

The 3Play Way - AI Dubbing

[00:00:00.11] JACLYN LAZZARI: Thank you so much for joining us for today's session titled "The 3Play Way, AI Dubbing." Before we dive into today's session, I'd like to introduce myself. My name is Jaclyn.

[00:00:13.23] I'm on the marketing team here at 3Play Media. And I'll be moderating today's session. So I'd like to take a moment to welcome our speakers today. We have John Slocum and Jesse Ariss. Thank you so much for being here today and presenting for us. And without further ado, I will pass it off to them. Thank you.

[00:00:35.19] JOHN SLOCUM: Thanks, Jaclyn. My name is John Slocum. I'm a Caucasian man wearing a black vest and a blue shirt-- excited to be here to talk about our AI dubbing offer. My experience brings me through a number of media-focused industries, most recently here at 3Play Media, where I lead the product function.

[00:01:02.33] I have some advertising media experience and data experience as well. I'm a recovering engineer and really excited to bring new products to market and learn about what market problems we can offer solutions to. So enough about me. I'll pass it off to Jesse.

[00:01:22.77] JESSE ARISS: Yeah. Thank you so much. So my name is Jesse, and I'm a white male, late-30s, early-40s, depending on who you ask. I've got short gray hair and a white button-up shirt.

[00:01:34.24] So I'm a product marketing manager at 3Play Media, and that means my job is all about making connections and figuring out how we can take all this cool great new technology that we're seeing every day and make it not just useful but a game-changer for businesses and organizations who want to get their content into a new market and improve accessibility at the same time.

[00:02:00.11] So in my job, I get to dive into our AI dubbing technology. And I've been seeing firsthand how it's changing the way that organizations are talking to the world. We're going to talk today about this new technology that's breaking barriers and how it can do heavy-lifting for your communication strategy.

[00:02:19.84] But to better tailor our session, let's start by understanding where each of you are in your AI dubbing journey. So we'll do a poll here. So if we can start the poll, the poll says, which of the following best describes your stage of the dubbing journey?

[00:02:41.00] Are you an explorer, so just starting out, exploring the basics? A voyager-- do you have maybe a little bit of experience, you're starting to do it more? You're known as a trailblazer, so you're the dubbing person who's always involved in those dubbing projects.

[00:02:56.87] Beyond that, are you a dubbing savant? And if we've got a few of those, I'd like to talk to you afterwards. And then finally, are you just a spectator? Are you sitting on the sidelines here? You just want to learn a little bit about dubbing?

[00:03:11.39] So we'll let this poll close in about five, six seconds here. And then we'll take a little peek. And I just want to make sure that the way we're talking to you aligns with where you're at in your journey. So let's show those results, please.

[00:03:27.82] OK, excellent. A really nice mix. But it looks like we have one savant, who I'm going to get in touch with after because I feel like I could learn a few things from them. But this is great. There is a bit of diversity here, and this webinar is going to be structured in a way that everyone is going to be able to get value, whether you're a beginner or you're my new future friend that I'm about to meet afterwards.

[00:03:51.88] All right, so next slide, please, yeah. So there's a few truths about dubbing. And one of the truths is that dubbing is hard, right? There's a need to be accurate, not just in what's being said, but also for little cultural nuances. Every language and culture has little things that are unique to them.

[00:04:12.86] We also need to account for timing of the spoken word. Is what's being said on the screen matching up as accurately as possible with what's happening on the screen? This is one of the reasons or these are some of the reasons that dubbing has always really traditionally been time-consuming and expensive. Now, at 3Play Media, we're experts in dubbing.

[00:04:38.42] We've been doing this for a long time, and we've learned a lot about localization and dubbing over the past several years. I'm going to let John talk about that in a few moments and share a lot about what we've learned and how we've built trust with our customers. But to see why AI dubbing is so exciting, I want to take just a quick look back yesterday to understand how far we've come and why I think, in a few minutes when we're done here, you're going to be as excited as me about the potential of AI dubbing.

[00:05:15.38] So on this screen, there is a timeline that represents 100 years ago, to 25 years ago, to today, with various thumbnails from video content and films that I'll be describing as I go. So I want to share a few historical milestones. And don't worry, I won't spend too much time on this.

[00:05:34.71] But I personally was very interested in learning about this. The 1930s was really the start of a new era in film and motion pictures. If you've heard of the movie M-- it was a German classic-- it was one of the first to be dubbed for a broader audience when it was released in the early 1930s. The transition from subtitles to dubbed content was a pivotal moment in making video content universally accessible.

[00:06:07.49] Casablanca, which we all are familiar with, was dubbed from English into multiple languages in Europe, like French, Italian, and German, and it saw a lot of success through that process. In fact, that success in Europe with Casablanca being dubbed caught the attention of a lot of folks who saw the power in using dubbing to transcend language barriers and enhance both reach and, of course, revenue, right? This is important.

[00:06:36.42] But there's a catch. Dubbing was incredibly expensive. It was an investment. It cost a lot of money, and it took a lot of time.

[00:06:45.09] Jumping ahead to the digital era, which I would call it, which is about 25 years ago, the turn of the century. I have now a thumbnail of two motion pictures. One is Run, Lola, Run, and the other is Crouching Tiger, Hidden Dragon on the screen. Now, if you've heard of either of these films, it's probably because of dubbing.

[00:07:07.86] Had these films not been dubbed, there's a chance they would have never been seen by an international audience, including me, right? Run, Lola, Run is one of my absolute favorite movies. It's so fast-paced. The cuts are so fast that subtitles probably wouldn't work as well as dubbing because we want to stay focused on the content that's being shown on the screen.

[00:07:29.31] Even 25 years ago before YouTube, filmmakers understood the value, and how to create attention, and capture that attention, and get folks engaged. So the price for dubbing 25 years ago came down by a factor of 10 from 70 years before that. The speed came down by a factor of 10.

[00:07:51.06] And this was because of the introduction of digital tools, digital workflows, and just general technology. Smaller films-- maybe you've heard of Amelie-- they could start to find budget to localize their content and bring it into new markets. And then finally, on the right, this is where we are today.

[00:08:09.91] This is what I call the AI Age. And this is obviously really exciting, some of the most incredible progress I've seen personally in my 25-year career in video and film. We're talking about reducing now the time and cost from 25 years ago by another factor of 10, which means we're now seeing what I would call the democratization of dubbing.

[00:08:34.50] So we're opening the doors for experiences and content beyond multi-million dollar movie budgets. And now, we're able to put it into the hands of almost anyone who wants to create localized, engaging content. I'm talking dubbing for all, right? I have examples here on the screen of an industrial training video and a fitness type YouTube video.

[00:09:03.51] This is incredible. Imagine what having this content, your training content, your marketing materials, your video on-demand library-- imagine having all your content in dozens of different languages and what that could do for your business or your organization, for the growth of your organization, for the growth of your business. And that's why I'm so excited, because we don't need to imagine that anymore.

[00:09:28.83] It's here. That's what we're talking about today-- AI dubbing from 3Play Media. So let's go to the next slide, and I just want to briefly cover the two sort of solutions that I've talked about so far today. So I've spent a lot of time talking about traditional dubbing. This is what you're seeing on the left here.

[00:09:48.76] This is the fully human process that involves transcribing, translation, and also having a voice actor perform the dialogue. Obviously-- I don't even have to tell you already-- you know this is time consuming. It's going to involve casting, recording, manual synchronization, making sure that the dialogue is accurate, matches what's shown on-screen, right?

[00:10:13.89] But it's incredibly accurate, and we know how well it works. Let's compare that, though, to the right, which is pure AI dubbing, which we've seen now for a few months in the market. There's some solutions out there which are starting to do dubbing using a pure AI solution.

[00:10:33.83] This uses artificial intelligence to automatically listen to what's being said, translate, generate dub dialogue, and use sort of a synthesized Siri voice, right? This has really sped up the dubbing process and brought down the cost significantly, like I mentioned. But there are some challenges with this as well.

[00:10:56.78] When left unchecked-- I wouldn't even leave my Roomba running in the other room without being home, you know? I don't trust robots just yet. They're awesome, but they're not there yet. And they've been known to make a few mistakes.

[00:11:11.90] We've all had a Siri mixup. If AI is reading dialogue, how does it know if the word is lead, L-E-A-D, like the metal, or lead, L-E-A-D, like what a leader does? And that's actually still a problem with pure AI solutions that today has not been solved for. We're working on it, but as an industry, that's still a huge gap.

[00:11:42.87] We think there's a better way. We think there's a better way to leverage the speed and the cost savings of AI with the accuracy of the traditional methods of dubbing. So let's explore how our hybrid solution solves for this.

[00:12:00.25] So we're looking at the state of dubbing, and if you're following along with me, you know that there's traditional dubbing, and you know that there's pure AI dubbing. And I think we can all agree that traditional dubbing is too slow, too expensive, while pure AI dubbing is just not quite there yet when it comes to both accuracy and quality.

[00:12:23.12] So what we've done at 3Play, in true 3Play fashion, is we've carved out a new segment where we've created a new category that combines the speed and the cost savings of artificial intelligence with the nuanced understanding of human oversight that traditional dubbing has without compromising on quality or cost. So we built our AI dubbing solution around the concept of HITL, which is human in the loop.

[00:13:00.63] This means that throughout the dubbing process, we have an actual real human who's going to ensure that every piece of content is aligned both culturally and contextually and still preserving the authenticity and accuracy of that original message. I'm going to let John speak about how we do all that, how it all comes together, because it's really, really interesting and innovative. And I do think you're going to leave this thinking about ways that this is going to be able to open new doors for you.

[00:13:34.20] So I do just want to add, we're not saying there is no place for traditional dubbing. We actually believe that traditional dubbing still does have a place for some types of content. Like, if you've got a independent film that's an artsy film where the tiniest nuances in the actor's voice matter, I think you're going to want to stick with AI dubbing.

[00:14:04.82] Anything complex with that, let's stick with AI dubbing to keep that original vibe intact, because we're not trying to replace it. We're just providing a new approach, a new way for you to take your modern digital content like your subscription video on-demand content, or your corporate training, or your advertising, or sales videos, and get that localized.

[00:14:27.89] OK, I've been talking a lot here about dubbing history and the dubbing landscape. But I do want to have another survey quickly. So far-- actually, if you could start the survey poll for us here. I'd like to know-- let's just see-- which hesitations-- here we go-- if any, do you have with trying AI dubbing?

[00:14:55.60] Why haven't you done this? Have you done it? Are you concerned about quality? That's response number one-- doubts about the AI's ability to produce natural-sounding dubbing. Option two, cultural nuances.

[00:15:10.24] Maybe you're concerned that AI may not be able to capture cultural nuances. Brand alignment-- is AI going to work with our brand's tone and image? Is it going to be good enough for our brand? Value and ROI-- does it actually make sense to invest in dubbing and the value that dubbing will bring to the business?

[00:15:35.13] Finally, technical concerns. We've heard from some folks that there are worries about compatibility, integrations, or issues with glitches, some of the things that I've mentioned before. Finally, the last option is no concerns. I'm ready to try it. And if you want to click that button too, you'd make me happy.

[00:15:55.89] Of course, there's another option, "other," where you can fill it in yourself. And hopefully by now you've all had a chance to make your choice, so we can close this poll and we can take a look at some of the results. OK, this is fascinating-- "doubts about the ability to produce sounds" seems to be the number one result with technical concerns being at the bottom.

[00:16:20.43] This is really interesting, and this is also sort of mirroring what I'm seeing when I'm talking to people in this industry. I think we can actually move ahead. And I'd like to actually move forward one slide here, please, and let's see what this actually looks like in action.

[00:16:39.18] Because I have a feeling, after watching this video, you might be not so much concerned about the quality of the voice. Why don't we press play?

[00:16:49.62] [VIDEO PLAYBACK]

[00:16:50.01] - Did you know are over 7,000 languages spoken globally, making dubbing a powerful tool for reaching your audience no matter where they are.

[00:16:58.44] - While dubbing has been around for a long time, the available options have been at two extremes.

[00:17:02.91] - Traditional dubbing-- expensive, time-consuming, and manual.

[00:17:06.29] - Wait, that's going to cost how much?

[00:17:08.97] - Pure AI solutions, quality just doesn't cut it, eroding trust in your brand.

[00:17:15.73] - That was bad.

[00:17:17.70] - Enter 3Play Media's innovative AI dubbing service. Our secret-- a human in the loop solution.

[00:17:23.68] - But hey, we know you have to see it to believe it. Roll it back.

[00:17:29.76] [SPEAKING SPANISH]

[00:17:31.08]

[00:17:37.14] - [SPEAKING FRENCH]

[00:17:38.64]

[00:17:41.85] - [SPEAKING GERMAN]

[00:17:43.33]

[00:17:45.53] - [SPEAKING FRENCH]

[00:17:47.32]

[00:17:48.27] - [SPEAKING GERMAN]

[00:17:49.69]

[00:17:54.67] - [SPEAKING SPANISH]

[00:17:56.07]

[00:17:56.90] - [SPEAKING JAPANESE]

[00:17:58.36]

[00:18:02.96] - [SPEAKING SPANISH]

[00:18:04.27]

[00:18:06.60] - Didn't that sound great? Our AI dubbing solution combines human-edited transcripts and translations with best-in-class AI voices to produce premium, cost-effective, and highly accurate output.

[00:18:18.06] - Plus, with a user-friendly platform, integrations, and fully-supported API, you can easily incorporate this process into your workflow. It's dubbing made simple.

[00:18:27.89] - Say "goodbye" to language barriers. Say "hello" to global engagement and brand expansion.

[00:18:32.81] - Ready to revolutionize your content?

[00:18:34.80] - [SPEAKING FRENCH]

[00:18:36.09]

[00:18:36.95] - [SPEAKING SPANISH]

[00:18:38.46]

[00:18:38.91] - [SPEAKING GERMAN]

[00:18:40.26] - Get started today with 3Play Media.

[00:18:45.80] - Did you know there are over 7,000--

[00:18:48.20] [END PLAYBACK]

[00:18:49.93] JESSE ARISS: All right, so that's just a little sneak peek at what it looks like in action. You may have noticed that the background noise was mixed back in, that multiple people were speaking, both on-camera and off-camera. There's some really cool tech there, and John is going to get into a little bit about how this all comes together.

[00:19:11.63] But before I pass it over to John and say "goodbye" to you folks, let's do one more survey. If you could please bring the survey up for us. So this just helps me understand where the market's at today. And I'm curious to understand your dubbing goals with AI over the next year. So the options are learn about what's possible with AI dubbing, identify an AI dubbing solution that can scale to support significant volume, transition or supplement subtitle localization volume with AI dubbing, transition from a self-serve AI dubbing volume to full service solution, and transition voice artist dubbing volume to AI dubbing.

[00:19:58.42] Where are we at? Are we thinking about anything here? We'll let this run for about five more seconds. But if I was to guess, we might still be in the early stages around learning what's possible and identifying an AI dubbing solution. Let's take a little peek and see the results.

[00:20:16.29] Well, that's great. I'm very excited to see that we're here to learn about what's possible. I think the sky's the limit when it comes to possibilities. So you're definitely in the right place, and I'm glad to see that you're excited about the next year when it comes to AI.

[00:20:34.86] OK, let's pass it over to John, because I think John is going to be able to tell us a little bit more about how it works and what the secret sauce is that goes into all this. Thanks, John.

[00:20:51.69] JOHN SLOCUM: Awesome. Thanks, Jesse. It's really helpful to know that I'm not the only person out there with an irrational fear of Roomba vacuum cleaners. So I've learned something already today. Appreciate that.

[00:21:06.13] But I want to talk about why we're excited about our AI dubbing process, and specifically the components that make up the process, that build upon 3Play's core competencies that we've been refining for over 15 years. AI dubbing takes a tech-driven approach to solving dubbing challenges that previously required, as you mentioned, scheduling voice actors, studio time, and localization specialists that results in a lot of time and money that AI dubbing can remove from the process, create efficiencies there.

[00:21:44.10] So I guess taking a step back and considering 3Play's DNA, right, we've done this before. We've solved similar challenges, similar problems by applying technology consistently. And looking at media accessibility challenges, right, we always take a tech-forward, tech-first approach to ask ourselves how we can apply technology to solve the problem in a more efficient way than the standard legacy solutions in market, right?

[00:22:16.86] We have 14 patents covering aspects of transcription, captioning, audio description, and real time captioning systems and methods. And we approach every market problem we see this way, right? That's the first question we ask.

[00:22:31.89] And the result there is creation of customer benefit. The efficiency directly translates to customer benefit. We've been refining support for automation workflows at scale for over 15 years with feature-rich APIs and integrations that mean we're meeting customers in their media workflows. We're not asking customers to take on a new process by meeting us in our workflow.

[00:22:57.59] So our intent with AI dubbing is to offer the same 3Play peace of mind that we offer to customers with 3Play captions. Our captions are guaranteed to be 99% accurate and ready to be published as delivered. We want our AI dub work product to land in the same way in our customer workflows-- that is, ready to publish as we hand it over to our customers.

[00:23:21.35] So enough about our background. Let's step through our process and step through specifically why we're confident that 3Play's AI dubbing will meet our customers' needs. So this visual describes our process and gives you an inside look at how we're producing AI dubbing. And I'll jump through each of the boxes here.

[00:23:45.09] We have seven steps described from ingestion, transcription, translation, the dubbing synthesis, step-mixing, back into the audio to be delivered to customer. But before we deliver, we do a QA check, and that's a pretty significant part of the step. And then we deliver.

[00:24:05.40] So 3Play customers will recognize some of these steps-- the ingestion of media, the delivery of the work product, and then the steps in between are how we're producing the AI dub. In 15 years working with AI and ASR, we've learned that AI is only as good as the humans training it. Garbage in, garbage out.

[00:24:28.13] For example, as recently as just a few years ago, a household voice assistant-- I don't want to say her name, because she might respond in the middle of the presentation here-- was still telling folks that dolphins breed with their lungs due to, I don't know, a misspelling, inaccurate training-- obviously, incorrect.

[00:24:48.32] But 3Play has processed millions of hours of transcription data and has state-of-the-art training in post-process correction at scale on AI output. We know this. This is our business. Hold that thought. Our AI dubbing process was designed specifically so that we can provide a service that is fast, cost effective, scalable, and provides the quality that our customers' audiences want and deserve.

[00:25:15.83] So let's take a few minutes to get through each of these steps and how we're ensuring the end result creates a quality viewing experience. So where it all begins-- our process begins with the ingest stage where our customers can upload files to us directly. We may be automating delivery of media files to the 3Play platform.

[00:25:39.68] It's a user-friendly ordering process. 3Play customers will recognize the ordering process that we've developed for AI dubbing immediately-- just a few clicks to place the order in our existing order flow. Customers can order AI dubbing today in the 3Play platform. Just let us know to enable it for you. So it's already there, right-- leveraging the same video platform integrations and media files that you have.

[00:26:05.84] If you've already produced captions, transcripts with us, you can order dubbing on those today. So from there, our AI with human oversight transcribes the content, ensuring every word is timed correctly with the video. It's important to note that the output of this transcription is a dubbing script that has accurate timing and identifies the speakers with best-in-class accuracy.

[00:26:32.10] This is where we start to see benefit from the millions of hours of transcription training and automated correction that 3Play applies in post-processing downstream from the AI. So the accuracy ensures that the source of truth for the translation to the dubbing script best represents the intended source content meaning. The timing is critically important in dubbing to ensure that speech aligns with lip movement.

[00:26:59.28] I'm sure we've all seen audio continuing past a speaker's lip movement in a dubbed video. That's jarring. That tells the audience that, yes, indeed, this is dubbed content. So we have processes in place to control for that.

[00:27:17.81] Speaker labeling is critically important to ensure the right voices are being synthesized. This can be challenging with multiple speakers. And AI gets speaker diarization wrong often enough to make this potentially very awkward.

[00:27:32.45] I did some research myself and created an AI dub of a recording of an internal meeting just to see how our tool set would support the process and how I could produce a quality dub. And I ran a fully automated version just to test. The automated speaker diarization that I was working with split my speech mid-utterance, resulting in the wrong voice being used to finish synthesizing my sentence, my phrasing.

[00:28:05.01] So it switched in mid-utterance, mid-paragraph on my own dub that I was producing. It's kind of awkward. So next up, the translation step. This is where the precise transcript is localized into a dubbing script in a target language capturing all the nuance.

[00:28:23.58] We're preparing the dubbing script translation with customer preferences in mind. So with the human in the loop, we look for each customer's dubbing translation profile that our customers can create and send with the dubbing order. We require a dubbing translation profile selection with a human in the loop translation. And we can store multiple translation profiles on the 3Play platform.

[00:28:52.81] So the customer can tell us which profile to use to get the right tone and target language that works for their dubbing audience. Worth a note, subtitle translations and dubbing script translations are very different. Subtitles benefit from different space and time constraints-- that is, a subtitle does not need to align with regard to timing of the actual audible speech in a video.

[00:29:22.23] A subtitle can remain on-screen after speech has ended so there's more time for the audience to read the subtitle. The dubbing script has to finish as soon as I'm done speaking. If the script keeps synthesizing after I'm done speaking, you'll hear my voice after my lips have stopped moving.

[00:29:44.79] That's what we see in the QA process. It's really obvious when you see it. So these are a couple of the things that we're trying to control for in the process, right? And then next up, we actually have to produce the voice.

[00:29:56.90] This is the voice synthesis. This is the core of AI dubbing where we synthesize speaker voices, multiple speaker voices, depending on how many speakers are in the source video, audio, and we take this right from the translated dubbing script. We read the timing that's passed in the translated dubbing script on each segment for each speaker.

[00:30:18.80] And this is where AI brings a new voice to life in the target language. We're using voice-matching and not voice cloning. So by using voice-matching algorithms rather than cloning, we're avoiding any consent and legal exposure that may be associated with voice replication.

[00:30:39.99] We've seen ungated access to voice cloning surface in recent news articles, whether it be voice cloning without rights, consent, to replicate popular celebrity voices, for example, or more nefarious examples from bad actors, including impersonation intended to mislead or deceive both targeted individuals-- we've heard about ransom demands in cases, worst case-- or entire populations, where mass communication will go out with a cloned voice.

[00:31:12.66] So we're using voice matching, right, which can be very close, and also close to the source speaker voice, and very high quality, but generally different enough to be obviously not the source speaker's voice. So to test this, I dubbed my own voice into English. And just as listening to your own recorded voice can be unsettling, my voice match was clearly not me.

[00:31:37.96] But it was still close enough to be more than a little awkward to my ear. It was fun, regardless. I had a good time with it, learned a ton. We're matching with about 50 different voice options for each supported language that we offer.

[00:31:54.61] We're including male and female in multiple regional dialects, with original and native speaker voices. We're offering 10 different languages in the 3Play platform today, with line-of-sight to triple that number of languages in the coming months. So if you are a 3Play customer and you want support for a language that you're not seeing in our platform, let us know.

[00:32:17.42] Reach out to your account manager, reach out to support. We can probably support that. We have line-of-sight to all major target languages that we see demand for.

[00:32:32.91] So the mixing step-- we're getting there, just two more steps after this one. The mixing step includes separation of audio layers. And those layers are voice, background, and/or music layers, and then remixing the layers post-voice synthesis. This ensures the final mix persists non-speech audio events that are often vital to supporting the intended audience experience.

[00:33:01.89] This could include a door slamming, a car horn honking, and, of course, music. So this step ensures that the audio experience is one that viewers will enjoy listening to and that will keep them engaged with the content.

[00:33:20.53] QA step. There's so much to say about this because there's a lot potentially happening here. Now, the accuracy of the transcript, the timing, the speaker labels, the quality of the translation to the dubbing script, our ability to incorporate regional and social dialects into the dubbing script all cuts down on what needs to be addressed in the QA step.

[00:33:50.69] So we may see dubbing videos that don't need much in the QA step, or we may see videos that need significant editing in the QA step. It depends. So any of you who have used or are using a pure AI dubbing toolset to produce your own dubs likely know all too well that you won't know whether you need to edit the translated segments until you synthesize the dubbed speech.

[00:34:15.43] That's where the speech rate issues may surface. Our model employs an algorithm, and this is fairly common with the dubbing tooling that we're aware of in market-- our model employs an algorithm that controls speech rate to fit a timed segment and will flag faster speech rates for review that may prompt edits. So we'll speed up speech to fit a translation.

[00:34:41.30] Accelerated speech rates tend to sound robotic or chipmunk-like, and this is the result of non-English target languages such as Spanish, German, or Japanese tending to require 20% to 40% more syllables and characters to express intended meaning from fairly concise

English source content. This may be OK. If we have time to synthesize the longer segment without accelerating speech or visibly exceeding speaker lip movement, then we might be all right with that target language translation.

[00:35:15.72] But if not, we'll need to edit the target segment translation to prevent a jarring audience experience, right-- that scenario where the speech continues past the speaker's finished lip movement. So our editors may not need to edit most or any segments, as I mentioned, depending on the source speech rates and the segment timing. But if they do need to shorten, they'll take particular care not to change the meaning of the translated segments and honor the translation profile preferences.

[00:35:48.19] So, customers, if you're ordering AI dubbing with us, this is where your translation profile preferences are really important to ensure that, as we're interpreting the more efficient target language translation, we're keeping your priorities in mind in that interpretation-- make sure we have the right speech rate, the right speech, and also are honoring your brand preferences, your product names, your target language translation formality.

[00:36:17.34] All of these things are critically important in producing the experience that you want for your audiences. Also, we may need to apply phonetic spelling to ensure the synthesis reads out correctly. For example, do you like-- I'm going to struggle with this one-- do you like "Were-Chester-Shire" sauce, or Worcestershire sauce? What do you prefer?

[00:36:42.47] What do you think synthesis is going to read out if you pass it "Were-Chester-Shire" sauce as spelled, right? Not sure. Voice match quality did-- we get the best voice matched? Should we try a native speaker voice or stick with the original speaker voice? Should we try a different voice that better matches the source content voice?

[00:37:06.80] Consistency from speaker to speaker, including speaker labeling, getting the right voices associated to the right source content voices-- volume and speaker changes, right? Expressivity. So do we see the right emphasis and expressivity that best reflects the source content?

[00:37:28.82] We expect a lot more focus on this feature in our roadmap, right, coming up. This is an evolving capability. We can approximate expressivity today with punctuation in the dubbing script and a few other tricks. But this will be a major focus area for the remainder of the year for us in delivering that experience that aligns with the audience expectation.

[00:37:51.44] Our goal is to deliver a product that the audience doesn't realize is an AI dub, right? Like, they may know, if they look for specific things, they'll be able to figure it out. But we want to remove the jarring experiences-- the lip synchronization, the pronunciation-- that make it obvious, right, and especially the regional and social dialects.

[00:38:14.76] So once we're satisfied with all that-- and, again, we may get through the QA step pretty quickly, or we may be spending significant time there. It depends on the complexity of the content. But once we're satisfied, we're ready to deliver. So what does that mean?

[00:38:33.25] It means we're sending you a work product. We're sending you an asset, the same way that we're sending a caption file, or the same way that we're sending an audio description file. So say "goodbye" to manual dubbing processing and "hello" to our platform integrations and workflows that you know. And these are all designed for an efficient and seamless user experience.

[00:38:55.79] You'll immediately recognize the dubbing order experience, as I mentioned. We have two primary dubbing asset deliveries. We can offer just the dubbed audio file or the final mixed dubbed video-- depends on what you need.

[00:39:10.73] If you work with a player that supports multiple audio tracks, an online video player such as a Brightcove, a Vimeo, Wistia, Kaltura, YouTube, for example, you'll just need the dubbed audio track. And you can load that into the secondary audio track associated with your video.

[00:39:29.63] If you need the final mixed video with the dub, we offer that as well. So we're eager to hear from our customers on how they're presenting dubbed content to their audiences and have a variety of methods we can offer to support the right delivery for your publishing workflow. If you have questions, raise your hand, chat with support. We're eager to hear from you.

[00:39:55.86] So before we wrap things up here, we have one more question for you, and then we can take a crack at some of the questions that you have for us. So if we could get a poll up here-- we're curious, what type of content are you currently dubbing or planning to dub? What might you be dubbing? Training and educational content, media and entertainment content, fitness and wellness-- exercise videos, yoga videos-- subscription video on-demand, marketing and advertising videos, promotional content, short clips, or something else that we're not thinking of here.

[00:40:37.52] Or are you not dubbing right now anymore, currently, and you're just here to learn? So we'll give you a couple of seconds here to respond to that.

[00:40:57.40] All right, if we can wrap that up. Great. And a lot of training and educational content. That is exciting. We're seeing a lot of activity from customers and folks interested in AI dubbing inbound to us trying to understand how they can deliver a more effective training program and training videos globally.

[00:41:23.84] So that's one we're hearing a lot-- not surprised to see that. Media and entertainment, also not surprised to see that-- a lot of variation here. So excited to dig into this.

[00:41:36.89] So next up, I think we're going back over to you, Jesse.

[00:41:41.05] JESSE ARISS: All right, good stuff. Thank you, John. I appreciate that. And I've been learning so much myself about this process. And John did a great job in explaining what's really sort of an in-depth process and how we simplified it for you, the customer, and just to

create this content that is top notch, top quality, but at a fraction of the cost and much faster than the traditional methods.

[00:42:08.05] I hope by now you're sort of starting to think, hey, are there ways that I can use this as an organization? We know that there are so many use cases out there for AI dubbing today. We know that we've been talking to a lot of folks when it comes to getting their marketing materials into new locations or making sure that their training is available to everyone who works at their company, regardless of where they're located or what language they speak.

[00:42:37.72] So this is really an incredible innovation when it comes to localization and accessibility. Really, the options are endless. So what I'd love to do is learn a little bit more from you, talk to you, sit down, see how we can work together, and see how AI dubbing can fit into your strategy of reaching an entirely new global audience. If you don't mind going to the next slide there for me.

[00:43:05.51] If you do want to get started, there is a link in the chat that we will be posting. Fill that out. If you want to mention that you joined the webinar, that will help us as well just to know what you've seen already.

[00:43:22.33] But we'll reach out to you. We'll start the conversation. We'll talk to you about the options, see where it fits. And if it does make sense, we can get started right away. So I'm very, very excited about that.

[00:43:34.67] Let's go to the next slide. And I think our audience has been quite good and engaged, John. And I think there's a few questions here. Is there any that we're ready to go with, maybe, Jaclyn?

[00:43:50.02] JACLYN LAZZARI: Yeah. Great presentation, guys. Loved it. Yeah, we do have some questions that came from the audience. So we can go ahead and dive right in if you guys are ready.

[00:44:02.55] JOHN SLOCUM: Ready.

[00:44:04.26] JACLYN LAZZARI: So we had a question-- what is the turnaround time from ingestion to delivery?

[00:44:12.10] JOHN SLOCUM: It depends on the process that our customers are able to choose in ordering the dub. So standard turnaround process should be about nine days. We can go faster than that. We can expedite different steps in the process to turn the dub around faster.

[00:44:41.66] And if you're familiar with the 3Play platform, you're familiar with transcription and/or translation expedited turnaround service options, those are the same process steps that we're incorporating into dubbing. So if we want to expedite some or all of those, we can get down to five days, for example.

[00:45:05.39] If we want to establish a process that's even faster, we can do that. We'll need to connect with you and understand what your process needs are, what the volume is, what the languages are. But it varies.

[00:45:19.69] We're talking about days of turnaround-- five to nine days. We could also take longer if you want, but we don't need to. And so it varies. It's up to you. And you'll have control over that.

[00:45:33.56] JACLYN LAZZARI: Thank you. And then we also saw this question asking, how closely do you reflect original inflection and tone with the AI dubbing product?

[00:45:45.78] JOHN SLOCUM: So that's expressivity. That's what I refer to as "expressivity--" bit of a tongue-twister. That's an evolving capability in the AI dubbing space, right? We're at market right now in that we can reflect some expressivity. We can manage expressivity with punctuation and some other tricks-- question marks, punctuation.

[00:46:14.54] There are different ways to break up the segments and timing that will reflect pausing back in the synthesis that are all essential in dubbing a lifelike soundtrack for the audience. But expect that to evolve. That'll be a roadmap focus probably over the next couple of quarters of ours as capabilities in-market and our technical approach continues to mature.

[00:46:46.51] This is evolving rapidly. So keep an eye on it. We have some now-- expressivity detection is another capability that we're looking at and thinking about how to incorporate into our process. That's even more challenging than manually updating expressivity.

[00:47:06.62] So a little bit now-- the voice artist dubbing, obviously, with humans representing expressivity in their dubs and studio actors, right, is still pretty far ahead of what AI dubbing can accomplish. But we're catching up quickly.

[00:47:28.02] JACLYN LAZZARI: Thank you for your answer, John. And then kind of along those lines, but slightly different direction, how does fast and slow speech influence AI dubbing results?

[00:47:41.49] JESSE ARISS: Can I speak to this one, John? Is that OK?

[00:47:43.08] JOHN SLOCUM: Absolutely.

[00:47:43.83] JESSE ARISS: Yeah, I think this is a great question because this is huge. This is what it's all about. Everyone speaks at a different cadence. I'm speaking fast. Sometimes I speak slow.

[00:47:56.88] And also, that compounds when you start to look at the complications of multiple languages. So some of us may speak German on this call. I have some friends who speak German, and a lot of English words that are maybe one or two syllables are actually four or five syllables in German.

[00:48:18.15] And the sentences just compound that further. So we're talking about a bunch of different things here when we're talking about speed. And this is where a hybrid solution really, really shows its value.

[00:48:31.05] So our translators that we work with for this process are experts in not only dubbing, but they're also native speakers of both languages that are being dubbed. And so we do extend to them a little bit of-- it's an art and a science. And that art, you just can't get from a pure AI solution.

[00:48:55.13] So an interpreter, a translator would understand what's being said, what the message is that's trying to be said, and look for a way to maybe get across that exact same idea synchronized with what's being said on-screen. This is one of the real advantages we have. And we're able to still maintain that accuracy or that commitment to the message or the idea that that original speaker is trying to get across.

[00:49:23.15] So to answer your question, how do we deal with varying speeds, slow speakers, long speakers, different languages that run long? The answer is with that human in the loop and that unique advantage.

[00:49:39.12] JACLYN LAZZARI: Love that, Jesse. Thank you for your answer. And now the next question-- someone asked, what's the pricing model? Can each of you or one of you speak to that?

[00:49:50.16] JESSE ARISS: Each of us should, and then and then we'll decide which answer we like the best, right, John? No, this is obviously something we can talk about when we meet. But just to set expectations, obviously-- what I talked about, I talked about the cost and the speed being reduced year-over-year by factors of 10.

[00:50:12.84] You can expect to see that kind of pricing-- very, very affordable pricing, much closer to the pure-- a little bit more expensive than the pure AI solutions just because of that human touch that you're going to get just a little bit-- but significantly lower cost than the traditional method of AI dubbing. So the cost savings are absolutely there. The value is there.

[00:50:41.91] And to get into specifics, it's going to depend on which languages you're dubbing into, how much content, how much video, what your library looks like.

[00:50:50.23] So there's a lot of variables there. But the short answer is it's going to be significantly less than the traditional dubbing that you may have been exposed to.

[00:51:06.80] JACLYN LAZZARI: All right, thank you. And I will just add, just in general, the pricing model is priced per minute. So I think maybe you touched on it, Jesse, but I might have missed it. Just wanted to mention that as well.

[00:51:20.84] All right, and then we have another question, kind of a specific one for video player. So they asked, can we send the final result to Brightcove or leverage the API connection that we already have with 3Play?

[00:51:37.55] JOHN SLOCUM: Yes. So what you'd be able to do in the 3Play platform is see the video file that you may want to dub-- if you've already dubbed a file from Brightcove in the 3Play platform, you'll be able to see that and order dubbing on that file. And then Brightcove supporting a secondary audio track means that you'll be able to load the soundtrack, the audio file back into Brightcove as your Spanish audio associated with your video, or German, or whatever languages you have there.

[00:52:24.11] The auto-post back to Brightcove is not something we've enabled yet. If that's part of the workflow that you want supported, let us know. That is a low lift for us to implement. So, yes, we can.

[00:52:40.32] If you looked at it today, it wouldn't be there. But we could probably have it for you in a number of days or weeks, depending on your needs.

[00:52:52.08] JACLYN LAZZARI: Great. Thank you, John. We have another question that came in kind of regarding the lip movements of the dubbed track. So they asked, can you alter the shape of the speaker's lips and mouth to reflect what they're saying?

[00:53:11.49] In other words, are there any sort of software "tricks," quote unquote, being used to adjust the facial expression when we're dubbing?

[00:53:26.34] JOHN SLOCUM: I will take that one. No, not yet. We're looking at technology to potentially do that in the future. Right now, we're working with just the audio associated with the video. So we are assessing the timing of the source content and synchronizing to that timing in the creation of the synthesized speech-- So doing a lot to match that timing up in how we're producing the target language voice.

[00:54:01.95] But we're not modifying the actual visible video to modify lip movement to coincide with the target language speech. Keep an eye on that one. That's probably a little bit further out than the work that we're doing on expressivity this year. But definitely, we have eyes on it and are researching that as well.

[00:54:28.26] JACLYN LAZZARI: Thanks, John. We have a question asking, can you simultaneously incorporate dubbing and audio description for a video? Or how would you go about that?

[00:54:44.50] JESSE ARISS: Yeah, that's an interesting question. John, I'm curious to hear how you would handle that as an expert in this.

[00:54:53.89] JOHN SLOCUM: That's a great question. What we'd want to do is mix the audio description and the dub into the target language sound file, right? And we'd want to, effectively, create one file, I think, that if it's a Spanish target language, we'd want the speech and the description-- well, maybe not if we had two Spanish.

[00:55:25.32] We have some options there. So whether multiple audio files or a final mix to incorporate the description and the dub into a new final mix video file, we could do that. We

have the technology to be able to do it. I think we'd also, as we often see in audio description use cases, be constrained by the video platform capabilities to support that. So it may vary. I'm sure we could get there from here. That's my answer.

[00:56:02.83] JACLYN LAZZARI: Great answer. Thanks, John. And we have time for one more question, and it's a good one. So let's finish it up. The question is, we are a movie distribution company. We also have our own channels that we are spinning up Spanish and Portuguese channels and would like a lot of titles dubbed. Does this technology work well enough for narrative movies and not just content, such as a talking head?

[00:56:33.99] JESSE ARISS: Yeah, sure, no, that's a wonderful question. I'm hearing it a lot. I think it really comes down to the type of content that you're creating as well as the languages that we're dubbing it into. This technology, this voice technology, it gets incredibly better every single week. We're seeing iterative progress in what we're able to output as a voice synthesis.

[00:56:59.74] It's getting better and better. Even as it stands today, there are certainly cases where we could see success with a fully synthesized voice for a feature film or a short film. Again, there's going to be a few things we're just going to want to watch carefully.

[00:57:16.99] But having that human in the loop and having that quality control and quality assurance that John was speaking to earlier is really going to dial it as far as we can. We could work with you on maybe doing a little sample so you can see how it sounds.

[00:57:34.22] But I do have a feeling that the technology is right there. We're right at the point where this could certainly be applied to your use case, absolutely.

[00:57:47.30] JOHN SLOCUM: Yeah, I agree. And we're testing movie production content now. And so we'd be eager to test the content that you have as well and work with you on it. And we see most often in tier one big distribution film production that voice artist dubbing is still the right solution.

[00:58:16.37] And we offer that as well. But we are testing the application of AI dubbing for movie production, post-production content. So bring it in, give us a shot at it. We'd love to work on that with you.

[00:58:30.24] JACLYN LAZZARI: All right. Well, thank you so much. That's all the time we have for Q&A. John and Jesse, thank you so much for a wonderful presentation. And thank you, everyone, for joining us and just asking wonderful questions today.

[00:58:43.84] And with that, I'll end it. Thank you so much again, everyone. And I hope you all have a great rest of the day.