
AWS
re:Invent

AIM201

Share information by removing language barriers with Amazon Translate

Watson G. Srivathsan

Senior Product Manager, Amazon Translate

AWS

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

COVID19: an information challenge

Information campaign and task

Speed – accuracy trade off

Sample application demo

Do more with Amazon Translate

Wrap-up



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

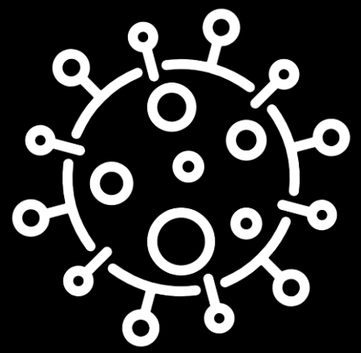
COVID-19:
an information challenge



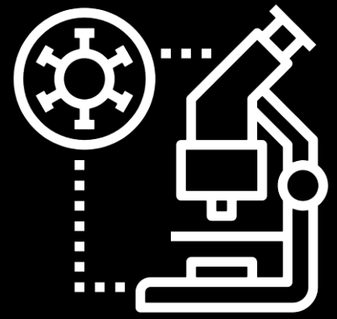
AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The challenge with COVID-19



Novel challenge

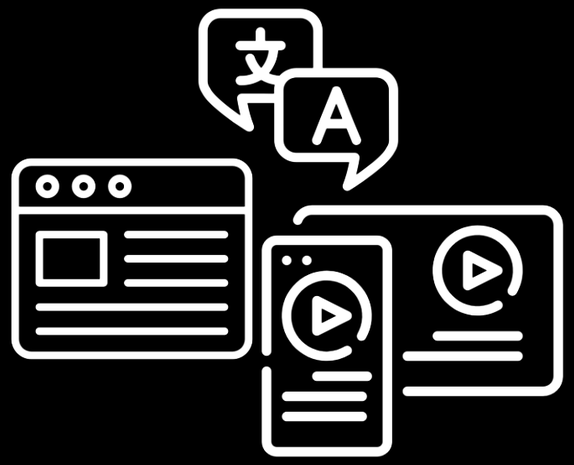


Research in centers of excellence



Global impact

COVID-19 information campaign task



Multi-language



Accuracy



Speed

Speed – accuracy trade off

Fitt’s Law



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Measurable goals

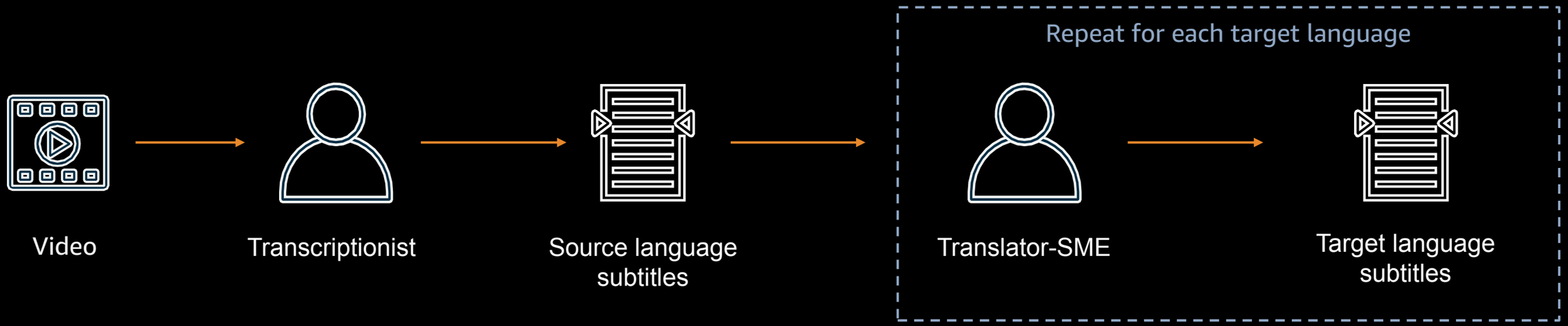
Speed: decrease the time to produce translated subtitled content

- Measure the overall speed of the end-to-end workflow including the time for the automation to run and for human edits
- Normalize by the length of the content

Accuracy: decrease the number of corrections that transcriptionists and translators need to make to the machine-generated subtitles

- **Use Word Error Rate (WER) as a quantitative metric; WER measures the number of corrections relative to the overall word count of the subtitles**

Manual subtitle workflow



High

Low

Automation with machine learning



AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Automate with AWS AI services

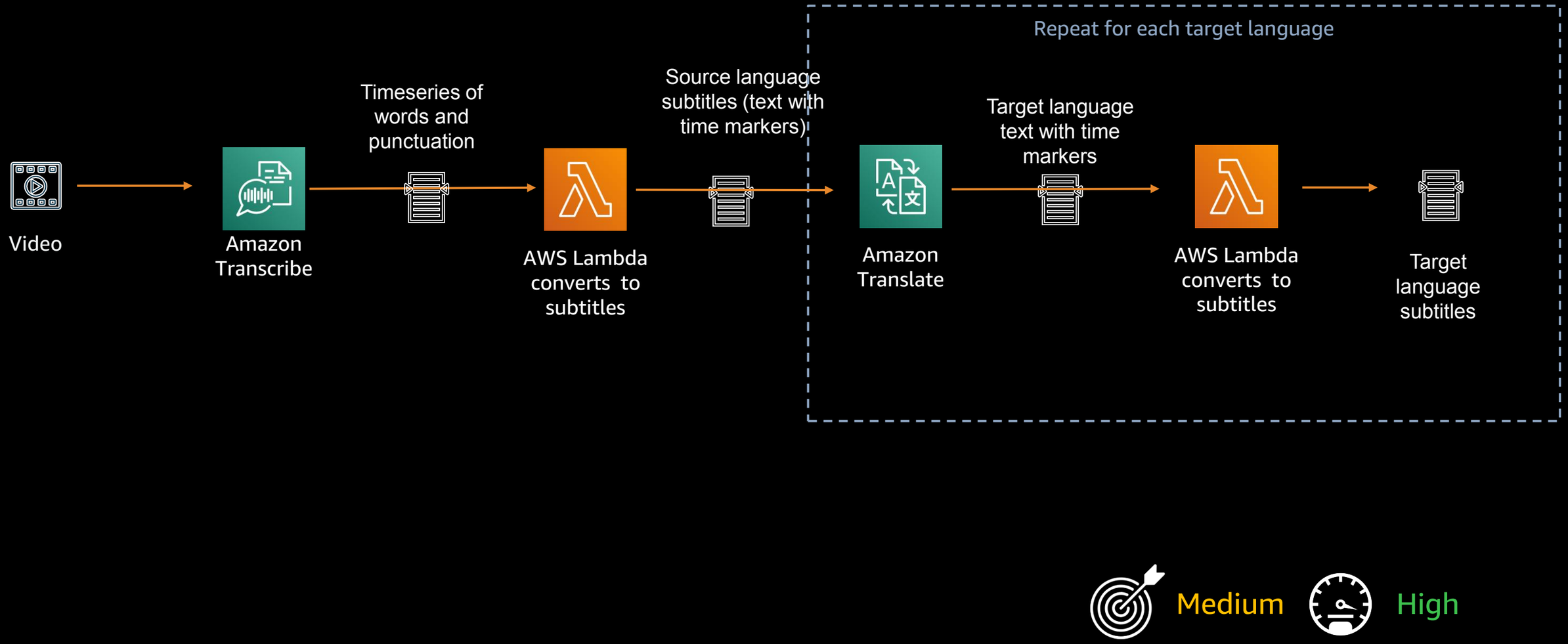


Amazon Transcribe



Amazon Translate

Automated subtitle workflow



Medium High

Common problems

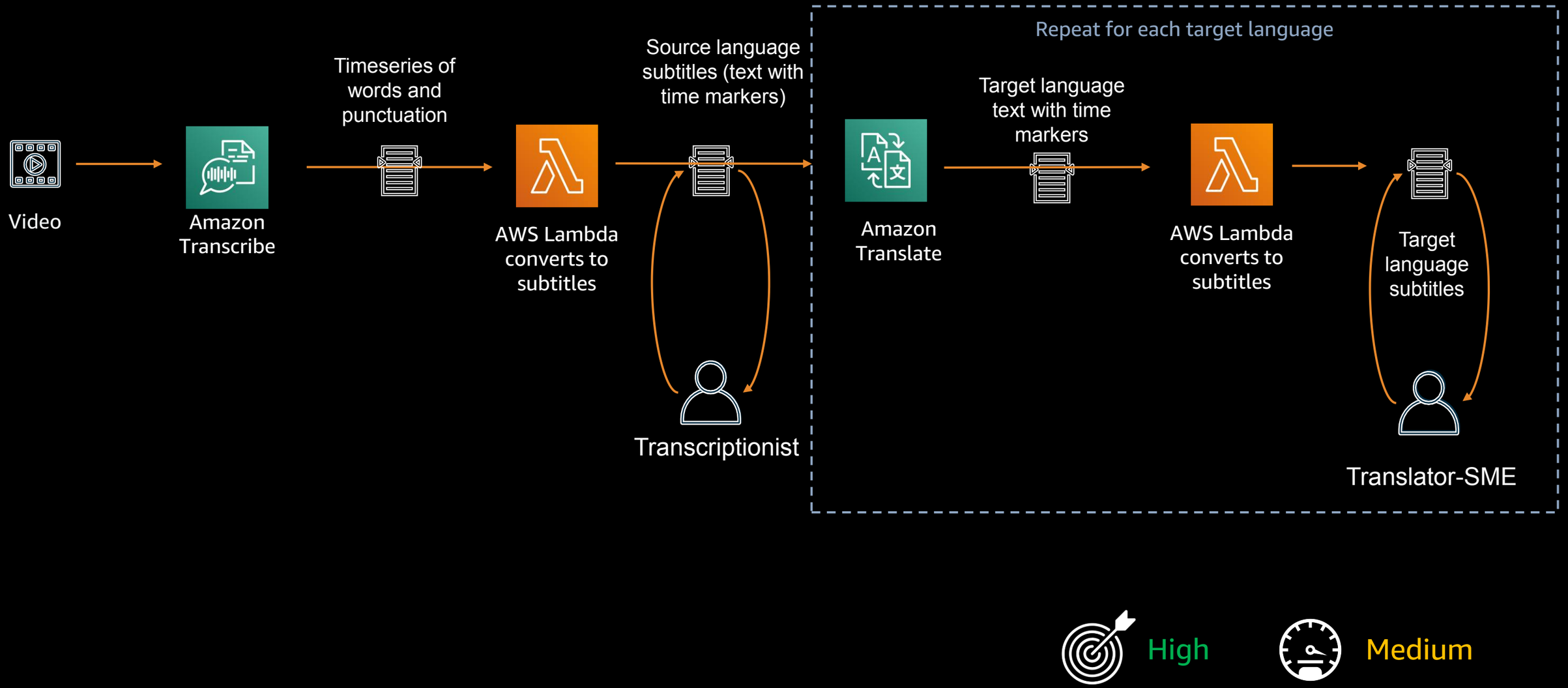
Machine transcription

- Domain-specific terminology
- New / trending words or phrases – COVID-19
- Mis-transcription of similar-sounding words or phrases
  - “1-to-many” <> “One too many”
- Mis-transcription of punctuation
- Speaker accent, dialect and speech patterns

Machine translation

- Errors propagated from transcription
- New / trending words or phrases
- Words or phrases that don’t have a direct translation

Human in the loop (HIL) subtitle workflow



Customization with AI



AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

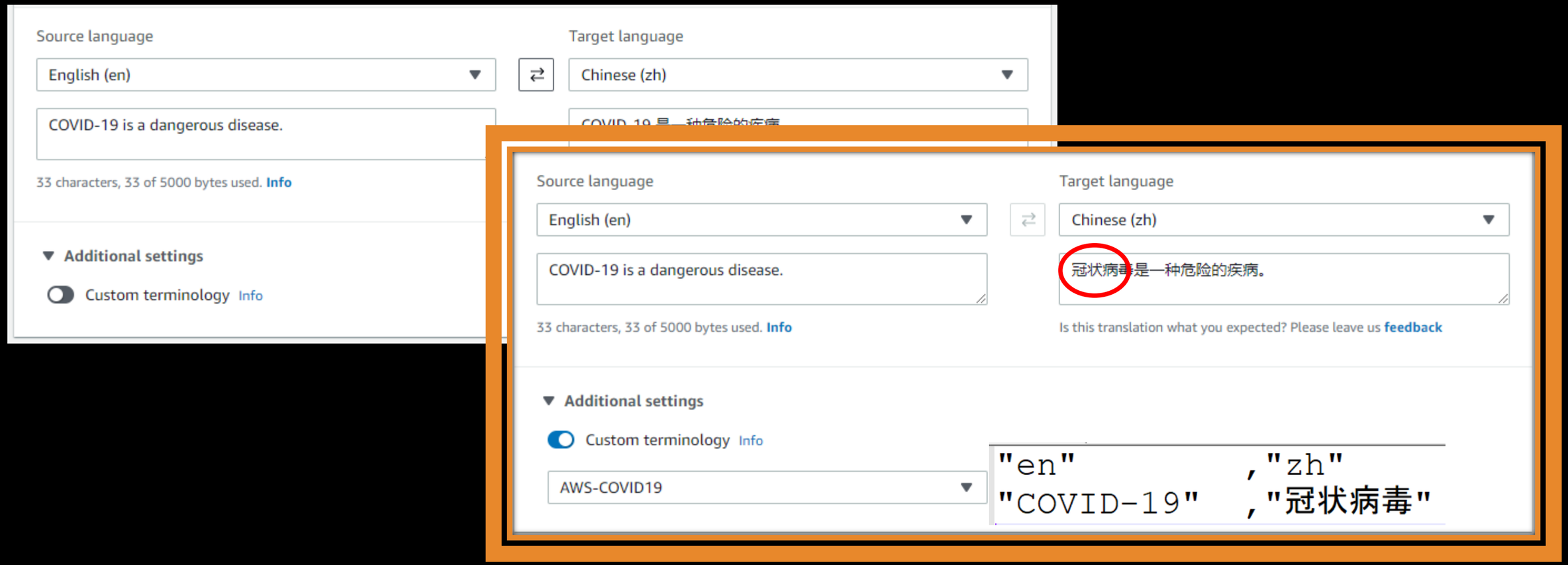
Amazon Transcribe custom vocabulary

- Automate corrections for common transcription errors
- Describe domain-specific phrases, phonetic information and desired output text

	Phrase	SoundsLike	IPA	DisplayAs
2	Kerkhove			Kerkhove
3	Zoonoses			
4	MERS			
5	SARS			
6	Zoonoses	Zoo-Noses		
7	COVID-Nineteen			COVID-19
8	COVID			
9	Corona-virus			Coronavirus
10	Corona-viruses			Coronaviruses
11	IPC			
12	PPE			
13	WHE			
14	WHO			
15	emergencies			
16	H-one-B			H1B
17	H-one-N-one			H1N1
18	N-ninety-five			N95

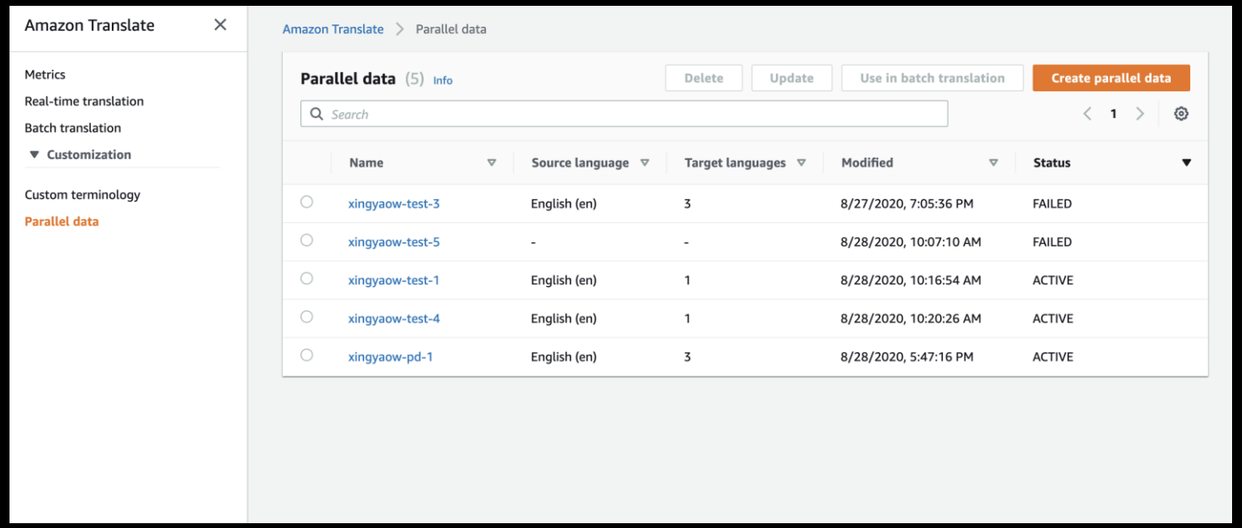
Amazon Translate Custom Terminology

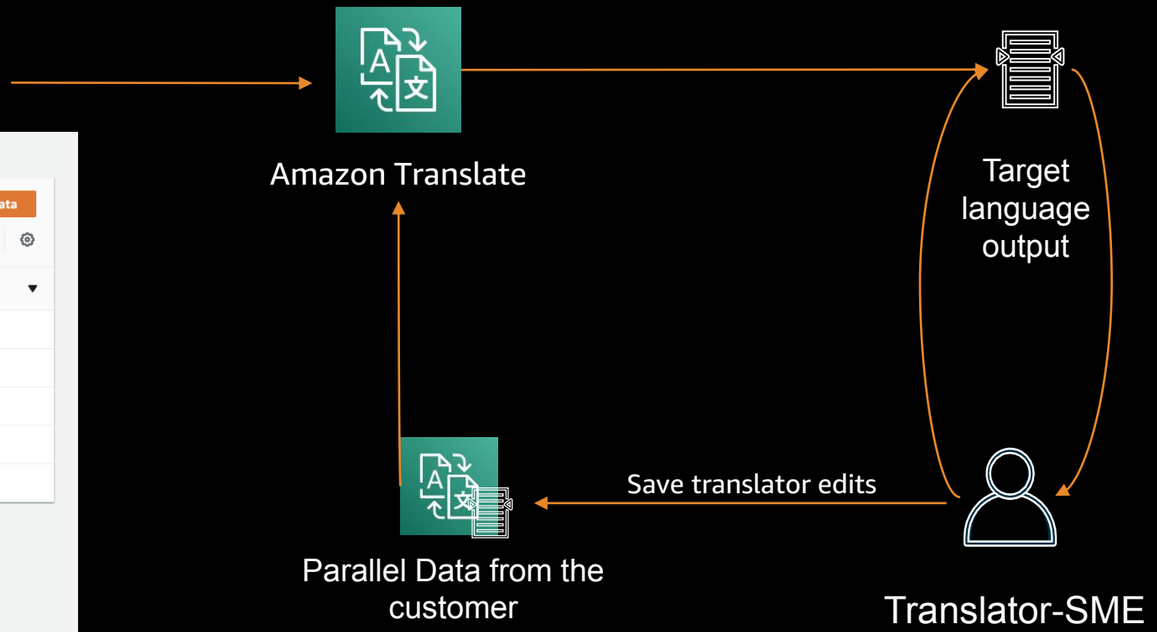
- Translate Named Entity, such as product name and organizations names the way you want using Custom Terminology.
- Describe domain-specific phrases, phonetic information and desired output text



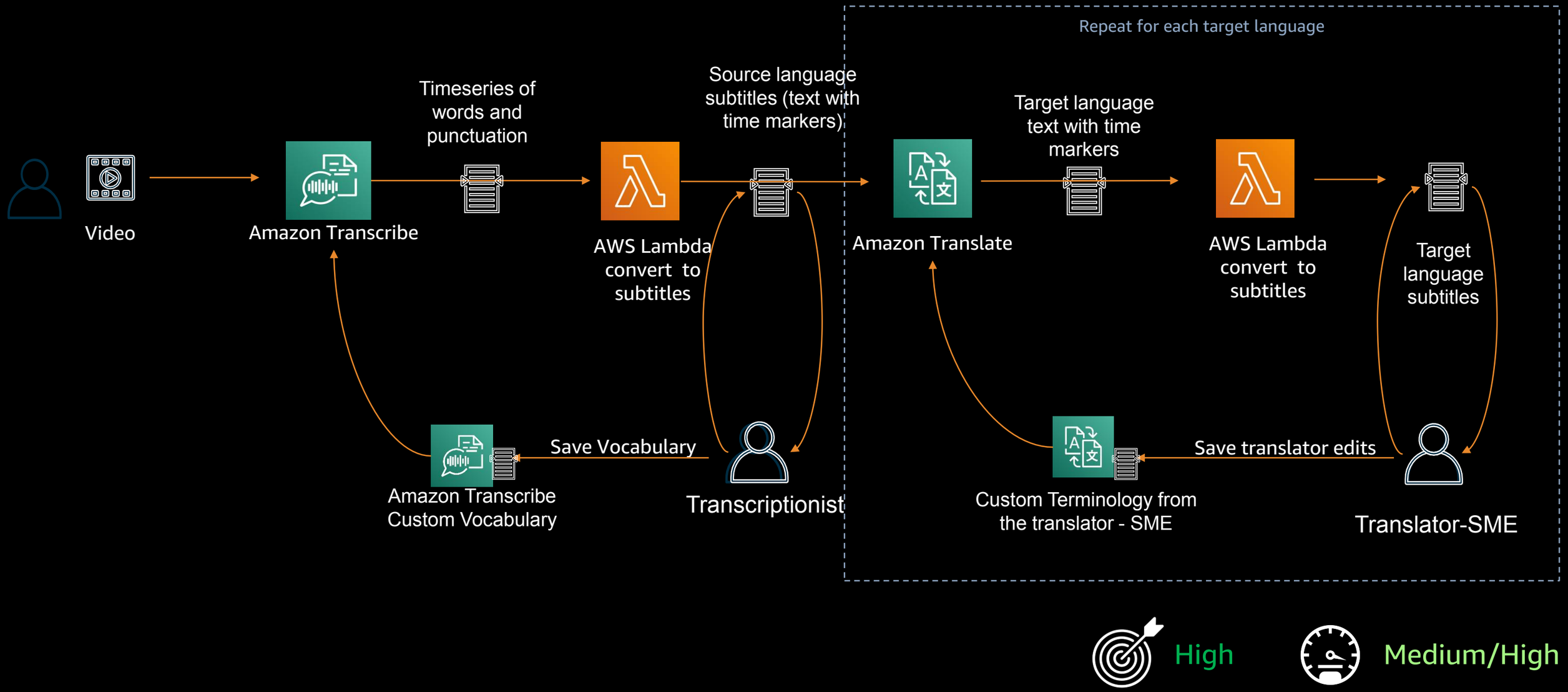
Amazon Translate Active Custom Translation

- Customize output without building Custom Translation model
- Active Custom Translation (ACT) enables translators to provide feedback to Amazon Translate to improve the output continuously





Adding Customization to HITL workflow



High

Medium/High

Summary

- Create an application that combines automatic translated subtitle generation with human in-the-loop review

- Capture the corrections so we can make use of them to improve the results of automation over time

- Improved the speed while maintaining accuracy

Speed		Medium/High
Accuracy		High

Dr. Gaya Manori Gamhewage



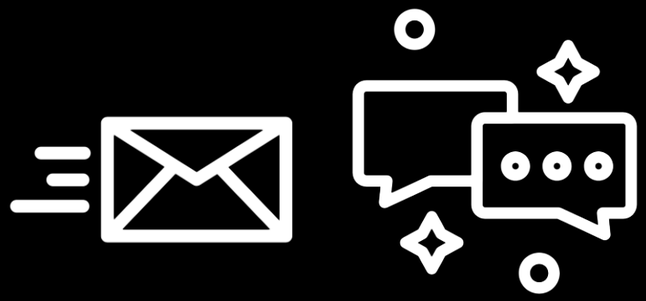
AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

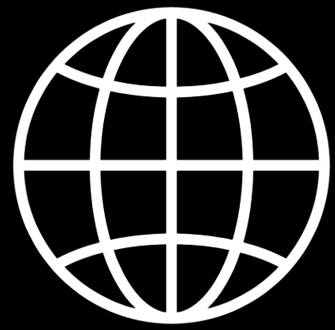
What else can you do with Amazon Translate?



Information



Email / chat communication



Text / research analytics

Next steps



Create multi-language captions for live streaming

https://github.com/awslabs/live-streaming-with-automated-multi-language-subtitling



Engage with a solution architect

Brainstorm use cases and get support for solution implementation



Connect with an AWS partner

Let AWS certified partners help you with the POC and build integrations



Get started with a free active account

Sign up for an AWS account and translate your content today

Thank you!



AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Please complete the session survey

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.