

Workshop: Amazon ML Services for Video Transcription & Translation

Niranjan Hira – Solutions Architect

David Ping – Solutions Architect

Jillian Forde – Associate Solutions Architect

Kashif Imran – Sr. Solutions Architect

Ro Mullier – Solutions Architect



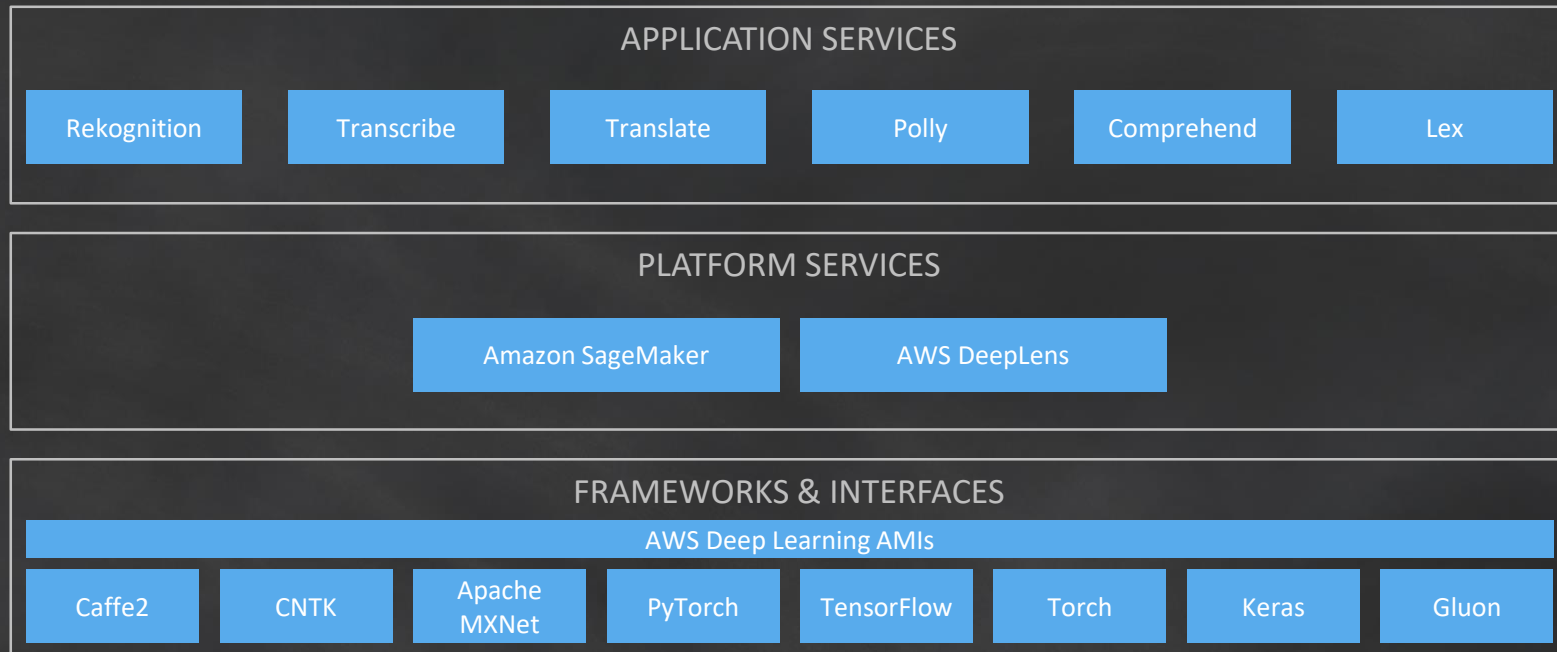
Pop-up Loft

Workshop Goals

- Introduction to relevant services
 - Amazon Comprehend, Amazon Translate, and Amazon Transcribe
- Generate subtitles in English and Spanish
- Considerations for a production workflow



The Amazon machine learning stack



The Instructions

[https://s3.amazonaws.com/mast-mast-3/
public/langws/2018/readme.html](https://s3.amazonaws.com/mast-mast-3/public/langws/2018/readme.html)



Amazon Comprehend

Natural language processing



Amazon Transcribe

Automatic speech recognition



Multiple
languages



Intelligent punctuation
and formatting



Timestamp
generation



Support for
telephony audio



Recognize multiple
speakers



Custom
vocabulary

Amazon Translate

Natural and fluent language translation (based on neural networks)



Real-time
translation



Batch analysis



Automatic language
recognition



Low cost

The Instructions

[https://s3.amazonaws.com/mast-mast-3/
public/langws/2018/readme.html](https://s3.amazonaws.com/mast-mast-3/public/langws/2018/readme.html)



Common Use Cases for Language Services



**Voice of Customer
Applications**



Knowledge Management



Education



**Customer Service/
Call Centers**



Semantic Search



Accessibility



**Enterprise
Digital Assistant**



Captioning Workflows



Information Bots



Personalization



Localization

[Amazon Transcribe](#) > Create transcription job

Create transcription job

Input [Info](#)

Transcription job name

Enter a name for your audio content job

Job names should be 1-200 characters long, consisting of uppercase and lowercase letters, numbers, and hyphens.

S3 input URL

Paste an S3 link URL to your audio content

Supported formats: mp3, mp4, wav, flac

Language

Specify the language

Format [Info](#)

Specify the audio file format that you are uploading

Supported formats: mp3, mp4, wav, flac

Media sampling rate (Hz) - optional [Info](#)

Specify the number of audio samples per second taken per second

Must be an integer between 8000 and 48000

[Cancel](#)[Create](#)

[Amazon Transcribe](#) > [Transcription jobs](#) > [lab-01](#)

lab-01

Details

[Download transcription](#)

Job name

lab-01

Status

In progress

Expiration date

Transcription job is in progress. Expiration date will become available once the job is complete.

Input URL

<s3://mast-mast-3/private/labs/transcribe-speech/bbc-one-minute-201804012144CT.mp4>

Output URL

Output URL is not available because the transcription job did not complete.

Code Samples

[View API documentation](#)

Audio conversion

JSON Request

```
{
  "TranscriptionJobName": "lab-01",
  "LanguageCode": "en-US",
  "MediaFormat": "mp4",
  "Media": {
    "MediaFileUri": "s3://mast-mast-3/private/labs/transcribe-speech/bbc-one-minute-201804012144CT.mp4"
  }
}
```

[Select](#)

JSON Response

```
{
  "TranscriptionJob": {
    "TranscriptionJobName": "lab-01",
    "TranscriptionJobStatus": "IN_PROGRESS",
    "LanguageCode": "en-US",
    "MediaFormat": "mp4"
  }
}
```

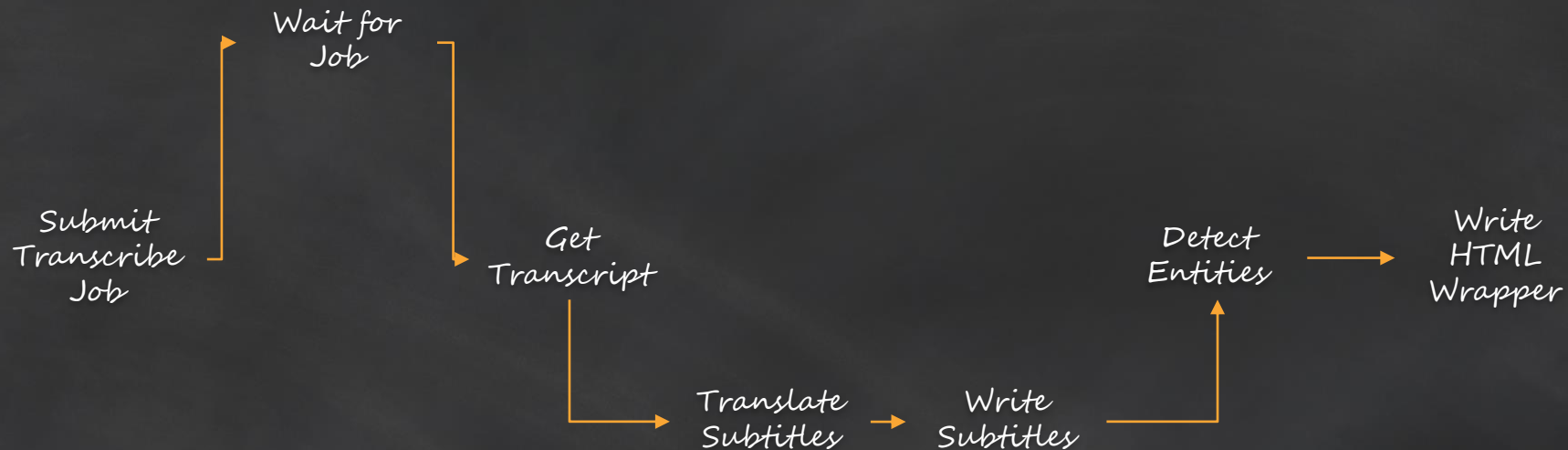
[Select](#)

Transcription Results (walk-through)

Amazon Translate

```
1. Translate.translateText({  
2.     "Text": 'Hello world',  
3.     "SourceLanguageCode": 'en',  
4.     "TargetLanguageCode": 'es'  
5. }, function(err, data) { ... });
```

Demo – Video Subtitles





[00:00:03,440]
[00:00:06,590]
[00:00:09,610]
[00:00:12,730]
[00:00:15,780]
[00:00:18,600]
[00:00:21,740]
[00:00:25,260]
[00:00:28,590]
[00:00:31,760]
[00:00:34,800]
[00:00:38,040]
[00:00:41,260]
[00:00:44,290]
[00:00:47,670]
[00:00:50,840]

This is BBC news, and these are the headlines China says it's defunct tiangong one space laboratory has re entered the earth's atmosphere on most of it has burnt up over the south pacific. Experts say the craft would have burned during its descent. Beijing is imposing import taxes on a range of US goods and new tally ation to donald trump's, tariff increases on steel and aluminium imports. Last month, China said it was planning tariffs on up to three billion dollars worth of US imports on north. Korea's leader kim jong un has attended a groundbreaking pop concert featuring South Korean stars. The event is the latest in a series of conciliatory gestures that appeared to mark a thaw in relations between the two sides. The leaders of the two Koreas are due to hold a summit later this month.

Entity	Type	Max Confidence	Count	
Last month	DATE	0.99	1	
this month	DATE	0.94	1	
Beijing	LOCATION	0.57	1	
Korea	LOCATION	0.67	1	
US	LOCATION	0.9	2	
BBC	ORGANIZATION	0.78	1	
China	ORGANIZATION	0.81	2	
South Korean	OTHER	0.75	1	
kim jong un	PERSON	0.86	1	
trump	PERSON	0.97	1	
Index	Avg. Confidence	Terms	Unsure	
00:00:03,440	0.97	20	0	
00:00:10,100	0.99	28	0	
00:00:20,090	0.97	33	0	
00:00:30,340	0.98	22	0	
00:00:40,190	0.99	35	0	
00:00:50,840	1	3	0	
	0% unsure	141	0	



[00:00:05,410] This is the BBC news and these other headlines. Politicians in the
 [00:00:08,490] United States in Europe are summoning executives from
 [00:00:11,590] facebook to answer questions about whether personal data
 [00:00:14,710] was misused to manipulate it. The recent elections,
 [00:00:18,150] facebookers said it was deceived by a British company, cambridge
 [00:00:21,350] analytica, that worked for president. Trump's election campaign,
 [00:00:24,450] among others. Emergency services in texas, a and
 [00:00:27,610] essentially device that went off in the city of austin, was not
 [00:00:30,810] related to a series of explosions this month. However,
 [00:00:33,870] the FBI has confirmed that two packages found it separates
 [00:00:37,070] a fedex delivery. Officers in the area on Tuesday
 [00:00:40,140] are connected to the earlier attacks.
 [00:00:43,450] A BBC investigation has found that girls from the and mars
 [00:00:46,500] arranger muslim minority who fled violence in their home
 [00:00:49,680] country on being trafficked into prostitution in
 [00:00:52,700] Bangladesh, hundreds of thousands of revenge and now live in huge
 [00:00:55,890] refugee camps where they're vulnerable to sexual exploitation.

Entity	Type	Max Confidence	Count	
Tuesday	DATE		1	1
this month	DATE	0.97		1
Bangladesh	LOCATION	1		1
British	LOCATION	0.58		1
Europe	LOCATION	1		1
United States	LOCATION	0.99		1
austin	LOCATION	0.95		1
texas	LOCATION	0.99		1
BBC	ORGANIZATION	0.97		2
FBI	ORGANIZATION	1		1
cambridge analytica	ORGANIZATION	0.69		1
facebook	ORGANIZATION	0.99		1
fedex	ORGANIZATION	0.75		1
muslim	ORGANIZATION	0.59		1
Trump	PERSON	1		1
Index	Avg. Confidence	Terms	Unsure	
00:00:05,410	0.88		17	2
00:00:10,440	0.91		25	2

What does the HTML look like?

1. `<video controls id="vplayr">`
2. `<source src="....mp4" type="video/mp4">`
3. `<track kind ="subtitles" label="English" srclang="en-us"`
 `src="..._en.vtt">`
4. `<track kind ="subtitles" label="Spanish" srclang="es"`
 `src="..._es.vtt">`
5. This browser does not seem to support the video element.
6. `</video>`

What is WebVTT?

- Web Video Text Tracks format
- <https://www.w3.org/TR/webvtt1/>
- “Borrows” from SRT (SubRip Subtitles)

```
1. 2
2. 00:00:08.490 --> 00:00:11.000
3.   Europa están convocando a ejecutivos de
4.   facebook
5.
6. 3
7. 00:00:11.590 --> 00:00:14.000
8.   para responder preguntas sobre si los
9.   datos
```



Pop-up Loft

Everything and Anything Startups Need to Get Started on AWS

aws.amazon.com/activate