

JOSH MILLER: All right. Thanks for joining us today on this panel on Implementing Accessible Video Captioning. My name is Josh Miller. I'm a co-founder of 3Play Media. We provide services for premium transcription, closed captioning, subtitling. We also view web captioning as an opportunity to kind of supercharge the viewer experience through some interactive search tools and things that are all based on this notion of having synchronized text.

We have a great panel here. We have some true web video experts who work with a number of different types of content and technology, and they've all had different ways of tackling the video accessibility challenge. So we've got Bill McCarty from Dell, Ali Daniali from T-Mobile USA, and Brad Ellis from Google and YouTube.

So I'm going to start by just going through a very quick overview of some of the applicable accessibility legislation that a number of organizations are dealing with today, and then we'll have a discussion with our panelists. And we'll certainly have time for questions from everyone.

So Section 508 is a fairly broad law that requires all federal electronic and information technology to be accessible to people with disabilities, including employees and the public. For video, this means that captions must be added to that content, whereas audio podcasts and audio files, a transcript is actually sufficient.

Section 504 entitles people with disabilities to equal access to any program or activity that receives federal subsidy. Web-based communications for educational institutions and government agencies are covered by this. And Section 504 and 508 are both from the Rehabilitation Act of 1973, although Section 508 wasn't added until the mid-'80s. And many states have actually enacted similar legislation to these two bills, but they're often called something different. But the wording is often very similar.

Next is the Americans with Disabilities Act from 1990. This covers federal, state, and local jurisdictions. It applies to a wide range of domains, including employment, public entities, telecommunications, and places of public accommodation. The ADA of 2008 broadened the definition of disability to actually be more in line with Section 504. So it actually is covering more situations.

The ADA is interesting because that's the law that was cited in the recent lawsuit against

Netflix. Netflix argued that the ADA applied only to physical places of public accommodation. But the ruling was that Netflix, because it's a global or has global reach, it is a place of public accommodation, even though it's this web-based entity. And so that in itself has some pretty profound implications for people publishing content online now. So now the broader your reach, which is obviously a good thing, there are other considerations that come with that.

The one that probably gets a lot more attention now is the 21st Century Communications and Video Accessibility Act, which is often referred to as the CVAA. This law, which was passed in October of 2010, expands the closed caption requirements basically that were originally set out for broadcast television to be expanded to the web. So basically any content that aired with captions on a network on television now also have to have captions on the web.

And there are a number of milestones that go along with that. More recently, a couple of big milestones were passed. One is now, going forward, anything that aired on television going online must have captions, assuming it had captions on broadcast, which most content does. So that is already in effect.

What also passes this idea of any clips. Sometimes you see shorter clips of videos that were online, say, on Hulu or YouTube or any of these distribution networks, you see these options to see a clip of a show, that also now has to have captions. There are some exceptions to that.

We actually just put out a brief on our website. It's free. If you're interested in more details of these milestones, definitely take a look. I put a link up here for it. Basically, the idea of, say, outtakes or edited scenes, those are not being included in this clips milestone yet. So there are some intricacies to the different milestones and then a number of upcoming milestones having to do with archive content.

So this is some data from the WHO. It's a 2011 report on disability, and it states that more than 1 billion people in the world today have a disability, and nearly 1 in 5 Americans age 12 or older experience hearing loss severe enough to interfere with day-to-day communication. So the interesting inclusion here is that the number of people requiring accessibility accommodations is actually really quickly growing, certainly relative to population growth.

So that's something to pay attention to. And of course, we kind of asked why. Well, a lot of it has to do with medical advances. People are able to survive a premature birth. People can survive accidents, and we certainly have an aging population. And these are all very good

things, but sometimes there are other effects of that that we need to pay attention to.

Another thing is we've been at war for now over the better part of the decade. So modern armor, for examples, allow soldiers to actually survive an injury at a rate of 10 times more than before. Again, this is really good stuff, but there are some consequences that the can come with that. Really what this means is accessibility is a pretty important issue and is something that needs to be considered as we're publishing content online.

There are also a number of benefits to adding captions and text with video content, especially the fact that the internet is a text-based entity in many ways. I'm not going to go into too much detail right now because I think we'll get a better view from our panelists. So let's dive into the panel.

I'm going to ask each person to introduce yourself by telling us the type of content you work with, the main audience for that content, who's paying for that content to be made available, and how much of the content, if you have any data you can share, is getting captioned. So, Bill, you want to kick us off?

BILL MCCARTY: Sure. Bill McCarty. I'm with the Dell Software Group, which is a business unit within the larger Dell organization. And we're primarily creating marcom content that's designed for business-to-business users. So we're educating people about our products, benefits, that sort of thing. It's largely pre-sales and then some post-sales as well for training type things. What were your other questions there?

JOSH MILLER: So you mentioned the type of content and the audience. So who's paying for the content to be produced?

BILL MCCARTY: Oh, yeah. We're paying for the content ourselves, so it's a part of that pre-sales marketing kind of thing. And then in terms of the amount of content we're having captured, we transcribe everything now and expose everything for captioning.

We've been doing that for about a year and a half now or so. Some of our legacy content is not yet transcribed, but everything new that we create has been transcribed for the last 18 months to 24 months.

JOSH MILLER: Great. And Ali.

ALI DANIALI: Hi. My name is Ali Daniali. I'm with T-Mobile USA. I'm part of the online communications team

that's within the corporate communications group. I work with a lot of internal content and videos specifically, and 100% of that content is captioned. And the reason for that for us is a large portion of our audience is our front-line employees in the stores. And they don't have speakers on their machines.

So the way that they connect to our intranet where most of these videos live-- without speakers, they wouldn't be able to really get and understand what's going on. So content from leadership, content that we create in house through our studio, user-generated content that's captured from our employees, everything goes through a funnel of getting closed caption and then put on our intranet. And all that cost is really part of the budget of corporate communications because we want to give the best experience to our employees of being able to consume the content. We know that they're using it because they do watch it.

JOSH MILLER: Great. Brad.

BRAD ELLIS: Hi. My name is Brad. I'm a product manager at YouTube, and I work on captions. So the kind of content we create in our audience is very broad. And we don't actually create the content, which is an interesting position to be in.

So the captions team at YouTube doesn't actually go in and type in captions for anything, but we build a platform that allows anybody to upload captions in 20-plus different formats, if they've created them for their content already, and then display those captions on all of our players, on all YouTube players.

And then we also build tools for people who are creating captions for their content on their own to easily and quickly create captions for their videos. And our goal is simply to make every video understandable to every user. A very long-term goal, but that's what we're aspiring to.

JOSH MILLER: Great. Thank you. So, Bill, for Dell and what was Qwest before, what's been driving this initiative to create accessible video?

BILL MCCARTY: Initially, I was looking at the 508 requirement, and that was one of the major drivers for making sure that we had accessible videos on the platform. We have a video gallery that we pushed out through qwest.com, which we call Qwest TV, and that was going to be a central point where all of our customers could go and find the videos they want to see.

So 508 was definitely one of the biggest initiatives to begin with, but then also is based around SEO qualities of having transcriptions available and making the videos searchable. So that

was a really big one. We obviously want to have our videos be able to be found by people and then drive traffic back to our site. So the transcription part of that really has become more of the main focus now that we have the captioning available for the 508 compliance.

JOSH MILLER: Great. And Ali, you mentioned the fact that people really can't hear the content in many cases. Was there a legislation piece to any of that? Or was it just pure function?

ALI DANIALI: To my knowledge, it was all pure user experience and the functionality of giving folks that-- EITs decided that they can't have a speaker. So we kind of worked with our constraints and wanted to make sure that whenever an employee that was in front of the internet, they had all the opportunities. And we considered a couple of options of just providing a transcript and having the video. But then that became kind of-- it didn't seem the right solution with the amount of space we had.

So early on, this was about-- we started doing this almost two years, three years ago. It was decided that we were going to just provide closed captions for all the video that was going to be put on our intranet. The amount of video was limited at first, but as time has gone, we've added more video, more types from different sources, especially as our current executives have a lot more appetite for video for consumption of internal use. So we're captioning everything.

The piece that's really helpful for us is with our online platform, and 3Play, they're very integrated in being able to submit via FTP and then have it appear on our online platform with closed captions. And the player's already there. So it's a great workflow.

JOSH MILLER: Great. So Brad, as a platform-- we had an image up here. Let's go back to that for a second. YouTube been ahead of the curve in terms of accessibility support for a while. Where is that coming from? What's driving that?

BRAD ELLIS: I think it all starts with Google's mission to make all information universally accessible. And I think Bill mentioned, too, that video it's important to be able to search, and transcripts help with that. So that helps us-- it helps search. It helps people find the content they're looking for.

But I think that the team that I work with is very motivated, just intrinsically motivated to help people understand all videos. And it's not just deaf or hard-of-hearing viewers, but second-language speakers. Translating captions into foreign languages makes the video understandable to the rest of the world.

I think Ali's example, where you have devices that don't have speakers, I watch captions at home if somebody is sleeping, and I can't turn the volume on. Or I'm on a train, and I don't have headphones, and I want to watch something, I need captions. So it helps everybody, and I think that's the primary motivation for the team and why we put so much effort into building technology.

JOSH MILLER: Yeah. And as someone who is now obsessed with closed captions because it allows him and his wife to watch TV while their newborn sleeps, so little things like that you never really think about, yeah.

BILL MCCARTY: It's interesting the secondary benefit, like you say, is that one of the things we've been able to do also is to play our videos on our lobbies in several of our offices. And secondary benefit is we have these transcriptions and now these captions so people can watch them because we don't have the volume up. We have multiple monitors playing different videos, which would conflict with one another. So, yeah, a lot of-- it's just interesting how you can find ways to do that.

JOSH MILLER: Yeah, it's great. Bill, you've actually been through a few iterations of the production process to publish video, to publish captions, and really getting that process down. Could you tell us a little bit about what that's been like and how accessibility in captioning has been considered in different parts of that?

BILL MCCARTY: Yeah. Initially, we began by captioning content that we're already had done. But quickly you learn that you need to build the transcription and the captioning process into your production workflow as well, which is really something to think about. So you want to make sure that, A, like Ali is doing, build some of that budget into your process so that you can do the closed captioning and then think about where your video is going to be deployed in terms of do you have localization initiatives, for example.

So you're going to want to make sure that you potentially build in budget for translation after the fact in closed captioning. And then one of the things we also discovered is how you need to think about what your on-screen content is going to be as well. A lot of people would think that a lot of callouts, text callouts, that sort of thing, are important to build it into a video. And they can be helpful, but that can also be very limiting, especially when you look at localization initiatives, because that's burned-in content that you can no longer translate.

So we found that we try to keep the on-screen text to a minimum and then do all of our translation transcription through the voice narration and stuff. So it's been an interesting process. And when you start to think about it, you build it into your production process, and it makes it very helpful.

JOSH MILLER: Yeah, definitely. And you remember, when we were talking, you mentioned something about how production quality is also something that you think about more than probably you did before. How does that factor in?

BILL MCCARTY: In terms of the production value that we build into the videos?

JOSH MILLER: Or even just the recording quality. So we were doing mic checks, things like that.

BILL MCCARTY: Yes.

JOSH MILLER: That's something that I think a lot of people don't think about that you've thought about.

BILL MCCARTY: That is an excellent point, yes. So, yeah, it's important that you be able to hear what's going on. And that's not just for the audience, but also for the transcription service you're going to use as well.

So you want to make sure that, yes, you are using good-quality microphone, people can understand what the person is saying. We do a number of product videos that rely on our product experts, and we've had to go back and kind of say, well, some of the product experts speak with such a heavy accent that maybe they're not the best person to do this presentation. So it's kind of the trade-offs that you have to make because you do have to be able to have that content transcribed effectively and then serve it up to your audience.

JOSH MILLER: Definitely. And that actually gets into something really interesting that YouTube does, which is this idea of auto captions, right? So if the sound quality is no good, guess what? Your auto captions are going to be pretty comical. But you've been doing a fair amount of work on making that more usable. You want to talk a little bit about what's going on with that?

BRAD ELLIS: Yeah, I'll start out. How many people here have watched a YouTube video with automatic options on? How many people thought that, oh my god, this is totally understandable and makes perfect sense, right?

So we know there are issues. But going back to our long-term goal of making every video

understandable to every user, technology is the only way that we can scale. We have over 80 hours of video uploaded every minute to YouTube. And to expect-- we can't hire an army of people to caption every single video that's there. It's ultimately up to the content creator.

And if the video uploader did not add captions themselves, we do the best we can to help somebody who needs captions understand what's happening and occasionally offering some comedy relief. So these are not actual automatic captions. If you look at this video and actually look at the automatic captions for this video, they're surprisingly good. But we do know that it's not perfect, and we have a lot of errors.

I just want to call out that we're always making improvements. And we had a big improvement in English automatic caption quality that we launched this summer, so every new video starting this summer. And then we've been rolling it out to old videos on the site. You should see a pretty stark difference between what you saw before and what you'll see now.

So this is just a pure before/after comparison where we went from "see pursue your project" to "super secret project." And we also support now identifying proper nouns, capitalization, and we're doing more work to make this better and better. But really I think that automatic captions are most powerful as a tool to help creators add 100% accurate captions.

We don't want it to stop with automatic captions. Hopefully, it will help people add high-quality captions to their video. And if the uploader did not do that, then it provides at least a minimum that we can improve over time.

JOSH MILLER: Yeah, makes sense. And I think Bill said it really well, that the clearer the person is speaking, it's easier for the transcriber, but it's also easier for the engine to actually be accurate. So there's a lot of ways that that can play out. Great. So Ali, you mentioned that you guys are captioning everything.

ALI DANIALI: Yes.

JOSH MILLER: Are you measuring the effect of that or the success of that at all?

ALI DANIALI: We measure different audiences with our online video platform. It provides us a lot of metrics on user engagement. And we also know what groups they're coming from like frontline versus our field service office, which do have speakers. And we notice that there's a lot more video being watched on the frontline. With the correlation that they don't have speakers and they are watching video, we just make the assumption that they actually are using the captioning

because it would be kind of useless for them without speakers.

So those are the kinds of correlations that my director wants to just know that we are reaching that certain group of viewers. And also we provide a lot of commenting within our intranet. So there's a lot of engagement outside of the actual player that we know that people are finishing it or actually engaged into it. That's a metric that we also captured.

JOSH MILLER: Interesting. And Brad, so maybe not metrics in terms of the success of captions, but you guys have some pretty interesting metrics in terms of why caption, in terms of the audience and who's watching videos. What kind of data points do you have that might startle people?

BRAD ELLIS: I would start out-- I had an interesting conversation with an uploader once, with a YouTube creator who said, one thing that YouTube needs to work on is getting outside of the US. You guys are only in America. But we actually have 80% of views on YouTube coming from outside of the United States.

And that's huge, and a lot of that is non-English. And captions are very important. Translating those is very important. It's a huge opportunity for growth. We see huge demand from non-English uploaders as well to get their content translated. So that's a huge motivator, yeah. I'd say the other surprising metric is going back to the 80 hours of video uploaded per minute. Thanks to automatic captions, we do have captions available for over 300 million videos, which I think would be very impractical to do without technology.

BILL MCCARTY: 3Play would probably like it, but you know.

JOSH MILLER: We'll do it tomorrow.

ALI DANIALI: Is that actually part of the search mechanism then? And whenever I'm searching for a specific video is the auto caption being actually used for that?

BRAD ELLIS: We don't use the automatic captions today. I hope that we will down the road. But, again, it's a trade-off with the quality. But if you upload captions yourself to any YouTube video, we do index that. That is searched.

We did an experiment with one partner a year ago and saw just by captioning videos in the same language-- they were English videos with English captions-- we did a scientific A/B test and saw a 4% increase in traffic in views and watch time on YouTube. And then imagine what that could be if you're making it accessible in more languages.

JOSH MILLER: Great. We like to use the analogy that a video without captions or text is the exact opposite of, say, a newspaper article when it comes to search. So a newspaper article is the best possible piece of content you can have because it's got your title, your byline, and the actual body of the article all automatically indexed because it's text.

Well, the video, if you don't tag it, if you don't put a title on it, if you don't do all the right metadata, and if you don't transcribe it, you've got none of that. That's kind of what a transcript can provide for a piece of video, is that it allows it to be more like a newspaper article and be indexed by a company that knows a little bit about search-- Google.

Great. I'm going to open it up to the questions from the audience. There's a microphone up front here. So maybe we can try to pass that around. And if it doesn't get to someone, I'll repeat the question. So we have any questions for our panel? In the back there.

AUDIENCE: What do you all see with people getting annoyed by inconsistency of the captions as far as crowd-sourcing corrections to those captions?

JOSH MILLER: So the question about audience annoyance in terms of crowd-sourced captions. Do you mean in--

AUDIENCE: No. More with automatic captions--

JOSH MILLER: Oh, oh, oh.

AUDIENCE: --that are machine generated and these errors. Is there any talk of giving users the ability to correct those errors?

JOSH MILLER: So that's interesting. So a question about giving a crowd the ability to correct the auto captions on YouTube, which I assume you mean if you don't own the video?

AUDIENCE: Right.

JOSH MILLER: Yeah.

BRAD ELLIS: So today we already make it really easy for the owner of the video to go in and make those corrections. So we have a lot of people who do do that. And we are definitely working on making it easier for other people to help do that.

We have some baby steps that we've taken in that direction, and we're working definitely more to make it easier to get other people to help you improve the quality of your captions as well at translate, yeah. So the short answer is yes. Not today, though.

JOSH MILLER: Other questions?

ALI DANIALI: I'd like to know how many of you are actually captioning for on any of your videos that you're doing? And how many of those are actually external facing versus internal captioning? External hands? Pretty much everybody is for external.

JOSH MILLER: Is anyone dealing with user-generated content? And are you captioning that?

AUDIENCE: We're both in education.

JOSH MILLER: OK, very interesting. That's cool.

AUDIENCE: It's all teacher-- faculty-generated content that's being generated.

JOSH MILLER: Yeah, faculty generated. That makes sense. Great.

BILL MCCARTY: Another follow-up question to that is so are you doing the captioning primarily for enablement of disabilities? Or is it also for the search engine optimization aspect of it? Any hands there in terms of search engine optimization? Yeah. Good. How are you embedding your transcripts? In the video pages? Or how are you incorporating that?

AUDIENCE: Well, we're using the--

JOSH MILLER: I think you're on yeah.

AUDIENCE: We're using the Kaltura platform, and that allows us-- that makes the content searchable. It makes it searchable at least through the Kaltura platform. It doesn't necessarily make it searchable through, say, a Google Search, though. But it is both for discoverability purposes and for accessibility.

ALI DANIALI: Nice. The previous session was about live streaming, and I asked a question about-- big panel-- about how many of them are actually doing captioning for live events. And they're all kind of like, well, no one's doing that. Anybody who's doing that? You guys are--

AUDIENCE: I actually have a question

JOSH MILLER: Yeah.

AUDIENCE: So we actually presented [INAUDIBLE] that this gentleman did a lot of work on. And basically, we ended up having to-- had to hard code the captioning in by using the decoder and then recapturing that and then restreaming that, which had a delay and everything else. And to our knowledge, there was no good solution to put the captioning into the stream that would then follow all the way down the pipe until it got to the player at the other end of the pipe.

So the only solution that really came up was to hard code it in and make an optional closed caption stream. So unlike broadcast, where you can on your set-top box turn it off or on, it was just either on or off, right? But with the lack of a solution, do you guys see that there is anything for live in that area? Because we can put the metadata in there and send it down the pipe, but if everything's stripping it on the way down or if no one pays attention to it, it doesn't do us any good, right? That's probably the reason why most people aren't doing it with live.

BRAD ELLIS: So we added support for out-of-band captions that you can do, and they're closed captions that you can turn on and off. So if you're watching the same screen, you can watch it with captions if you want. I think there's still a lot of work to do to make the entire integration from actually recording it to streaming it on YouTube work more smoothly.

Right now we work directly with the transcriptioning software so that as the transcriptionist is typing in the captions, they pipe that directly to YouTube. And then we distribute that.

AUDIENCE: And that's happening on the live product?

BRAD ELLIS: Yeah, yeah, for live streams.

AUDIENCE: OK.

BRAD ELLIS: If you just search for live streams captions YouTube, it's in our Help center, and there's more information.

JOSH MILLER: So YouTube probably has some better support for live captioning than other platforms. It's very platform dependent, though. Some of the captioning companies that offer live services will offer a little box that can be opened up kind of floating on the screen for those viewers. But I believe you have to send out a separate link for people to go to that page to open up that box. So it's still not perfect, but at least it's a solution where you don't have to worry about two versions of the video.

BRAD ELLIS: Some questions over there.

AUDIENCE: Just on that, on the live streaming component. Andrew Spaulding from Ooyala. We actually support captions with live streaming, and we used to do that with one of our customers here at BYU Broadcasting. So the CEA-608/708 metadata that you can actually put in the stream for broadcast, many of the hard ring coders will actually embed that into the live stream.

So you take that as part of the HOS straight into iOS, for example, and from iOS 5 up, supported the embedded 608/708 metadata for closed captions. And then there's other players on the web, so you can also convert that into the Flash player's metadata for delivery to the Flash players online as well. So it is possible to embed multiple languages into live streams.

BILL MCCARTY: Huh, pretty cool.

AUDIENCE: Question for Google. Do you support for live stream CEA-608/708 captions embedded in the stream?

BRAD ELLIS: Not today. I do think that's something we should do eventually. But right now, we just support captions that are sent directly from the transcriptionist to YouTube.

AUDIENCE: Gotcha, yeah.

BRAD ELLIS: But I think-- I mean, as I said earlier, I think there's a lot of room to make it very, very straightforward so that we can just integrate with any streaming platform.

AUDIENCE: The CEA-608 really is the broadcast standard. So if you want live streams from broadcasters, you pretty much got to support that.

BRAD ELLIS: Yeah.

JOSH MILLER: Great. So live is definitely one topic that comes up with captioning. The other one that I think is pretty relevant is mobile. Do you guys want to talk for a minute about how you're handling mobile?

ALI DANIALI: I guess I could talk about it since we're a carrier, and we kind of wonder why are we worried about not-- our frontline folks since we're a mobile company and they could use our device. And the big thing for internal video is security. A lot of the content is proprietary. A lot of it is

executive.

That whole challenge is something that we're working on to bring a stream that's being sent from our intranet, give them options so that it can be consumed, not only on the closed network, but also via their own device in a secure manner. I don't know if we'll then not provide captioning when it's something-- your own device. That's some question that we've got to figure out, but probably will include it because it's already part of the desktop version.

It's definitely top of mind. Every meeting we're talking about, well, why aren't we doing this versus this? And it's a lot of-- our IT and bandwidth infrastructure. One thing that I was very surprised with working at T-Mobile was whenever we wanted to send out a video to employee handsets, there's a very large sensitivity on putting lots of data on the public cellular network. It's not even allowed for T-Mobile for our own employees to send data because they don't want to mess up the public cellular network. So that's another challenge in itself.

JOSH MILLER: You want to talk a little bit about what YouTube does, too?

BRAD ELLIS: Yeah, so mobile, mobile, mobile, mobile. I just moved to San Francisco three weeks ago from Tokyo. I lived in Japan for the last six years. And I feel like from the moment I moved out there, everybody's like, mobile is really important. And here, this is like pre-- maybe like the iPhone just came out.

But everybody was like, mobile? Are you talking these little feature phones? You can't really watch a video on this. Nobody's going to watch a movie on their phone. And in Japan people already were doing that. They were still kind of crappy phones compared to what we have now.

And today, over half of our traffic in Japan and Korea is mobile already. And the US is catching up. We're going to be half some time soon. So mobile is really important and supporting captions on all devices. So mobile tablet, TV is something I think it's part of the requirements that Josh was talking about earlier. And so we've focus very hard on making sure that all of our platforms support captions.

I think that one of the difficulties is the fragmentation that we have in the markets. We have so many different phones and so many different versions of operating systems, applications, mobile web browsers, but we're seeing everybody catch up. And I think in the long term, I'm looking forward to the day where we say 100% of all places support captions, and everybody

will be able to watch a video with captions no matter where they watch it.

JOSH MILLER: Great. Other questions? Yeah?

AUDIENCE: Is there anything special that you do in terms of design or form factor for mobile captions versus for captions on TV or on PC?

BRAD ELLIS: I think that-- I'm going to take this. Design form factor, I mean, obviously, the size of the screen is different, so the default font size and whatnot we should take into account. All the settings look very different. You can't fit-- Josh had that screenshot up there of how we show all of the available settings on desktop.

I think there's room to make this prettier even on desktop, but imagine putting all of that on one screen on mobile. It gets pretty crazy. So there's a lot of adjustments that we do to make things easier on mobile devices. But I think ultimately our goal is to show captions as the creator, as the uploader intended them on all devices.

So positioning should work on all-- like if they position their captions to show speaker identification, then that should work on all devices, any color or anything-- anything else specified that a caption should work across all devices. And I don't think that really changes too much, depending on what device you're watching it on.

BILL MCCARTY: I'd be curious, the gentleman from Ooyala, correct? How do you address that? Is that a platform-specific thing? So how do you address that in the delivery of mobile captions?

AUDIENCE: First we take a DFXP file from someone like 3Play Media or Dotsub, for example, and we use that for our desktop platforms. For mobile, we'll actually transcode that and translate that into a web VTT format for HTML5, which is much more compatible across to a broader range of smartphones and tablets today.

And generally, we do provide the positioning capability that are embedded inside of the stream itself. So we don't provide our publishers' customers to customize if they want a blue background or a red background, but we provide our publishers the ability to change the positioning, the background color, the font, the font color, is it off the top of the screen, the bottom of the screen, what font are they using.

If they don't provide that, then the device itself often has guidelines. And then they will have some predefined setting. So if you were to watch a stream on an Android device with a native

video player or a stream on your iPhone or iPad and you turn on the closed captions, then there is a default that the system will use.

JOSH MILLER: And one thing to point out about what he's describing here is that that's an automatic thing that Ooyala does, is it automatically generates the mobile friendly version. YouTube does the same thing, but not all platforms do that automatically for you. So it's a nice little feature.

AUDIENCE: As a publisher here and a customer, this is one of the most frustrating aspects of publishing video, because a lot of it depends if you're using, for example, an Ooyala player that has an SDK that is intelligent. That's one thing if you're relying on the platform. Then it's kind of hit and miss.

We deliver HLS to Xbox, Roku, iOS, Android, and the captioning, what it looks like on the screen is very different. Apple does a very good job of handling CEA-608 and web VTT captions. It's a core part of their OS. iOS 7 is actually great. Android is-- no offense, Google. It's a mixed bag what you're going to get.

So it depends on what player you're using and what platform you're on. Because what's coming with-- you mentioned some of the legislation. Part of that legislation is giving the user control over what those captions look like on every single device.

And as a publisher, we're sitting there going, well, is that my job? Is that the platform? Who's the platform? Is that Ooyala's job? Is it Apple's job? Is it Google's job?

Right now there's a lot of finger pointing, and we're not sure whose job it is to make those captions look pretty and user configurable. And that's a requirement that takes effect this coming year. So no one's in compliance that I can see except Apple.

BRAD ELLIS: I can't speak about compliance and requirements, but our goal is to make sure that every YouTube player-- and you're right. It depends on the platform. If you're just sending HLS to a random platform, it depends on what they support. Any player that we control should have positioning, all of the formatting options, all of the user settings. But, yeah, I think everybody recognizes that with this huge range of devices and platforms out there it's hard, but I think in the long run it should get better.

JOSH MILLER: There was a question towards the back. He's got a question? Yeah.

AUDIENCE: For internal-facing content, is it safe to only caption as needed, as requested? That's one

question.

JOSH MILLER: Do we have a lawyer in the room? Laws are laws. They're there to be interpreted and to be enforced or not. So it depends on-- certainly if you have someone requesting it, now legally you absolutely have to do it or else they have quite a strong argument against you from just like an ADA standpoint. Otherwise, at some point, it's a user design, or a user experience decision.

AUDIENCE: And the second part of my question is for public-facing content, I understand that moving forward, you are captioning everything. But you also have content that's already there. How quick of a timeline do we have to make everything that we have public facing captioned?

JOSH MILLER: Well, CVA, so for television content, that's actually pretty clearly outlined as to what the rules are in terms of how old the content is and when it has to be captioned. So for entertainment content, that's basically been written. It's a great question for the situation where-- either Ali or Bill, have either of you gotten requests to have stuff captioned that wasn't captioned? Have you dealt with that ever?

ALI DANIALI: We've captioned everything, so we put it in our budget. We know that there are some secondary benefits to it, but we decided right from the beginning that's just part of having video, we're going to have captions. We had a lot of video in the past that we would burn on CDs or DVDs. But once we had an online video platform, everything moving forward was captioned.

I've seen some of the numbers, and it's not that expensive, honestly. It's a benefit that gives a consistent user experience. When people need it, they can use it. And if it's not their primary need, then they have it as a secondary need. That's at least for us so far how it's been.

BILL MCCARTY: That's a good point. I think we talked about that earlier, is that it's hard to find a good reason not to do it, to be quite honest with you. It is such a minimal cost, especially when compared with if you're doing high-end video production or even mid-range video production.

It just is something that is very easy to build into a budget. But to address that question, we do have a lot of legacy content that has not been transcribed or captioned. I haven't got any requests for it, so I don't really know the answer to that question. But you can bet I'm going to go back and find out because I do want to make sure that we're compliant.

JOSH MILLER: We've seen organizations who will literally look at the content in terms of views, and that's how

they prioritize it at least. So the content that gets viewed a lot, the legacy content that gets views a lot that's where they start. And they over time work through it.

AUDIENCE: That's a great idea. Yeah. Thank you.

BRAD ELLIS: I'd say that I hate to talk about compliance or requirements and all of this. Really, when it comes down to it, every creator, every YouTube creator that I've talked to about captions, I say, do you caption your videos? And they'll say, oh, I have people in my audience, my fans, some of them have emailed me and said I really want captions on your videos. Or I have a friend who is deaf, and I know I should caption my stuff for them, but it's so much work, or I don't have time. There's always an excuse.

And I really think that it's not as hard as people make it out to be. It can be really hard. Like if you-- there's a range, right? Everything from automatic captioning to every bell and whistle you can imagine. But perfection is the enemy of good in this case. Something is better than nothing.

Making video accessible to people who need captions is really important. And I want to encourage everybody who has power over this to make their videos accessible, to add captions, and to not focus on the excuses or the reasons why not to do or what's required, but how you can have the biggest impact and reach the most people.

JOSH MILLER: One thing that's actually a really interesting thing about YouTube, Google, when it comes to search results page rank, one of the factors that's taken into account is kind of this fuzzy idea of user experience. So time on site, all these things, the better the user experience for a particular website, that actually makes a difference in your search results.

It's not talked about a lot, because it's not as obvious, because it's not, say, text on the page that can be directly indexed. It's not as tangible, but that actually is taken into account in that algorithm. And captions is a great example of how you can improve the user experience because you can basically guarantee if they can't understand or hear the video and there are no captions, boom, they're gone. They're off the site. So it's little things like that that can actually start to make an interesting difference in your traffic.

ALI DANIALI: One thing--

JOSH MILLER: We have time for one-- yeah, go ahead.

ALI DANIALI: One thing that I want to mention, too, is after you've been doing captioning for several years, the question will come up as how do I archive all of this, especially if you've been using a video platform, an online video platform, where you've submitted all these files, video, and it's not kind of all flat. There comes a time, if you want to switch between video platforms, the whole question is how do I move this stuff? How do I archive it so that-- especially for a public company, you've got to keep everything together for discovery. So that's a challenge that we're having. And it'd be interesting if any of you have any sort of input on that.

BILL MCCARTY: Haven't run into that yet.

BRAD ELLIS: OK.

JOSH MILLER: Well, YouTube allows you to download the captions, which is probably unique for some platforms.

BRAD ELLIS: Yeah. The exact issue is?

ALI DANIALI: Just after it's been submitted to online video platform, whether it's YouTube, Brightcove, Ooyala, wherever, it lives on that platform. But after, let's say, five years, you need to archive this stuff, or you need to put it somewhere, or you need to move it from one platform to another, how do you get all this stuff moving to the right place so you can keep your captioning?

BRAD ELLIS: Our strategy at YouTube, we have an API so that you can add captions to your videos. You can download those captions in several different formats. If you go to youtube.com, you can download the captions for your videos in-- I think we have three different formats there. But through the API, you can do it in many more.

BILL MCCARTY: I think that's one of the benefits of using a third-party company like 3Play. That's not a plug. But in that they are the source for your captions and the APIs that companies have available, then Ooyala, Brightcove, whomever you happen to be working with should be able to ingest those captions and make them transportable between locations so you're not working with hard text files or files stored on a server or something.

JOSH MILLER: Yeah, that's great. We have time for one more question. All right. Yeah, go ahead.

AUDIENCE: Thank you. My name is Andreas Damianou from the UN Web TV at the UN. We have six official languages, and we try to have captions in as much readers as possible. But the amount

of content that we have is so much that the cost is tremendous.

And we tried to see if there's any kind of automated method to-- either through speech to text that we could be shown that could be accurate enough to be acceptable to everybody.

Especially in diplomacy, every words counts, and it could create diplomatic incidents in case it's translated incorrectly. Is there any way to automate this process to have so much content in different languages closed captioned?

BRAD ELLIS: YouTube supports automatic captions, speech-recognition captions in 10 different languages right now. But I think one part of your question was, is they're accurate--

AUDIENCE: Accurate.

BRAD ELLIS: --to the point where-- I think in an UN video, accuracy is very, very important.

AUDIENCE: Yeah.

BRAD ELLIS: I go back to what I said earlier that we see automatic captions as a bare minimum. Again, something is better than nothing when it comes to accessibility. But if you need 100% accuracy, it's more of a tool, a stepping stone to reduce the cost for creating those captions and not a replacement for it, today at least. Maybe in the long run, maybe some day.

JOSH MILLER: For what it's worth, that's what we-- we take speech recognition as the-- we use that as the first step in our process, and then we clean it up with humans. And there's not a single file that goes through our system where someone's expecting accurate captions where that clean-up process doesn't take place.

So we don't believe that you can get consistent accuracy without a human. Speech recognition is a great tool. It's a starting point. If you're looking for consistent accuracy, it's just not there yet.

BRAD ELLIS: Maybe 10 years from now 20 years?

JOSH MILLER: Yeah.

BRAD ELLIS: 50 years? Someday we'll get the *Star Trek* computer.

AUDIENCE: Thank you.

JOSH MILLER: Great. Well, I'm going to wrap up. Our panelists have been nice enough to offer their contact info if you have follow-up questions. Thanks for joining us today. Thank you to your panelists.

BRAD ELLIS: Thank you.

ALI DANIALI: Thank you.

[APPLAUSE]