LILY BOND: Hi everyone, and thank you for joining this webinar entitled "The Nuts and Bolts of Captioning and Describing Online Video." I'm Lily Bond from 3Play Media, and I will be presenting today, along with Owen Edwards, who is a Senior Accessibility Consultant at SSB BART Group.

We're going to cover several different aspects of captioning and describing today. So we're going to go through what are captions and the audio description, why should you caption and describe your videos, how do you create captions and descriptions, where do you publish captions and descriptions? And then, we'll leave time at the end for Q&A. And as I mentioned, please feel free to ask those questions throughout.

So to begin, what are captions? Captions are time-synchronized text that can be read while watching a video. They're usually documented with a CC icon. And they originated as an FCC mandate for broadcast in the 1980s. Captions assume that the viewer can't hear, so they convey all relevant sound effects, speaker identification, and other non-speech elements.

So there's an image on the screen here of a man grabbing his computer and visually looking upset. And the caption for that would be [SOBBING MATHEMATICALLY], which is a sound effect, and not spoken word. And relevance is really key here. So for instance, if someone is walking down the street with keys jangling in their pocket, you wouldn't need to include that sound effect, because it's not really critical to the plot development. But if it's a horror movie, and someone is behind a door, and there are keys jangling trying to get in, keys jangling would be a relevant sound effect in that case.

And just to go through some of the common terminology around captions and to clarify the difference between captions, subtitles, and transcripts. Captions assume that the viewer can't hear the audio, so they're time-synchronized and include sound effects. Whereas, subtitles assume that the viewer can't understand the audio, so they're really about translating the content. And they're also time synchronized.

And then, transcripts include a plain text version of the audio in the original language. And they are not time-synchronized. And they are sufficient for audio-only content.

OWEN EDWARDS: Thanks, Lily. And then, what is audio description? Audio description is, in some ways, the exact opposite of what captions are. And it's certainly much less well-known. Narration of audio description is defined as narration added to the soundtrack to describe important visual detail that cannot be understood from the main soundtrack alone.

So it's really intended for people who are blind or low vision, who are trying to understand the whole content of a video purely from the soundtrack. And there has often been confusion around this, because it's gone by different names. It's sometimes called audio description, sometimes video description, sometimes narrative description, and sometimes simply description.

It's starting to be seen more broadly. There's a couple of places that are starting to show up. And one that's become quite well-known is on Netflix. And I have a little image here where it shows a Netflix video being played. And the words, letters, pop out from a white background, then turn red, Netflix."

Similarly, a video from the royal National institute of the Blind in the UK, which has additional audio, "A man sits alone in a hospital waiting room" to describe what's going on. And it's really very similar to the director's commentary that many of us may have experienced on a DVD, but intended specifically to understand what's going on in the content that can't be understood purely from the soundtrack.

LILY BOND: Great. So now, we're going to cover some of the legal requirements and benefits of captioning and audio description, to answer the question, why should you caption or describe your video? So I'm going to talk briefly through the accessibility laws in the US. The first major accessibility law in the US was the Rehabilitation Act of 1973. And there are two sections that really apply to video content.

Section 504 is the first one. It's a very broad anti-discrimination law that

requires equal access for individuals with disabilities. So this applies to Federal and Federally-funded programs. And then Section 508 was introduced much later in 1998 to require Federal communications and information technology to be accessible. So this applies to Federal programs, but it's often applied to Federally funded programs through state and organization laws.

And then, closed captioning and audio description requirements are written directly into Section 508, but are often applied to Section 504. And Section 508 was refreshed in January. And the requirements will phase in in January of 2018. And these will reference the WCAG 2.0, the Web Content Accessibility Guidelines. And Owen is going to talk through those shortly.

The second major accessibility law is the Americans with Disabilities Act of 1990. This one has five titles. Title II and Title III apply to online video. Title II impact public entities, and Title III impact places of public accommodations, which does extend to the private sector. And so one major question around the Americans with Disabilities Act is, what constitutes a place of public accommodation?

So originally Title III was intended to-- the ADA was written in 1990, before the internet was anywhere near as prolific as it is today. So when Title III was written, it originally applied to physical entities like, for instance, wheelchair ramps on a building. But now, it's being extended more and more to the online sector.

And a couple of major lawsuits with captioning have noted Title III of the ADA. Netflix was sued by the National Association of the Deaf in 2012 for failing to provide closed captions for most of its watch instantly movies. And this was the first time that Title III of the ADA was applied to an internet-only business.

The court ruled in favor of the National Association of the Deaf. And Netflix settled, agreeing to caption 100% of its streaming content, which set a profound precedent across industries. And so Hulu and Amazon have also recently settled with the National Association of the Deaf, agreeing to caption

their online content. And as Owen mentioned, Netflix has also agreed to provide audio description on many of their watch instantly videos.

In higher education, the National Association of the Deaf is in an ongoing lawsuit against Harvard and MIT, who they sued for providing inaccessible video content that was either not captioned, or was inaccurately or unintelligibly captioned using automatic captions. And in 2015, the Department of Justice submitted a statement of interest supporting the National Association of the Deaf in this case, but we are still waiting on a decision. And the outcome will have huge implications for higher education.

And then, the final major accessibility law in the US is the 21st Century Communications and Video Accessibility Act, or CVAA. This applies to both captioning and audio description. So for captions, the CVAA impacts online video that previously appeared on television with captions. So any of the clips or full length video content that you see on something like ABC's website, HBO's website, all of that has to be captioned under the CVAA. And both "straight lift" clips and montages have to be captioned as well.

Audio description has slightly different requirements that phase in between 2010 and 2020. Currently, the top 60 TV markets are required to describe 50 hours per calendar quarter. The next phase-in is in July of 2018. And the goal is for there to be 100% audio description on television by 2020. So now, Owen is going to go through the Web Content Accessibility Guidelines.

OWEN EDWARDS: Right. So the Web Content Accessibility Guidelines, or WCAG 2.0, are international guidelines developed by the W3C, the same organization that developed all of the standards around the internet and the way that information travels over the internet. And this version was really intended to be much more testable than Version 1.0. And it has some subsections or different levels-- levels A, double AA, and triple AAA, along with a number of numbered success criteria.

The levels denote increasing complexity of accessibility, where level A is a base level, level AA is a more broad level, and level AAA includes some features that are pretty unusual to see in the commercial market. And so

most interpretations of ADA have focused on level AA. And as Lily mentioned, the Section 508 refresh, which is going through this year, which was passed in January and will go into effect January 2018, synchronizes Section 508 with WCAG 2.0 level AA.

Level AA requires conformance to both the AA success criteria and A success criteria. And so specifically around video, prerecorded video, the requirement at level A, the base level for WCAG, are that there should be captions, and there should be either a transcript or audio description.

A transcript allows a screen reader user or another blind or low vision user to read through the content of the video at their own pace, but doesn't convey all of the information that's included in the soundtrack-- some of the noise, some of the atmosphere of the video, some additional information that's in the soundtrack. And that's, in general, why level AA, the most common level that people are complying with, there's a requirement for both captions and audio descriptions. At that point, a transcript isn't sufficient for people who are blind or low vision.

So that means that you can meet level A by having captions and audio description and also meet level AA. And then, at level AAA, there are additional accommodations that are possible. There's a requirement that you have captioned and transcript in audio description, but also an extra video, which includes a sign language translation of the video to supplement the captions or to complement the captions.

LILY BOND: So given all of these laws, many of which were written before online video was as popular and as large a part of our everyday lives as it is today, we often have to turn to the courts to see what they believe about how the ADA, in Section 504, should apply to the online sector. So there are just a couple of quotes here that I think really document their push for accessibility in online video.

In the case of the *NAD versus Netflix,* Judge Ponsor said, "Excluding businesses that sell services through the Internet from the ADA would run afoul of the purposes of the ADA." And then the Department of Justice, in the

case of *The National Association of* the *Deaf versus Harvard & MIT,* in their letter of interest said, "The United States respectfully submits this statement of interest to correct Harvard's misapplication of the primary jurisdiction doctrine and its misunderstanding of the ADA and Section 504.3."

Beyond the legal requirements, there are many other benefits of captions and description. I'm going to go through the benefits of captions, and then Owen will cover the benefits of description. The main benefit is, obviously, accessibility. There are 48 million Americans living with hearing loss. And that's growing, due to medical advancements. But captions also provide a lot of benefits for people who are not deaf or hard of hearing.

The Office of Communications in the UK conducted a study where they found that 80% of people who were using captions were not deaf or hard of hearing. So it provides better comprehension for all of you, as in the case of things like a thick accent, if the content is esoteric, if there's background noise, if the viewer knows English as a second language. Captions also provide the flexibility to view videos in sound-sensitive environments-- I apologize for the typo on that-- like the office, the library, or the gym.

And we're often seeing that captions provide a lot of flexibility in social video. So more and more companies are defaulting to playing videos without sound, when you scroll onto them on your mobile device or desktop. So Facebook, for instance, when you scroll onto a video on your newsfeed, it autoplays without sound. And there's no way for anyone to know what's going on in that video, unless there is text showing on the screen. So captions are really beneficial in that environment.

Captions also provide a great basis for video search. So once you have a transcript, you can create an interactive search environment. And MIT did a survey of their students who were using interactive transcripts, and 97% of users said that they felt interactive transcripts enhanced their experience.

Captions also help with SEO, or search engine optimization. Adding captions to your video allows Google to understand the contents of your video, because they cannot watch your video. They can only read the content that

you have provided. And adding captions to YouTube videos led to a 7.3% percent increase in views in a study done by Discovery Digital Networks.

Captions also provide a great basis for translation. So once you have English captions, you can translate them to create multilingual subtitles and reach a global audience. And captions are also really re-usable. So once you have captions and transcripts, you can use them to create things like infographics, study guides, blog posts.

The University of Wisconsin found that 50% of students were repurposing transcripts as study guides. And we've also heard of professors transcribing lectures from a semester and then using them as a basis for a textbook. And of course there are many legal requirements, so there's always the benefit of complying with the law.

To speak a little bit more to engagement and better comprehension, we conducted a research study with Oregon State University's eCampus Research Unit last year to learn how and why students use closed captions for learning. We find that 98.6% of students find captions helpful. 75% of students use captions as a learning aid. And the number one reason students use captions for learning is to help them focus. So those are some great tangible stats to help get buy-in for why captioning is so beneficial for all students and all viewers, not just those who require it as an accommodation.

OWEN EDWARDS: OK. And then the benefits of audio description-- again, the key benefit that we're talking about here is the accessibility of video content. With video content becoming so much more prevalent online, and particularly in education settings, there's a barrier for an estimated 22 million Americans, or 10% with vision loss.

But the other areas where it's beneficial to add audio description-- particularly in an education setting, we're seeing a lot of educators talking about the benefits to people with autism, other autism spectrum disorders, and other learning disabilities, that it helps them understand the emotional and social cues that are present visually, but maybe don't come across

specifically and need to be highlighted. That's something that the audio description can provide, which is necessary to understand exactly what's going on in a scene.

And then similar to captions, but in a different setting, it allows flexibility. An audio description added to a video can allow the consumption of that video or the viewing of that video in an environment where your eyes are either distracted, need to be somewhere else, it allows more of a radio consumption of a video, or the ability purely to listen to it.

Also, listening to language is an important part of language development. So hearing a spoken description of something that's visual allows learners to understand that link between what's going on visually and what they're hearing described. And also, there's a benefit for many learners, many types of learners-- well, there are many types of learning. But between 20% and 30% of students retain information best through sound. So it can act as a reinforcement of what they're seeing when consuming a video to really emphasize what the key important parts of a busy scene or a scene that contains a lot of information. The audio description will capture those.

And then, finally, as we've mentioned, there's growing legal requirement. And with description previously being less common and less well-known and, in part, related to that transition from level A WCAG up to level AA where, previously, a transcript was sufficient, we're seeing more and more that, as level AA compliance is required, audio description is going to be a requirement for videos.

LILY BOND: Great. Thanks, Owen. So we're going to dive into how you actually create captions and descriptions, starting with captioning. One DIY method for creating captions is to use YouTube, which I highly recommend for their timings alone.

The first step is to transcribe your video, which often takes five to six times real time. And remember that you need to include things like relevant sound effects, and speaker identification, and any other non-speech elements. And once you have your transcript, you can upload it along with your video on

YouTube and click Set Timings. And this actually allows you to create a time-coded caption file without having to do any of the time coding yourself, which can be a huge pain. There is also a lot of room for sync errors, if you are setting timings manually.

You can also start with YouTube's automatic captions, which are not accurate enough for compliance by any means, or for really being able to understand whats going on in the video, but they do allow you to edit the audio captions in the interface itself. So you could also start with the YouTube automatic captions and edit them there.

And then, once you have a timed caption file from YouTube, if you need a different format for another video platform, you can use a caption format converter. There are several out there. We have a free caption format converter on our website at 3Play Media. But that allows you to create a caption format that might be otherwise very difficult to create by scratch, because not all of them are that simple to read or that understandable to the average viewer.

When you are creating captions from scratch, it's important to follow best practices for caption quality. There are some standards out there, although there are no real legal guidelines, except for the FCC's caption quality standards. But there are many standards out there. The DCMP Caption Key is a great resource.

But some standards that we feel are really important-- spelling should be at least 99% accurate. And you should include grammar, for readability. Speaker identification should be consistent. So you don't want to use Speaker One in some places, and the name of the speaker in another place. You want to include relevant sound effect, as I've said several times. You should use punctuation to improve readability.

And then, there are two different options-- there's verbatim transcription and clean read transcription. So if you're transcribing something that was scripted, you want to make sure you're using verbatim transcription, because every um or other insertion was scripted purposefully, and you want to make

sure you capture all of those.

Clean read is something you would use-- for example, this presentation, I have noticed that I've said "um" several times. And it's easier for someone to read a transcript or a caption file that doesn't take into account all of those word insertions. You want to make sure that you have one to three lines per caption frame, with no more than 32 characters per line. It's best to use a non-serif font format.

The sync should be perfect, so it should be time coded exactly. No caption should be less than one second in duration. The placement of the caption frame should not obscure other important visual information onscreen. And you should note things like silence or music playing. And those frames can drop off the screen after a few seconds.

I mentioned automatic captions. I want to talk for a second about accuracy rates. When we talk about transcription, accuracy is really, really important. So we expect continuous improvements in automatic speech recognition. But currently, automatic speech recognition accuracy rates are anywhere from 50% to 80% accurate, depending on the technology, the background noise, the audio quality. And even 95% accuracy isn't sufficient for accurately conveying complex material.

So if you look at this chart, you can see that inaccuracy compounds over time. So if you look at a typical sentence of eight words, a 95% accuracy rate means there will be an error every two and a half sentences. And word-to-word accuracy issues really multiply the chance of each word in a sentence being incorrect. So you can see here how those accuracy issues compound as you get from each word to the next word.

I wanted to go through some common ASR errors. So I'm going to play a short clip of an automatically created transcript, along with the original audio. Then, I'm going to point out some of the errors and why they came up in automatic speech recognition. So I'm going to go ahead and play this quickly.

[AUDIO PLAYBACK]

- One of the most challenging aspects of choosing a career is simply determining where our interests lie. And 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'm 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 wild that he had introduced to me.

[END PLAYBACK]

LILY BOND: So there were several ASR errors in there, automatic speech recognition errors. Some of the things to point out are punctuation. So the automatic speech recognition didn't take into account punctuation, hesitation words, word removed. Speaker changes were not captured. There's no speaker identification. No non-speech elements were captured.

And then, there are a lot of acoustic errors. So for instance, one of the speakers said New England Aquarium. And the speech recognition generated "new wing of the Koran," which actually sounds similar. And you can see where that automatic speech recognition error came from, but a human wouldn't have made that error. Similarly, one of the speakers said

forester, where the real word was four story.

So a lot of these acoustic errors come up. And a lot of important clarifying words also end up with errors. For instance, did and didn't make a huge difference in a sentence. But a lot of times, speech recognition won't pick up on that type of error.

OWEN EDWARDS: And then, as far as creating audio description, there are fewer platforms and fewer systems in place, really because audio description is less well-understood, less common, and it's such a growing technology. Really, what we're emphasizing when we see people creating their own original video content, we very much encourage people to include description in the video production stage. It's by far the least expensive way to create description. And also, it can be tied in with the production in a way that weaves it through the actual content, the audio of that video, while still potentially being something that could be turned on and off.

And another way to describe video that isn't being produced, isn't new production, really, there are multiple steps there. And that requires somebody to sit down and go through and write a script of description of the most important parts of the video that need to be described, and then to align that into the timeline. It then needs to be edited to make sure that it will fit within available pauses, quiet areas of that video. And then, a voice artist needs to voice it. And it's then mixed with the original audio of that video.

A third way that we have seen people creating description is actually to create a separate version of the video where they've added extra time to that description. In some cases, there isn't enough time available to describe the key important features, and it's necessary to actually create a separate cut of the video which extends those pauses, or provides some additional time to put in that description.

One of the areas where-- it's relatively rare that people, at this point, are actually supporting the ability to turn on and off description. But right here, we're going to look at an example of a video that has had audio description added to it. First of all, we'll go through and hear the introduction to this video

with no description. And people can experience it. Maybe they look away from the screen and imagine what it's like to hear this introduction, purely through the audio, without vision. And then, we'll turn on the description that's been added and see how that changes the experience of it.

So I'm to go ahead and play that video. Whoops. There we go. We're going to go ahead and play that video. And as I say, it starts off without description. And then, we'll see it again with the description.

Ah. Come back. Here we go.

[VIDEO PLAYBACK]

- At the left, we can see a-- the-- we can see-- at the right, we can see the-- the hat snarlers. And everything is safe, perfectly safe. You know, evil.

[SWOOSHING]

Are you hurt?

[END PLAYBACK]

OWEN EDWARDS: So from just that audio track, it's hard to understand it. Now we're turning on the description, and we'll replay that same section.

[VIDEO PLAYBACK]

- Titles appear under the rippling water that floods a grungy tile floor. The Orange Open Movie Project presents "Elephant's Dream." An old man's face is reflected in the water.

- At the left, we can see, though we can see. At the right, we can see the

hat snarlers. Everything is safe, perfectly safe. You know, evil.

- He looks up.

- The old man shoves young Emo to the floor of a bridge, as a flurry thick black cables zips by. The cables plug themselves into banks of sockets that line the walls of a steel chasm. The ever-changing web of cables moves on.

- Are you hurt?

[END PLAYBACK]

OWEN EDWARDS: Hopefully, that shows how much easier it is to understand a video where you can't see the visual content, by having that additional description added into it. But it's a complicated process to add that audio description in. So the question is, what's good enough for description? And at this point, there are no specific standards included in WCAG, coming from the FCC, or within CVAA regulation.

Guidelines exist, especially the DCMP's description key, similar to the captioning key that Lily mentioned. And there are a number of good description companies that have internal best practices. So the description coming out of those companies is a very good quality product.

But what we're watching-- in the same way that that automatic speech recognition for captioning is not producing good enough quality, in many cases-- is that, since the CVAA requirement that broadcast television will increasingly have the audio description, there is the potential that there may be lawsuits in the future related to the definition of what good enough is and that that will drive a board of standardization of what will be good enough in audio description.

LILY BOND: Great. So now that we've covered the how and the why of captioning and describing, we're going to talk a little bit about where you actually publish captions and description for online video. For captioning-- captioning functionality is built in to most video players across devices at this point in time, although it did originate as a mandate for broadcast television.

Most players and platforms have caption compatibility. Some are more advanced and allow for things like placement of caption frames. There are FCC requirements for caption control on video platforms. So most video platforms will allow you to change the display of your captions, like the color and the size, whether or not they are outlined, whether they have a background, that kind of thing.

The main limitation we see with closed captioning is for social video. So there are social platforms that don't have closed captioning functionality built in, like things like Instagram and Snapchat. Some of these more social video sharing platforms.

There are a number of ways to publish captions. The most common is probably what we call a side car file. So this is the way you would publish captions to most video platforms and players like YouTube, Vimeo, BrightCove, Kaltura, most of the major online video platforms. You would upload a caption file to associate with your video, and that'll play along with the video. And that allows you to turn the captions on and off using the CC icon. You can also encode captions onto the video. You would use this for things like kiosks or offline video. And so a lot of the kiosks at airports, for instance, would have encoded captions that are directly encoded into the video itself.

And finally, open captions are actually burned directly into the video and can't be turned on or off. And so you would want to use that, if you're distributing a single video asset that you want to make sure has the caption functionality built in. And if you're using an integration between a captioning vendor and a video platform or player, all of this becomes trivial because, for the most part, video players and captioning companies that integrate allow the captions to just post back directly, without having to download and upload caption

formats.

But if you are dealing with caption formats, I wanted to give a little insight into what they look like. So I mentioned the potential of converting caption formats instead of doing it yourself. You would definitely want to do that if you're using something like an SCC caption format. So SRT is a very common web-based format. It's used for YouTube and other players, like Wistia. And it's pretty easy to understand.

So on the right, we have an SRT file, which shows the caption frame number 1, followed by the beginning and end time codes-- here, it's zero seconds and four seconds-- and then the text contained within that caption frame during those time codes. And on the left is an SCC file, which contains the captions for the exact same file, but it uses hex codes. So SCC is a lot more difficult to understand and create from scratch, which is why we highly recommend converting caption formats. But SCC is required for a lot of broadcast use cases and for some online video platforms.

OWEN EDWARDS: And then, as far as which players support description, right now, relatively few of the large player platforms support the idea of a second audio track for description. That demonstration I gave you a while ago included a way to turn off, switch over to a different audio track, which included that description.

But relatively few other platforms support that right now. It's coming. It's increasing. But what we're seeing right now, the easiest way to implement it, is to create two copies of the video. One is the original one, and then one has the audio description instead of the audio track. And in some ways, that's a little bit like where Lily mentioned about open captions burnt into a video. You're creating a second copy of a video that has those descriptions built in to them. We already have a description track added in.

Some players do support the idea of a text track description where it's really a text track handled like captions, with the idea that a screen reader, which is the device, or the assistive technology, that people who are blind or low vision may be using to access websites and web content. The idea was that that would then lead it out.

So far, we're seeing issues with that. That's really a technology that isn't fully mature, to depend on the screen reader to read that out, because it's very hard to control whether that read text overlaps with speech in the original video content.

And we have seen some text-only merged transcripts. And another exciting new area and, particularly, 3Play Media has announced just earlier this month, an audio description plugin which allows description to be added as a supplement to existing platforms, to things like a YouTube video, a Brightcove Video, or various other videos. If the platform itself doesn't support description, an additional plugin can be added to a web page to make that description available, to allow the user to control it, turn it on and off to adjust the volume. So there are some exciting features that are coming through, particularly around that idea of plugin.

Once we start talking about the need for audio description, the need for people who aren't just expecting captions as the accessibility accommodation for video, then there are some other issues that come up, which we, at SSB, have a lot of experience around web accessibility, in general, the accessibility of the content and controls within web pages, as well as within other software, mobile apps, but particularly within online video.

There are people using screen readers that I mentioned, who are blind or low vision, but there are also people with mobility issues who are using things like Dragon Naturally Speaking to control their interaction with a web page, or with a video player, or maybe use keyboard-only access, people who are unable to use a mouse, who may just use keyboard-only access to control the video playback. And so this has implications for the video player itself, that it is accessible to all those different areas.

And I say keyboard-only access, low vision people, who need to go to see the controls, people who use screen readers, people who use voice input. And then, how do you turn on that audio description? That it's a separate track or switchover to that other video, if it's a separate video.

And one note that I've added is the idea that autoplay is a very big problem,

particularly for screen reader users. Earlier, we mentioned it's becoming more and more common on social platforms. Videos will autoplay when they scroll into view, but they typically autoplay with the audio muted. The main reason for that, or a big reason for that, is that, for screen reader users, if audio plays automatically when you get to a page, it's very difficult to navigate to those controls within the page to pause the video. And so that's actually a violation of WCAG, to have an automatic play, without an easy way of getting to control that video and pause it to mute it. If it plays automatically muted, then that's not an issue.

So this is just a general list of some players that are out there. It's not a comprehensive list, but there are a couple that are available which really showcase all of the accessibility requirements. The Able Player that I showed on the previous page and the OzPlayer have a lot of features in there for accessibility. But we're not seeing them as widely adopted as those more common platforms like YouTube, Kaltura, JW Player, Brightcove, and Akemi Player, a couple of other smaller players.

Those players are increasing their accessibility, but some of them rely on or don't have the specific support for a description track. And that's where, as I mentioned, something like a plug-in or an add-on may be necessary to provide that description track, now that WCAG 2 level AA requires an audio description track rather than just a transcript.

LILY BOND: So before we get to questions, I like ending on this quote, because we've talked a lot about the specifics of creating captions and descriptions. But I think the reason for all of this is to remember that there's still a lack of access. As we said earlier, there are 48 million Americans living with hearing loss and 21 million living with visual disabilities.

And the first goal is to expand and guarantee access to all individuals with disabilities. And so this quote from the Department of Justice really speaks to that. And it says, "Access to information and communication technologies is increasingly becoming the gateway civil rights issue for individuals with disabilities." And it's true. We have a long way to go.

Let's open it up for questions. To start out, Owen, this question is for you. Are there instances where you can describe everything in the video itself, or where a description isn't necessary, or a description is not even possible?

OWEN EDWARDS: Right, that's a great question. So WCAG itself specifically points out that-- and I'm going to quote directly here-- if all of the information in the video track is already provided in the audio track, no audio description is necessary. There are certainly videos that are created. And in fact, we, at SSB, create our own training videos where, in the production process, we intend for that main audio track to describe everything that's important in the video content, so that there's no need for supplemental description. The video is considered self-described.

So that's certainly possible. And there aren't clear guidelines on in which situation that is considered acceptable. We can certainly give guidance on that, on a case-by-case basis. But it's really a matter of considering what of the content-- is there something that would be missed, if you couldn't see the content?

And then, the separate case are the situations where you can't add description successfully, or there isn't a possibility of doing it. There are certainly videos that don't have gaps in the audio to insert that additional description, and there isn't a clear way to deal with that situation.

And WCAG added the level at level AA. At level AAA, there's a feature called extended description where the video can be paused. The description happens, and then the video resumes. It's been a little confusing, but that's considered a triple AAA requirement because, really, it's a solution to a AA problem. The AA problem is that video that has too much speech for there to be spaces.

So in order to address that, as I mentioned earlier, we see people making a second cut of the video where they're essentially performing that extended prescription, where they pause the video content, or have some extra background video that can be spoken of. And we haven't seen specific cases where something's been highlighted where it wasn't described. And the

argument was, well, it's not possible to describe it. But that potentially could be a legal liability if something was not described.

LILY BOND: Great. Thanks, Owen. Someone is asking, you mentioned Facebook, how does captioning increase views for Facebook and other social media platforms? So Facebook started implementing captioning just over a year ago. And they have done a little bit of research into how captions have impacted viewer engagement, and they found a couple of interesting things.

One was that the vast majority of people did not like videos audio playing with sound, which is why they turned the sound off, although they also, obviously, were in violation of the WCAG requirements, as Owen said. But they also found that adding captions increased viewer engagement by over 11%, I believe. I will find that exact stat and send it along afterwards.

But when videos do autoplay on your newsfeed without sound, captions really help get viewers engaged in your video and draw them into something that they otherwise would likely just skip passed. So captions are actually really, really important for getting that engagement on Facebook.

Another question here is, can you share some best practices for audio description?

OWEN EDWARDS: Right. So again, I would refer to the DCMP description key, which really breaks down how description should be created in terms of what needs to be described and what style should be used. In general, if there's onscreen text, that needs to be read out if it isn't included in the soundtrack. But that's really a matter of, is it text which is there to convey information? It wouldn't be necessary, for example, if somebody was speaking and there was a road sign. But it would be necessary if the name of a speaker and maybe their job title appeared on the screen. So the DCMP description key is a great open guideline to the best practices around description.

LILY BOND: Great. Thanks, Owen. Someone else is asking, can these techniques work for videos that have already been posted to sites like YouTube and Vimeo? That's a great question. So for captioning, you can always add captions after you've published a video, particularly to video players like YouTube and

Viemo. Both platforms allow you to upload a caption file, so it'll just be associated with your video once you publish that caption file.

And it will never republish your video or require you to publish a new video. Audio description, it's kind of dependent on how you choose to publish. YouTube and Vimdo do not really allow for a secondary audio track anyway. But if you are using something like a captions plug-in-- sorry, an audio description plug-in, you could certainly add that for a video that already exists on YouTube. Owen, I don't know if you have any additional thoughts on the description side of things there?

OWEN EDWARDS: I mean, certainly, there isn't a mechanism to add description onto an existing video, except for these new plug-ins that are coming along, like 3Play's. There is also a research platform called You Describe, which allows people to describe, in fact, including by extended description that I described.

So that's youdescribe.org. And that is a great way to add description to existing video. We're not seeing that widely adopted in production systems right now, but it does give ways to describe it. And then the 3play plugin allows description of the video that exists.

LILY BOND: Great. Thanks, Owen. Someone else is asking if we are concerned about SEO. And we have a video with no voiceover and only music, would adding audio description help improve our SEO? So that's a great question. And it really depends on how you are publishing that audio description.

So SEO, or search engine optimization, really draws from text. And if you are publishing audio description, for the most part, it's another audio file, which doesn't contain a text alternative. So Google wouldn't read the audio track, just like it wouldn't read the audio of a video with spoken word.

You could certainly publish a text version of the audio description in the description of your video, or for the very few players that have the ability to use a web VTT description track, that would also help with SEO. But the use case is really small there. And I think that the main SEO benefit from the description would come from adding it to the description of your video itself in the video player. Owen, is that right?

OWEN EDWARDS: Right. I mean, that's a really good point. And that's part of why there's been a lot of discussion around the idea of whether description is better done as an audio track or as a text track. But inherently, the people who are looking to consume it, looking to get the benefit of it, want to hear it as audio. So it's usually a recorded audio track. There have been ways to do it with a web VTT track that's spoken by the screen reader. I touched on some limitations around that, but it would have SEO benefits.

So really, if you're concerned about that, I think really the best solution there would be to include an audio description track as a piece of recorded audio, but to also include the transcript, which combines both the captioning and text version of that audio description track.

LILY BOND: Thanks, Owen. Someone else is asking, in higher education, if a professor uses content that is not their own and does not include captions, if we contract a vendor, what are the legal considerations of adapting someone else's videos? What are the necessary steps to get such videos captioned?

So that's a great question, and it's something that a lot of people are concerned with is, are they violating copyright law by republishing a video so that they can include captions and make it accessible? I will say that captioning videos for accessibility purposes in higher education is often considered fair use, but you would want to consult with your legal counsel at your organization to make sure that they are comfortable with that.

However, if you're using a vendor, some vendors-- and certainly 3Play Media, we actually provide a captions plugin, which allows you to add captions without having to republish or edit the original video at all. It's just a very simple-- one line embed code that includes both the YouTube embed and the caption embed. And it'll play the captions along with the video.

It does work for several other video players as well. And that allows you to publish captions to a video to make it accessible without worrying about copyright. And the audio description plug-in would be very, very similar. So it would allow you to publish description to a video without having to republish

and get in trouble with copyright law.

Another question here is-- sorry, just looking through them quickly. Owen, maybe you can speak to this. Is automatic playing of anything on screen an issue for cognitive disabilities?

OWEN EDWARDS: Right. Right. That's a great question. And absolutely, there is. If you talk to people who produce videos, or in general, create websites, of course, something that is intended to grab your attention, which is typically advertising, is often put in there intentionally to get somebody's attention, but can cause problems with people with cognitive disabilities.

The reason that it isn't specifically called out in WCAG is that WCAG 2.0 itself doesn't have a huge focus on cognitive disabilities. And so that isn't necessarily highlighted. But that's right. Autoplay is a very distracting issue. It's something we really discourage site creators from doing. And a future version of WCAG may well call that out specifically around cognitive disabilities. There is an update to WCAG in development, which may well highlight that.

LILY BOND: Great. Thanks, Owen. And I think we have time for one final question. Someone is asking, what is the acceptable accuracy rate or quality requirements for educational video content? For captioning, the generally assumed accuracy rate is 99% accurate or higher.

As I said earlier, errors really start to compound when you get below 99% accurate. And the first errors to go are really clarifying words, like did and didn't, which can be a big issue for educational content, when students are relying on those captions to understand and to learn. So 99% or higher is the generally assumed rate for captioning, although I will say that there are no clearly specified accuracy rates in any of the laws, which is something that we are hoping for more clarity on. Owen, do you want to speak to the quality requirements for audio description?

OWEN EDWARDS: Sure. And really, that there aren't clear guidelines, or actually really clear ways to measure quality of audio description. There are certainly things to avoid, which are things like stepping on the audio in the main video content,

particularly dialogue. And there are things to avoid, like giving away things that are coming up in the video. But there haven't so far been test cases that say this kind of description isn't good enough. So really, we're recommending that people go to reputable lenders of description where they have a quality level that certainly exceeds those requirements.

LILY BOND: Great. Thank you, Owen, so much, for taking the time to present with us. Your knowledge and expertise is much appreciated. So thank you, for being here.

OWEN EDWARDS: Sure. And if I can, I noticed a couple of people have been asking about the DCMP description key. The website for that is wwwdescriptionkey.org. That's all one word, descriptionkey, and I'll post that into the chat window.

LILY BOND: Perfect. And we will include that resource in the follow-up email as well, so that everyone receives that, along with the recording and side doc. And thank you, everyone, for being here and for asking such great questions. And we will see you next time. Have a great day.