

It's 11:00 AM on the East Coast, United States,
so we're going to begin today.

This is a webinar called, "Getting Practical & Tactical.
Digitizing court reporting with new processes."

Thank you so much for joining us.

Just remind you guys that we do have
the Chat window or the Q&A box open.

Please leave your questions,
and we will try to get to them at the end.

If we cannot get to them,
please email us, and we will answer
them as soon as we can to you guys.

I'm moving on.

My name is Benjamin Jaffe.

I'm the Digital Training and
Development Manager at BlueLedge.

Today, I'm joined by Tony Sirna,
Senior Customer Success Manager and
Legal Strategist for Verbit and Lisa Dees,
Program Manager for Justice AV Solutions.

Thank you both for joining us today.

Thanks for having us.

Thank you, Ben.

Thank you. This is our agenda today.

We're going to start with our welcome
and introduction. What we're doing now.

We're going to move on to technology

used at the deposition or at the proceeding.

Then we're going to look at the workflow
after the proceeding.

What is similar and what is different to
stenographic and other methodologies
of court reporting, and
how does that work flow work inside of
an agency or a courthouse with this newer technology.

Then we're going to look at transcription tools.

Tools that will help, so be it ASR or other tools.

Then we're going to take time for
live audience questioning and answering at the end.

As I already said, please add it to the QA box.

We will be sending out a
survey/quiz at the end of this within
48 hours to allow you to be
eligible for AAERT CEU credits,
so look for that email.

We will be asking for those
back in a timely manner so that we
can issue you your certificate
and you can get your credit.

Thank you again for getting those back to us timely.

Starting off today, we wanted to really
discuss what is the role of
a court reporter and what is
their job and to think

about this in the most basic terms.

To start, it's the administering the oath and the acting of that neutral third party in the room.

The person who is taking down the record and isn't subjective to what's happening in the case.

They basically manage the proceeding to get that clear record.

They act with ethics and professionalism to preserve the security of that record.

They want to make sure that what is said is what gets down on paper, and they want to capture that record and preserve the chain of custody. Who said it?

Who gave the exhibits?

How did they ultimately get to the courthouse?

How did they ultimately get transcribed? Then they collect and deliver that record either to the agency or the court system, or in some cases, directly to the client to move it forward in the process.

Then the second thing is what does the client really care about?

By the client, I mean the judge.

I mean the attorneys who are buying the transcript.

I mean the courthouse. I mean the agency.

They really care about a high quality, accurate, and timely record.

What I mean by that is,

is it formatted correctly?

Are the words accurate?

Does it match grammar and spelling correctly?

Is it delivered in

the time frame that that client expects?

That's really what they care about.

They don't care the methodology at which it got to there,

as long as those three things happen.

Now, how to make that happen.

Well, technology used at

the proceeding really affects

the quality of the transcript in the end.

Lisa is going to guide us

through some of that technology, and I'm going

to work with her on some of these next slides. So, Lisa.

All right. We're going to cover just a general overview,

if you will, on some of

the important key factors of

capturing that quality record.

We're going to speak first about in-person proceedings.

The ones that we get in our car, and we go and we set up

our kits, and we're all together in the same room.

There some essential basics

that every court reporter

should have in their reporting kit.

We're talking about digital court reporting

here, obviously.

The right microphones for the right environment,
a multi-channel interface, interface
the microphones into the laptop or computer,
a backup recorder of some sort,
multi-channel recording software,
audio monitoring devices,
and don't overlook power management.

What I'd like to do is go into
a little bit more detail on each one of these.

I could speak to you for about an hour,
maybe 10 hours on microphones.

There's a lot to know about microphones.

I think the first takeaway

I'd like to share with everybody
is know what you're
purchasing when it comes to microphones.

Know that you have the right type of
microphone for the right environment.

Types of microphones start
out with dynamic and condenser. A
condenser being an electric-charged microphone, and
a dynamic is that microphone that does
not require phantom power that we hear about.

Know your pickup patterns.

There are many, many types of pickup patterns.

The two basic are omnidirectional and cardioid.

Omni meaning all is going to pick up all the way around the microphone, which is not always the best when we're looking for channel separation.

Then there are a dozen different variations of cardioid.

Cardioid meaning it's a little bit more of a defined pick up with cancellation where you don't want to pick up sound.

There are many styles of microphones.

You have gooseneck microphones that are going to protrude up from the desk and actually face the speaker.

You have boundary-type microphones that are going to sit flat on a table top, handheld microphones, lavalier microphones.

I'm telling you these just to give you an idea the things that you need to know and understand, and you can always contact a professional who can help you with it.

Remember that this is going to be in a portable kit most likely, so the weight of the microphone is definitely something to take into consideration.

The type of cables. You have XLR cables.

You can use Ethernet cables.

It could be 3.5 millimeter cables.

Make sure you're choosing something of quality,

something that's going to handle

being tossed around in a kit.

Again, know what options are out there.

Yeah. I'll agree with you a lot, Lisa,

and then I'll say that it's not always price determinant,

but definitely when you

cheap out, especially on your microphones,

that becomes evident in the record.

You want to make sure that you're buying things that are

in a professional level quality of the microphone,

not getting it in

that lower tier that is made for the consumer.

We are professionals, and we

need professional level equipment.

Agreed completely.

The multi-channel interface is the device we

use in between the microphones and the computer,

and this is where we will plug

our multiple microphones into, and it will

actually deliver it in a format that the computer can

take and then record into our multi-channel recording.

Really look for an interface

that's a minimum of four mic inputs.

The example we have in the picture here

uses Ethernet to connect the microphones.

This is a six channel.

Five to six mics really seems to

be the most common used in the field,
but definitely a four mic minimum.

You want to find something that's durable.

You're going to be throwing this in and out of a kit,
in and out of your trunk.

If you get something that's made out of plastic,
it's not going to hold up.

It's going to be a disposable device.

So look for something that's lightweight,
but made out of metal,
little more durable.

I will tell you from experience,
my opinion is that these interfaces that are completely
controlled by dials on the front can
be dangerous because as you're moving it around,
those dials get changed, buttons get pushed.

So if you can find one that it is software
driven rather than a hard-dial driven,

I think you'll get a better result in the field.

This is not a recorder.

This is just an interface.

There are multi-channel recorders out there.

Hard appliances that don't
require the interface of a laptop or computer.

So that's not what we're talking about here.

We're actually talking about
a multi-channel interface to the computer.

Yes, and those multi-channel recorders certainly can connect and they can be interfaces, but they aren't necessarily.

If you're looking at one of those, I would urge somebody to really take a deep dive in it and check how many of the channels or how it passes the audio to the computer system.

Many of the ones I've looked at say they're at 4, 6, 8 channel, but that's the recorder.

There are oftentimes only two channel on the interface, or sometimes they cannot record and interface at the same time.

So I just very much urge people to really do their homework on their equipment, especially such an essential device as this.

Backup recorders.

AAERT best practices says that we should have redundant recorders, and that means two independent recording systems.

We typically refer to them as our primary recorder and our backup recorder.

You want to look for a device that has sufficient storage.

Some of them can have limitations to the size of the cards that can be put in them.

You want to make sure that you're not going to run out

of recording space in the middle of a hearing.

You're looking for sufficient running time.

What I mean there is

if this device will only run on

batteries, do your testing,

do your due diligence, and know how long of up time,

recording time you're going to get on that device.

On the batteries, makes sure you're

starting with fresh batteries,

and really and truly many of them are coming with

an alternative power source to run them.

I've had several that'll run power

right off the USB cable connected to my laptop.

For larger rooms,

consider having multiple devices.

Set one up on the judge's bench;

set one down on a podium so that

you're adequately recording your backup.

Again, this is independent from our primary recording,

so it doesn't necessarily have

the advantage of the multiple microphones

that we've placed in the room.

I think you made a great point there about

battery life and testing your devices.

I would do that with all your devices,

be it your computer,

your backup recorder, your interfaces

if it has a battery.

Anything that has a battery,

you should really test how long will it last

so you have the best understanding

of what your limitations are if you

do get stuck without power or if the power

goes out and you have to continue your proceeding.

I've been in rooms where the power

just goes out, and the attorneys want to continue.

That's important that you

have the ability to do that.

Can your laptop, can your microphones,

can all of that keep going,

or can you make a quick switch

to enable that to keep going?

Very good point. Software. Your software, it

needs to be a multi-channel recording software.

It needs to be able to interface with

as many recording channels as

your interface device is

going to deliver to the computer.

Basics in the recording software,

you're looking for integrated, time-stamped note-taking.

Why is this important?

Because it's our job as a reporter to give the

transcriber as much information as we can.

Designate your speakers, what they're saying.

Be prepared to do a playback.

The more detailed our notes are

that we take, the faster we can find a playback point.

Make sure that it's got plenty of visual cues that

remind you and confirm

that you're recording throughout the process.

It should be easy to use.

Note-taking should be a smooth, streamless process.

It should be as easy as writing an email.

Confidence monitoring, so important to the quality of

this recording that we're handing off to transcribers.

It's our job to ensure that we are taking

down the absolute best quality recording we can.

A stellar recording.

The only way we can do that is to

actually monitor not what

the mics are picking up necessarily,

but what is being recorded.

So confidence monitoring is going to allow us to listen

to the recorded audio a couple of seconds delayed

after it's been written to disk

so that we know the quality of the audio is good,

and if we have to interfere at

any point to ask somebody to speak up,

we can do it quickly. The hotkeys.

We don't want to be weighted down with a lot of typing.

Make sure it's got a tool like

hotkeys or Quick Notes that Notewise has here so that we can do keyboard shortcuts, put in an entire paragraph with one stroke.

Not everybody gets to choose their software.

Sometimes your agency or your courthouse is going to already have software. They're going to have a contract. They may want you to employ their software because that's streamlined with their system.

So if you are given a software or you have to become part of a group that already has an established software, make sure you learn it.

Read that manual and learn it deeply. Learn all of the functionalities, how to troubleshoot it.

Practice, practice, practice.

Yes, exactly. Practice.

I always say learn to troubleshoot.

Now, some programs, if you just turn them off and turn them back on, they'll have reset. Others, you need to restart your computer.

Every program has their own unique way it was coded and unique way it operates.

So please learn it. Read the manuals. Ask for help. Watch the videos. Take the trainings.

They're all up there online.

JAVS does a great job putting videos,

but all the companies out there put

their videos out so you can watch them and

learn what is out there

so you can understand that software deeper and better.

Yeah. Become a mini expert.

Audio monitoring.

We're going to talk about two different sources

for the audio monitoring.

You need a good quality set of headphones or earbuds

or the style that you're most

comfortable with for your confidence monitoring.

We were having a conversation the other

day, noise-canceling headphones,

probably not the best choice for this situation

because as the reporter,

we're having to monitor

the live room and

our confidence monitoring at the same time.

But the quality of the headphones can definitely

have a direct correlation to

the quality of your recording.

If your headphones are of poor quality,

then you're not hearing

the recorded audio in its true version,

and you're going to be making adjustments to the audio

that's actually going to ruin your recording.

So a good set of headphones, and they

don't have to cost \$200.

A \$30 set of

Skullcandy's are going to give you excellent quality.

Something that's overlooked most often are speakers.

If we're a mobile reporter

and we're asked to do a playback,

do we really want to have to rely

on the speakers that are on our laptop?

Because usually, they don't have a lot of

amplification, and they're not the best quality.

So you might want to consider a nice USB

attached speaker so that you can

project your playback through the room.

As Lisa mentioned about noise canceling,

that is different than sound isolation.

If your headphones have good sound isolation,

you see in the picture here,

all of the padding,

the separation between the ear

and where the speaker inside the headphones are,

that's isolation. That's blocking the room out

versus doing the mechanical way

which is actually recording from a mic,

and it's basically producing

the inverse of that sound back

into the headphones to noise-cancel it.

Noise canceling can cancel what is in the room that you are recording versus isolation that's giving you the true pickup of what your microphones hear.

So I encourage people to look for isolation, not for noise canceling.

We're in total agreement on that one.

Something overlooked quite often is a power management system.

Part of our job as a reporter is to not become an intrusion to that room.

So we need to look for trip hazards.

We're laying out mic cables.

The last thing we want to do is have half a dozen power cords plugged into an outlet behind us, creating not just a trip hazard but a trip wall.

Just bring a power strip with you.

Look for one that's a surge protector.

You're going to be investing money in your equipment.

A simple investment in a surge protector will help.

Most of them come with On and Off buttons.

Make sure that button is lighted.

You want to be able to look down instantly and know if somebody has turned off your power strip.

Most of them these days are coming
with optional USB ports.

I know Ben has some more to say about that,
but definitely look for a USB port.

We're always having to charge your phone.

It could be powering your backup recorder.

She said USB power is becoming
more and more important on these devices,
so don't buy your cheap convenience-store power adapter.

Those can fail at any point.

I'm sure we've all been in the car and tried to
plug something in, and oh, it's not working.

Okay. I'll go spend three more dollars
or five more dollars,
but that's not what this is.

We are professionals.

We need you to get the professional quality products.

The other thing that Lisa talked
about with the power strip is,
either get one with a nice long cable
or bring an extension cord with you.

Because during my days as a videographer,
sometimes you would be across
the room from the nearest power outlet
because that's the attorney's seat,
and you're in his office,
and he wants to sit at that particular seat, and

you don't sit near the power.

So you have to be amenable

and come into the room prepared, so power.

Then, for trip hazards,

have gaffer tape. Be able to tape it down if you need to.

It is not the same thing as duct tape.

Please, please, please

do not put duct tape down on people's nice carpet.

The glue will stick.

You'll never get it up.

It'll be a bad scene.

Gaffer tape has a special glue in it that releases,

so please, please, please. I know it's an investment.

I know the rolls are more expensive,

but that is the proper tape.

Throw one in your bag and always have one.

Again, like we said at the beginning,

these are just basic

minimums that we're talking to you about.

This is not the only thing you're

going to have in your kit.

Let's talk for a minute about remote proceedings.

These have become very popular.

In fact, I think most proceedings are happening

using some type of

video-conferencing technology these days,

and I hope everybody

is having a good experience with them.

I've been hearing a lot of really positive feedback.

There's a minimum set of requirements for the hardware that we need to look at for this as well, and we'll touch on it here slightly.

We're going to talk about the importance of a headset, your microphone.

You still need to have a backup recording, because that's a AAERT's best practices.

You still want to use your multi-channel software, and you will need an additional speaker.

Microphones. The requirement here, a little different.

If you already have a multi-channel audio interface and you've invested money in that and good quality microphones, you can use it for this.

You're only going to need the one microphone because you only need to mic yourself.

But there's no need for you to go out and buy another microphone unless your microphone doesn't have a Mute button.

You definitely going to want to Mute button.

You don't want to interrupt or be an intrusion into the proceeding with them hearing the clicking of your keyboard, your dog barking in the background, so it needs to have a Mute button.

If yours did not come equipped with the Mute button,
then a simple Amazon purchase, 20-\$25 USB microphone.

They've actually been good quality.

I've checked several of them,
and the sound is good on them.

Know your laptop, what ports you have to connect to for
the microphone, and make sure that
your microphone's connection type
is going to match what you have available.

You've got your microphone in
your standard 3.5 millimeter.

You may need a USB port.

There are headphone-microphone combinations,
although usually they don't have a Mute button.

But know your equipment, and know
your laptop, and know what they need to work together.

Agreed, Lisa.

Backup recording. We have a few more options here.

Of course, we can use our handheld recorder.

We can hold it up to a speaker.

Know that you're going to still get the clicking of
your keyboard recorded with everybody's voices.

You could do a phone-based recording.

There are many phone soft applications that will
allow you to call into a phone number on
a video-conference platform rather than just connecting.

So this would be in addition to connecting via

the video, and let it record.

To be honest, most of these video-conferencing platforms allow for recording themselves.

Let it record.

You need your redundancy.

We need to maintain those best practices.

If need be, that recording can be used during the transcription process.

It's a great referral.

If something happened with your primary recording, the sound wasn't quite there, you could always refer back to your backup recording.

Yeah, and I would just urge that you make sure you're using a separate connection method.

If you are connecting via the Internet, let's say, with your laptop, then dial in from your phone or use your cellular Internet.

Because if your power goes out, the attorneys on the other side aren't going to know that both of your connections are lost.

They may see that you've gone away, but they're not going to realize to stop the proceeding.

So it's great to have two methods of connecting so that if one goes down, you're still connected.

You can say, excuse me,

I need a second. Because that's very important.

It's why the redundancy is

such a strong topic with AAERT,

and we're talking about in this presentation today.

Recording software. I highly recommend that we

maintain a multi-channel format with Notewise,

because we can direct connect to

the video-conferencing platform to take in the audio,

and it's also mixing in my mutable microphone,

and getting that added security that

comes with that multi-channel recording platform.

It's the technology that we use

whether it be in a proprietary audio format file;

that is a layer of security right there,

or if it's the technology used to encrypt

the multi-channel format within

a non-proprietary file system.

Again, it's an added layer of security.

Direct connect to the video conferencing, if you can, and

it is so important.

I know I touched on it,

but you need to include your audio.

Your audio is a part of the record,

so when you're swearing in a deponent or you

need to interrupt or they're asking for a playback,

you need to be a part of that record as well.

Anything else that you want to add, Ben?

I just wanted to say, as Lisa said earlier,
become a mini expert on your software, on your platform.

Each one connects slightly differently.

Talk to your company because most of them,
either before the pandemic or since,
have come up with a technological way to do this
that is superior than air
gapping it or doing it with a phone and a microphone.

There are new ways that this have been done
that are superior to that, shh.

Everybody in the room, don't talk.

I got my microphone and my phone.

So really become an expert on your software.

I can't advocate that enough.

Agree.

Now we're going to talk about at the proceeding.

Doesn't matter what preceding your in,
it's about the environment.

If you're in a meeting room, you're in your office,
you're remote, it's about controlling your environment.

Don't do this at a public park
where kids can be running around or in your dining room
where maybe your dog might bark.

If you're going to be a professional,
be in a professional environment.

Now, you can't always control
your environment like a courtroom,

so you have to understand how to mitigate the issues.

If you're in a big echoey room,
you have to move your microphones closer.

If you're in a remote office like this,
maybe a headset will isolate you
from your family or from others,
who may be speaking.

I know we all have others in our houses these days,
and so you have to come up with tools to do that.

Remember, the digital reporter is
really there for proceeding management,
chain of custody, and they are
that secure conduit to transcription.

Making sure that that audio record gets transcribed,
gets to the agency,
gets to the client,
and is delivered to the client in
the way that client wants it to be.

As we spoke about earlier, high quality,
reliable, accurate, timely transcript.

As a digital court reporter,
it's good to be trained.

Either you can be trained through
a mentorship program or through formal training.

BlueLedge, who I work for,
we do formal training.

We just did an agreement with a company called ed2go,

and now we're training through their network of more than 2,000 colleges and universities so you can take classes on digital court reporting where you are, across the country or across the world.

There's also the need for experiential training, so it's great to take a class especially one that is AAERT approved. They have a whole list of approved programs.

But getting that experience in the field, I talk about it a lot.

Like if you go to school to be an accountant, you get a business degree.

Okay, I'm going to go work for an accounting firm.

You don't get those \$10 million tax returns right off the way.

You might do the most simple.

You just EasyFile.

Let me fill out the form for you. Or you're looking over someone else's shoulder,

Hey. Do you mind if I watch when you do that \$10 million tax return?

I want to understand what is different.

That's experiential training.

That's very important, seeing what a proceeding is like, and then making sure you're adapting to the remote environment that we're in today.

It seems like this is going to last even when we get back to normal. I think there's going to be a new normal of remote mixed in with real proceedings. So getting training, working with your companies, understanding your product, that's all really important. That also leads to certification. Making sure that you are representing yourself with those letters after your name. You're recognized by court reporting agencies, courthouse, and government entities as a profession in your field. You're committed to professionalism. You're committed to the ethics and the standards of the industry. AAERT's certification for digital court reporters and electronic or legal transcription is very similar to the NCRA or NVRA transcriptions for their court reporters, except for that reporting and transcribing is two different certifications, whereas for them, it is a singular pathway. But that doesn't make it any less relevant and important to obtain if you're going to be doing those functions. Moving on. Now I'm going to hand over to Tony to speak about the digital transcription

after the deposition and the workflow.

Thank you, Ben, and thank you, Lisa,
for that session we just completed.

This next session is going to primarily focus
on digital transcription process,
both in the differences between
that and traditional stenographic means
and also take a look at some of
the assistive technologies that are used on that,
including automatic speech recognition.

The goal of that is to provide
a more general understanding of how ASR
works and particularly how it
fits in legal transcription.

There is some misunderstanding out there,
primarily because I don't think
there's enough education on it,
and so that's the goal of this section.

We're going to first start with the next slide; Ben,
if you could change it, which is essentially this,
though the processes may be somewhat different,
at the end of the day,
the modality of delivering
a transcript is going
to be still about the record integrity.

That applies regardless of modality,
whether it's voice writing,

stenographic, or electronic court reporting

and that is the focus of this.

Lisa's presentation was focused on how you

ensure the integrity of

the capture of that record and the best way to do it;

the subtleties of even using the right microphone so

you don't mix something down.

This will be on how you get to that integrity using

technology tools along with

the people who do the transcription.

This next slide is going to be

just a quick overview of what

the steps are across the board; Lisa, Ben,

and I have went over this earlier,

and some of the differences between

a stenographic method and the digital reporter, so just

so you understand there is

a difference and there are some distinct phases.

In the middle of the presentation from left to right,

you'll see a group of arrows,

capture step through sign off and delivery.

The big difference here is that it's

the roles of who's doing what that changes.

It's still the same thing.

You will have a scoping processor,

a proof process. You will have

an initial rough draft that

your transcriptioners will work on.

It has to be quality-controlled
and certified and assembled
with exhibit pages and indexes.

Then finally delivered to the law firm.

Now, above that graphic is the stenographic message,

which you'll notice is the capture phase

or the taking-down phase,

which Lisa talked about in

digital, is basically using

keystrokes representations of words, as stenographers do,

and also utilizing CAT software to basically translate

that keystroke symbolism into readable English.

That's controlled by a stenographic professional.

From there it'll go through a

scoping process or go to a scopist.

This is a person who takes that raw,

rough draft and brings it up to

that next level of quality.

Once that done, an ASCII file or a text file,

usually used with production software,

is passed onto a proofer

who does exactly that,

does the proof, the final edit,

the QC to make sure that

the document is at a very high level of quality.

From there it can be sent to an agency.

Agency will do some quality control internally.

They may do some of

the production around the different title pages,

add the certification to

that with the reporter's signature,

create the index, package

it up, and deliver it to the customer.

Essentially, that's what we've

been doing for quite some time.

Below that, you have the digital reporting message.

The big difference here, if you just simply

look at capture and rough transcript,

just who's actually doing it and how it's being done.

Instead of taking down,

as you would it stenographically,

you're capturing either video

or audio of the proceeding itself

using tools Lisa talked about

with the same goal of getting the verbatim down.

That's going to be your electronic court reporter.

Your digital reporter in a courtroom,

it'll be a courtroom monitor.

That person will create

and the output, which is a media file,

they will add the annotations,

the notes, the various events,

the exhibit markings through the software,

but that's their role.

From there the file will then go most likely back to the agency for utilization and transcription.

Now, with an ASR model,

the raw media file is

put and uploaded through a ASR software;

we'll get into that in more detail,

which then outputs a rough transcript.

A rough transcript in the sense that it still has to go to professionals, and it still has to be scoped.

There has to be a proof level or an edit level.

The terminology may be different. In digital, it's legal transcriptionist and proofer or editor.

It could be the agency.

It can be a vendor you're using, but

it's a complete digital media file that it's worked on.

From there it would

go back to the agency as an ASCII file or Word file, RTF, whatever the agency is

using to create the final transcript in its production, and the same exact thing happens.

A quality check has to be certified.

It'll have the reporter's signature.

They'll stamp on there and put it into a consumable format.

That can be a PDF document with your exhibit pages, beginning, end pages, index list.

However the agency chooses to serve
its product, or what the client needs,
and deliver it either electronically in PDF,
delivered online,
and separately and/or with synced video for example.

Some people request this,
particularly as they prepare to go to trial,
with tools like Summation.

That is really the difference between the two process.

So while the steps the same,
it's really the roles,
it's how it's being done,
but fundamentally they do have similar processes to them.

Most important to know is
your media files are going to be MP3 files;
they're going to be M4A;
it could be For The Record,
which is a specific software utilized in a lot of courts;
it could be Lisa's job system.

So there are numerous formats that can be
supported with multichannel or
sometimes just single channel,
depending on how it's being recorded.

This is just to give you the high level of it.

Now, let's talk a little bit about
that second phase of getting that raw, rough transcript.

We'll explain the similarities with stenography later,

but essentially it's that initial version of the transcript you get after you take it down and you capture it.

Legal speech recognition has some very unique aspects to it.

What is the application that you're going to use it for?

When we often think of speech recognition, we're used to what we use every day, which is Siri or Alexa or even Google.

They can be frustrating at times, but they are a form of speech recognition designed for a specific purpose.

Legal dictation which lawyers may use, is going to be a different type of speech recognition than what you would use to transcribe audio files, so it's important to understand the differences.

Word accuracy is important, but also context is key, and we'll get into that as it relates to automatic speech recognition, is that legal is a specific domain with a specific lexicon.

How things are said matter. Words obviously matter.

So a speech engine is also going to try not only get the basic words: he, she, they, the, but also what's the best contextual words used based upon the context of what the preceding is.

Another aspect of ASR that I think people aren't aware of is, as with stenography, you will have a general dictionary for a job, which is all terms you need, and then a specific dictionary, and Lisa will get into this later, that's used for this specific deposition or hearing. It could be proper names, things of that nature. Similar in automatic speech recognition, the software has a general dictionary, so to speak, of all terms, and then through what's called adoption or adapting or learning, can understand new words through contexts. Very simple, a deposition notice may have an attorney's name; you may have a speaker list. So similar to the stenographic, what the ASR software will do is be able to learn and use new words and have a specific model for your deposition versus all depositions. So we're going to cover a couple of these points now. Ben. The integrity of the record and the quality of the transcript in the digital process starts with the quality of the audio, and it finishes with the professionalism

and expertise of the people.

It's not an exclusive or all technology or all people,
but it starts with those two elements of that.

In ASR, a multi-channel file type,
like Lisa talked about,
is very good because you can isolate speakers easier.

Also important is very good backup.

Audio backup is just standard practice,
but it's also used by the ASR and by the people who
do the transcribing to
further narrow down on the speech said.

So transcription with ASR tools,
a couple of things to understand.

The technology is 24 hours a day,
seven days a week,
365 days a year.

It's a machine. It's not a human.

However, it's not perfect, and it improves continuously.

You'll see industry estimates for
word accuracy is at about 90 percent, maybe less.

But again, that depends on the type
of speech you're utilizing.

It's raw output. I was reading a study,
basic human error in understanding and
writing is about 4-5 percent. Very good
transcribers have that range.

Machine, that was about 12 percent.

That's just a statistic we know that's going to happen.

The difference is the machine can process

16 million words in

an eight-hour period and get you

the volume and scale to get to that rough draft.

Processing time. Why is this important?

ASR processes segments of audio.

That's called parallel processing.

So instead of processing

a whole audio file from start to finish,

a five-hour deposition, it's just taking

chunks of that to break it

down to process more efficiently.

Roughly, five minutes of audio can be

converted to raw ASR output in 2-3 minutes.

Again, it's the parallel processing

and the power of the computing itself.

A couple of other aspects

to talk about are speaker diarization.

ASR engines, this is very specific to legal,

looks to separate speakers

to create understandable sentences,

so speaker diarizations.

That's done through voice recognition

or pattern recognition,

so the person who said this line is

distinct from the person who said that line.

So for the transcriptionist later on,
that means this line is The Court,
that line is by Mr. Smith,
by Mr. Johnson, or The Witness.

Short answers: Q/A, yes, or no,
believe it or not, that's difficult
and a challenging mathematical model.

But the ASR also will break it up to short answers.

There's a number of different things
that legal transcription

has that it's trying to do to get the words right,
correct utterances and who that was
assigned to and also punctuation.

Because if it understands the punctuation,
it can complete the full sentence. Next slide.

Couple of things, and I think
this is about all we need to know in
the industry for non-technicians.

I'm not a technician.

I'm a lay person in this industry.

But there's two different types of speech recognition.

My speech technologists was
disagreeing with me a little bit

this morning, our developer,
because they have a different way in looking at it,
but this actually comes from a very good article

I can send you a reference to.

You have what's called directed dialogue,
automatic speech recognition, and
natural language speech recognition.

Very simple.

We use directed dialogue speech recognition everyday.

You call your bank. You do automated banking.

The automatic banker says, what can I help you with?

The computer states that
press one or say access my account.

That's directed dialog.

It's a complex model,
but it's a simpler form of dialogue
because it's really just
listening for key words and selecting words in a script.

You may have the same thing when you call
any call center to try to even get your computer fixed.

Now, we come into
more complex automatic speech recognition,
where it's not just word commands we're looking for,
but we're really trying
to understand what's actually being said
in the context of a given speech.

This is a deposition;
this is the person speaking;
this is what they said,
is much more complex type of algorithm or
model than select one or select two.

It's a more sophisticated form, and it's trying to understand real conversation.

So if you ever hear the word deep learning networks, natural language processing, just understand it in the context of it's a type of ASR technology that's looking to process language as it happens in real speech, natural language.

A couple of other points I wanted to cover with you.

I'm sorry. My slide is getting cut off here.

One moment, please.

Let's talk a little bit about some of the misperceptions out there, and this is from an educational perspective.

Complex speech to text, and that's, you take that audio file. It goes to an engine where. You get a raw output.

It's about context.

Context is very different for us humans versus for a computer and computer software.

Interestingly enough, and I was doing some research on this and some reading to get some very simple terms that we can use to explain the differences.

You may know, but there's a particular region of the brain in the left hemisphere we have called the Broca region.

This was found back in, I think, the mid-1850s
by a scientist and surgeon
who recognized there's this little part of the brain
here that has a big responsibility for speaking,
for how we process speech.

It's tied to other parts of the brain,
and those parts of the brain are tied
to all our five senses.

So very simply, we learn to speak through our experience
through all our senses and
the interaction and the environment we're in.

The brain is not a computer.

It's something very different.

So our speech and the way we understand it,
and we do this intrinsically,
is based upon social cues.

We're reading people's faces. It's why
attorneys want to be in the deposition room really.

They want to see that person. They want to read them.

The context of what's being said.

Is it a medical malpractice case?

Have we experienced that medical malpractice case before?

The vocal, the tone.

We do it really quick, and we do it very intuitively.

But we do it slowly.

We can only do one deposition a day as individuals.

Artificial intelligence, it's really using

math and mathematics to replicate
some of the logic of a human in the way we understand.

It is far from being a human and intelligent.

We use artificial intelligence just to
say we're trying to
replicate some of the logic of human intelligence.

It's based upon predictions,
statistics, probability, and data.

It gets the input for that.

Whereas we hear it, feel it, taste it,
sense it, it gets the input from the media file.

Within that media file,
it listens for tone,
it listens for vocal patterns,
it tries to fade out ambient noise,
but it also reads context.

So what's context?

Words, speaker lists,
specific lexicons from medical malpractice,
for example, the deposition notices.

So this content all contributes to
the mathematical model that
predicts what the word will be.

Not all automatic speech
recognition software performs the same.

They have different functions.

We talked about a Alexa, Siri, Google Home.

They are complicated. They do try to recognize natural language, how we speak, but they're more directed and simple commands.

Do this. Get directions. Turn on.

Not necessarily great at broader contexts.

I had an argument with Siri the other day driving.

Just had no idea what I was talking about.

It could be my New York accent.

Domain specific, legal, medicine.

This is what we're really talking about today.

In layman's term, what the algorithms are trying

to understand is this is a deposition,

the discussion is about XYZ, and the speaker is Mr. Smith.

So it's understanding context.

Tony, I just want to remind

you we have about ten minutes.

I do want to try and get a couple of questions in,

but I know you have lots of important content.

We'll go through this quickly. How does

the ASR model do it?

There is three types of models.

There's an acoustic model.

An acoustic model converts a sound wave to a phoneme.

A phoneme is the basic unit of speech.

There's 44 of them in English.

I have an example, cat, hat; ca, ha.

So it tries to do that.

From that, it creates

a chain of letters which become words.

It then goes to a linguistic model,

which is how likely a word was said.

Very simply, that linguistic model is simply predicting,

if it hears, hi, my name, most of

the time after that is going to be is.

The contextual model, very simply,

if you load up a deposition notice,

the speaker list loaded up with the file,

is getting context clues from other words.

Very simply put, is

it whether I do this or the weather outside?

Well, if there's the word weather plus forecast

then I know the correct spelling

is weather, and it's weather,

and that's how the machine thinks. Next slide.

ASR in legal transcription compared to stenographic.

Verbatim is recorded live on live stream as it happens.

Same exact thing with shorthand reporter.

Shorthand reporter can report it live or post.

They're just using a different method.

It converts sound to phonemes,

the basic units of speech.

Digital reporter converts it

to a symbolic representations.

Lisa will cover that in a second.

Using an AI model,
it converts it to speech,
CAT software converts the keystroke to English.

Both use specific and domain types
of dictionaries and processes. Next slide.

I'm going to have Lisa address this very simply.

But if you just look at it from

a model on the left is ASR,
does it in digital. It records it.

Sound gets converted directly into words.

You can see the timestamp.

Lisa, I'm going to let you talk a little

bit about the shorthand for about a minute,

and how you see some of the ways of the symbol processing.

Machine shorthand, we're writing phonetics.

Phonetics in the form of syllables.

You have beginning phonetic,

a middle phonetic, and an end phonetic,

which is going to create your syllable.

Each of the keys on

the shorthand machine represent a sound.

So similarly to where ASR is

taking and creating those sound waves into sounds,

we do the same thing with a shorthand machine.

The example here shows you

how I've written down phonetically.

In this example it's using

some briefs and phrases that I'll touch on in a minute,
and it's converting it to words.

Now, if a stenographer to get speed and accuracy is
going to try and come up with what
they call briefs and phrases,
and that's a way not to have to write
every syllable out phonetically,
but to group words together
or take a large word and have a single stroke.
But it all comes down to sound and phonetics.

Thank you, Lisa.

Thank you.

Well, let's put this all together.

By the way, I don't know if Lisa introduced herself.

She's a stenographer by training,
and did legal stenography as well.

If we put this whole process together
with ASR, with the technology,
with legal transcription,
the electronic reporter will capture the sound.
How you capture it is very important.

That file then gets loaded to a speech engine,
and it is broken down using the models
into a sequence of words.

Phoneme, which is that basic unit, becomes a word.

Ca, ha; cat, hat.

Those words are then broken into sentences,

and we try to identify those sentences
with who actually said it,
based upon their voice,
and other audio aspects.

We break those sentences down into using punctuation.

So punctuating make sentences.

We have the basic phoneme. Now we have the word.

Now with the punctuation, we have the sentence,
and that sentence is assigned to the speaker.

Not going to be a one-word sentence. A response to Q/A.

Did you do this? No

Or it could be a more complex colloquy that's going on.

It's capturing speech at its basic element.

It's using mathematical models, and

it's using context to produce

that workable draft very quickly,

which then goes to the human.

So the human has to make this audible.

The human has to make this a

final last transcript with their professionalism.

Last slide, I just

want to talk about a couple of key points here.

This is wrapping up from the previous conversation.

Getting the transcription process

in digital to final quality,

the end product that goes to

the attorney with the integrity needed.

Digital who's capturing the proceeding,
it's a CER or trained digital court reporter,
court monitor, electronic court reporter.

Who does the first transcript output?

If you're using ASR,
that very raw transcript
that comes out will be the machine.

From there, it goes to
essentially a scoping level or a level 2 edit,
then goes to approval or level 2 and 3 edit.

It can be done by the agency
or a transcription partner.

It is fundamentally assembled,
quality checked, and put
together the same way they would do with stenographic.

Couple of important considerations if you're considering
using automatic speech recognition.

Context matters. Please provide,
and this is to whichever engine you're using,
depending on what they have, context, paper context.

The deposition notice has a lot of information on it.

Proper names, any type of lexicon, or dictionary.

Because remember, it's recognition.

It learns and gets better because it's using
other words to confirm other words.

Whether [inaudible 00:54:56] , what's the weather outside?

I see weather plus forecasts.

Oh, this must be the weather.

So you can apply that to legal proceeding.

Lisa talked about main and backup audio files.

Very important not only

for having your backup in NDR application,

but also used in digital with

automatic speech recognition because

speech recognition can have a main file.

That's the one you used to take it

where proceedings started out off the [inaudible 00:55:24].

Then the backup file because if this part of

that main audio that isn't exactly clear,

the transcriber can use the technology to

perhaps listen to a backup audio just to see

if they can get it better or the channel.

So the attorney leaves the podium,

walks to the jury box,

there's a backup mic at the jury box.

Well, he may fade on the main channel,

but you can pick him up over here or her over here.

Other things that are used in

speech recognition in a

software flow transcript templates,

a little different rather than typing them out.

CAT softwares have this.

These tend to be predefined and preformatted.

They can be a TXT, ASCII, Word, RTF.

So a mixture that can be supported as well.

Inaudibles.

Just real quick word on inaudibles.

I hear this a lot in the industry.

Inaudibles are going to happen to humans.

Mishears are going to happen to humans.

Computers will go through the same thing.

So the most important thing about

inaudibles is we use backup tracks,

we use multichannel tracks,

and we use fundamentally professional of

humans to isolate and resolve those.

Ben, thank you very much,

and I hope this will provided you

some background and education material.

Oh, I think it was great, Tony and Lisa.

I think it's great to talk to

our industry about what technology is

and give a realistic view of how it works

and where it is in context to what is happening today

and what will be happening in the future

and then how to make it better.

We do have a couple of

questions and only a couple of minutes left.

Tony, this first one you sort of covered,

but I just wanted to say it, is

would loading exhibits or notices,

or proper names, technical terms into your ASR before
the proceeding assist in
its learning and assist in making a better transcript?

Loading any documentation would help us.

Now, how the exhibit specifically work out,
I'm not sure how the engine would help that.

But I can absolutely get an answer on that,
particularly if it has names of other things.

But any content that you can load
will certainly go to building
out that specific dictionary.

Great. I think we have time for maybe one more question.

Lisa, we had a question on getting
experience as a digital court reporter, becoming a CER,
and getting experiential training
to become better in the industry.

I know you've been doing this for a while,
and you've been training for a long time.

What are your takes on that,
and then I'll follow up after you?

Well, and I'm sure you'll touch on getting
some type of formal education is definitely a start.

Mentorship is invaluable.

Finding somebody that you trust that has experience
that's going to take the time
to let you sit in with them.

Practice and experience, that's

going to be your training.

If you can't find somebody to mentor you, you know what?

Call AAERT.

They are really happy to help.

They can give you some contact information for people.

I mean, that's what we're here for.

That's what they are here for.

I was going to say something similar about AAERT. Join.

Become part of the membership.

Then look at who's in your area who can help you.

Membership in the organization is

an important part of building this network.

I also say that the agency, the courthouse

who you may work for

for your first job is really important.

Talk to them about what their experiential training is.

Some of them have videos.

You can watch their own training programs.

They allow you to go out and

watch other proceedings happening,

or they have different methodologies to accomplish this.

So I would definitely say that is

important. As far as becoming a CER,

the experiential part is important

but also learning the methodologies and

the best practices the AAERT provides.

That comes down to your initial training and

your continual training within the organization,
either through formal training like a training center,
or a college that's doing it,
or your informal training linked with the learning and
understandings of best practices and policies
and procedures from AAERT, knowing the ethics,
knowing the professionalism guidelines.

Those are what become important in that process.

With that, my clock says that we have hit an hour,
so I do want to thank everybody for joining us today.

Thank you so much, Lisa.

Thank you so much, Tony.

Thank you, Verbit, for allowing us to
do that and for our own organizations
for allowing us to bring this information to all of you,
for JAVS and BlueLedge.

Thank you all. Please look for
the survey and quiz after this.

It should be coming in the next couple of days.

Answer that back to us quickly
so you can get your CEU credits,
and then you'll apply that at AAERT through
your membership in
the professional development area of your account.

I appreciate everybody joining us today,
and any questions that were not answered,
we'll get back to you directly.

Thank you very much.

Thank you.

Thanks.