Media services: Use AWS AI services to automate captioning & subtitling

**SPL-246 - Version 1.0.9**

© 2024 Amazon Web Services, Inc. or its affiliates. All rights reserved. This work may not be reproduced or redistributed, in whole or in part, without prior written permission from Amazon Web Services, Inc. Commercial copying, lending, or selling is prohibited. All trademarks are the property of their owners.

Note: Do not include any personal, identifying, or confidential information into the lab environment. Information entered may be visible to others.

Corrections, feedback, or other questions? Contact us at [*AWS Training and Certification*](https://support.aws.amazon.com/#/contacts/aws-training).

**Lab overview**

In this lab, you prototype a video workflow with Amazon Web Services (AWS) artificial intelligence (AI) services. The workflow will automatically generate captions, alternate language subtitles, and alternate language audio tracks. To accomplish these tasks, you will use Amazon Transcribe, Amazon Translate, and Amazon Polly.

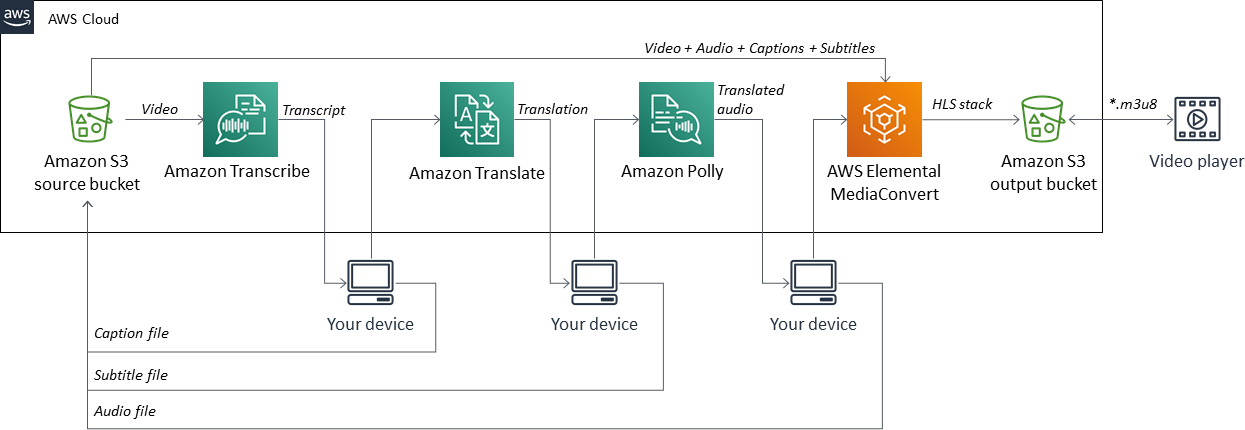
WORKFLOW AND REQUIREMENTS

For this lab, you have an existing video asset with audio. You want to increase the audience and meet regulations by including captions, alternate language subtitles, and alternate language audio tracks. You want this system to be automated, but first, you must prototype the workflow.

Requirements:

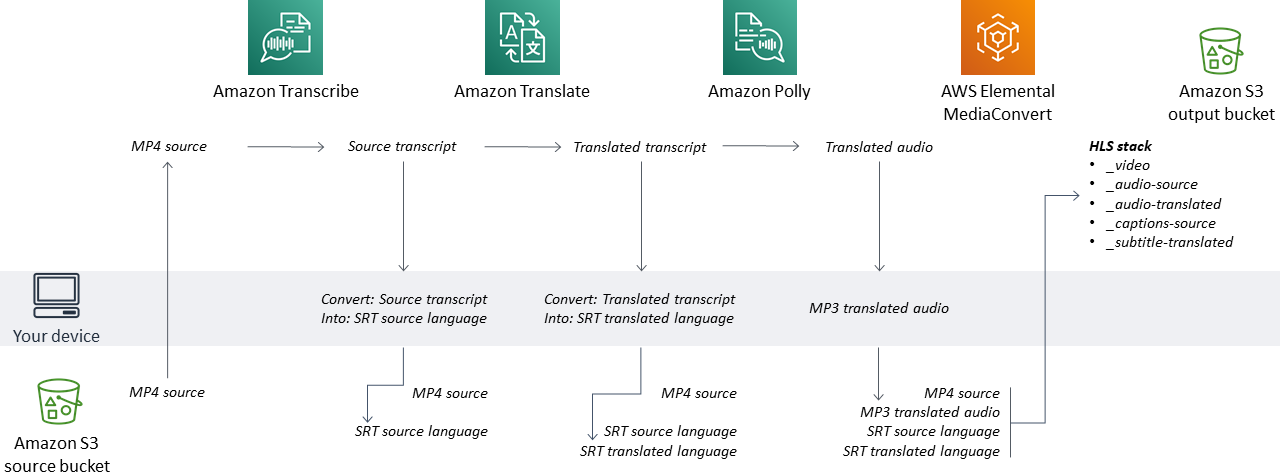
* Use a video with audio as the input
* Transcribe the audio using an automated service
* Use the transcript to create a captions file in an accepted format
* Translate the previously created transcript
* Use the translation to create a subtitle file in an accepted format
* Use text-to-speech to generate an appropriately timed audio file of the translated transcript
* Combine the original video and audio, captions file, subtitle file, and generated audio file into a single Apple HTTP Live Streaming (HLS) stack ready for streaming

Based on what you know about the AWS AI and media services, you defined a workflow for your process that looks like this:



*Workflow for this lab*

You will use AWS AI services to automatically generate the necessary components for creating captions, alternate language subtitles, and alternate language audio tracks for the original video. You will then stitch the components together into an Apple HLS stack using AWS Elemental MediaConvert.



*Dataflow for this lab*

TOPICS COVERED

By the end of this lab, you will be able to:

* Construct a process to automate the generation of captions, alternate language subtitles, and alternate language audio tracks using three AWS AI services – Amazon Transcribe, Amazon Translate, and Amazon Polly
* Generate a transcript of a video using Amazon Transcribe
* Translate the transcript into an alternate language using Amazon Translate
* Transform the alternate language transcript into alternate language audio using Amazon Polly
* Construct an Apple HLS stack using AWS Elemental MediaConvert
* Validate that audio and captions/subtitles work as intended in a video player

TECHNICAL KNOWLEDGE PREREQUISITES

To successfully complete this lab, you should be familiar with basic navigation of the AWS Management Console and be comfortable editing scripts using a text editor.

**Start lab**

1. To launch the lab, at the top of the page, choose **Start lab**.

**Caution:** You must wait for the provisioned AWS services to be ready before you can continue.

1. To open the lab, choose **Open Console**.

You are automatically signed in to the AWS Management Console in a new web browser tab.

**WARNING:** **Do not change the Region unless instructed.**

COMMON SIGN-IN ERRORS

**Error: You must first sign out**



If you see the message, **You must first log out before logging into a different AWS account:**

* Choose the **click here** link.
* Close your **Amazon Web Services Sign In** web browser tab and return to your initial lab page.
* Choose **Open Console** again.

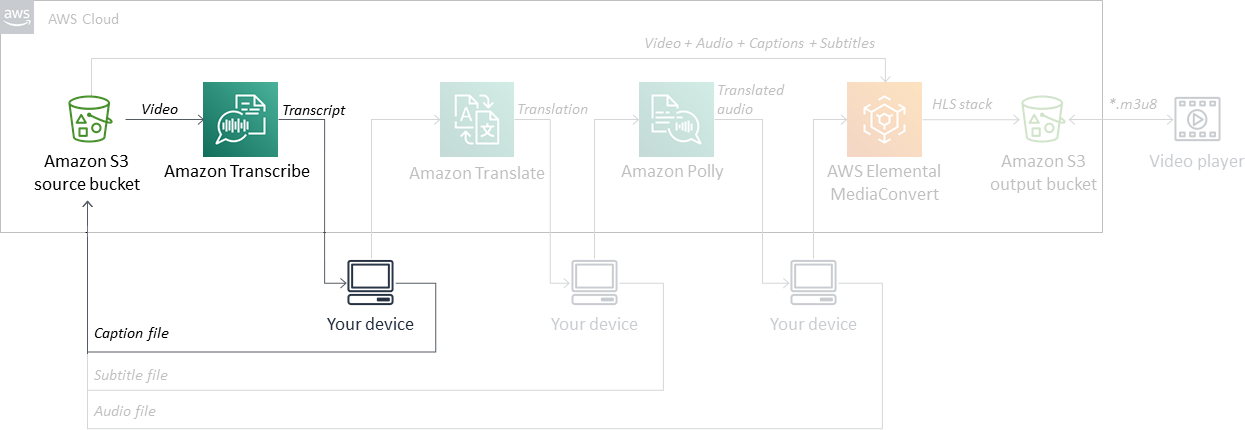
**Error: Choosing Start Lab has no effect**

In some cases, certain pop-up or script blocker web browser extensions might prevent the **Start Lab** button from working as intended. If you experience an issue starting the lab:

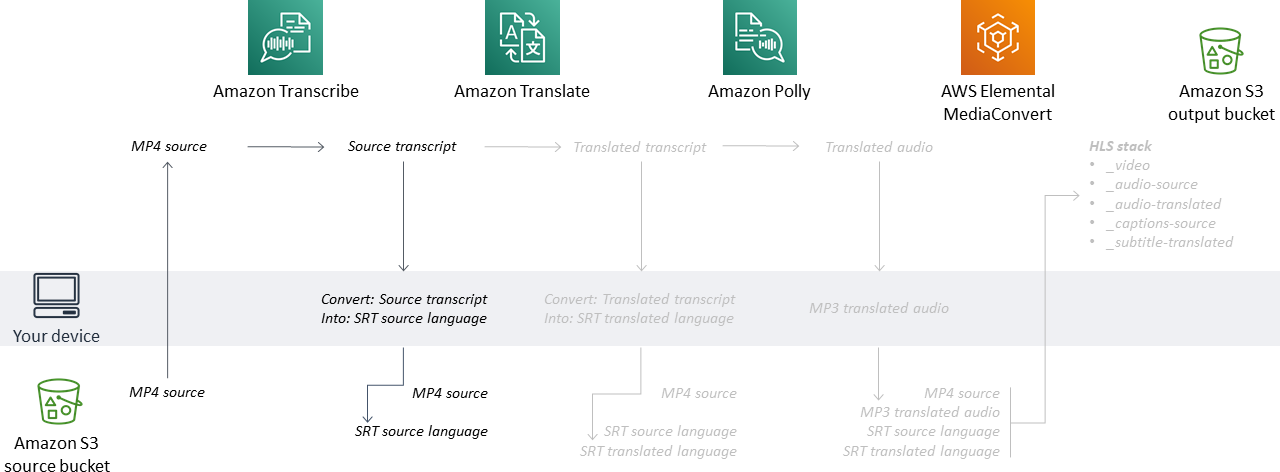
* Add the lab domain name to your pop-up or script blocker’s allow list or turn it off.
* Refresh the page and try again.

**Task 1: Generate a transcript of a video using Amazon Transcribe**

Based on the requirements, the first step in the process is to transcribe the video with Amazon Transcribe.



*Workflow for Task 1*



*Dataflow for Task 1*

PREVIEW THE TEST CLIP

1. In the **AWS Management Console**, on the **Services** menu, enter

s3

 into the search field.

1. Click **S3** in the results.

The **Amazon S3** home page displays.

1. From the list of S3 buckets, select the **source-lfxxxx** bucket.

**Note:** Each bucket in Amazon Simple Storage Service (Amazon S3) must be uniquely named. During this lab’s setup, the AWS CloudFormation template created a bucket named *source-lfxxxx*, where *xxxx* is a unique alpha-numeric identifier.

1. Click the file **test-clip-00m03s-1080p30.mp4**.
2. Click **Object actions** > **Download as**.
3. View the video in an mp4 player.

**Note:** The video is ideal for prototyping a workflow. It has high-quality video, a short duration, and a clear speaker.

1. Copy the **S3 URI** to your clipboard.
2. Paste the path into a text file, which you will use in the next section to tell Amazon Transcribe where to find the test video.

The pasted value should look similar to this: *s3://source-lfxxxx/test-clip-00m03s-1080p30.mp4*

GENERATE A TRANSCRIPT

Now that you have a clear and concise video for prototyping, the next step is to use Amazon Transcribe to create a transcript of the video.

1. In the **AWS Management Console**, expand the **Services** menu, and enter

transcribe

 into the search field.

1. Right-click on **Amazon Transcribe** in the results and choose **Open link in new tab**.

**Note:** On macOS, use ⌘-click.

1. Go to the new **Amazon Transcribe** tab.
2. On the **Amazon Transcribe** console page, view the left pane by selecting the menu icon .
3. In the left pane, select **Transcription jobs**, and then click **Create job** to create a transcription job.
4. In the **Job settings** section, configure the following:

* **Name:**

transcribe-test-clip

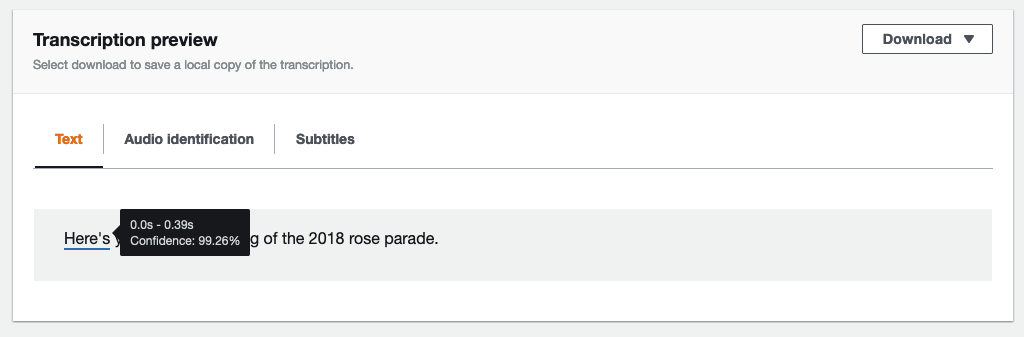
* **Language:** **English, US (en-US)**

1. In the **Input data** section, configure the following:

* **Input file location on S3:** Path to the video copied earlier, such as: *s3://source-lfxxxx/test-clip-00m03s-1080p30.mp4*

1. Click **Next** and then **Create job** to start the transcription job.
2. Wait for *approximately 45–60 seconds* for the transcription job to complete.
3. Under **Transcription jobs** section, select the job you just completed **transcribe-test-clip**.

In the **Transcription preview**, you should see the transcribed text.



1. Pause on the individual words, and you will see the transcribed text as well as additional metadata for each transcribed word, including:

* start\_time
* end\_time
* Confidence

1. Notice the *start\_time* for the first word and the *end\_time* for the final word in the sentence. This information provides the duration of the audio in the video clip. You will use these times later in this lab.
2. Copy the text from the **Text preview** section:

Here's your official opening of the 2018 rose parade.

1. Paste the text into a text file. In the next section, you will translate this text with Amazon Translate.

OPTIONAL: DOWNLOAD AND VIEW THE FULL TRANSCRIPT

For the purposes of prototyping, you only need the text in the **Transcription preview** section; however, in an automated process, your application would be working with the generated JSON data. You can download a copy of the data and review it.

1. In the top right of the **Transcription preview** pane, expand the **Download** drop down and select **Download transcript**.

* Save the file in a known location.

This JSON output is what you would use to create an automated workflow. While JSON is ideal for machines to read and is useful for automation, since we are doing the work to prototype the workflow in this lab it can be challenging for a human to read raw JSON. You can use an online JSON to YAML converter (such as, https://www.json2yaml.com/) to create a more human readable text, for example:

---

jobName: transcribe-test-clip

accountId: '111222333'

results:

transcripts:

- transcript: Here's your official opening of the 2018 rose parade.

items:

- start\_time: '0.04'

end\_time: '0.43'

alternatives:

- confidence: '0.9746'

content: Here's

type: pronunciation

...

- start\_time: '2.93'

end\_time: '3.37'

alternatives:

- confidence: '0.9998'

content: Parade

type: pronunciation

- alternatives:

- confidence: '0.0'

content: "."

type: punctuation

status: COMPLETED

TURN THE TRANSCRIPT INTO A CAPTIONS FILE

To overlay captions onto the video, you must generate a captions file that informs the video player what text to overlay on the video during what times. During the prototyping for this lab you will do these steps by hand; however, in an automated routine you would want to create a small script that copies and pastes the various values and formats the text into a captions file.

**Note:** Throughout most of the world, the terms *captions* and *subtitles* are used interchangeably. In the United States and Canada, the terms are often referenced as:

* **Captions:** Same language as the audio with additional, significant information from the audio track, such as speaker name or [sound information].
* **Subtitles:** Translation of foreign audio or clarification of audio (such as for a thick accent or noisy environment).

Functionally, both captions and subtitles are the same, as they both are characters that are overlaid onto video.

Many captions formats are available. This lab uses one of the simplest formats for captioning – a SubRip Subtitle (SRT) file. An SRT file contains the caption number, timing information, and text to be displayed.

To turn the transcript from Amazon Transcribe into a valid captions file, complete the following steps:

1. In a text editor, create a new file, and copy this (and only this) text into it:

1

00:00:00,000 --> 00:00:00,000

Replace me with desired text

1. Save the file as

test-clip-00m03s-1080p30.mp4\_eng.srt

 in a known location.

**Note:** If you use Notepad on Windows or Text Edit on Mac those text editors will force the file extension to a be something other than \*.srt. If this happens to you, close the text editor, navigate to and rename the captions file to the proper file extension, and then open the \*.srt captions file with your text editor of choice.

1. Replace the text with the transcription text from Amazon Transcribe.

Your file should now look like this:

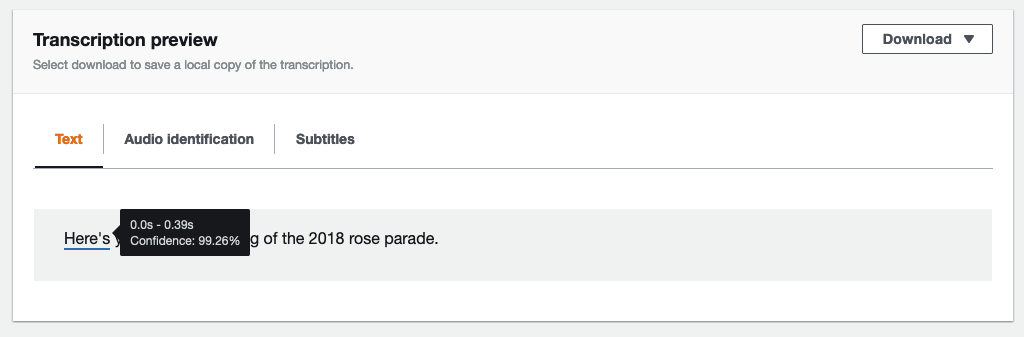
1

00:00:00,000 --> 00:00:00,000

Here's your official opening of the 2018 Rose Parade.

Now, you must add timing information to the file so that it starts and ends when the audio does.

1. Navigate to your browser tab for **Amazon Transcribe**.
2. To determine the *start\_time* and *end\_time* for the audio, view your completed Transcription job, go to the Transcription preview section, and pause on the first and last characters.



1. Return to the text editor where you created the SRT file.
2. Replace the start\_time and end\_time in the SRT file with the times displayed in the Transcription preview. You must convert those times to the time format standard for SRT files:

* *start\_time == 0.04s == 00:00:00,040*
* *end\_time == 3.37s == 00:00:03,370*

Your SRT file should now look like this:

1

00:00:00,040 --> 00:00:03,370

Here's your official opening of the 2018 Rose Parade.

1. Save the SRT file.

UPLOAD THE SRT FILE

1. Navigate to the browser tab with **Amazon S3** open.
2. If not already there, navigate to the top level of **source-lfxxxx**.
3. To upload your SRT file, first click **Upload** to open the upload dialog.
4. In the dialog box, select **Add files** navigate to and select the file

test-clip-00m03s-1080p30.mp4\_eng.srt

, and click **Open**.

1. In the lower-left of the dialog window click **Upload** to upload the file.

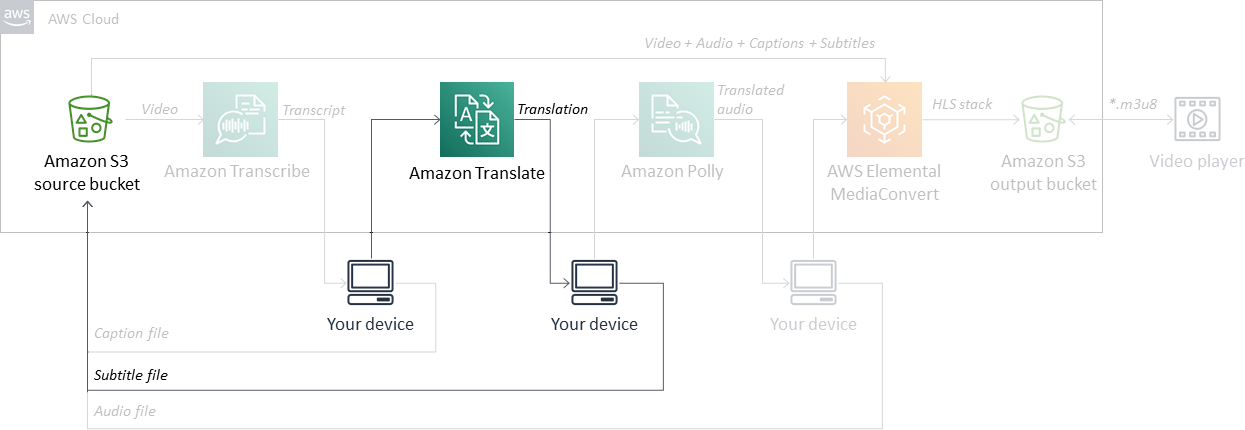
Great work!

You successfully used Amazon Transcribe to automatically create a transcript of a video file. You also converted the transcript into an SRT file, which can be used to create a closed caption overlay onto the video.

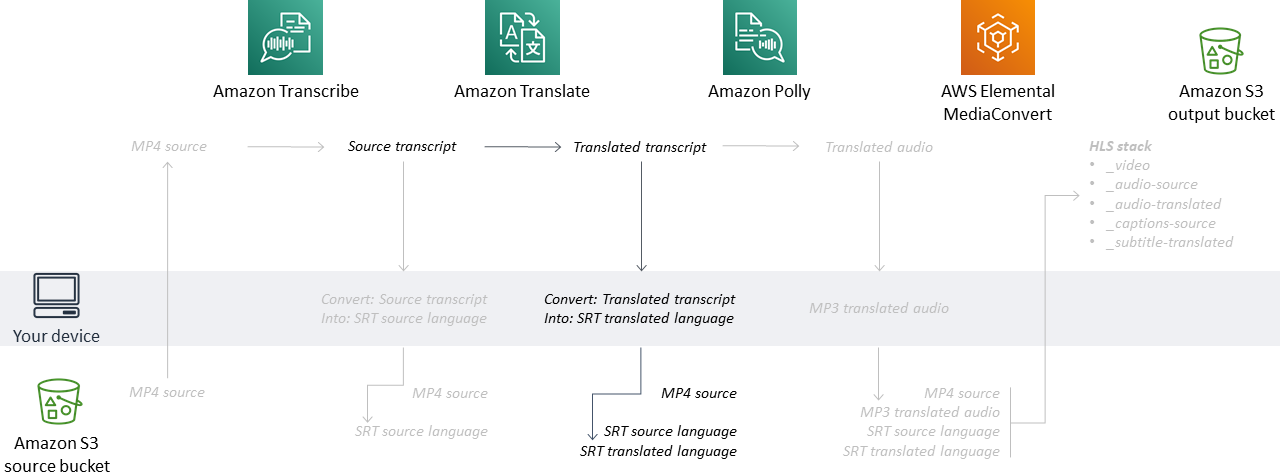
**Task 2: Translate transcript into an alternate language using Amazon Translate**

TRANSLATE THE TRANSCRIPT

Following the lab requirements, the second step in the process is to translate the video with Amazon Translate.



*Workflow for Task 2*



*Dataflow for Task 2*

1. In the **AWS Management Console**, expand the **Services** menu, and enter

translate

 into the search field.

1. Right-click on **Amazon Translate** in the results and choose **Open link in new tab**.

**Note:** On macOS, use ⌘-click.

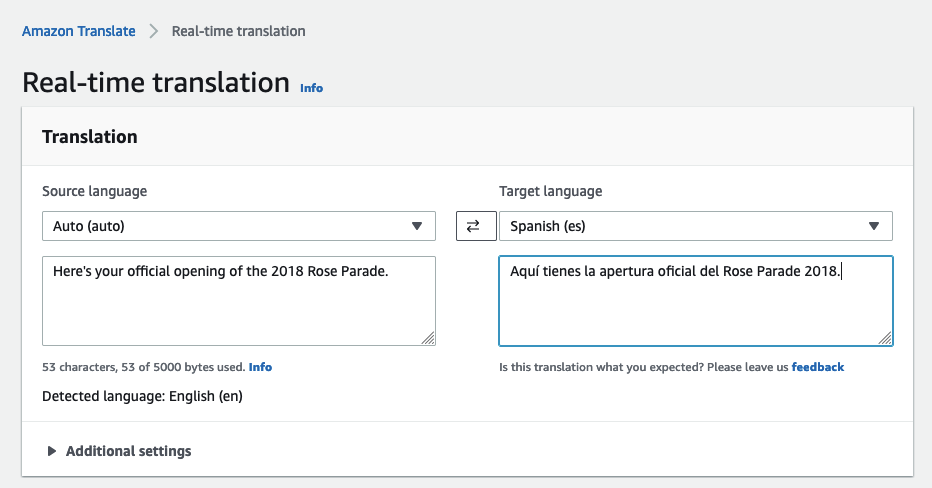
1. Go to the new **Amazon Translate** tab.
2. On the **Amazon Translate** console page, in the left pane, select **Real-time translation**.
3. Leave **Source language** as *Auto (auto)*.
4. For the Source language text field, paste in the transcribed text generated by Amazon Transcribe in [Task 1](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-1).

* Input:

Here's your official opening of the 2018 Rose Parade.

1. For **Target language**, select *Spanish (es)*, and watch the text be translated almost instantly.

* Output: *Aquí tienes la apertura oficial del Rose Parade 2018.*



You can also expand the  **Application integration** section to see how to configure the **JSON request** and parse the **JSON response** for later automation of this process.

HOW TO USE TRANSLATION IN YOUR WORKFLOW

The requirements note that using a translation service is too expensive and time consuming for your application. For this lab, you will simply copy and use the translation created by Amazon Translate. However, in the real world, you might need to meet regulatory requirements or quality standards. Therefore, you would typically have the translation validated by a native language speaker before adding it to your application.

In this case, the *Aquí tienes la apertura oficial del Rose Parade 2018.* output is technically accurate, but the direct translation misses some syntax and capitalization. In addition, given the informal nature of the clip, the voice might be too formal. Amazon Translate provides a good starting point. Looking at the output for this lab, a native speaker who reviews it might modify the translation as follows:

*Esta es su inauguración oficial del Desfile de las Rosas del 2018.*

While you might not use the Amazon Translate output verbatim in your final product, using the service can save you a significant amount of time. You can have it do the bulk of the translation. Then, a language expert can use the output to create a high-quality end product on a reduced timeline.

In this lab the goal is to prototype a fully-automated workflow; therefore, you will use the output from Amazon Translate.

CONVERT THE TRANSLATED TRANSCRIPT INTO A SUBTITLE FILE

To create a subtitle file you would repeat almost the same process as you did in [Task 1](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-1). As a reminder, previously you:

* Created a subtitles file.
* Added the desired subtitle text.
* Set the timing window for the subtitle text.

1. In a text editor, create a new file, and save as

test-clip-00m03s-1080p30.mp4\_spa.srt

 in a known location.

You will use the same base format and the same timing information as you did in [Task 1](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-1).

1. Copy the completed subtitle information into the SRT file:

1

00:00:00,040 --> 00:00:03,370

Aquí tienes la apertura oficial del Rose Parade 2018.

1. Save and close the SRT file.

**Note:** As seen in [Task 1](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-1), if you use Notepad on Windows or Text Edit on Mac those text editors will force the file extension to a be something other than \*.srt. If this happens to you, close the text editor, navigate to and rename the captions file to the proper file extension, and then open the \*.srt captions file with your text editor of choice.

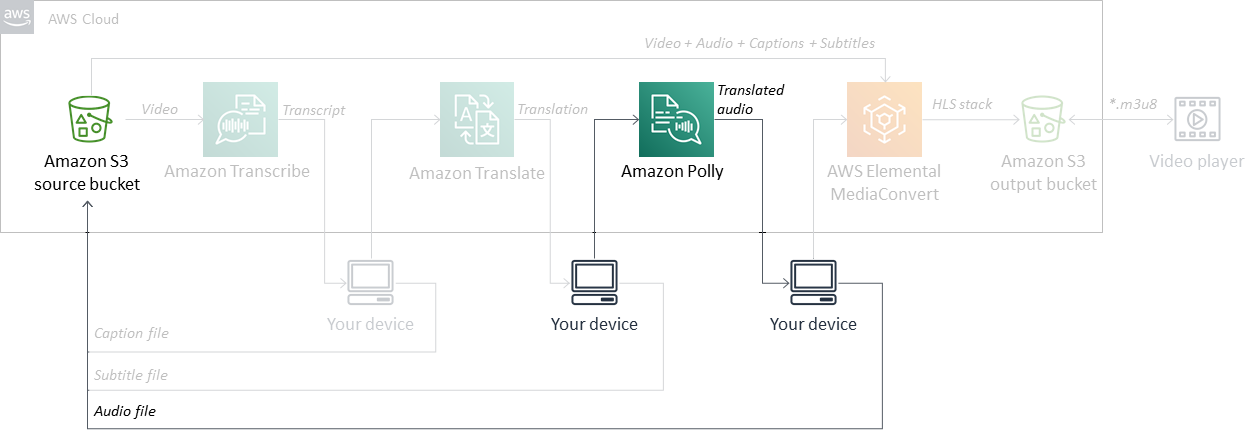
1. Navigate to the browser tab with **Amazon S3** open, and, if not already there, navigate to the top level of **source-lfxxxx**.
2. Upload the new file.

Great work!

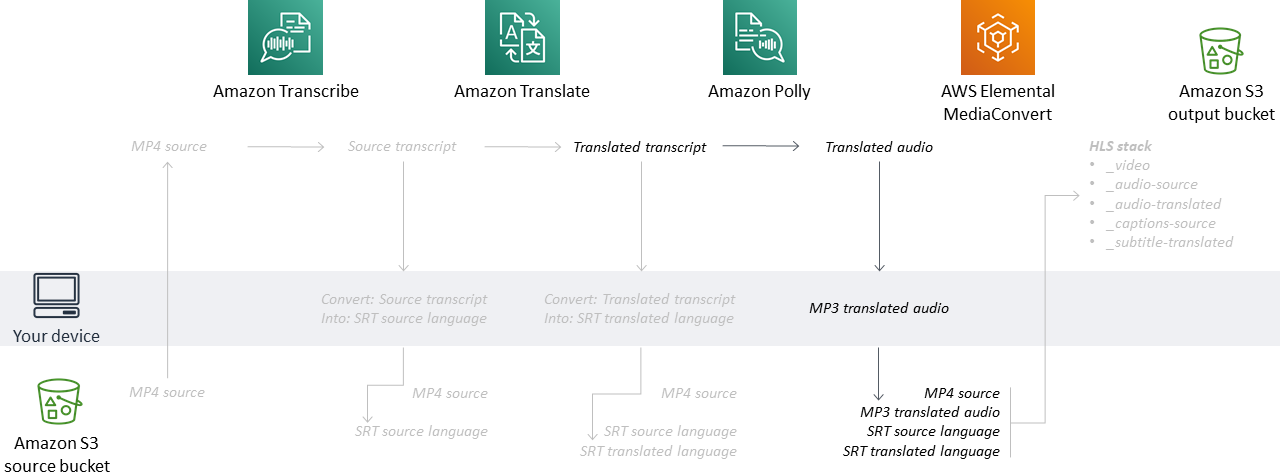
You successfully used Amazon Translate to automatically create a translation of a transcript. You then converted the translation into an SRT file, which can be used to create a subtitle overlay onto the video.

**Task 3: Transform the alternate language transcript into alternate language audio using Amazon Polly**

Following the lab requirements, the third step in the process is to take the translated transcript and generate alternate language audio with Amazon Polly.



*Workflow for Task 3*



*Dataflow for Task 3*

CREATE ALTERNATE LANGUAGE AUDIO

1. In the **AWS Management Console**, expand the **Services** menu, and enter

polly

 into the search field.

1. Right-click on **Amazon Polly** in the results and choose **Open link in new tab**.

**Note:** On macOS, use ⌘-click.

1. Go to the new **Amazon Polly** tab.
2. Select **Try Polly**
3. Copy the validated translated text used in [Task 2](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-1):

Aquí tienes la apertura oficial del Rose Parade 2018.

1. For the **Input text**, paste the text into the text box, and select the following:

* **Engine:** Standard
* **Language:** Spanish, US
* **Voice:** Miguel, Male

1. Click **Listen**

While the audio sounds good, it’s longer than the original audio. If you click **Download** and look at the properties for this audio, you can see that the audio is longer than the original audio, which is 3.33s. For this application, an extended duration is not acceptable. You must find a way to shorten the audio duration without losing any of the content.

1. Since this file is too long, delete it. You will be creating a file in subsequent steps with the correct duration.

CREATE PROPERLY TIMED ALTERNATE LANGUAGE AUDIO

To have more control over the output, you must use Speech Synthesis Markup Language (SSML).

**Note:** To learn more about SSML, click the **Info** link next to the **SSML** rocker. However, studying SSML commands is not a requirement for this lab.

For this lab, you want to increase the audio speed property. To do this, you can use the [supported SSML tag](https://docs.aws.amazon.com/polly/latest/dg/supportedtags.html) for *Setting a Maximum Duration for Synthesized Speech*: *<prosody amazon:max-duration>*

1. To determine what the max-duration should be, navigate to the browser tab with **Amazon Transcribe**, look in the Transcription preview area, and determine the difference between the *start\_time* of the audio and the *end\_time* of the audio.

In this case, *start\_time = 0.04s* and *end\_time = 3.37s*. Therefore, the max-duration allowed is **3.33s**.

**Note:** First you will read through how to do the process, then you’ll add the completed values.

The basic construct for SSML is that any speech will be wrapped in a *speak* tag, and other tags are added inside the speak tag. To create everything you would follow a process that looked like this.

* In **Amazon Polly** select the SSML tab, click in the text box, delete any existing text, and paste the translated transcript: *Aquí está su inauguración oficial del desfile de rosas 2018.*
* To set the max-duration, you would wrap the text to be spoken with this tag: *<prosody amazon:max-duration=“3.33s”></prosody>*
* To denote this is the phrase you want spoken, wrap the whole thing with this tag: *<speak></speak>*

Now it’s time to use **Amazon Polly** to generate the Text-to-Speech with our desired tags:

1. Navigate to the browser tab with **Amazon Polly** open.
2. Turn the rocker for SSML on, delete any existing text in the **Input text** field, and paste the following text:

<speak>

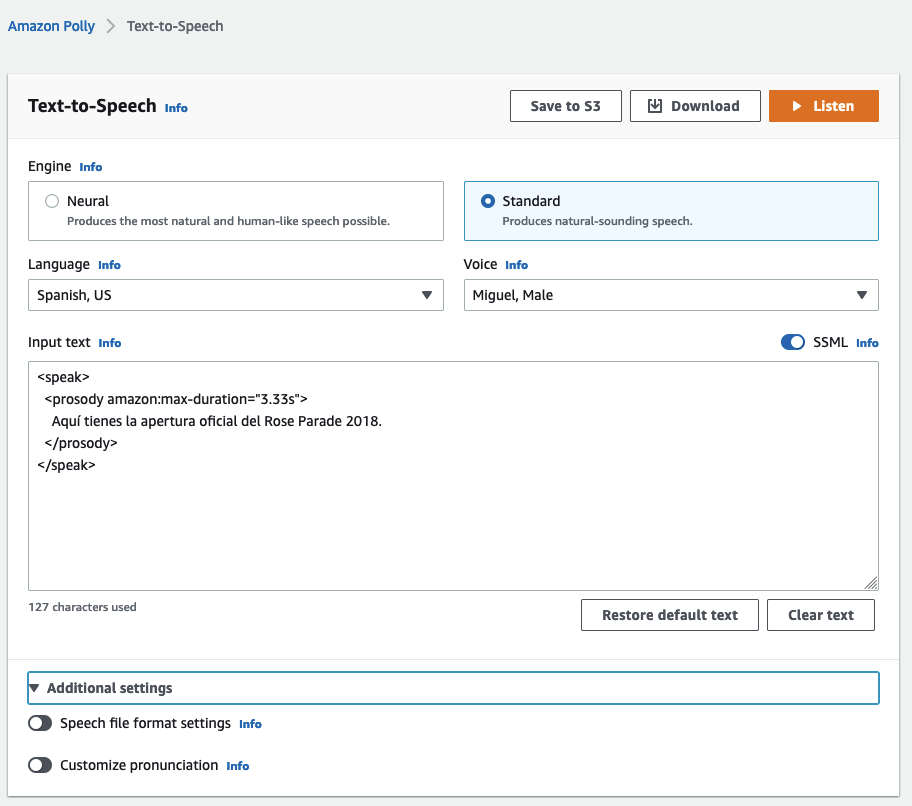
<prosody amazon:max-duration="3.33s">

Aquí tienes la apertura oficial del Rose Parade 2018.

</prosody>

</speak>

Your final configuration for Amazon Polly should look like this:



1. Now if you click **Listen** the alternate language audio duration matches the original audio’s duration.
2. Click **Download** to download a copy of the alternate language audio.
3. Rename the downloaded file to

test-clip-00m03s-1080p30.mp4\_spa.mp3

1. Navigate to the browser tab with **Amazon S3** open, and navigate to the top level of **source-lfxxxx**.
2. Upload the mp4 file that you just downloaded.

Your *source-lfxxxx* bucketshould now have the following four files:

* *test-clip-00m03s-1080p30.mp4* – Provided with the lab
* *test-clip-00m03s-1080p30.mp4\_eng.srt* – [Task 1](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-1)
* *test-clip-00m03s-1080p30.mp4\_spa.srt* – [Task 2](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-2)
* *test-clip-00m03s-1080p30.mp4\_spa.mp3* – [Task 3](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-3)

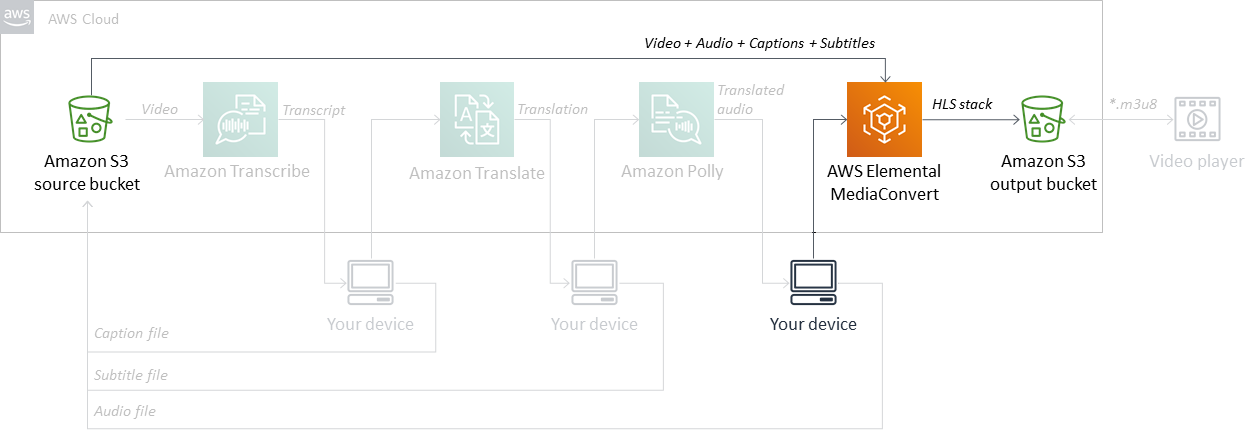
Great work!

You successfully used Amazon Polly to automatically create an alternate language audio track from the alternate language transcript. You then downloaded an MP3 of the audio, which will give users an option to select an alternate language audio track to play for the video.

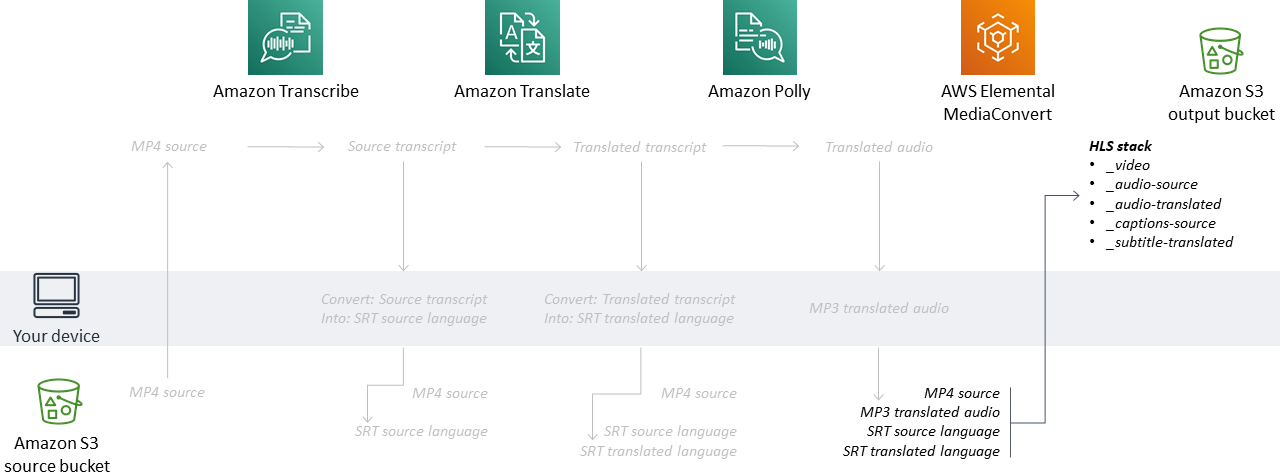
In addition to doing all of that, you also made sure that the alternate language audio track exactly matches the timing of the original audio track!

**Task 4: Construct an Apple HLS stack using AWS Elemental MediaConvert**

Following the lab requirements, the fourth step in the process is to take each of the components generated by the AWS AI services and create a streaming output using AWS Elemental MediaConvert.



*Workflow for Task 4*



*Dataflow for Task 4*

1. In the **AWS Management Console**, expand the **Services** menu, and enter

convert

 into the search field.

1. Right-click on **MediaConvert** in the results and choose **Open link in new tab**.

**Note:** On macOS, use ⌘-click.

1. Go to the new **AWS Elemental MediaConvert** tab.
2. On the **AWS Elemental MediaConvert** console page, click the menu icon .
3. In the left pane, click **Jobs**, and then select **Create job**

**Note:** There are a large number of steps in [Task 4](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-4), none of your work can be saved until you click Create at the end. Do not click Back or refresh in your browser or you will lose your current configurations.

If you for some reason lose your configurations you can start from configured file by going to the [Bonus task](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#bonus) for how to import a MediaConvert template and you can modify the template file in the *solutions-lfxxx/downloads/* folder to reference your Account ID and reference the files from the *solutions-lfxxxx/source/* folder.

CONFIGURE THE INPUTS

You currently have four files to use as inputs:

* Original video and audio file
* Alternate language audio file
* Captions file
* Alternate language subtitle file

These input sources will be used to create an output suitable for streaming.

ATTACH THE VIDEO INPUT

1. Once in the **Create job** window verify that in the **Job** section **Inputs >> Input 1** is selected.
2. In the **Input 1** section, select **Browse**
3. In the **Choose a location** pop-up, select the following:

* **S3 bucket:** source-lfxxxx
* **File:** test-clip-00m03s-1080p30.mp4

1. Click **Choose**

Your video is configured as the input.

CONFIGURE THE INPUT’S ORIGINAL AUDIO

1. On the **Create job** page, scroll to the **Audio selectors** section, find Audio Selector 1, and configure the following:

* **Use as default:** Checked
* **Selector type:** Language code
* **Language code:** English

Now, the video and original audio inputs are configured.

CONFIGURE THE INPUT’S ALTERNATE LANGUAGE AUDIO

You currently have a singular audio track configured. You must add another audio track for the alternate language audio.

1. On the **Create job** page, in the **Audio selectors** section, click **Add audio selector**
2. For Audio Selector 2, configure the following:

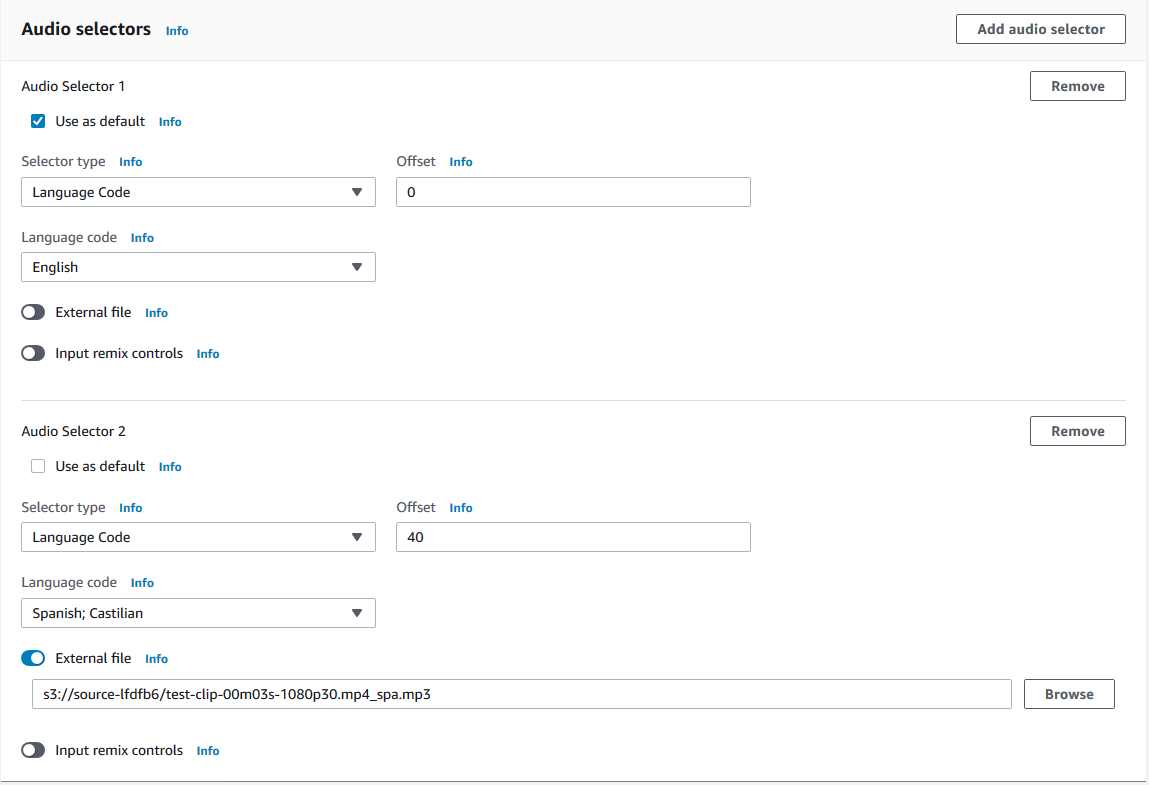
* **Use as default:** Unchecked
* **Selector type:** Language code
* **Offset:** 40
* **Language code:** Spanish; Castilian
* **External file:** On

1. To point to the external audio file, select **Browse** and configure the following:

* **S3 bucket:** source-lfxxxx
* **Location:** test-clip-00m03s-1080p30.mp4\_spa.mp3
* Click **Choose**

And for one last detail, the generated audio is only to the duration of when the speaker is speaking. While the original audio track lasts for the duration of the test-clip, the generated audio is only the duration of when the speaker is speaking. Since you noted in [Task 1](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-1) that the speaker doesn’t start until 0.04s (40ms) into the video clip, you must offset the start of the alternate audio, too.

The fully configured **Audio selectors** section is shown here.



The video, original audio, and alternate audio inputs are configured.

CONFIGURE THE INPUT’S CAPTIONS

Next, you point to the caption file’s location.

1. On the **Create job** page, scroll to the **Captions selectors** section, and click **Add captions selector**
2. For Captions Selector 1, select the following:

* **Source:** SRT
* **Source file:** **Browse**

1. In the Browse dialog box, configure:

* **S3 bucket:** source-lfxxxx
* **Location:** test-clip-00m03s-1080p30.mp4\_eng.srt

1. Click **Choose**

The video, original audio, alternate audio, and caption inputs are configured.

CONFIGURE THE INPUT’S SUBTITLES

Next, you must point to where the caption file is located.

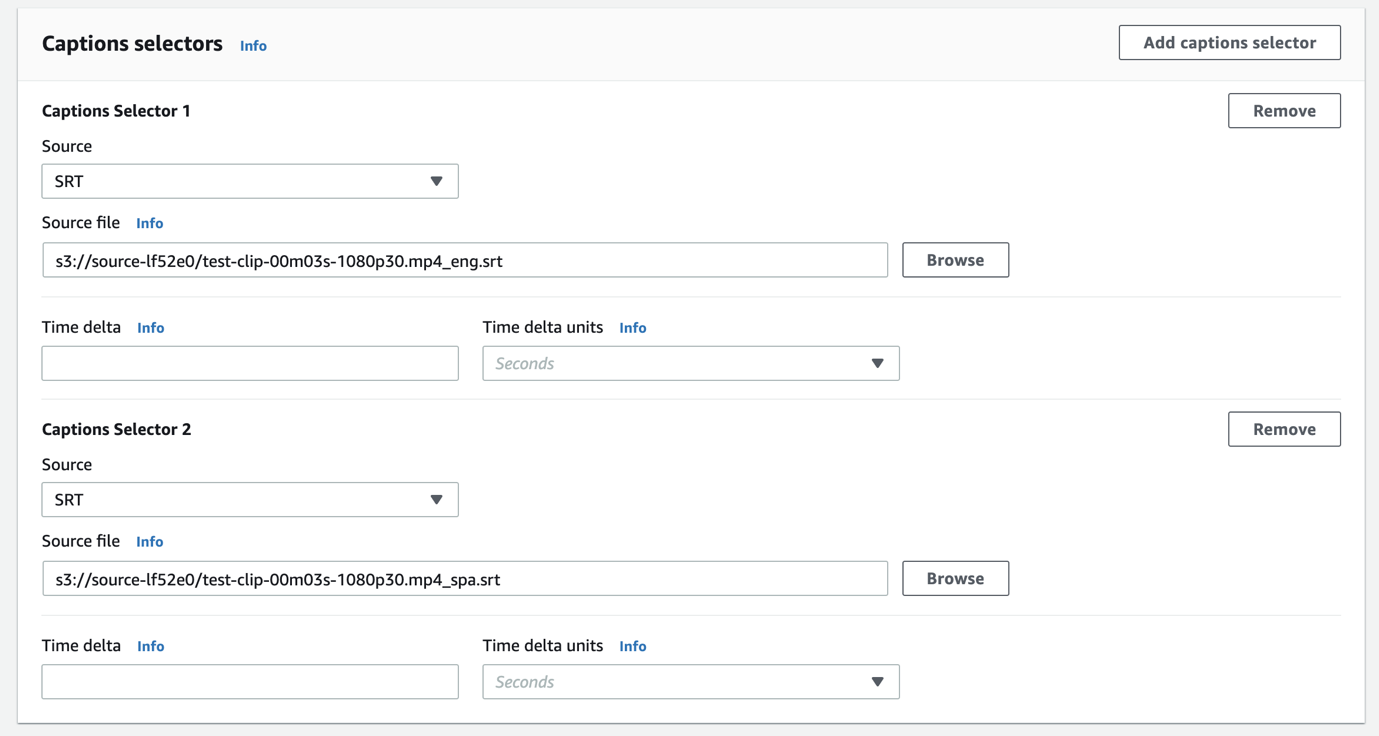
1. In the **Captions selectors** section, click **Add captions selector**
2. For Captions Selector 2, configure the following:

* **Source:** SRT
* **Source file:** **Browse**

1. In the Browse dialog box, configure:

* **S3 bucket:** source-lfxxxx
* **Location:** test-clip-00m03s-1080p30.mp4\_spa.srt
* Click **Choose**

The fully configured **Captions selectors** section is shown here.



Congratulations!

Your inputs are configured – video, original audio, alternate audio, caption, and subtitle inputs.

CONFIGURE THE OUTPUT GROUP

Now that the input components are configured, you can set up a streaming output group. You will create a stack of outputs for adaptive bitrate (ABR) streaming using Apple HLS (HLS). HLS is easy to configure and usable by most streaming devices.

1. On the **Create job** page, scroll to the top, locate the Job box in the left column, find **Output groups**, and select **Add**
2. In the **Add output group** pop-up, choose **Apple HLS**, and then click **Select**
3. In the **Apple HLS group settings** section, in the **Destination** section, select **Browse**
4. In the **Choose a location** pop-up, complete the following:

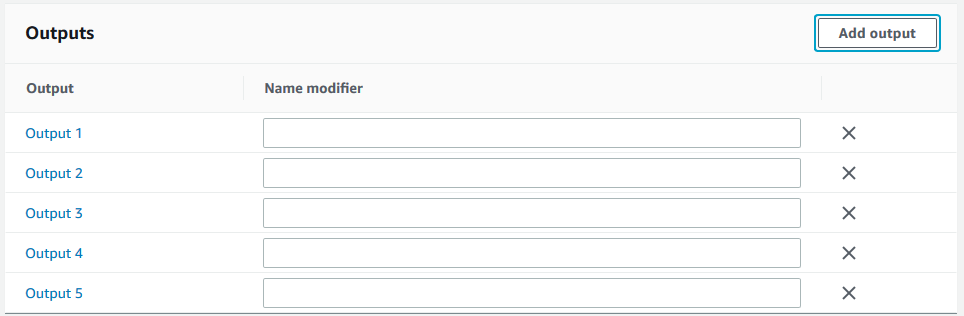
* **S3 bucket:** output-lfxxxx
* **Location:**

test-clip

* Click **Choose**

You must configure the Apple HLS ABR stack to have a total of five outputs.

1. Scroll to the **Outputs** section and click **Add output** four times until there are total of five outputs.



1. Change the **Name modifier** for the outputs to the following:

* **Output 1:**

\_video

* **Output 2:**

\_audio-eng

* **Output 3:**

\_audio-spa

* **Output 4:**

\_captions-eng

* **Output 5:**

\_subtitles-spa

The outputs should now look like this:



You are ready to configure each of the outputs.

OUTPUT 1 (\_VIDEO) – CONFIGURE THE OUTPUT’S VIDEO

For the first output, you only want video – no audio, no captions, just video. And for the video you output, you want to use the high-quality codec options available with AWS Media Services products.

1. On the **Create job** page, in the left column, find the **Output groups** section, find **Apple HLS**, and select **Output 1**.
2. In the **Encoding settings** section, select **Audio 1**.
3. To remove the audio from Output 1, select **Remove audio**

Next, you create a high-quality video output. In the **Encoding settings** section, **Video** should be automatically selected.

1. Configure the following settings:

* **Video codec:** MPEG-4 AVC (H.264)
* **Resolution (w x h):**

1920

 x

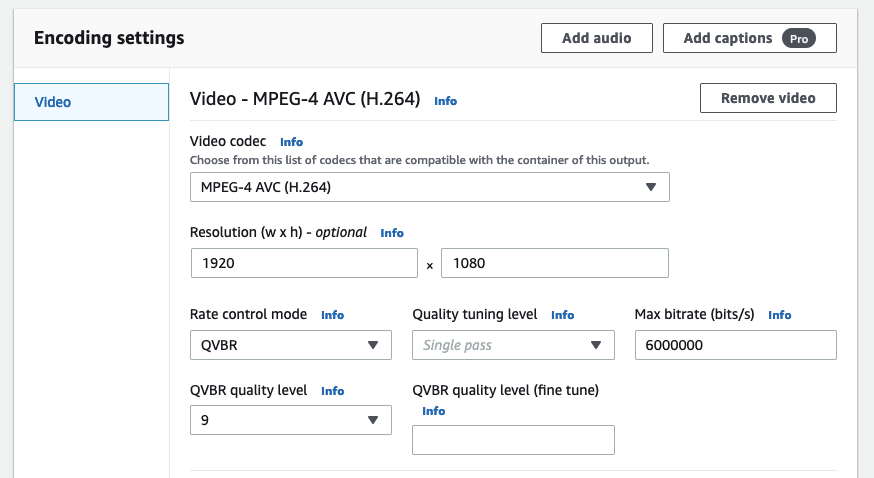
1080

1. Scroll down, and then configure the following additional video settings:

* **Rate control mode:** QVBR
* **Max bitrate (bits/s):**

6000000

* **QVBR quality level:** 9



The video output is set to create a singular, high-quality video output in the HLS stack.

**Note:** Outside of the lab environment, you would want to create multiple video output options to better handle scenarios where you would need an adaptable bitrate. Typically, you would create video outputs of 1080p, 720p, 480p, 360p, 240p, and 144p, as well as any other video output options that are valuable to your end users.

OUTPUT 2 (\_AUDIO-ENG) – CONFIGURE THE OUTPUT’S ORIGINAL AUDIO

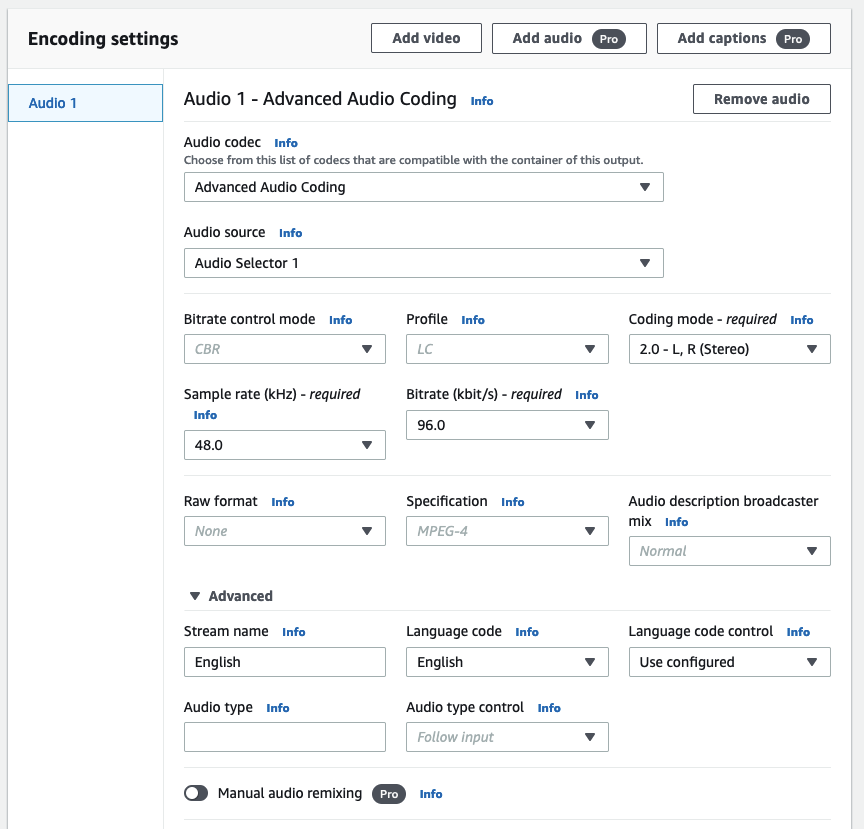
For the second output, you only want the original audio – no video, no captions, just audio.

1. On the **Create jobs** page, in the left column, find the **Output groups** section, find **Apple HLS**, and select **Output 2**.
2. In the **Encoding settings** section, select **Video**.
3. To remove the video from Output 2, select **Remove video**
4. In the **Encoding settings** section, confirm that **Audio 1** is selected.
5. Scroll up to the **Output settings** section and under **Audio** for **Audio track type**, select **Alternate audio, auto select, default**.
6. Scroll back down to the **Encoding settings** section.
7. Expand the  **Advanced** section, and configure the following settings:

* **Stream name:**

English

* **Language code:** English
* **Language code control:** Use configured



The first audio output is now set to use the Audio Selector 1 from the input’s configuration, which is the audio from the original video file. But the real key is that your end users can now choose between that option and an alternate audio, which you create next.

OUTPUT 3 (\_AUDIO-SPA) – CONFIGURE THE OUTPUT’S ALTERNATE LANGUAGE AUDIO

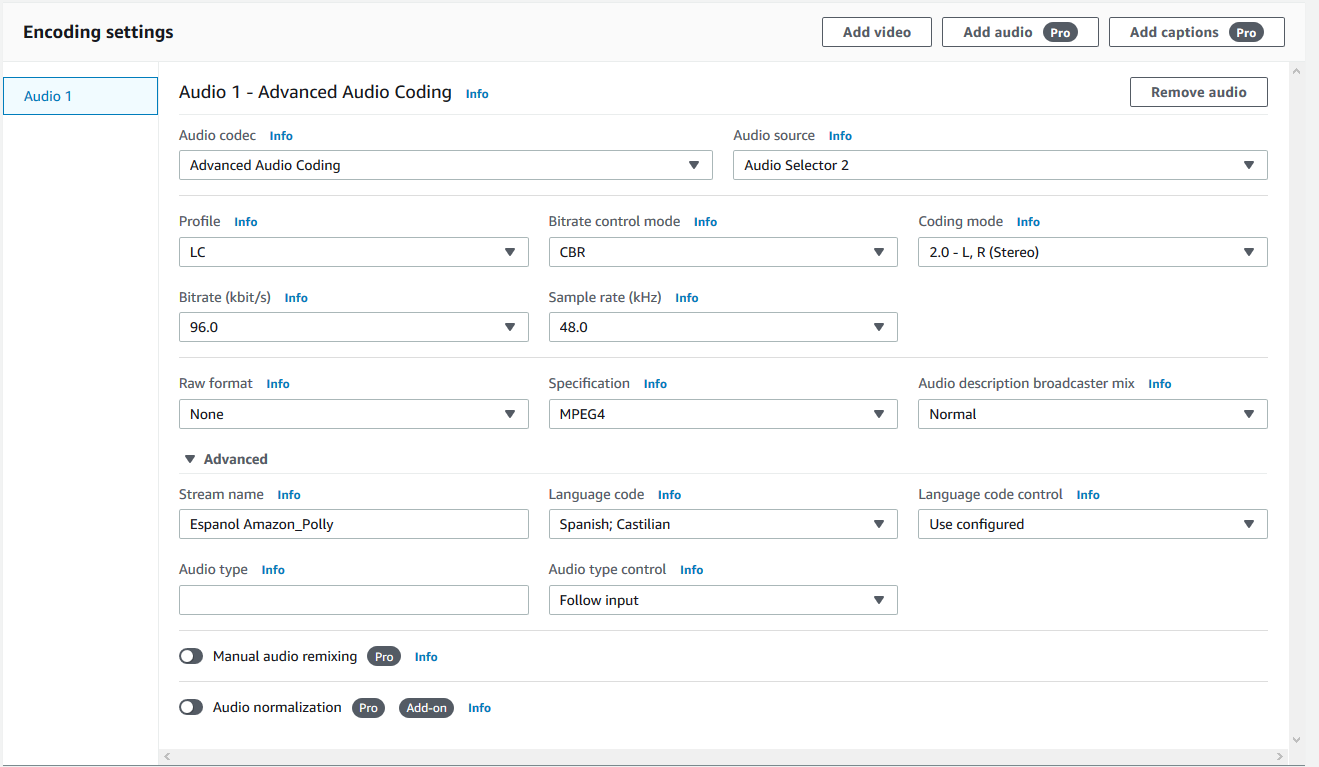
For the third output, you only want the alternate language audio – no video, no captions, just audio.

1. On the **Create jobs** page, in the left column, find the **Output groups** section, find **Apple HLS**, and select **Output 3**.
2. In the **Encoding settings** section, select **Video**.
3. To remove the video from Output 3, select **Remove video**
4. In the **Encoding settings** section, confirm that **Audio 1** is selected.
5. Scroll up to the **Output settings** section and under **Audio** for **Audio track type**, select **Alternate audio, auto select, not default**.
6. Scroll back down to the **Encoding settings** section.
7. Set **Audio source** to **Audio Selector 2**.
8. Expand the **Advanced** section, and configure the following settings:

* **Stream Name:**

Espanol Amazon\_Polly

* **Language code:** Spanish; Castilian
* **Language code control:** Use configured



The second audio output is now set to use the Audio Selector 2 from the input’s configuration, which is the MP3 file generated by Amazon Polly in [Task 3](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-3).

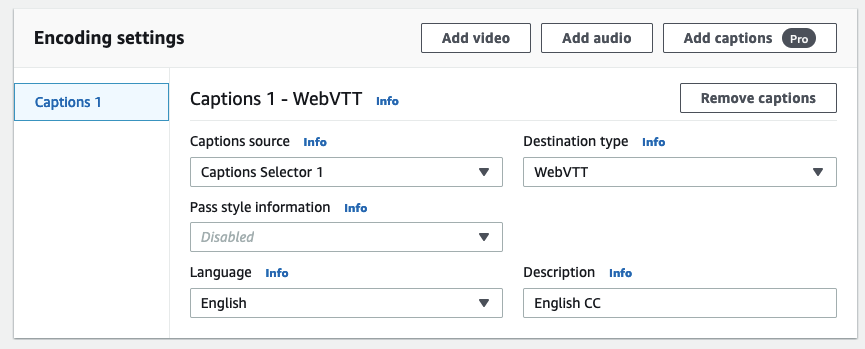
OUTPUT 4 (\_CAPTIONS-ENG) – CONFIGURE THE OUTPUT’S CAPTIONS

For the fourth output, you only want the captions – no video, no audio, just captions.

1. On the **Create jobs** page, in the left column, find the **Output groups** section, find **Apple HLS**, and select **Output 4**.
2. In the **Encoding settings** section, confirm **Video** is selected.
3. To remove the video from Output 4, select the **Remove video** button.
4. In the **Encoding settings** section, confirm **Audio 1** is selected.
5. To remove the audio from Output 4, select the **Remove audio** button.
6. In the **Encoding settings** section, select **Add captions (Pro)**
7. In the **Encoding settings** section, confirm **Captions 1** is selected.
8. For **Captions 1**, configure the following:

* **Captions source:** Captions Selector 1
* **Destination type:** WebVTT
* **Language:** English
* **Description:**

English CC



The first captions output is now set to use the file configured by Captions Selector 1 from the input’s configuration, which is the SRT file you created based on the transcript from Amazon Transcribe in [Task 1](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-1).

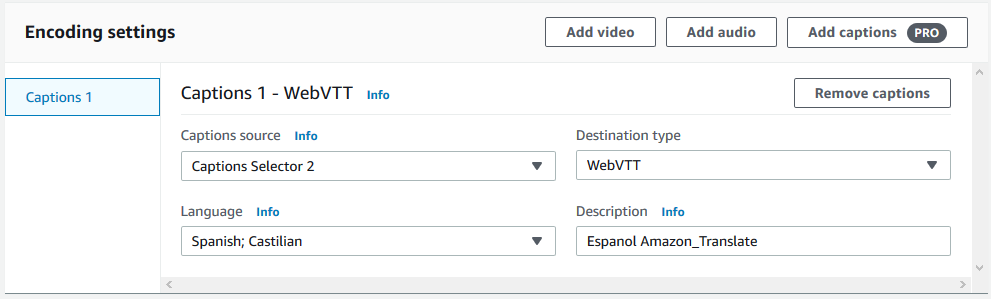
OUTPUT 5 (\_SUBTITLES-SPA) – CONFIGURE THE OUTPUT’S SUBTITLES

For the fifth and final output, you only want the subtitles – no video, no audio, just subtitles.

1. On the **Create jobs** page, in the left column, find the **Output groups** section, find **Apple HLS**, and select **Output 5**.
2. In the **Encoding settings** section, confirm **Video** is selected.
3. To remove the video from Output 5, select the **Remove video** button.
4. In the **Encoding settings** section, confirm **Audio 1** is selected.
5. To remove the audio from Output 5, select the **Remove audio** button.
6. In the **Encoding settings** section, select **Add captions (Pro)**
7. In the **Encoding settings** section, confirm **Captions 1** is selected.
8. For **Captions 1**, configure the following:

* **Captions source:** Captions Selector 2
* **Destination type:** WebVTT
* **Language:** Spanish; Castilian
* **Description:**

Espanol Amazon\_Translate



The second captions output is now set to use the file configured by Captions Selector 2 from the input’s configuration, which is the SRT file you created based on the translation of the transcript that you put into Amazon Translate in [Task 2](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-2).

Congratulations!

You successfully set up an HLS stack to allow end users to select the audio and captions/subtitles that they want to hear and view in a streaming video.

CONFIGURE THE JOB SETTINGS

Moving down the section on the left side of the Create job window, the final section to configure is Job settings.

1. In the **Job settings** section, click **AWS integration**.
2. For **Service role control**, choose **Use an existing service role**.
3. For **Service role**, select **MediaConvertAccessRole**.

**Note:** AWS Identity and Access Management (IAM) roles grant permissions to operators (human or software) to interact with each other and perform their necessary functions. The MediaConvertAccessRole was created as part of the lab setup. MediaConvert assumes this role to access other AWS services.

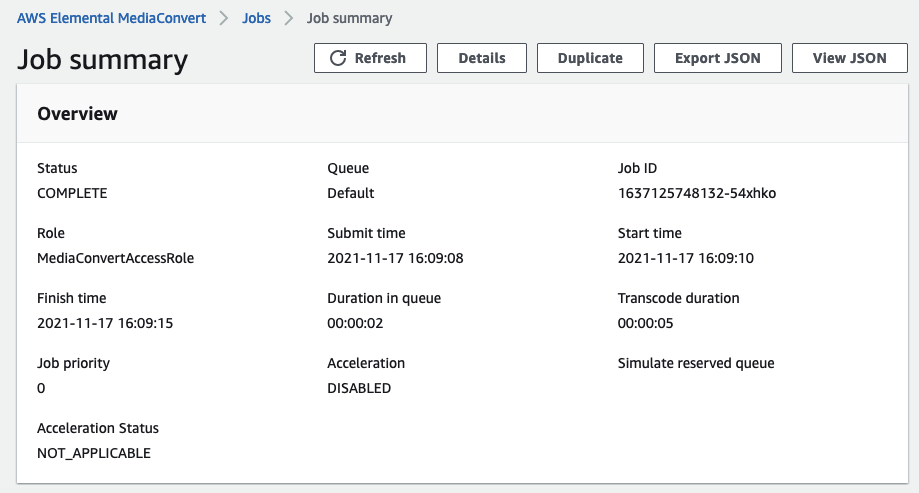
START THE MEDIACONVERT JOB

Your MediaConvert job is now ready to go. All you have to do is:

1. Scroll to the bottom of the **Create job** page.
2. Click **Create**

This MediaConvert job should execute quickly. In about 10–15 seconds, click the **Refresh** button.

You should now see a **Job summary** page that looks like this.



Congratulations!

You successfully set up your first prototype to use AWS AI services to automate global video captioning, subtitling, and audio tracks.

SAVE YOUR WORK

One of the biggest mistakes you can make is to not save your work. Best not make that mistake today! Fortunately, you can easily save your configuration from the **Job summary** page.

1. Click the **View JSON** button.
2. Copy the JSON contents to a new text file and save it in a known location.

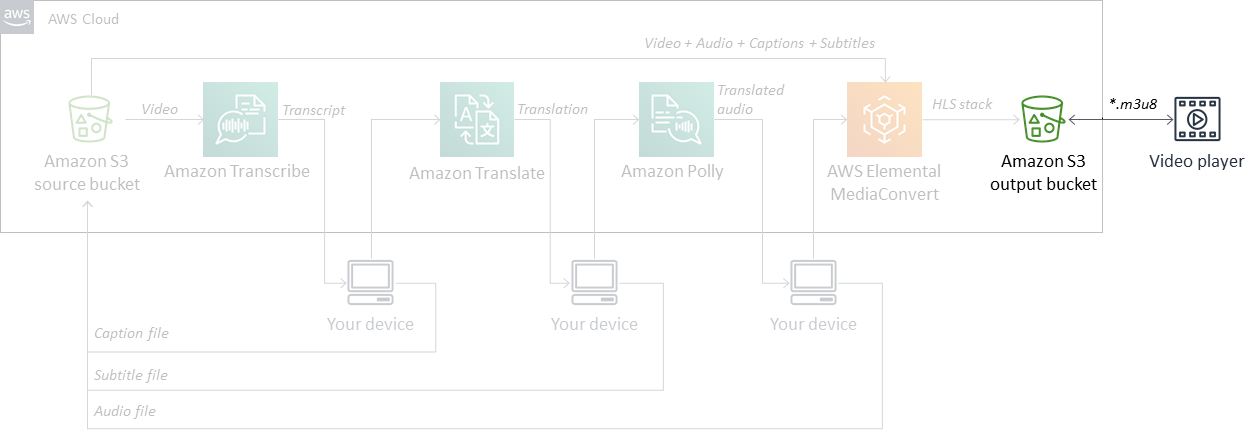
You can now do any of the following:

* Save your settings for archival purposes.
* Open the file to view the JSON created by the job.
* Use the JSON to create a template for quickly creating similar jobs in the future.
  + If you’re interested in this, check out the [Bonus task](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#bonus) at the end of the lesson.

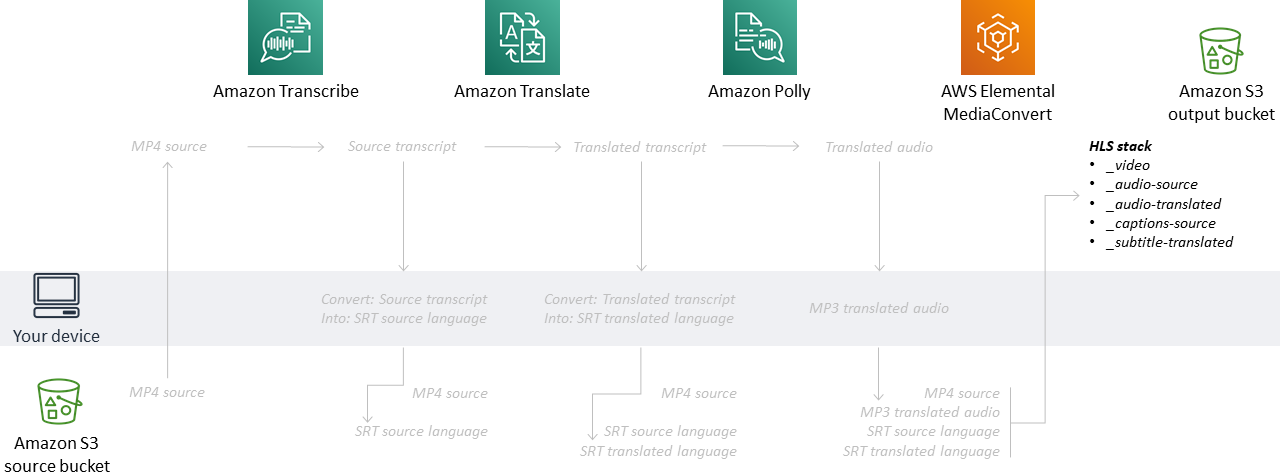
At this point, you probably want to see if this worked. Continue to [Task 5](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-5) to find out.

**Task 5: Validate that audio and captions/subtitles work as intended in a video player**

For Task 5, you use a video player to view the HLS stack you created in your S3 output bucket.



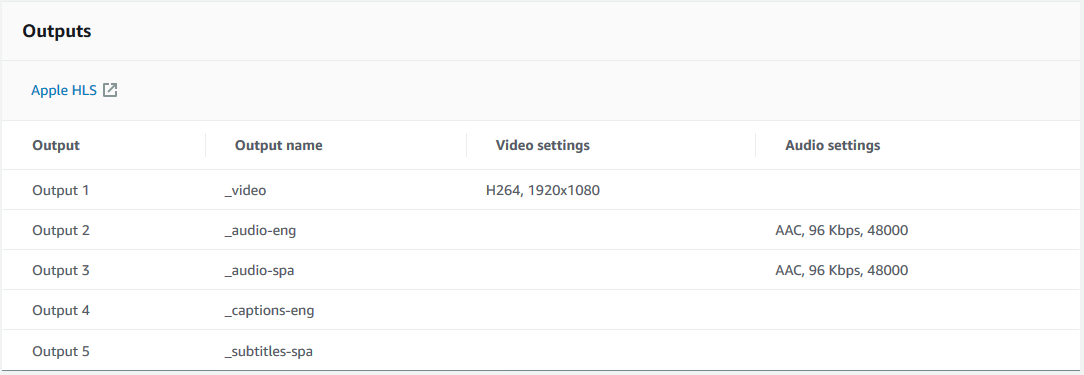
*Workflow for Task 5*



*Dataflow for Task 5*

After completing [Task 4](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-4), you should see that the MediaConvert job has completed and looks like this.

**Note:** If the MediaConvert job doesn’t show as completed, click the **Refresh** button.



1. In the Outputs section, click **Apple HLS** to go to *output-lfxxxx*, where the files from the completed job are located.

This will open a new tab pointed at the location on Amazon S3 where you told MediaConvert to send its output. You should see 11 files generated by MediaConvert to create an HLS stack.

* .m3u8 - Manifests, small text files that reference and list the locations of the actual media segments
* .aac - audio files
* .vtt - captions/subtitle files
* .ts - Transport Stream file containing the mp4 video encoded with h264

Because this prototype is only approximately 4 seconds long, there is only one segment of each of the audio, captions, and video files. However, for longer videos, you would see additional segments, 00002, 00003, and so forth. And the manifest files would point the video player to exactly which files to load from where and when.

CHANGE SETTINGS AND PERMISSIONS FOR PUBLIC VIEWING

The purpose of the \*.m3u8 manifest files are to serve as a header and to point a player to the components the player wants to stream.

S3 is set by default to ensure privacy and security.

To allow a player to access other files referenced by the .m3u8 manifest in the same S3 bucket, you must create a [Cross-Origin Resource Sharing (CORS)](https://docs.aws.amazon.com/AmazonS3/latest/dev/cors.html) policy.

1. Select the **Permissions** tab.
2. Scroll down to the **Cross-origin resource sharing** section, then click **Edit**
3. In the text box, paste:

[

{

"AllowedHeaders": [

"\*"

],

"AllowedMethods": [

"GET",

"HEAD"

],

"AllowedOrigins": [

"\*"

],

"ExposeHeaders": []

}

]

1. Click **Save changes**

Your S3 bucket is now configured to allow the test player you will use to access all the files in the bucket.

**Note:** This lab uses the JWPlayer video test player. If you have a preferred video player, you can use it in the following steps. To do this, you must make sure that your preferred test player’s URL is allowed in the CORS policy above.

To make the files publicly accessible:

1. On the **Objects** tab, select all the files by checking the box by **Name**, and from **Actions** select **Make public via ACL**.
2. Click **Make public**

All the selected files can now be accessed by anyone on the web.

**Note:** You typically would not want an end user directly accessing an S3 bucket. This is an acceptable step for testing, but in a production use case, you would want your S3 output bucket reflected publicly via a content delivery network (CDN), such as Amazon CloudFront.

1. Once all the files have successfully been edited for public access, select **Close**

VIEW THE VIDEO

It’s time to stream some video!

First, you need the URL for the top-level header in the HLS stack.

1. From the **Objects** tab for the *output-lfxxxx* bucket in S3, select the **Name** column twice to sort by alphabetical order.
2. Select the top-level header file for the HLS stack

* This should be the shortest-named item in the list: *test-clip.m3u8*

1. Copy the **Object URL**, which should look something like: *https://output-lfxxxx.s3.amazonaws.com/test-clip.m3u8*

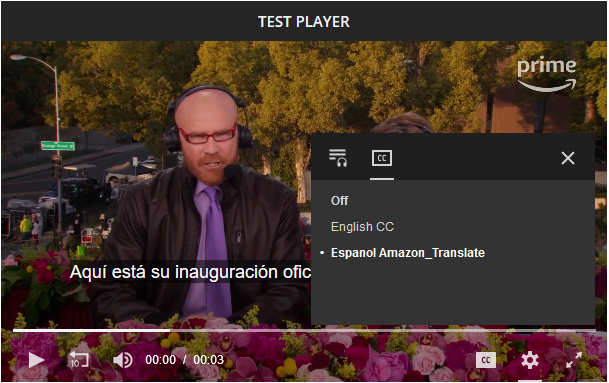
Next, you must load the URL into a test player.

1. Go to [JW Player Stream Tester](https://developer.jwplayer.com/tools/stream-tester/).
2. In the JW Player Stream Tester for **File URL**, enter the object URL you copied from S3.
3. To play the HLS stream, click **Test Stream**
4. Play with the settings () for the streaming video player.

* Make sure that the audio is on:
* Select the  and toggle on/off the captions and subtitles
* Select the  and toggle between the original and generated audio

Try some of the following:

* Start-then-Pause the video, change the captions to subtitles, and then play the video again.
* Start-then-Pause the video, change the audio and then play the video again.



**Conclusion**

 Congratulations! You successfully:

* Constructed a process to automate the generation of captions, alternate language subtitles, and alternate language audio tracks using Amazon’s AI services – Amazon Transcribe, Amazon Translate, and Amazon Polly
* Generated a transcript of a video using Amazon Transcribe
* Translated the transcript into an alternate language using Amazon Translate
* Transformed the alternate language transcript into alternate language audio using Amazon Polly
* Constructed an Apple HLS Stack using AWS Elemental MediaConvert
* Validated the audio and captions/subtitles to work as intended in a video player

**End lab**

Follow these steps to close the console and end your lab.

1. Return to the **AWS Management Console**.
2. At the upper-right corner of the page, choose **AWSLabsUser**, and then choose **Sign out**.
3. Choose **End lab** and then confirm that you want to end your lab.

**Additional resources**

* For more information about AWS Training and Certification, see [*http://aws.amazon.com/training*](http://aws.amazon.com/training/).

For more information about AWS Training and Certification, see [*https://aws.amazon.com/training/*](https://aws.amazon.com/training/).

*Your feedback is welcome and appreciated.*  
If you would like to share any feedback, suggestions, or corrections, please provide the details in our [*AWS Training and Certification Contact Form*](https://support.aws.amazon.com/#/contacts/aws-training).

**Bonus task: Create a MediaConvert job template from a successful job**

In [Task 4](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#bonus), you successfully created a MediaConvert job, and clicked the Export JSON button to download the Job summary for future use. In this section, you learn how to create a Job template in MediaConvert from a successful job to allow for easy creation of a new job using the same settings.

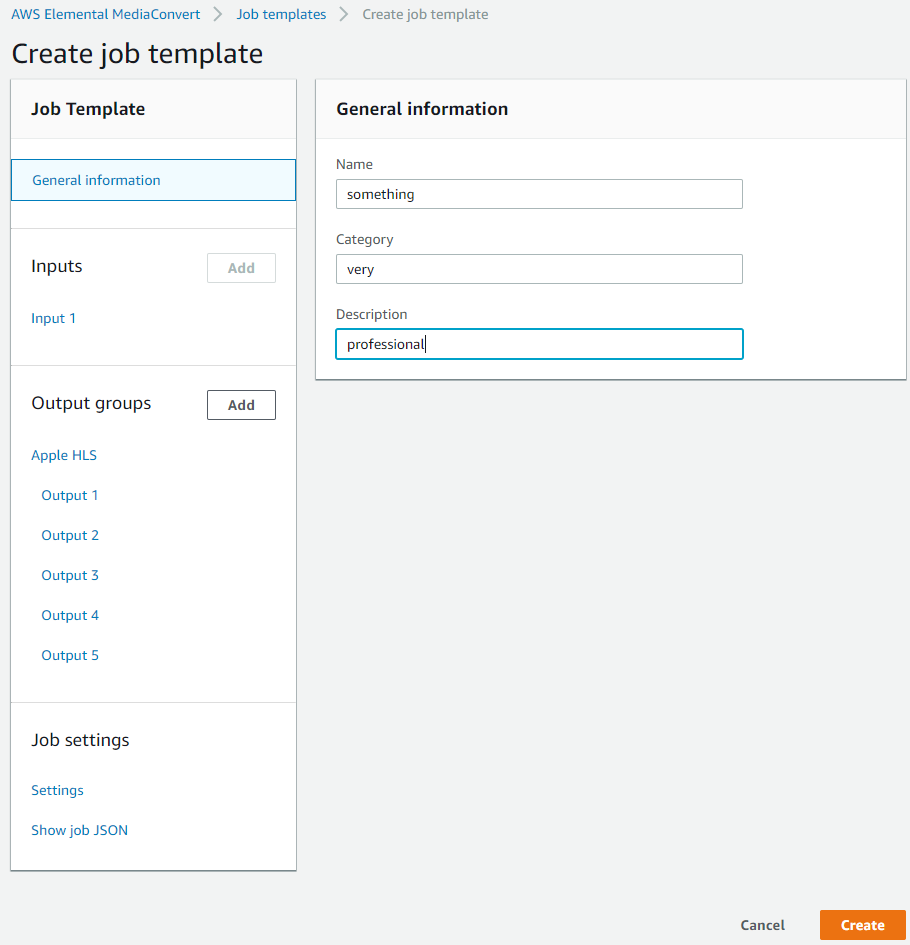
1. If not already there, navigate to the browser tab with MediaConvert.
2. In the left pane, click **Job templates**.
3. From the Job templates window, select **Import template**

* Open the JSON file you downloaded earlier to export the MediaConvert job. It should be a file that looks like *aws-mediaconvert-job-111122223333-a1b2c3.json*.

**Note:** If you did not download the JSON from the job in [Task 4](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#bonus), use the file located in *s3://solutions-lfxxxx/downloads/*. The only caveat is that you must replace the values *${ACCOUNT\_ID}*, *${SOURCE\_BUCKET}*, and *${OUTPUT\_BUCKET}* with the values for your account and bucket names for this template to work within your own account.

1. In the **General information** section, enter the values that make you smile for:

* **Name**
* **Category**
* **Description**



1. Click around in the template to on the inputs and outputs to confirm that the settings you imported match the job you successfully created in [Task 4](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#bonus).
2. When you are satisfied, click **Create**

Now, you can easily use the template to create a new job.

1. From **Job templates details**, click **Create job**
2. In the **Inputs** section for **Input 1**, point the new job to the source video file.
3. In the **Job settings** section, assign the service role you used earlier to create the job.
4. When you are happy with the configuration, click **Create**
5. After 10–15 seconds, click the **Refresh** button.

Next, you follow the same procedures shown in Task 5 to validate the audio and captions/subtitles work as intended in a video player.

1. Continue to [Task 5](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-5).
2. Already completed [Task 5](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#task-5)? Go to the [Conclusion](https://labs.skillbuilder.aws/sa/lab/arn%3Aaws%3Alearningcontent%3Aus-east-1%3A470679935125%3Ablueprintversion%2Fspl-246%3A1.0.9-0e45d8c6/en-US#conclusion).