32, 44.1, and 48kHz sampling rates, as those are the sampling rates supported by DAT tapes. Although probably you can get away with one that doesn't support 32kHz, since it's not likely this sampling rate would get used for digital recording, as the standard was CD quality audio (16-bit/44.1kHz).

If you don't already have recording software on your computer, Audacity is free and used in many audio transfers in archives and libraries. Ensure that the bit-depth and sampling rate of the recording session matches the characteristics of the incoming DAT information. The bit-depth is likely going to be 16-bit, and the sampling rate probably 44.1 or 48kHz. DAT could also record at 12-bit/32kHz, but it's unlikely this configuration would be used for a studio audio recording.⁸²

From Audacity, export the file onto either WAV or AIFF, again ensuring that the settings for bit-depth match what you recorded.

CD and CD-R Transfers

⁸² Ibid.

I assume that most people reading this have a good level of comfort with importing information from CDs onto a computer or hard drive. The main point I want to address here is regarding potential instances where an audio CD is going to be used to regenerate a set of digital master files. Not all programs that import or rip CD audio are created equal. If the intention is to replace digital master files for archival purposes, you will want to use a program with robust error correction capabilities. iTunes, for example, is not the best bet. A good solution for PC users is a free program called Exact Audio Copy (EAC).⁸³ EAC has many features that allow it to cleanly import CD audio on a more consistent basis, such as slowing read speed when encountering error and automatic detection of drive failures. For Mac users, the free program X Lossless Decoder (XLD) is similar. It also features robust error correction features. XLD checks imported files against the AccurateRip database to make sure your files are bit-accurate to the original files.⁸⁴ These programs also have features that allow reporting of any errors do occur.

Inventory

While inventory creation can take time, in this case it might not have to as the label might not need to track very many additional data fields. In the context of the discussion of multiple stakeholders, one of the most important elements to track would be who else has worked on the release. Having information in one place providing who mastered the record, where it was pressed to CD and/or vinyl, what digital distribution companies have handled the release, and contact information for all of these would be very valuable. Whereas you might not need as

⁸³ Exact Audio Copy, "Features of EAC," accessed May 6, 2014, <http://www.exactaudiocopy.de/en/index.php/overview/features/features-of-eac/>.

⁸⁴ Kirk McElhearn, "Four Reasons to Rip CDs Without iTunes," Macworld, February 23, 2011, accessed May 4, 2014, http://www.macworld.com/article/1158059/four_reasons_rip_cds_without_itunes.html.

much assistance from an inventory to locate a file, if that file is corrupted the inventory would help identify other possible locations of a master quickly.

Since time is often an issue, identify the minimum amount of items that would be useful in the context of your label. If staff is an issue you may be able to identify a volunteer or short-term contractor.⁸⁵

Cloud Backup

Owning a second copy of a work can be much cheaper than might be expected. This might not require periodically buying external hard drives or other storage solutions that the label then has to manage. Cloud solutions are becoming very affordable, and require little maintenance by the user. The context of a small label is ideal for some of the cheapest options available. At the moment, Amazon's Glacier cloud storage is the least expensive choice.⁸⁶ It could potentially be very useful as a last resort insurance. Glacier is ideal for files that do not need to be accessed often, and when they do need to be retrieved immediate delivery is unnecessary. When a stored file is requested, it becomes available for download within 3 to 5 hours, and is available for 24 hours. This amount of time should be more than sufficient for a situation in which a label wants to replace a master file.

The storage is also very robust. Redundant copies are kept in geographically separated locations, and integrity checks are made periodically. An integrity check involves making sure bit rot does not impact the stored digital objects. Algorithms called checksums are used, which generate values that could be considered digital fingerprints. If even one bit changes, the algorithms are built so that even a small change in the file will change the resulting fingerprint

⁸⁵ Some kind of metal/punk/preservation nerd perhaps (*cough cough*).

⁸⁶ "Amazon Glacier," Amazon Web Services, accessed April 24, 2014, <https://aws.amazon.com/glacier/>.

value. Checksum outputs are verified against past results, and if there is a discrepancy the file that has changed is removed, and replaced by a copy of the redundant file that has remained the same.

Depending on location the cost for storage alone is \$0.01 or \$0.011 per GB per month. Requesting access to files can generate other expenses. Upload and retrieval jobs cost \$0.05 per 1,000 requests. The first GB of data retrieved from storage every month is free, but after that it is \$0.12 per GB. In the case of replacing one lost master file, or all of the material (artwork, masters, press scans) associated with a release, 1 GB should be a sizeable fraction of the total size. Given the fact that none of the labels reported major losses, more than one retrieval a month would be rare. Even if the entire archive had to be retrieved, for example in the event of a flood, fire, or “heavy metal terrorist attack,”⁸⁷ it would still be relatively inexpensive and less time consuming than replacing this material piecemeal by searching out all the possible copies held by others, which copies might themselves be problematic.

To do a quick example of what actual costs might look like, I’ll invent an example.

- Dan Finn Records has 250 hours of music between all its releases.
- Just to get to the higher end let’s say all of the releases have 3 masters; CD at 16-bit, 44.1kHz, and high quality digital and vinyl masters each at 24-bit/48kHz.
- The size of all the master files is 676.76 GB.
- Monthly storage cost is \$6.77.
- A better option might simply be to keep the best quality digital master, as that file can be used to generate CD-quality or vinyl masters.
- Storing 250 hours of digital audio at 24-bit/48kHz costs \$2.59 per month.

⁸⁷ Clay Marshall, email.

Other options exist but given the size most of the labels' catalogs I have talked to, they are not big enough to make other services as affordable per GB.

Communicate With Artists

Mastering houses and pressing plants may purge their copies of master material after as little as a year if there are no plans to do more work on a release. One simple step to take is to communicate to artists explicitly that the label does not make long-term preservation of the masters a priority, and that if the artist wants those digital masters to be around for them to show their grandkids, they had better take a role in that themselves. If the label has no interest or ability in trying to preserve material, it would benefit the artists to let them know and let them make their own decisions and preparations regarding preservation of their work.

Moving Forward, an Ideal Workflow

While going through an entire back catalog for preservation purposes is a large commitment, it should be easier to incorporate preservation practices moving forward on new projects. The above types of preservation activities should be able to be integrated into normal production workflows without too much hassle.

Below is an example of points in the production process where steps can be taken to improve preservation practice. As such I am calling it an ideal workflow from a preservation standpoint. I've structured this list as a sort of archivist wish list. The preservationists and music historians of the future will love it if you can incorporate these items, but it is of course always up to you what makes sense for your context and capabilities.

1. Recording contract/agreement – One place where preservation activity can be defined is in the recording contract. If you were already writing up documents to assign rights and payment policy, this would be a great place to incorporate the roles and responsibilities of preservation for the respective parties. For instance, if you are interested in incorporating certain preservation steps, but are unsure if the cost will be easily offset by the sales of the label, then there could be sections of the contract that can assign responsibility in the contract to the band. In the very least this is where you can communicate to the band some of the preservation risks, and that you are not taking preservation measurements per se. Alternatively, if you want to offset cost of cloud storage for masters, you can offer the band the chance to have their stuff stored for a small fee or percentage of sales to cover that expense.

2. Recording/Mixing/Mastering – When a band goes to a studio, have the information for that studio entered into the inventory tracking the multiple stakeholders. Suggest to bands that they try and save a copy of all elements they find important. Tell them to ask for copies of the session files and mixes if they might be interested in saving more than a master. Alternatively, if you are involved in the recording process as well, come to an agreement with the studio defining how long they will save copies of the session files, mixes and masters, and to warn you before any deletions are made. They should be able to deliver this material to you before deleting it if you want it.

3. After Mastering – Get a copy of the master files when they are made. Keep each version if there is more than one (i.e. CD, vinyl, higher quality digital). Store these with the

artwork files and any other documentation relating to the release that is relevant. Keep one copy onsite the way you normally would. Upload the files to a cloud backup service for emergency purposes. Again, to save money, at this stage it may only be necessary to save one or two versions. If there are three masters, one higher quality digital master, one CD quality, and one vinyl, it may only be necessary to save the higher quality digital master. The other two can be reconstructed from this version. However, since making a vinyl master is more complicated than adjusting the bit-depth and sampling rate of the file, storing that master in the backup storage may be useful as well. Recreating a vinyl master would require sending the digital master back to a mastering engineer, which will likely be more expensive than the cost of storing the file in a cloud service.

4. Pressing Plants – If a vinyl is being produced, try to keep the master lacquers after the pressing. Having a copy of the lacquers will be very useful, as they can last a long time given proper conditions. Vinyl's continuing popularity should translate into a long useful life for lacquers. If you do not have room or money to pay for them immediately, try to make an agreement that the plant will hold on to the lacquers for a certain amount of time, and that they will advise you before they throw them out. Similarly define the parameters for the duration they will store the digital master files, again advising you before making a deletion.
5. Digital Distribution – If entering into a new agreement with a digital distributor, find out their policies regarding data storage and retention of masters. Also, outline in the agreement what will happen to your data if the business arrangement comes to an end. It

will be useful to define a certain time period where you can decide how much of what you've given to the distributor you want transferred back to you.

6. Represses or Other Work After Initial Release – While revisiting a release it will be useful to check your version of the master files. Even if the master you need to use is still with the mastering engineer or pressing plant, it provides an opportunity to ask for a copy of the digital files used along with other work in the event your file has been misplaced or corrupted.

CONCLUSION

The repercussions of digital technology on our ability to keep cultural productions available to future generations are not yet well understood. Oceans of digital information have been generated in every sphere of cultural behavior, and production of digital media is only increasing. If left alone, the likelihood that digital material will survive centuries or even decades is uncertain at best. Many problems still want satisfactory solutions, and as technology continues to rapidly evolve, more problems continuously arise.

I am certain that this period in music history will be of intense interest to music historians and enthusiasts in the future. Digital technology has allowed people to listen to music they would otherwise have never heard of. You no longer have to be part of a regional scene and go to shows to discover local artists. People all over the world can follow what is happening anywhere else on the planet if they know where to look. The technology has also allowed more artists to create great recordings without the assistance of major label record deals or large recording budgets. While major labels' budgets have given them the ability to implement advanced preservation

strategies, those budgets and those strategies are not wholly necessary for taking effective steps to ensure the survival of all kinds of digital music.

This is an extraordinary era for independent music. The breadth of choices available of not only artists to listen to but also of ways to listen to music has completely reshaped the culture of discovering and sharing music. The idea that this amount of choice could disappear in the not-so-distant future is as alarming as it is avoidable. While it is true that not everything can survive, we can take steps to increase the longevity of the music we think is important. Archiving is a form of curation, just as surely as labels curate their releases when choosing projects. Releasing music is a label's active documentation and promotion of the music they want to share with others. There is no need to limit the impact of that promotion by allowing the content to fade away. With all of the artistic passion and strenuous effort that goes into the production of these releases, it will be a great loss to have so much succumb prematurely to digital decay. By making smart and effective choices now, we can give the listeners of the future the chance to hear who we were, and to feel how hard we rocked.