

LILY BOND: Welcome, everyone, and thank you for joining this webinar, entitled Closed Captioning Online Video Clips for FCC Compliance. I'm Lily Bond from 3Play Media, and I'll be moderating today. I'm joined by my colleagues, Josh Miller, the VP of Business Development and one of the co-founders here at 3Play, and Andrew Schwartz, our Senior Research and Development Engineer.

We have 30 minutes slated for this presentation. Josh and Andrew will talk for about 20 minutes, and then we'll leave time for a Q&A at the end. For an agenda, Josh is going to start by giving an overview, and then he'll go through the FCC and CVAA legal requirements for captioning, followed by solutions for captioning online video clips. Then Andrew will do a walk through of our Video Clip Captioner and go over best practices for captioning video clips. And then, as I said, we will have a Q&A at the end.

So before I hand it off to Josh, we're going to do a quick poll. So the question is, how many online video clips do you have? And you can select 0 to 100, 100 to 1,000, 1,000 to 5,000, or over 5,000. I'll give you just a minute to answer that, and then we'll see the results. Great. So as you can see, it's pretty much all over the board, honestly, so that's really great to know as we move forward. Thank you. And now, I will hand it off to Josh.

JOSH MILLER: All right. Great. Just to give everyone a bit of background who we are, we provide a more efficient and scalable solution for closed captioning, subtitling, transcription. We started with a project at MIT involving speech technology and built a solution that aims at taking what has really been a laborious and, in many cases, expensive captioning process and really make it more like a web service but still retaining all of the key quality components.

We're able to effectively address accessibility requirements, along with video search and navigation needs that are a major part of video publishing and consumption online. We've served over 1,000 customers across media, education, corporate, and government agencies. Our core services are around creating closed captions and subtitles.

We've also built a number of tools to make the overall workflow far easier. That includes everything from out-of-the-box integrations with video platforms to other tools to address various parts of the caption publishing process using technology. So in some cases, you might have a transcript already or a script, and in some cases, you might even have captions to

begin with in some form or another.

So as many of you already know, the CVAA has gone into effect and mandates that any full-length content that was aired on television with captions must also have captions when aired online. There are a number of milestones and deadlines that expand these requirements as well. For example, as of last week, a new deadline went into effect stating that any archival content going up online must also have the captions added within 30 days of being published online, and that window will shrink next year down to 15 days.

Another requirement of the CVAA has to do with captioning quality, and these guidelines have been established by the FCC to make sure captions are good enough to accomplish their goal. This includes items such as general accuracy, proper timing and placement-- and the idea of placement is to avoid obscuring other relevant text on screen.

As of last week, each programmer is actually required to certify that they are following these quality guidelines. The FCC also established a set of best practices for captioning vendors in order to achieve the stated guidelines for programmers. Here at 3Play Media, we do adhere to the guidelines and can certify compliance, if needed.

So let's go a little deeper on these captioning quality requirements, since this is actually one of the newer requirements that went into effect. The first part is really all about text accuracy and making sure that the captions match what's actually spoken in the dialogue. And this is referring to original language, English or Spanish.

It must not be paraphrased. It should be actually what is spoken. This is especially important for recorded content because there's not a whole lot of leniency for recorded contents, as it's done in advance. There is leniency for live captioning just because of the time constraint.

The captions must coincide with what's being spoken as much as possible, and they need to show up at a speed that can be actually read by viewers. So for content being edited for re-broadcast, captions would also have to be edited for accurate synchronization.

Captions should cover the entirety of a program or a film. It shouldn't just be part of the title. The FCC states that captions should not block important visual content on the screen or other information that is essential to understanding a program's content when the closed captioning is on. For example, a news segment might feature an interview that displays a graphic with the name of the person speaking. The caption should therefore be moved to a location on screen

that does not obscure that graphic.

And one of the big upcoming milestones for the CVAA has to do with clips. Up until now, all of the rules around captioning have applied to full-length content. The deadlines to include clips within the captioning requirement have been pushed out a bit. In fact, they were really established just last year. But now, these deadlines are fast approaching.

So clips are defined as edited segments of content that derive from the longer piece of content. This could be what's called a single excerpt, or straight lift clip, or a montage, which is also sometimes referred to as a mash-up. Single excerpt clips are simply a straight lift of contiguous content from the longer form content, whereas montages are created by pulling pieces of content together, either in order or not, from the longer form content. So even with montages, the idea is that each piece of the montage does actually exist in the full form content, meaning it's not a piece of content that's been edited where you have the same visual but a completely different audio track. That's not what's being talked about here.

So here is a timeline view of some of the more significant deadlines related to the CVAA. The clip deadlines, highlighted in blue here, start to go into effect January 1 of 2016, so about nine months from now. At that point, only the single excerpt clips will be required to be captioned. And then, one year later, montage clips will also be mandated.

So that brings us to how do we actually address this new requirement? In some cases, a programmer might have thousands of clips, as we just saw with the survey. And it could even be 1,000 clips for a single show because over time, they add up very quickly. So that could make things very challenging in order to basically go back and get all of these captioned.

So there are a few options. One is to actually go back and re-caption all of the clips from scratch-- certainly an option. It's in some ways very easy because you already have a captioning workflow, but that's going to be probably pretty expensive and time consuming.

You could certainly edit the longer form captions-- have someone sit down with the right program, edit all the captions, avoid redoing what's already been done. But now you actually have to staff that process. So you have to have someone sit down, take that on. That's probably a pretty expensive ordeal as well.

Or there's a new option, and that is what we're going to introduce next. This is our Video Clip Captioner, which we're officially announcing today for the first time. Video Clip Captioner is

based on patent-pending technology and provides an automated solution to extract all these clips without having to redo them. And it's basically going to use what you've already done to give you the clip captions. So it'll be fast. It'll be low cost.

We've actually already been conducting testing with Fox and a number of other programmers to help us get it ready for prime time. We're very excited to show it off. So with that, I will turn things over to Andrew, who is the lead R&D engineer on the product, and he's going to walk you through how it all works.

ANDREW SCHWARTZ: All right. Hi. So as Josh mentioned, we're talking about our Video Clip Captioner, and the problem here is that we have a clip that could either be a straight excerpt or a montage, and we need to use-- we've developed an algorithm that will automatically identify what segments of that clip came from the parent and at what time.

So the algorithm that we have is going to work on both straight excerpt as well as montage clips. So just to go through a little bit in brief on how this algorithm actually works, what we're going to do is we're going to break up the child and the parent into smaller segments. And we're going to find what we call a fingerprint for each of the segments. And that is basically just trying to identify some-- compute some identifying information of that segment. And then we do a scan from the child fingerprint into the parent and determine where we can find a match so that this algorithm is then going to figure out from this process-- iteratively figure out where these segments in the child came from in the parent. As we can see in this example, we see the match here, and then we can note down what were the captions in the parent and then translate them into the child segment.

The algorithm that we're building-- we're building this into the workflow into our account system, so we're going to support a user-driven as well as an automated interface. So you can either interface with this product using our account system with the graphical interface on the web, or you can also use API integration, or you can upload things to an FTP server, and we'll pull those automatically, and we can perform this process for you.

On the output side, we can give you notifications when the Clip Captioner is finished, when they're completed. You can download the captions, just like with our regular services, in any number of formats that we support.

And you can also, in addition, for this, you'll receive a detailed report on each file. This will include some confidence information on the automated algorithm, as well as a couple of

additional pieces of information, such as if there were regions in the file that weren't matched, or if maybe we suspect that it was actually the wrong file that was uploaded-- if, for example, you uploaded the wrong child asset that was attached to a different show instead of this one.

So I'll just give you a quick rundown on how this looks like on our account system just first because it's a little bit easier to follow through this way. So if you're in our account system, which we have a little picture of here on the screen, you can use-- so for the first workflow, this assumes that you already have the captioned files, but we also support a workflow if you don't have the captions for the parent file. But if you do, then you will use the Import Captions feature on our website.

And then, a dialog box will pop up to help you import the caption files and also upload the associated media file with that caption, and that's all pretty quick. Once that file is in our system, you can then order a Video Clip Caption using the button that's shown. And another dialog box will pop up, and you can upload any number of child assets-- that is, any number of clips that belong to that parent media file-- and those will upload. And then, it will go through the process, and it will begin to automatically generate the captions for that file.

Here, we can see an example where I've uploaded eight files that were all attached to the same parent file. And the interface will show you just this timeline of where the child files came from within the parent file. So that's just a convenient interface so that you can organize your files that way and see where they came from.

If you don't have the parent captions, then that's fine too. You can just simply use our standard transcription service to generate those captions. Or if you have a transcript, you can use our alignment service to generate captions from the untimed transcripts, and that will all support Video Clip Captioning as well from that.

So I'll talk a little bit about what is the content-- the kinds of content that this will work on. As I've mentioned, this will support both straight excerpt and montage clips. There's not going to be any particular limits to the number of clips inside a montage. It can be 10. It can be hundreds of clips from a single parent. The clips in the montage can be in any order, or they can even be repeated. The algorithm supports that.

As for the caption content itself, it can be in English. It can be in Spanish. And that's good for the FCC compliance ruling. It will support other languages. Any language that we support for our existing translation service will be supported in this.

Also, we will retain any captioned framing and positioning requirements, and if desired, we can rerun caption positioning. We have an algorithm that will detect text on the screen and will move captions around in order to comply with the ruling about obstructing text or other necessary information on the screen.

A couple of notes about the limitations of the algorithm. So this is an entirely automated process, which means the algorithm is going to assume that the input captions are correct. It assumes that the timings are accurate. If the input captions are mistimed or slightly out of sync with the audio, that will generate possibly some problems in the output.

However, what that will do is that will generally reduce the score, which you will get as part of the report, and so you will be able to see that these files have a low score, and you can review the captions to see why they have a low score. And you'll see that there were some timing issues, and then we can either upgrade that to full transcription, or we can talk about how to fix that.

In addition, if there's any additional content in the clips, then we cannot generate the captions for that using this video clip captioning service. So if, for example, you add sound effects to the clip, or if maybe there's extra narrative-- for example, if you say next week, we'll see the following-- if that wasn't captioned in the original parent asset, then we cannot generate the captions using Video Clip Captioner.

So in order to make the most out of this product, there are a few best practices that we recommend. The first is to use accurate parent captions in preferably run-time based formats. And I say this because there are certain details about broadcast formats, such as SCC, where they actually adjust the caption frame times for character limitations and so on.

What that does, though, is that actually makes those caption frames out of sync with the audio. And in most cases, that's fine, because the algorithm is robust to a couple of variations like that. But the more of that missed timing that you introduce in the input caption file, the more likely it is that that will cause a couple of issues in the output file. It's definitely recommended to use input files of good audio quality, but it's fine to save bandwidth by diminishing the video quality because as long as the video file is still intact, then we can process with it.

This algorithm will work best, as I've mentioned already, on content with minimal editing. That applies to straight excerpt and montages without additional transition sound effects. We can

deal with some amount of that, but again, it's just the more additional content you put in, the more editing that's been done to a file, the more likely it is that the algorithm might not know where it came from because, in fact, it wasn't present in the parent.

But on most of the input files that we've demoed with with some of our test customers, sound effects have been fine. We've been able to deal with it. It's just something to be aware of as you're using this product. One particular thing to note is if a clip has additional sounds such as background music that did not exist in the parent, that could be potentially problematic.

LILY BOND: Great. Thank you, Andrew. Again, feel free to type your questions directly into the questions box of the control panel, and we're going to get started with those right now. So the first question-- Andrew, maybe you want to take this one-- is how long does this take?

ANDREW SCHWARTZ: Right. So the algorithm actually runs rather quickly-- I'm going to say roughly one-tenth of the time of the video file itself. That is obviously going to depend exactly on queueing time and how many files we're processing in our system at the moment.

But once the file is picked up by our system, it's going to be about one-tenth of the time of the file. That applies to the-- the longer the parent file-- but once you've done this for a single parent-- let's say you've ordered one clip from a show. Then any subsequent clip you order from that same parent will be very quick because we recycle that same resource, and we don't have to do all the computation again. So it's really quick.

LILY BOND: Great. And Josh, a question for you here-- how much does the Video Clip Captioner cost?

JOSH MILLER: So we're trying to keep the pricing pretty low. The pricing is based on usage, so it's all based on the duration of the clips that we're creating captions for. The full description is on our Pricing page on our website, if you want to check it out, but it basically comes down to about \$1 a minute. There are a few other stipulations, but basically, that's what it comes down to. So it's pretty inexpensive.

LILY BOND: Another question here-- maybe both of you can take this. If the algorithm is automatic, how do we know that the resulting captions are accurate?

ANDREW SCHWARTZ: OK, I'll start with that. So as I mentioned, there will be a report that can be delivered with each file, and you'll have access to that as long as the file is in our system. The report's going to include an overall score for the file, as well as a couple of breakout sub-components of what

went into that score. So that will all be available to you.

We'll work towards supporting workflows where we can notify you on how many files maybe fall below a score, and that can happen if, for example, you accidentally upload the wrong file attached to a parent, or if maybe the file was edited heavily and you didn't realize it. And all of those-- we have some good testing to indicate that we can detect those conditions, and we can alert you to it. Other than that, we'll give you the score, and you can see that. We'll have a review process for the captions and can review the accuracy.

LILY BOND: Thanks, Andrew. Do you have anything to add, Josh?

JOSH MILLER: I think that's exactly right is that, basically, we'll be able to really get a sense for what's coming out well and what's not, and we'll be sharing that pretty openly with everyone. And there will certainly be ways to fix the captions or redo them if needed.

LILY BOND: Great. So I see a question about source captions that might not be properly synced, and someone is wondering if they're able to manually make changes to the captioning file to edit the in and out times.

JOSH MILLER: So we don't provide tools for someone to make their own edits to closed captions. We do have services to do this. We have an alignment service. Depending on the type of content, it might work really well for it. It's actually another automated, very inexpensive service. So that would be an option. Otherwise, it would be up to you in terms of fixing any sync issues.

LILY BOND: Great. So I see a few questions up here about the specific requirement for captioning video clips. Josh, do you maybe want to talk a little bit more about who specifically is required to caption their video clips?

JOSH MILLER: Yeah. There are a number of laws. We actually have a webinar on this that talks about the requirements and exemptions for this law. It's on our website. Most broadcasters need to caption all of the content that airs on television, which means it applies to any of that content, certainly, online. The main rule has to do with-- basically, it says if it was aired on television with captions, it therefore needs to have captions online.

There are certain exemptions, though, for when content may or may not need to have captions when broadcast on television, and that tends to be things like-- it's a new network, or it's less than two years old or not generating enough revenue. So at this point, most stations are required to add captioning.

LILY BOND: Thanks, Josh. And I see some questions about a link to that. I'll add a link to the Resources about FCC exemptions in the email that we send with the recording. So I see some questions about live video. Josh, do you want to reply to whether or not captions are required for live clips? I'm happy to jump in on that too.

JOSH MILLER: Yeah. So it's kind of similar to the previous question. Basically, in most cases, captioning is required for live content when shown on television. So similarly, when that content is put up online, the answer is yes. The way the FCC and CVAA have talked about it is that the captions need to be as good or better than what was shown on television. That even has been modified slightly to basically say it needs to be properly in sync.

So the expectation is going to be that the live captions actually get synced up better for the on-demand version when shown online. If it's a straight-to-web live event, that's different and does not fall under the CVAA.

LILY BOND: Thanks, Josh. And just to specify, for video clips starting in July of 2017, live and near-live clips will have to be captioned-- so that's, again, live when they were broadcast on television. So I see another question about FCC's requirements for user-uploaded video clips. I think Josh just touched on that, but again, if it's a straight-to-web clip, it does not fall under the FCC requirements. So Andrew, a question for you, again, about how it works-- would this Video Clip Captioner work for music videos?

ANDREW SCHWARTZ: Yes. The algorithm will work on music videos as well, as long as the clip that you're dealing with-- if the audio matches that in the parent, then it should work just as well as with any other kind of content. So music videos should not be a problem.

LILY BOND: And kind of similar in terms of its capabilities, how does it handle mixed language clips? Will it do the translation or caption in the mixed languages?

ANDREW SCHWARTZ: So this service will not do the translation by itself. Basically, what you give us as input is what we're going to give you back as output. So if the input captions are entirely in Spanish, we'll give you a Spanish file as output. If there are mixed languages in the input captions, then whatever sections of that file were in the clip, we will give it back in the same combination of languages. If you need to translate it, that's a different service, which we also support.

LILY BOND: Thanks. A question about formats-- so someone is asking if there's any limitation on file size,

and what type of video or audio format you recommend submitting.

ANDREW SCHWARTZ: So we support a large number of different video formats, although we do have a best practices page for our preferences and what we recommend-- mostly open formats and forms that are supported by general media players. There's not very-- I don't believe that there's an explicit limit to the file size, other than just if it's too large to transfer. But basically, it can be a gigabyte, and that's fine. Did I cover-- what else was there?

LILY BOND: Yeah, I think that covered it. Thanks. What about if someone is worried about privacy with their files? They're wondering if they could send the video files on a portable drive or whether or not there's any privacy limitations to our platform.

JOSH MILLER: So we have secure methods for uploading into our system. The actual account system is a quite secure password-protected system that only-- each customer only has access to their own account, so it is a very secure system. Depending on the scenario, it would probably make sense to have a more specific conversation offline. We're happy to do that. We've gone through a number of different security audits in the past. It's something that we can certainly take a look at.

LILY BOND: Thanks, Josh. Someone is wondering how clean the closed captions are between scenes in a montage clip. Andrew, maybe you want to talk about that.

ANDREW SCHWARTZ: Yeah. So between scenes is generally going to be-- if there are going to be errors, it's generally going to happen there, and that's definitely something that the review process is going to try and catch. That's also something that gets caught in our automated scoring algorithm.

So if there are captions that are kind of over the edge between scenes, that will get caught, and you should generally be notified about that. But in general, in the tests that we've seen, the results are all pretty good, and it tends to capture the frames correctly. Obviously, this is dependent on the frames being timed correctly in the input as well.

LILY BOND: Great. Thank you. Someone is wondering what happens if the captions were added by the broadcaster, and the clips are generated by the production company? Does the person supplying-- so I guess the question is, really, do you absolutely have to supply both the source and the clip files?

ANDREW We need both the source and the clip files in order to do this process. They can come from

SCHWARTZ: different sources, but somebody needs to upload them into our system and associate them with each other before we can proceed. But the caption files can come from-- can be captioned by somebody else. You can import them in any number of formats, but we do need both the parent and the child files in order to do the processing.

JOSH MILLER: And the only thing I'll add is that if it's a huge hassle to get the source, we can always caption the clips from scratch. That's fine. That's always an option. This tool is meant to save time and money. We do have very fast turnaround options. It's just this clip captioning tool that is more automated using the source cuts the price down by quite a bit.

LILY BOND: Great. Thanks, guys. I think we have time for one more question. And then, again, we're happy to answer any more questions via email or at NAB. Feel free to stop by our booth. So the final question I see here is about the caption file formats. What types of formats can you take, and what types of formats can you produce?

JOSH MILLER: So we can import most major caption formats-- so certainly SCC, CAP, WebVTT, SRT. In fact, the whole list is actually on our website. There is a page dedicated to Video Clip Captioner on our website now. DFXP, STL-- basically, most of the major formats that will be used by this community should be covered. So if there's one that we don't seem to support, that's something that we can certainly talk about.

In terms of output options, we actually support probably over 50 different caption output options, so we should be able to have you covered in terms of what we can publish out. And one thing that's worth noting is regardless of whether we create the captions ourselves or import the captions from someone, we make all of those output options available, and they're always available for download.

LILY BOND: Great. Thanks, Josh. And thank you, everyone, so much for attending. Again, we will send out an email with a link to the recording and to the slide deck from this. And of course, feel free to reach out with any other questions. Thank you, Josh and Andrew, and thank you, everyone, for being here. Have a good day.