

; 10/30/13 3:24 PM

;

;;;Learning Times 131030

>> My name is Laura Hertz Stanton the director of preservation services at the conservation center for artifacts.

We are a nonprofit regional conservation center based in Philadelphia and we specialize in paper-based materials that do also have some experience with audio visual materials that's why it's been great to work with heritage preservation.

So, thank you.

>> Thank you.

Before we move on to topic today, let me just quickly review what you can expect at the conclusion of this webinar. To officially complete this course we just ask for a few things, first that you register so you are in our system.

We ask that you watch all five of the webinars in the course whether you're showing up live or watching the recordings.

Wednesday November 6.

So, we had to make a few adjustments to the way we award our certificate of completion. We began that the course we had absolutely no idea how many people would go -- would not only go through the effort of watching but also completing all of those homework assignments. On average we have about 00 people finish each course and you might imagine we are very close to hitting our ceiling on our postage budget it.

In an effort to reduce costs we are going to ask to you help us just a little bit for those of you who complete the course you will receive an e-mail notification that includes your name, number of instruction hours and other pertinent information about the course. This will in essence serve as your certificate and proof of your achievement.

As always you will receive digital credential.

We hope that this is satisfactory alternative that you understand why it's necessary.

But with that I will say we do know that that -- that 8 1/2 by 5 1/2 piece of paper has become important to some of you.

We want to do our best to accommodate those folks who have found it really important.

So there are two additional options we can e-mail you an image of your certificate that you can print yourself or if absolutely necessary we can print it and mail it to you like we have done in the past.

In today's homework assignment you'll notice that the second question after your contact information will be a spot to denote how you want to receive your certificate.

Make sure to fill that out to let us know.

As in course past the final assignment is the evaluation.

You really look forward to hearing your feedback.

If you feel more comfortable doing so anonymously you'll notice at the end of this

evaluation there is an opportunity to provide anonymous feedback.

If you are not interested in earning or completing the course officially which means you haven't been doing the homework assignment I'm going to ask you if you could please do that last assignment fill out that evaluation because that would be incredibly helpful to us.

The course webpage it will remain up and continue to hold all of those presentation, resources, transcript and after the homework is due the course will have concluded, will start the process of posting the webinar recordings to that page so you can share with your colleagues or go back and look at everything that you've gone through already.

What's next.

Shortly following this webinar we will send you an e-mail with links to all the webinar reportings so that we hope will happen today.

It will include recordings and links to homework assignments you'll have everything in one place.

Again all the materials are due November 6, one week from today shortly following that deadline pull down the links to the homework assignments replace them with recordings of the webinars on the course webpage. Staff at heritage preservation will begin the process of logging all the homework and tracking attendance.

Once we have that logged usually takes about a week.

We will send you an e-mail notification and online credentials.

Then if you haven't done so already consider signing up to become a member of the online community.

Remember ship is free does gives you access to postings which is great way to continue some of these conversations.

As always if you have questions please feel free to e-mail us or call us.

With that out of the way let's move on to our topic.

I am pleased to introduce today's instructor Stephanie Renne.

Stephanie is archivist and audio visual at -- she has worked in variety of academic and public libraries archives and nontraditional library settings around the globe.

Her accomplishments and experiences are extensive including the management of the arms disarmament in international security library at the university of Illinois. Development of independently managing the Pacific based institute archive at spa know ma college in Claremont, California, where she was also responsible for building a new campus facility.

I know I'm missing a lot of stuff in there.

But I'm going to go ahead and move this out of the way and hand things over to you.

>> Great, thank you.

I am happy to be here and happy to see so many people from all over the globe.

Welcome.

Like I said I work at George  
blood audio and video.

We do digitization, reformatting  
of music, film, video  
collections for libraries and  
archives all over the globe.  
Today I just want to talk to you  
guys about understanding  
reformatting options and  
providing access to your  
collections.

I was just going to begin with a  
bit of an overview of the audio  
digitization workflow process  
here.

As you can see by the graph,  
this is just a visual  
representation of the  
digitization process that we go  
through here at the studio.

Mainly just going to cover  
details, audio in this course,  
given the timeframe but videos  
just as extensive and complex  
and I'd be happy to be a  
resource in the future for  
questions pertaining to that as  
well.

So I'm going to go through the  
particulars of audio  
digitization and practice.

Cover ring the standard, many  
metadata standards right now I  
think most of you are dealing  
with collections that you're  
trying to figure out what tools  
to use.

Also kind of what happens during  
the digitization process.

Just want to provide an overview  
for you.

So when we receive and process  
materials there are three  
different types of files that  
can be created to make up  
digital archival set.

They consist of the preservation master, the use and access copy and web accessible copy.

All of these depend on the situation of your institution and wants and needs for the project.

But we'll begin first by looking at preservation master.

A preservation master is the most important file to manage and as such it should be rarely accessed.

It's meant to provide a copy of the original but in digital form.

It refers -- typically -- the standard is 96 -- 24 bits which refers to the bit depth and sample rate for an audio file.

Sometimes that is done in 44-1-16 bit.

You can think of 200 hertz to the pulse code modulation in an audio file.

Which is the digital representation of sample analog signals.

So you can compare this to the DPI of a TIFF file.

The amplitude and it's similar to how TIFF file would document its range of colors.

Broadcast wave format and I'll be discussing the broadcast wave later in the course.

So key advantages of preservation master, we process it in to a broadcast wave format which is most widely used format.

It has a high resolution than 99% of sources.

File format for audio data it adheres to the EBU or European broadcasting union technical recommendations and standards.

It's better than most playback chains and derivatives can be easily created from this file. If you can think of it like a sound TIFF it's kind of a way to visualize what a B-wave file is. The preservation master comes with some difficulties as well. There's no standard storage medium.

The data types are often expensive to maintain because the files are larger so there are -- they are too big to house on a CD-ROM.

Online storage it requires ongoing maintenance and Internet delivery often impractical for this type of file.

It's just too large.

The advantages of having a preservation master over its derivative is that it will provide you with a basis that you can use your file.

Typical solution that we provide for some of these problems is that give 96-24 hard drive to digital library.

Requires enterprise level storage often you might need I.T. staff to help with that. And from there it can be migrated to hard drive when available or you can also put it on a gold CDR if it's small enough you can put it on a CD-ROM or LTO-3 data tape to keep it long term.

A main derivative of the preservation master is the use and access copy.

The key traits -- of use and access copy readily accessible in user friendly format will make CDs or DVDs to have on shelf in a library.

It's good enough to substitute if you lose the preservation master but important element to remember once you're pulling from a preservation master file you lose some of the information in the file.

Often information that you can't necessarily hear to the human ear, but because it's pulled from a larger file, it doesn't have the same breadth and depth to the sound.

Some key difficulties of use and access copy we'll look at CD-audio versus CD-ROM.

Now both of these mediums are -- can't last long term.

It will deteriorate in some form.

A CD-DA is a digital audio CD. It has a pure serial read so you can't reread it to correct errors even small transient errors.

Where as CD-ROM is more speck for based it's provided at data so it can be reread but it requires computers, software, particular OS system to retrieve the information.

CD-DAs are more widely playable but CD-ROM,are more reliably played.

Depending somewhat on the preservation master.

It will be provided on CD-DA for near universal playability.

And in video form we provide it on a DVD.

It's always important to have multiple copies you can put one copy on a gold CD-DA or one on a green.

These are different standards of the YD.

Often we'll put preservation



master on CD-R -- excuse me ha  
gold CD-R and the user and  
access copy on a green, gold  
CD-R is a bit of higher quality.  
Our third format of file is  
web-accessible copy.

It depends on rights of whether  
or not you're able to the RA and  
AAC you'll see on the power  
point here, RA refers to real  
audio which is a streaming  
format.

AAC is the advanced audio coding  
file which is standardized  
completion scheme for digital  
audio.

That was designed to be the  
successor railroad of MP3 as part  
of Mpeg two and four  
specifications.

MP-is used as a web accessible  
copy or WMA.

To have real audio, RA or AAC  
you have to have access to those  
rights.

It depends on your needs.

You might have restrictions to  
put things online beyond your  
institutional ability.

But sometimes you want to give  
further access to your material  
and a web accessible copy can  
provide that.

It won't be quite the same  
standard of sound as you would  
get in a preservation master or  
a user access copy.

But it can still provide  
different resource to your  
material.

One of the benefits I think of a  
web accessible copy is the  
accessibility obviously, says it  
in the name.

That's a large part of  
preservation.

Just access to materials.

I think I accidentally pressed something here.

Sorry, technical difficulties.

>> No worries.

>> Apologies, thank you.

There it is.

Just to summarize, preservation master used to be provided on analog process from -- to further format as things were developed.

But now subjects we're in the digital realm we stick with 96-24 preservation master file which can then be moved to DCD, ROM, hard drive for storage or LTO data tape.

A CD-R it can be then made in to CD-DA or CD-ROM depends on your choice.

And the resources that you have to give to your collection.

Ultimately it stands to point out that digital is not forever.

The idea of preservation is often thought of as the care taking of old archival material found in someone's dark attic that has to be retrieved from a back room.

But the practice of digital preservation is more largely understood in the field of that of providing access to information as well as history.

The American library association PARS definition of digital preservation, PARS is the -- is cohort of American library association specializes in preservation, I think I provided the link to their website at the end of the course or on resources page I believe.

It's preservation formatting section.

Digital preservation combines

policies, strategies and actions to ensure access to reformatted and born digital content regardless of the challenges of media failure and technological change.

The goal of digital preservation is the accurate rendering of authenticated content over time.

The accurate rendering of authenticated content over time, look at preservation in terms of access to content versus that of medium.

So audio visual media deteriorates rapidly.

Play back machines for most if not all in analog material have become obsolete.

It's considered to be dead.

It was not considered to be forever as a result we need to be conscious of the constant change and challenges that face the practice of digital preservation.

This uncertainty considers value of access to information documented whether expected or unexpected in recordings.

The catch is regardless of the challenges of media failure and technical change.

Makes migration a way of life.

What do we consider when discussing migration.

How frequent, how long will your materials last.

What determines when you should migrate.

Often it's due to format obvious lessens which can't be as big of a problem as we know how to move forward in formats.

But carrier obvious lessens is a big problem if you want find anything so what then is

obsolescence.

Consider a CD-R versus LTO tape.

A CDR has good option for user access copy but comes in to use in somebody's hands often.

An LTO tape is something that can last a little longer but you need some I.T. support, we term that as enterprise level storage.

Boils down to the ability that your institution has to support your media and also the limitations, can you handle having I.T. level storage or dealing with a bunch of drives on a shelf.

The term "Bunch of drives on a shelf" was coined at the Vermont folk life center.

It's chief and fast and familiar we can all use hard drive.

But the issue with hard drives is they die easily and easily erased.

LTO tape has high density, high resolution but it's I.T.

intensive.

Has short lifestyles and complex machine dependency.

Often need I.T. support staff to help with the back up of LTO tape.

C Ds are cheap-ish and widely available but there's lots ever handling to migrate and metadata there is none except for the label.

On an enterprise class hard drive it's a fast, preferred solution but it needs technical staff and can often be expensive This is a large server type hard drive that you could use for back up of your materials.

What is all of this say.

Well, I.T. is getting even

cheaper, ever more quickly.

Also gets obsolete.

Over multiple migrations you have to plan ahead for the life cycle costs.

Any future time to support the digital content because decisions you make today are governed -- governed by that future ability.

This is an important point.

If you can't support the migration choices you make now further down the road then you're unable to support the continued digital preservation of your material.

I think before I go in to standard sections it looks like we have a few questions that I'll try to answer for you guys. If there's any confusion.

>> Hi, Stephanie this is Laura. There does appear to be a few questions a think a lot is due to like in any field all of the acronyms that end up floating around.

When you were -- some of the questions related to that more just terminology.

One was talking about the preservation master you mentioned PCM.

And what does that stand for. PCM is.

>> Stands for post code modulation.

It's essentially the digital representation of sampled analog symbols.

It's the way that if you would look at a wave symbol.

Picture a wave symbol for audio signal the ups and the down, is that is post code modulation of that.

>> Other questions in that same vein.

Some folks wanted to know, I know Jenny did a little bit explanation in the chat box but what actually is LTO?

>> LTO is lineal tape open. It's actually magnetic tape data storage technology.

It was developed in the '90s and it was an alternative to proprietary formats.

It's used as especially for back up for -- we use it to back up our files essentially we have things on the server then we'll back it up to LTO tape.

It's a reliable back up option.

>> We also have some questions about your -- with the preservation master the term 96/24 people were looking for a little bit more clarification there as well.

>> Sure.

96/24 is the sample rate and bit depth referred to in an audio file.

The sample rate is essentially the frequency.

Defines the number of samples per unit of time taken from an audio signal, continuous sample. So the sample rate is the frequency.

Bit depth, that refers to -- both of these elements are part of post code modulation that I mentioned before.

But it's the number of bits of information in each sample.

So it corresponds to the resolution of each sample.

There's -- in this case this is 24 bits per sample.

If that's a clear set up a little more.

>> A few more questions for you then we'll let you get started again.

One question was, how do you convert other formats to the wave files, the person was asking, how do you convert, for example, WMA to a wave format. If there is an easy explanation for that.

Generally our transfer engineers would be doing that process.

You would use converter, particular computer program that would convert the WMA to the wave.

It's similar to how from analog to digital you have an analog to digital converter which is essentially a machine that you plug in that connects the analog to the digital.

Something that makes the files talk to each other.

I think there is sources that you can use for that process. I'm not positive all the ones that we use in the studio here.

>> I lied I do have one more question.

You were talking about your access copy, use copy, of course your preservation master.

The question that we had is when you're doing your migration or conversion which copy would you use in the future, is that why you're making the preservation copy as well?

>> The preservation master copy would serve as the go-to copy. Used access copies and web copies are made to provide access.

The preservation master copy is similar thing to taking your film and putting it in a freezer

to preserve it.

It's not a copy that you want to be using that is what you make the use and access copies for. They develop as derivatives the preservation master.

The preservation master aims to be the closest to the original, the idea of it as to provide a digital representation of the analog signal as close as possible to its original.

>> We have more questions but we'll hold those.

>> Okay, great.

I guess I'll just move on to highlighting some of the standards.

I realize that probably throwing quite a few terms your way and seeming to gloss over some things that standards and ideas are difficult to explain in just a couple of minutes.

As introduction I hope this can provide you with resources in a way to begin looking in to how to process material.

We also welcome questions or calls to the studio, happy to be a resource in the future.

Looking at the practice of digital preservation there are important standards that include BWAV files which is broadcast wave that we use for preservation master.

There is Bext and INFO chunks that are part of metadata process within a BWAV file.

ID3 tags are referred to metadata.

There's 57 guidelines this is the audio engineering society guidelines.

These are all things I'll be going through in depth as much



as I can.

And checksums can launch in.  
So the B WAV is a derivative of  
the WAV file it's a broadcast  
wave.

It was released in 1992 as part  
of windows 3.1.

It is part of the RIFF standard,  
which is Microsoft's resource  
interchange file that also had  
derivatives, you've probably  
heard of AVI files probably most  
common on this list.

In addition to RMI and RDI.

The B WAV we use at George blood  
audio and video.

All of our files are created  
from a single original capture  
file.

That specific file info is then  
gathered from filemaker pro  
database which is where we house  
the metadata then we use LINUX  
command line audio utilities to  
to -- excuse me, to read the  
information.

The original file is split in to  
these three different types of  
preservation master access copy  
or web copy that specifics of  
those that I briefly mentioned  
before the preservation masters  
often processed as a wave.

Provided in 96 kilohertz, 24  
bits with Bext metadata attached  
to it that is housed within the  
file.

That is provided for CD burning  
most often.

Some of the audio utilities that  
we use in digitization process  
include Sox, sound exchange.

It's considered termed as the  
Swiss Army knife of sound  
processing.

Used for sample rate and format  
conversion.

This could be something that could be used in the WMA to WAV the question that was asked earlier.

Then there is the Libsndfile. It contains a program that gives a lot of useful info about the files with Bext embedding you can check that out at Mega-nerd.com I provided these links on the resource page as well.

BWAV file have two mandatory WAV chunks.

The FMT it describes the content of a wave file.

This is speaking to the metadata, the Bext chunks that are attached to BWAV file.

The format chunk it includes descriptions of the format, the number of channels whether it's monoor stereo.

Sample rate and bit depth which were 96/24 as I referred to before.

The streaming info.

The data then there is no compression.

Many other things that you can include in the audio data portion, multi-channel format. 64 bit audio.

This includes information on the list, you see here.

Any new info field can be defined.

An application shouldn't ignore any chunk it doesn't understand so there are common register info fields like an artist comments, copyright, genre, name, some examples of these optional items.

It includes info about looping the wave form during play back.

It's useful in data is used in

samplers but rarely holds value in the preservation.

But PEAK audio is one of the file format programs that we use and inserts sample chunk in every wave file.

It's just included as part of that particular program.

It allows a quick expansion of any other header chunks and program called Wavlab inserts pad chunks in all saved WAV files.

one of the most important elements of the BWA is the

next file and vocabulary more are for most fields.

It includes the description, the originator reference, origination date, origination time reference and coding history.

It limits the data chunk to post code modulation or MPEG formats. Both of which I mentioned pulse code and modulation before, MPEG I can go over in a little bit so here is an example of a next chunk that we would show in an audio file the description with the name and song that is played.

The originate Cher is where it came from.

The origination reference which is the client.

The date and time, at the bottom is the coding history.

You can see at the beginning there's PCM which is the pulse code modulation here.

96/24, bit and sample rate.

We document whether we process it in stereo, dual mono, the machine that's used.

And the programs that are used as -- in conversion.

Moving on, obviously BCAF is a surrogate of the file.

There are some problems with normal WAV files which is why BCAF is created they have proprietary chunks which means, for example, PEAK is one of the only apps that would read the WAV chunk.

they don't always ignore the saw push husband chunk the info chunk or the junk chunk that I mentioned before.

Efforts to be taken to write the most basic WAV file because the simpler it is the more interoperable it will be.

There are some programs that exist to strip extraneous chunks from your WAV files.

After conversion.

WAVtrim removes superfluous chunks and SOX is command line application that does many audio utilities.

Can download that from sourceforge.net is a great resource.

I also see some conversations about BWF metaedit that you can download that from sourceforge.net as wellment we use BWF metaedit our quality assurance after a file is processed we look at the metadata from the original to the converted file and compare between the two.

A WAV file, it has a lack of real estate format extensible support.

It was part -- there was windows 2000 update to the Spec that supported higher sampling rates and greater bit depths with

multiple channel greater than stereo audio.

It's just -- best to avoid it if you can.

It's difficult to support it down the road.

B WAV, there is some problems in the implementation of it.

-- few pro audio applications that embed metadata, for example, peak 6, adobe audition and Wavlab are professional audio operations.

Peak 5 the old version or they can't embed this same type of metadata.

That is held in the BEXT chunk.

W wave is mostly geared toward broadcast applications and has some limits for info preservationists.

Honestly it's just best to keep it simple, avoid extensible formats and know your software, stick with similar versions.

I don't know if I should maybe pause again for few questions, it looks like there are large questions over there.

>> It would be great to address some of the questions again. Some are more technical while others have been more general sorts of questions.

An draw from South Carolina had asked, do you think that there is any alternatives to the WAV files for storing audio since they're just so huge?

And are you familiar with, I know I'm just going to say it FLAC style format.

And would you be able to comment on that.

>> Sure.

As far as preservation master it's really -- we women a WAV

given the amount of sonic information that is covered. I think that it's probably best to, if you don't have as much space that's what we provide access copies for but storage is becoming a lot more cheaper, hard drives and the like. It's actually not extremely difficult to store your files when they're in WAV format. I think part of it is prioritizing what you're preserving.

>> Some people I think their heads are spinning a little bit with all this information because it is a ton of information.

And I guess I'm wondering and also few other people were wondering, say that they don't feel comfortable using one of these converters or they might not have access to it for one reason or another, some people have asked if they can't convert things to WAV files is it something so important that they should send this out to vendor to do right away or is it something that as long as they're storing what they got, are they safe, should they worry

>> At the very least stick to standards that you know in the field keeping things in cool, dry environment with proper temperature and controls, if you can't digitize right away then you have to care for the AV material itself.

But the fact is that a lot of these materials are deteriorating and time is not something that's on our side. I think the sooner the bet Cher

is partly the idea of prioritizing some of your material.

>> Great.

The question, just wanting to confirm what she was hearing was right.

The question was, can you confirm that the two required WAV chunks are the FMT and the data and although BEXT is optional that you think that part is actually most important?

>> I would say -- the FMT and data chunks are considered just part of the metadata standard that are mandatory.

It's true that BEXT chunk is considered to be part of the definition of the fields.

I guess might be a little bit misrepresented here.

But I would recommend all three.

>> Did I have another question here I wanted to ask.

Are you going to touch on metadata a little bit more later because one of the questions was, do you have software that you would recommend for adding metadata to WAV files.

>> Sure.

I guess there might be a little bit of confusion then.

A BEXT chunk is referring to metadata fields.

Right now in standard broadcast WAV file theism discussing the chunks are referring to different fields of metadata that can be added to the file.

It's a way of breaking down how metadata is attached to an audio file.

>> Great.

Then one more question then I do see a lot of questions coming in

about video format as well.  
Maybe those are questions that we can reserve for the very end since you are saying that you're go to focus most low on audio here.

Just want to let people know that we're not missing their questions there.

The one question I did want to mention was about raid drives and what do you know of those potentially for storage.

>> We use a raid drive for some of our larger storage.

As I'm not the I.T. department at the studio I am not quite as knowledgeable about that el meant.

But we back things up to LTO tape and use raid drives for a large part of our storage since we have so many projects coming in and out.

>> I'll let you get back to it, certainly field more questions here at the end.

>> Okay.

So I guess moving forward ID3 tags is another type of metadata.

Similar to what I was mentioning the BEXT chunk.

ID3 tags are most widely compatible.

It has limits, though, I'm sorry Looks like -- pardon me.

There are several different versions of the tag implementation.

The metadata is embedded either at the beginning or end of MP3 files depending on the version.

It's not for certain types of files like WMA they each have their own tagging format.

This website is actually



incorrect I apologize on my resources page this website went down.

On the resources page you'll see connection to ID3 tags standards also source.

I'll explain a little bit more about what's involved with this type of metadata.

Before it's most widely compatible but it's least capable because it has limits of characters and size.

There's lots of flavors of ID3 tags.

There's version 2 which is the most capable but it's tricky to support.

Applications are not really responsible with this type of metadata often pry pry Terry reasons are involved -- proprietary reasons are involved so transferring between the material is a little difficult. 2 is the most popular version used.

It stores a tag at the beginning of the file.

2.4 hasn't caught on yet as a successor which stores at the end of the file.

I guess just to drive the point home, we worked with Princeton university on a project where we were creating master files and streaming files from lectures.

The metadata they sent us was incomplete so as we were processing the audio they asked us to get dates and info from the beginning of the lectures to import that data then edit the date fields.

When we sent them back the audio files it turns out that they were seeing something different

in the metadata.

The dates were changing.

The reason why that was happening we embedded the files with a program which embeds all four types of ID3 tags.

Different tools look at different versions of metadata by default.

So something that you have to be aware of which tags are part of a file when you're embedding, which tags are involved with the file when you're editing.

If you look at -- here is an example the difference ways that you can see ID tags in iTunes or real player just the way that the metadata is December played differently as you can see in iTunes they use version 2.3. Real player uses a different version.

Which displays the information differently.

You can take a look at the resources link at the end that provide on ID3 tags for more information.

Winamp is a windows program that can be used to view version one and two.

ID3 version two is a command line tool for writing extracting and reracing version one and two tags.

Another standard that is developed, was developed by audio engineering society is the AS57, 2011.

It was standard that was set out to develop a vocabulary to describe both digital and analog audio visual elements which uses an extensive mark up language.

Lengthing the document to the physical objects.

PB core is the -- itfuls  
organized asset of specific  
fields that can be used in  
database applications.  
It utilizes data model for media  
cataloging as management  
systems.

It was based on Dublin Core  
version 2.0 was released in  
2011, it provided free with  
creative commons licensing.  
Interesting application and used  
as a standard in the field.  
Provides metadata standards.

PB core offers different  
elements of metadata as you can  
see by represented in the map.  
I guess maybe before moving on  
to checksum it looks like there  
are some curiosities.

>> I understand how difficult  
it is to distill complex topic  
in to an hour and half.

Of course more questions are  
going to come up.  
I guess for those of us that  
might really end up, maybe not  
doing this work ourselves but  
outsourcing it to a company or  
work with our I.T. department or  
other things like that to do  
this sort of work.

Do you have any suggestions for  
resources that we can turn to  
how we can take a look at our  
phone collections trying to  
assess what we need to do some  
of the options for conversions  
to go, when speaking with a  
vendor with our I.T. department  
be more informed on that  
language.

>> I sound like I'm talk Ink  
the technical hierarchy of the  
digitization world.

But most vendors are really  
approachable, we often act as

resource just to answer questions of things in this process.

There is a website I was just at oral history conference there was project called oral history in the digital age.

That if you want to check that out online there are some resources for smaller institutions.

And ideas of what you can do in house.

I think mainly if you're interested in digitization process it depends on the machinery, need to have the resource of the machinery and essentially equipment that will convert your analog materials to a digital format.

A lot of the elements of what I'm talking about is essentially discussing the technical details of metadata that we add on to files that can be embedded in files.

I think that something that's possible to do, you just need to -- just learning about the process of it and having the amount of staff to devote to it. Essentially it takes time, everything that's analog has to be processed in realtime to transfer to digital format.

Our transfer engineers sit with the analog material from start to finish to convert it to digital format.

I'm not sure if that helps to answer the question.

>> Absolutely.

I think in addition to the resources and ideas that you gave in our chat box other people are offering up a ton of

great resources as well.

The other question along the same lines that I wanted to ask is also lot of what I'm hearing is that you need to be consistent in what you're doing and document what you're doing. Do you have any other recommended resources for people that might want to look at developing a policy or a plan in relation to reformatting. Maybe some samples that they can look to.

>> We do have on our website a resource page that provides introduction to audio digitization and the processing of that.

I think that -- I could provide more resources for the website for you guys after the course. I'm sure that there is extensive list.

>> That would be fantastic.

>> That may be easier than me listing them at the moment.

>> Really appreciate that.

>> One question that always comes up when we are talking about these sorts of materials and the reason that we're really having to do all this conversion is because things become obsolete so quickly.

Some people are saying that a little afraid to commit to one digitization format or technology then it becomes obsolete so quickly.

Is there anything that we can kind of do to kind of prepare ourselves for this or make it easier down the line.

>> I think the expectation is that -- the idea of digital is that technology and media will

change so being able to to roll with the changes is important migration becomes a way of life. You're going to have to take in to account that you will need to continue to convert files down the road.

I guess that's partly why we recommend keep things simple, keep with machines that process things in the correct way then using consistent software programs throughout.

Like as I was discussing with the ID3 tag versioning issue. Means something to be considered and I think migration is the main point except that migration is a part of preservation.

>> Which is so hard for us as preservation people to trially embrace.

We would love to just put things in a box have it be done not necessarily think about it as this continual process.

Mindset on all sides I think in some respect.

>> There's the idea, I mean, digital preservation is a bit of a different concept, digital is not for ever and as things are -- we look at analog material it has a particular life span.

There are some, for example, some glass disks that we've received that we're unable to process because they have deteriorated so badly.

I think that the idea of digitizing material to save it before it's gone.

And speaking to the content I guess a lot of the technical details of what I'm discussing is -- we take processes, complex

processes like this and transfer process to ensure that we're giving as true of representation as the original -- what was originally recorded.

Because we're not here to edit it or change it we are here to preserve it.

That's the idea of sticking with a lot of these standards that have been developed in the field and that we use here at the studio.

>> Great.

We have a little less than 20 minutes left.

I'll hold any more questions until we're at the end.

>> I guess moving forward to a Checksum formula, you can take a look what's on the slide here.

Looks extremely complex and bunch of mumbo jumbo this is auto sevenly what a Checksum algorithm looks like.

This isn't something we'd be passing on to your end.

But this is the technical data integrity of a file.

Essentially after we process -- after we take, for example, an analog reel and convert it through an analog to digital convert Cher is connected to the computer which is using the software program that is creating the digital file, the Checksum formula is essentially an algorithm that we run that compares the first file to the last file.

If you even -- we can use this mathematical equation as an example.

If you change one digit in the algorithm, this is a mathematical representation of

an algorithm if you change one digit in the value you can see how it changes the end result. I point to this because change of a single digit in a Checksum gives a very different value. It isn't subtle.

The probability of having -- there are 3.4 times, ten to the 348 possible values.

That can be created and Checksum algorithm.

The point of the Checksum is to compare the two.

It's just simply pure math behind the file.

The chance of two different files having the same Checksum is -- does not happen.

There are three different major types of Checksums.

There is an MD5 which is a message digest.

And SHA-1 and SHA-256 these are technical terms but it is 128 bit value that is converted to make it easier for humans to read at the bottom example you'll see that the complicated algorithm that I show before that is how it's represented in Hex with 32 place values.

When they're looking at a file after it has been transferred. We attribute a Checksum to that file to highlight uniqueness. We're looking at it in binary that would be what the Checksum algorithm would look like where as hex is much easier to read and compare.

If you were to change one value it would completely change the Checksum.

Basically that's saying, looking at the data integrity of a digital file if any element is



changed in that file it will be a new unique file.

The Checksum is a way to highlight that this is the one and only version of this file. So I guess some conclusions, I would just say that it's important to note that there's really no magic solutions, vendors are here to help and we're all in this field trying to preserve material together so the idea of digital preservation we use establish tools, upgrade cautiously and just realize that every solution is temporary but it's something that steps that often have to be taken in order so that we don't lose the content that is on analog. I know that was overview of a lot of really technical things that sounded complicated and confusing.

I'd be happy to answer more questions.

>> Sure.

Always glad when there's time for a few more questions. Try to keep them a little bit more general.

We did have some very specific questions in there as well. One of them, we were talking about audio throughout all of this we did get a lot of questions about are there standards for video preservation and where might some folks turn for that although that might be another opportunity to add more links on the website later. If there is any insight you have to that I'm sure people would appreciate it.

>> Sure.

I can provide more resources

online looking at video preservation as well.

Right now I think standards for video are still being developed at the moment, George blood, the president of our company looking -- working at the library of congress to write a white paper on digitization standards for video preservation which highlights essentially a lot of standards that I'm talking about are technical details of the back end of how we're processing things to ensure adheres as closely to the original representation of that material.

Standards for video preservation are quite extensive they're still being developed.

I guess we're learning about this process as we go and library of congress would be good place to start.

They have a lot of resources on their website then I will certainly provide more links to an overview for video preservation and some different resources that you can look online.

It might be easier just to provide a list of links after the webinar for people to reference on their own time.

>> Great.

I think Marcia in northern California had really interesting and good question. Maybe a bit more theoretical in nature but she was asking for what your opinion is, what do you see as the next method for preservation after digitals there something coming up that we should all be on our toes

for?

>> Oh, Lord.

[ Laughter ]

I couldn't tell you.

I mean, at the moment I think we're just looking use the resources that we have with the amount of knowledge that we have and share the knowledge that we have.

Within the community, I think. Large part of preservation is about access and about the conversation especially being able to use this for, I can see throughout the entire time that I've been presenting people are exchanging ideas and links and helpful processes of how things are done at their particular institution.

I think those ideas, the more that we talk to each other in that sense the better we'll be able to exchange these ideas and help support preserving intellectual material.

>> Great.

Marsha just commented again that she's heard of DNA computers as well as light beams which I just can't even wrap my head around.

Those sorts of things.

That is really all the questions that we have right now.

Can throw things back to you, Jenny.

>> Sure.

Let me put up the assignment.

If you were feeling a little lost don't worry, it should be pretty easy to fill out.

Let's see, I have link to the page as always if you're watching in a group just ask that your group leader write down everybody that is watching

so that we have better idea of everyone who is with us today. And it looks like I'm going to give just few minutes for our group folks to post their names just got few more seconds to throw in any last minute questions for Stephanie. We'll just give that a few minutes.

>> I'd also, I can I think my contact information online, too? I would gladly accept questions over e-mail if people want to ask things about specific projects that they have or want help with ideas.

For resources of things. After providing the links online.

Also people can feel free to contact me here at the studio and we also have lots of very intelligent transfer engineers that work with these technical processes every day.

>> It does look like we have one last minute question from Catherine.

She is curious, any thoughts on AIFF versus WAV files for audio files.

>> I think that to be honest I'm not sure how to answer that. I suppose AIFF I think potentially is a proprietary format.

And WAV is more adaptable.

I think AIFF which I think you can run in to some issues when you have file format that is proprietary.

Just in terms of transfer conversions but using things between different computers.

>> Okay.

In addition folks can feel free

to e-mail I've posted my e-mail  
we can forward your questions  
along to Stephanie if anything  
comes up.

Again keep an eye on the  
website, I've been tracking all  
of your fantastic links and  
Stephanie has more links to give  
me for resources so do go back  
to the webpage I'll attempt to  
add those as quickly as  
possible.

It looks like we are done.  
Again the deadline for all that  
homework assignment is one week  
from today, November 6.

Thank you all so much for  
logging in and coming with us  
over the past couple of weeks.  
I have to thank all of our  
speakers who aren't necessarily  
here right now, you were all  
fantastic and Stephanie, you  
were fantastic as well.

>> Thank you very much for  
having me.

I appreciate you all attending  
my session.

>> Also CCHA thank you guys so  
much, thank you Laura for being  
on board to help us with all  
those questions.

>> Great it's been a pleasure  
working with you.

>> All right, everyone, have a  
fantastic afternoon and happy  
Halloween.