JACLYN LEDUC: So thank you all for joining this webinar, entitled Quick Start to Captioning. So hi. My name is Jaclyn, and I will 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. You can reach me at Jaclyn-- that's spelled J-A-C-L-Y-N-- @3playmedia.com if you have any further questions after this webinar.

So today we will be covering the following topics. The first question we'll cover is, what are captions? Followed by, how do you create captions, where to publish captions, why should you caption? And then we'll finish up with, who is 3Play Media? Who are we? So finally we'll finish off with a Q&A, like I mentioned, at the end with about 5 to 10 minutes left.

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 they're 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 then.

Captions are for if the viewer can't hear the audio. So they include relevant sound effects, speaker identifications, and other nonspeech elements to make it easier for the viewer to understand who is speaking and what sounds are occurring in the environment.

An example of this would be if you're watching a show, and someone is opening the door. And you can visually see their keys jingling. In that case, you wouldn't need to caption that sound effect. But if the keys are jingling offscreen, you would have to include that as a nonspeech element.

While closed captions are used for prerecorded video, live captioning is used for events happening in real time, like this webinar or meetings, fitness classes, and even conferences. Live captions ensure that all your live events are accessible to deaf or hard-of-hearing people, as well as it makes your content more engaging.

Live captions are usually treated by an automatic software or by a stenographer. In this webinar, we do use a live stenographer. There might be a slight delay in live captioning as the computer is processing

the words or the stenographer is typing. So just keep that in mind when using live captioning.

It is important to distinguish between captions, subtitles, and transcripts, as they all mean something different. Captions assume the viewer can't hear the audio. They are time-synchronized. And they include relevant sound effects, like I said. You can spot if the video has captions when you see that CC icon.

Subtitles, on the other hand, assume the viewer can hear but can't understand the audio. The purpose of subtitles is to display the dialog on screen. Like captions, subtitles are also time-synced.

Transcripts are a plain-text version of the audio, so it's not time-synchronized. And really, it's best when used for audio-only content, like podcasts. In the US, the distinction between captions and subtitles is important. But in other countries, like Europe, these terms are used more synonymously.

The next question I want to answer is, how do you create captions? So let's get into that.

There are a few ways to create captions. You can Do It Yourself, the old DIY method; use 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. You'll need plenty of time, as this method can take much longer than the length of the video, maybe even up to five or six times longer than the video itself. If you're doing one-off short videos, this method is definitely more doable. But it could be very costly at scale. So if you have a lot of video content, this might not be sustainable.

The second way to caption your video is to start with ASR. You can use YouTube's automatic caption generator to create captions. So YouTube uses speech recognition technology that aligns your transcripts 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 English. But if there's any type of background noise, poor audio, multiple speakers, or even thick accents, the number of caption errors can actually be quite high. So you will have to comb through the transcripts from YouTube and make edits to ensure that everything is accurate. Once complete, you can add the captions to YouTube videos. Or you can actually export the caption file to use in other applications.

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

First, a file goes through ASR technology, much like YouTube's transcripts or what YouTube uses to create their transcripts. And that produces the rough initial transcript. This likely needs a lot of edits, though, due to the nature of automatic captions being super inaccurate.

So from there, one of our professional editors with all different areas of expertise goes through and corrects the file to ensure accuracy. And then we have a third round of quality assurance review. A QA manager conducts a final review of the transcript and captions to ensure we meet our guaranteed 99% accuracy rate.

Now let's talk a little bit about quality standards.

When it comes to captioning quality, it's super important to follow best practices. The industry standard for spelling is a 99% accuracy. 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, then you'll want to ensure your captions are verbatim. So if you turn on the captions for your favorite

TV show, you'll want to see the "ums" and the "uhs" included in the captions because they are intentional. And those are part of the scripted dialogue. However, for lectures or live captioning, a clean read is preferable, meaning you'll want to eliminate the 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 nonserif. You should also ensure they are time synced and last a minimum of a second on the screen so that it gives viewers enough time to 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'll want to make sure the captions disappear after a moment or two so that they don't confuse the viewer into thinking the dialogue is still going on. You can always check out the Described and Captioned Media Program, the FCC, or WCAG, W-G-A-G, to review captioning quality standards.

When you use ASR technology, the accuracy rates are pretty bad, like I mentioned. A lot of ASR errors make sense acoustically, but not linguistically. I'll show you an example of a transcript captured by ASR so that you can see firsthand what it looks like. Listen closely to the audio and compare with the words on the screen and see if you catch the errors. You can type any errors you notice in the chat window. So I'm going to play it now.

[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 the childhood interest.

[MUSIC PLAYING]

- 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 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 there. So my entire childhood was really in wildlife and in wild places.

[END PLAYBACK]

JACLYN LEDUC:

OK, so I'll stop it there. So one of the issues here is, as you probably saw, the lack of punctuation. In this transcript, there are few periods and lots of incorrect capitalizations, which makes for difficult reading. Another issue is that hesitation words are not removed in this transcript. So they spill over into other words and cause inaccuracies. Speaker changes and speaker IDs are also not captured.

So these errors would be harder to catch if you were just listening to the transcript. But when you're reading it with the errors, it really doesn't make sense. So in this example, when the speaker says New England Aquarium, the ASR picked it up as "new wing of the Koran." And when the speaker said forester, the transcript read "four story." So the thing is a human wouldn't make these errors if they were doing this transcript. That's why human-made or human-edited captions are tremendously more accurate.

Most living-- excuse me. Most live captioning accuracy is 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 as well. For the best accuracy in live captioning, you'll want to have a strong network connection, have good-quality audio, have little to no background noise, a single speaker speaking at once-- so even if you have multiple speakers, trying not to talk over each other-- and then clear speech and pronunciation.

Now let's dive into how to publish captions.

There are many ways you can 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.

Another way is to encode captions onto the video. For example, these are found on kiosks or offline videos. And encoded captions can also be turned off and on.

Open captions are burned into a video. And these cannot be turned off or on. For social media videos on Instagram or Twitter, adding captions as a sidecar file is not possible. So open caption encoding is the best way to overcome this barrier and keep your social videos 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 everything more streamlined.

So why should you caption your content? There are many reasons why you should caption. The biggest is for accessibility. And I have on this slide the words "accessibility" and then, underneath it, A11y. And A11y, which we say as "ally," represents the 11 letters between the A and the Y in "accessibility." And it also represents being an ally for accessibility.

There are 40 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. Captions help to make your content accessible to them.

Captions are also great for engagement. They are necessary if you want to make sure your videos are comprehensible without sound. 41% of videos are incomprehensible without sound or captions, which means that if someone doesn't have headphones on them, they might not watch your video.

92% of consumers are watching the videos-- who are watching video on mobile with the sound off due to environmental situations. So 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, then the viewer might just be glossing right over them.

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

A research study from the *Journal of the Academy of Marketing Science* found captions improve brand recall, verbal memory, and behavioral intent. So it's clear that captions are valuable for marketing purposes as well.

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

OK. So now let's get into the accessibility laws around captioning. So

there are several, and I'll go over each one in some brief detail. So the Rehabilitation Act of 1973 was the first major accessibility law in the US. It has two sections which impact video accessibility.

There is Section 504. It's a broad antidiscrimination law that requires equal access for individuals with disabilities. And it applies to federal and federally funded programs. Section 508 requires federal communications and information technology to be made accessible. Section 508 refresh references Web Content Accessibility Guidelines, or WCAG 2.0, as we call it. So what's unique about the Rehabilitation Act is that closed captioning and audio description requirements are written directly into Section 508.

The next law is the Americans with Disabilities Act, also known as the ADA, which was the second major accessibility law in the US. It has two sections that impact video 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 like a doctor's office or a library, a hotel, a restaurant, and other public places.

The context of a place of public accommodation has been tried in many 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 closed 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 television with captions has to be captioned when it goes online, including video clips and trailers. As for audio description, the CVAA is phasing in audio description requirements by this year.

One other thing to know of are the standards that should be met to mitigate the risk of legal action. So I'm talking about the Web Content Accessibility Guidelines, or WCAG 2.0. These are the international standards and best practices for web accessibility. It's important to note that there is a WCAG 2.1. But for the time being, WCAG 2.0 is currently what's referenced in lawsuits and legal recommendations.

WCAG is the international set of guidelines, like I said, helping to make 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 of compliance. Level A is easiest to maintain. It's the least stringent. Level AA is what most people are aiming for. This is the mid level of standards. And then level AAA is most comprehensive, the highest accessibility standard.

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

The W3C, who puts out the WCAG guidelines, does suggest that any new websites should follow WCAG 2.1 since those are the most inclusive and mobile-friendly guidelines. To be compliant with WCAG, you are required to caption prerecorded video in order to be level A compliant. And then you have to caption live videos to be level AA compliant.

And I just want to 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 a lot of their online course videos. This was the first time accuracy had been considered in legal ramifications for closed captioning. The lawsuit represents a violation of Title III of the ADA. And it has extended the requirements to the internet. The

outcome will have huge implications for higher education.

And some other lawsuits happening at higher ed are towards UC Berkeley, Penn State, and Miami University. And they have been sued or have entered into consent decrees in regard to inaccessible video.

Another lawsuit to know, which I briefly mentioned earlier, is the *National Association of the Deaf versus Netflix.* And it 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.

Now a little bit about who we are at 3Play. We are a video accessibility company and we're spun out of MIT in 2007 and are currently based in Boston.

We started out offering captioning, transcription, and subtitling services. We also offer audio description, a service for blind and low-vision people. And we recently released a live automatic captioning solution as well.

We have over 2,500 customers spanning industries in higher education, media, government, e-commerce, fitness, and enterprise. And some of our customers include P&G, T-Mobile, MIT, and the IRS. Our goal is really to make the whole captioning and video accessibility process much easier. And 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. We have an easy-to-use online account system where you can manage everything easily from one place. We have a number of different options for turnaround, anywhere from a couple of hours to over a week-- so whatever fits your needs.

We have different video search plugins and integrations for captioning and audio description that help simplify the process of making your video accessible. And really what we're working toward is being a one-stop shop for captioning, description, transcription and subtitling, and

video accessibility as a whole.

Besides our services, we also offer tons of free resources all centered around online video and video accessibility so that you can feel empowered to take on accessibility initiatives on your own. On our website, you'll find weekly blogs, free white papers, how-tos, checklists, research studies on the impacts of video accessibility.

We also offer monthly webinars, sometimes several webinars a month, where we bring in accessibility experts to share their knowledge. So you can register for those for free, like this one, right on our website. And we now offer a free video accessibility certification that is available to take now.

You can go to-- I'll redo the link. It's 3playmedia.com/certification/. And you can go there to join our 3Play network and then also access the certification. And we'll be sure to include that link in tomorrow's follow-up email.

OK, so thank you so much. So that concludes the presentation. But we will now go on to Q&A. OK. So we have a question asking, how does integrations work? So I will explain that.

Integrations link disparate systems or platforms to make it easy to share information and build workflows between the two. So our integrations are engineered to make the captioning process just 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. So it makes it really, really easy for you. Integrations save you a lot of time. So that's the benefit of having access to those integrations.

OK. Another question is, are there quality standards for how captions should look? So I went a little bit over the quality standard. So yes, there are definitely quality standards for how captions should look. With accuracy, the FCC states that captions must match the spoken words in

the audio to the fullest extent. This includes preserving any slang or accents in the content and adding the nonspeech elements.

For live captioning, there is some leniency. Captions must be synchronized. They must align with the audio track. And each capture frame should be presented at a readable speed, three to seven seconds on screen.

So completeness is also important. Captions must run from the beginning to the end of the program and not drop off. So the captions must also be placed so that they don't block other important visual content. So those are some of the quality standards. But you can always visit the DCMP, the Described and Captioned Media Program. You can go to the FCC. Or you can look at WCAG for some more quality standards if you ever need a reference.

OK. Someone asked, how does the live captioning process work? And we'll make this our last question. So first, you create a live event in any of our integrated live stream video platforms. Then you schedule live automatic captioning in 3Play for your corresponding live event.

So after that, you stream your live event. Your captions will display directly in the video player or through an embed code. And then you can download, edit, or upgrade the live transcript. And you can access the final transcript for editing, upgrade to the full transcription, or even order more services on the transcript once all is said and done.

OK. All right. So that is it for questions. I went a little over time. Thank you for bearing with me. And thanks so much everyone for joining us today.