JACLYN LEDUC: Thanks so much for joining this webinar entitled "Quick Start to Captioning." So my name is Jaclyn, and I'll be presenting today. I'm a content marketing specialist at 3Play Media, where I write content on all things video accessibility, such as the legal landscape, accessibility best practices, and trends. So you can reach me at jaclyn@3playmedia.com-- that's J-A-C-L-Y-N@3playmedia.com-- if you have any further questions after this webinar.

So let's begin. Today we will cover the following topics. The first question we'll cover is, what are captions? And then we'll follow that with, how do you create captions? Where do you publish captions? Why should you caption? And then finally, we'll talk a little bit about who 3Play Media is. And then, like I said, we'll finish off with a Q&A at the end.

So let's get started. The big question here-- what are captions? Closed captions are time-synchronized text that can be read while watching a video, and are usually noted with a capital CC icon. Captions originated as an FCC mandate in the 1980s, but the use has expanded to online video and internet applications since that time.

Captions are for if the viewer can't hear the audio. So they include relevant sound effects, speaker identifications, and other non-speech elements to make it easier for the viewer to understand who is speaking and what sounds are occurring within the environment of the video. An example of this would be if you're watching a show and someone is opening the door and you can visually see that their keys are jingling in their hand, you wouldn't need to capture that. But if the keys happened to be jingling offscreen, you would include that as a non-speech element.

While closed captions are used for pre-recorded video, live captioning is for events happening in real time-- like this webinar-- or a meeting, fitness classes, or even virtual conferences. Live captions ensure that all live events are accessible to deaf or hard of hearing individuals, as well as they make your content more engaging. Live captions are usually created by an automatic

software or by a human stenographer. In this webinar, we use a live stenographer. There might be slight delays in live captioning as the computer is processing the words or as the stenographer is typing.

It's important to distinguish between captions, subtitles, and transcripts, as they all mean different things. So captions assume the viewer cannot hear the audio. They are time-synchronized, and they include relevant sound effects, as I mentioned before. So you can spot if a video has captions, often, when you see that CC icon.

Subtitles, on the other hand, assume the viewer can hear, but cannot understand the audio. So subtitles' purpose is to display the dialogue on screen. Like captions, subtitles are also time-synchronized. Transcripts are a plain text version of the audio. They are not time-synchronized, and really, they are best when used for audio-only content, such as a podcast. And in the US, the distinction between captions and subtitles is important, but in other countries, like in Europe, these terms are used more synonymously.

So the next question I want to answer today is how do you create captions? There are a few ways to create captions. You can do it yourself-- DIY method. You can use speech recognition, also known as-- I'm sorry, excuse me, automatic speech recognition, also known as ASR, or you can use a captioning vendor.

One way, if you have the time, is to manually transcribe the video yourself. So that is the DIY method that I mentioned. So for this, you'll need plenty of time as this method can take much longer than the actual length of the video itself-- potentially even five to six times longer. So if you're doing one-off, short videos, this method is definitely more doable, but it can be very costly at scale.

The second way to caption your video is to start with automatic speech recognition known as ASR. You can use YouTube's automatic caption generator to create captions. So that's why I have YouTube's logo up here on screen.

So YouTube uses speech recognition technology that aligns your transcript with the audio and breaks it up into correctly-timed caption frames. It does a pretty decent job when the video has high-quality audio and clearly-spoken words. But if there is any type of background noise, poor audio, multiple speakers, thick accents, the number of caption errors can be quite high.

So you will need to comb through that initial automatic transcript and make edits to ensure that everything is accurate. So once complete, you can add those captions to your YouTube videos if you're putting your videos on YouTube, or you can export the caption file for use in other applications.

And like I said previously, you can outsource your captioning jobs to a vendor, like 3Play Media. Our process combines technology with humans in a careful and strategic way so that we kind of have the best of both worlds. And I'll explain that a little bit.

So first a file goes through ASR technology, which produces a rough transcript. So this likely will need a lot of edits due to the nature of automatic captions. So the second step is one of our professional editors, each with different areas of expertise, goes through and corrects that initial transcript file to ensure caption accuracy.

And then we have a third round of quality assurance review. A quality assurance manager will conduct a final review of the transcript and captions to, again, ensure we meet our guaranteed 99% accuracy rate. So that's how 3Play does it.

Now I want to talk a little bit about caption quality standards. When it comes to captioning quality, it's important to follow best

practices. The industry standard for spelling is a 99% accuracy, as I mentioned in previous slides. So 99% accuracy, though close to perfection, means there is still a 1% chance of error. So in a 10-minute file of 1,500 words, this leniency allows for 15 errors total.

If your video is scripted content, you want to ensure your captions are verbatim. So if you turn on the captions for your favorite TV show, for instance, you would want to see things like the "ums" and the "uhs" that people are saying because those are an intentional part of the scripted dialogue. Those were written into the script.

However, for lectures or live events like this one, a clean read is preferable, meaning you'll want to eliminate filler words for clarity. Each caption frame should be around one to three lines, with 32 characters per line. The best font to use is non-serif. You should also ensure they are time-synchronized and last a minimum of one second on the screen so that viewers have enough time to actually read it.

Another key thing to keep in mind is caption placement. Typically, captions are placed in the lower-center part of the screen, but should be moved when they are in the way of other important text or elements in the video. As for silences or long pauses, you want to make sure the captions disappear after a moment or two so that they don't confuse the viewer into thinking that the dialogue is still going on. So you can always check out the Described and Captioned Media Program, the FCC, or WCAG-- that's W-C-A-G-- to review captioning quality standards.

So when you use ASR technology, the accuracy rates can be pretty error-ridden. A lot of ASR errors make sense acoustically, but not linguistically. So I will show you an example up here on the slide of a transcript captured by automatic speech

recognition. So as I play it, listen closely to the audio, and compare with the words that are showing on the screen and see if you can catch any errors. And then if you feel compelled, type those errors into the chat window and show me what you find. So I'm going to play the video.

[VIDEO PLAYBACK]

- One of the most challenging aspects of choosing a career is simply determining where our interests lie. Now one common characteristic we saw in the majority of people we interviewed was a powerful connection with a childhood interest.

- For me, part of the reason why I work here is when I was five years old growing up in Boston, I went to the New England Aquarium. And I picked up a horseshoe crab, and I touched a horseshoe crab. And I still remember that, and I still-- I love those types of engaging experiences that really register with you and stick with you.

- As a child, my grandfather was a forester. And my childhood playground was 3,600 acres of trees and wildlife that [? he had ?] introduced [? to me. ?] So my entire childhood [INAUDIBLE] wildlife, and in wild places. It just clicked.

- When I was a kid, all the cousins would use my grandparents' driveway--

JACLYN LEDUC:

So I got some great answers coming in. I saw [? Aaron ?] says, there was no period after "why," no paragraph line break for new dialogue. [? Kelsey ?] says, no paragraph break or period before "for me." Someone says, "Koran" question mark. "Forester, not four story." [? "It had ?] no punctuation."

Yeah, so these are all-- yeah, Stevie says there's no distinction when different speakers start. Yes, great observations. These are all 100% right.

So one of the issues here is a lack of punctuation, as you all pointed out. In this transcript, there are few periods and incorrect capitalizations which make for difficult reading. Another issue is that hesitation words are not removed, which spill over into other words and cause inaccuracies. Also speaker changes and speaker IDs are not captured-- so all these things you pointed out.

So these errors would be harder to catch if you were listening to the transcripts. But when you're reading it with the errors, it makes no sense. So in this example, when the speaker says "New England Aquarium," ASR picked it up as "new wing of the Koran."

And when the speaker said "forester," the transcript "four story." And I think one of you specifically pointed those two things out. The thing is that a human wouldn't make these errors. So that's why human-made or human-edited captions are tremendously more accurate.

So most live captioning accuracy can be over 90% accurate, but noise level, accents, or connectivity issues can affect the

accuracy. Much of the same quality standards for closed captioning apply to live captioning. So for the best accuracy in a live captioning environment, you want to ensure that you have a strong network connection, good audio quality, little to no background noise, one single speaker at a time, and then clear speech and pronunciation.

Great. Now let's dive into how to publish captions. There are many ways to publish your captions. The most common way is through a sidecar file, which is basically a file that stores the captions so they can be associated with the corresponding video. So when you upload your caption file on YouTube, you are uploading it through a sidecar file. These types of captions give the user control to turn the captions on and off as they need them.

Another way is to encode captions onto the video. For example, these are found on kiosks or offline videos and can also be turned off or on. Open captions are burned into a video and cannot be turned off or on.

So for social media videos on Instagram or Twitter, adding captions as a sidecar file is not possible. In this case, open caption encoding is one of the ways to overcome this barrier and to make sure that your social media videos are captioned and accessible.

Lastly, integrations are simply a publishing process for captions. It's a preset workflow between your captioning process and video publishing process to really make the process overall more streamlined.

So why should you caption your content? We're going to go over some of the reasons why you should caption, some of the benefits of doing so. So the biggest reason and why I have it up here first is captioning your videos makes them accessible. I

have on this slide the words "accessibility" and then "a11y" underneath it. And the "a11y" represents the letters between the a and the y in accessibility. And it also represents being an ally for accessibility.

So you'll see this word a11y pronounced ally. You'll see it used a lot in place of accessibility. So there are 48 million Americans with hearing loss, which is about 20% of the US population, and 360 million deaf or hard of hearing individuals around the world. So captions help make your content more accessible to them.

So captions are also great for engagement. They're necessary if you want to make sure your that your videos are comprehensible without any sound. And one stat shows that 41% of videos are incomprehensible without sound. So they need captions. Because if someone doesn't have headphones, or they can't watch your video, then without captions, there's no way they can watch your video at all. So captions are important for that reason.

92% of consumers are watching video on mobile with the sound off due to environmental situations. If they're in a quiet space, if they forgot their headphones, if they're at work or multitasking, people prefer not to watch videos with sound. So if your videos don't have captions, the viewer might just be glossing right over them.

Video accessibility has tremendous benefits for improving SEO, or Search Engine Optimization, the user experience, your reach, and your brand. So a study by Liveclicker found that pages with transcripts earned an average of 16% more revenue than they did before transcripts were added. And according to Facebook, videos with captions have a 135% greater organic search traffic.

So I have onscreen here captions improve brand recall, verbal memory, and behavioral intent. A research study from the

Journal of the Academy of Marketing Science found that captions do improve these things. So it's clear that captions are valuable for marketing purposes, too.

Video accessibility also benefits students tremendously. We conducted a nationwide study with Oregon State University where we surveyed students to see how and why they use captions. And the results prove that captions truly do help students learn. 98.6% of the students surveyed found captions helpful.

In addition, 65% of students use captions to help them focus. And 75% of all students who use captions, not just those who are deaf and hard of hearing, use captions as a learning aid. So all of these are really great benefits of using captions and kind of provide reasons behind why we all should be captioning our videos.

So what are the laws that require captions? This next section will be covering some of the accessibility laws that cover captions for online video content.

The Rehabilitation Act of 1973 was the first major accessibility law in the US. It has two sections which specifically impact video accessibility. So Section 504 is a broad anti-discrimination law that requires equal access for individuals with disabilities, which applies to federal and federally-funded programs.

Section 508 requires federal communications and information technology to be accessible. The Section 508 refresh references Web Content Accessibility Guidelines, or also known as WCAG 2.0. So what's unique about the Rehabilitation Act is that captioning and audio description as well requirements are written directly into Section 508.

The next law is the Americans with Disabilities Act, or the ADA, which was the second major accessibility law in the US. It has two

sections that impact accessibility. Title II applies to public entities. And Title III applies to places of public accommodation, including private organizations that provide a public accommodation. So that could be a doctor's office, a library, a hotel, a restaurant, and many more places are included in that classification.

The context of a place of public accommodation has been tried in lawsuits in regards to how it impacts internet-only businesses. And in several cases, Title III has been extended to the online space. For example, there were suits against Netflix both in regards to captioning and audio description. And in both cases, the outcome was that Netflix had to provide accurate captions for their streaming shows and audio description for their Netflix originals.

The third major accessibility law in the US is the 21st Century Communications and Video Accessibility Act, or the CVAA. For caption requirements, the CVAA applies specifically to online video that has previously aired on television. Any online video that previously appeared on TV with captions has to be captioned once it goes online. As for audio description, that's another topic. But the CVAA is phasing in audio description requirements by 2020.

So one other thing to know of are the standards that should be met to mitigate the risk of legal action. The Web Content Accessibility Guidelines, or WCAG 2.0, are the international standard and best practice for web accessibility. It's important to note that there is a WCAG 2.1 version, but for the time being, 2.0 is currently what's referenced in legal recommendations.

WCAG, as I mentioned, is the international set of guidelines helping to make digital content accessible for all users, specifically users with disabilities. It outlines best practices for making web content universally perceivable, operable, understandable, and robust.

So WCAG has three levels-- level A, level AA, which is what most people are aiming for. This is the mid-level of standard. And then level AAA is the most comprehensive high standard.

Most laws and lawsuits mention WCAG 2.0 compliance. So for now, that's what is legally recommended. Only if a law explicitly states that the web developers have to adapt to the newest WCAG version do they need to make their content 2.1 compliant.

The W3C does suggest that new web sites follow WCAG 2.1 since they are more inclusive both for desktop and mobile. So to be compliant with WCAG, you are required to caption pre-recorded video for level A compliance and caption live video for level AA compliance.

And I want to quickly highlight some accessibility lawsuits before moving on. In the lawsuit the *National Association of the Deaf versus Harvard and MIT,* the universities were sued for failing to caption and for having unintelligible captions on their online courses.

This was the first time accuracy had been considered in legal ramifications for captioning. They were using YouTube captions perhaps without editing them. This lawsuit represents a violation of Title III of the ADA and has extended the requirements to the internet.

Another lawsuit is UC Berkeley, Penn State, and Miami University. All are other examples of schools that have been sued or have entered into consent decrees in regard to their inaccessible video.

One last lawsuit I have here is the *National Association of the Deaf versus Netflix,* which was the first web accessibility lawsuit. Under Title III of the ADA, the court ruled that Netflix is considered a place of public accommodation and therefore

needed to make their content accessible.

So I know I've come up on time here. I'm almost done, and then we'll be right on to Q&A. So just a little bit about who 3Play is. So we are a video accessibility company spun out of MIT in Boston in 2007. And we're currently still based in Boston.

We started out offering captioning, transcription, and subtitling services. Now we also offer audio description, which is a service for blind and low-vision individuals. And we recently released live auto captioning as well.

We have over 2,500 customers spanning higher education, media, government, e-commerce. Some of our customers include P&G, T-Mobile, MIT, and IRS. Our goal is really to make the whole captioning and video accessibility process easier, and then I'll get into how we do that on the next slide.

Like I said, our number one goal is to make video accessibility easy. And we do that in a number of ways. So we have an easy-to-use online account system where you can manage everything from one place. We have a number of different options for turnaround, anywhere from a couple of hours to over a week-- whatever fits your needs.

We have different video search plugins and integrations for captioning and audio description that helps simplify the process of creating accessible video. And what we're working toward is being a one-stop shop for captioning, description, transcription, and subtitling and just video accessibility as a whole.

Besides our services, we also offer a lot of free resources all centered around online video and video accessibility so that you can be empowered to take on accessibility initiatives at your own organization. So on our website, you'll find weekly blogs, free white papers, how-tos with checklists, research studies on the impact of video accessibility in certain environments, such as

education.

We offer webinars like this one, sometimes several webinars a month, where we bring in accessibility experts to share their knowledge. And we also offer a free video accessibility course that is available to take now. You can go to 3playmedia.com/certification to join the 3Play network and access that course. So we'll be sure to include that link in tomorrow's follow-up email and in the chat window if you're interested.

So thank you for bearing with me. We're going go on to the questions and comments. Well, let's see what questions we have rounded up.

So someone asked, can you please restate the number of characters per line and number of lines that are recommended, as well as the font size for pre-recorded caption?

So each caption frame should be around one to three lines with 32 characters per line. The best font to use is a non-serif font. And as far as font size, let me revisit that slide. I'm not sure if there is a font size specifically. But it might just depend on which tool you're using. But the font should be readable. It should be large enough that one can read it on the screen. It shouldn't be too, too small. But readable fonts in most video players allow you to adjust the size. So just make sure that it's readable to the human eye and that you don't need a microscope to read it.

Someone else asked, someone just said, did you say sans serif font? Yes, a sans serif font is best for captions.

Someone also asked, if you've created captions using YouTube, is there any way to download or export your caption video as an open caption video to post on social media?

So the answer to that is it depends which platform. Most do not

allow you to upload a separate caption file. So you would have to use a service to encode the captions into a video. So with YouTube, you can export your captions as an SRT file, WebVTT, or an SBV file. But those are not going to be encoded in the video.

If you want to post that alongside or post that with your social media videos, the social platform has to accept one of those caption files. If not, you might have to find a service that will encode the captions or embed the captions into the video for you. So I hope that helped answer your question.

Just looking at the other questions we have here. So someone asked, how does an integration work?

So integrations link disparate systems or platforms to make it easy to share information and build workflows between the two. So 3Play Media's integrations are engineered to make the captioning process much easier. As I mentioned, we integrate with most leading video platforms. Our integrations allow you to select the files you want captioned directly from your video platform or cloud storage. And so, I mean, overall, integrations just are kind of there to save you a lot of time by streamlining the captioning process.

Somebody asked, can you talk about what a person can expect when they order transcription, captions, audio description, and then how to download them? The way it works is you would upload your video to 3Play Media. We have an account system, a user account system, that you would log into, upload your video. And then you would select your turnaround time and the language you would like to have captioned in or that you would like the service in.

Once your captions are complete, you will receive an email and will be taken to the file to download it in over 50-plus formats. So whatever file format you need, you can download it in. And then

also as I had said, you can also use an integration to reduce the number of steps there so, again, streamlining that process, making it easier on yourself.

And then I'll probably do one more question just because we are over time. Let's see. How does the live captioning process work?

So first, you create a live event in any of our integrated live stream video platforms. Then you would schedule live automatic captioning in 3Play for your corresponding live event.

The next step would be to stream your live event. And then your captions would display directly on the video player or through an embed code. And then, finally, you would download, edit or upgrade your live transcript, which we recommend for accuracy's sake. And then you can access the final transcript for editing, upgrade to full transcription, and then also order more services on the transcript, if you like.

So that is all the time we have today. If I missed your question, again, please feel free to email me. My email is jaclyn, J-A-C-L-Y-N, @3playmedia.com. And I'll be happy to get back to you. Again, thank you so much, everyone, for joining me today. And I hope everyone has a great rest of the day.