

Amazon Artificial Intelligence

Analyze Voice, Text, Video, Image

Chandra Lingam

Cloud Wave LLC

Natural Language, Image, Video Analysis



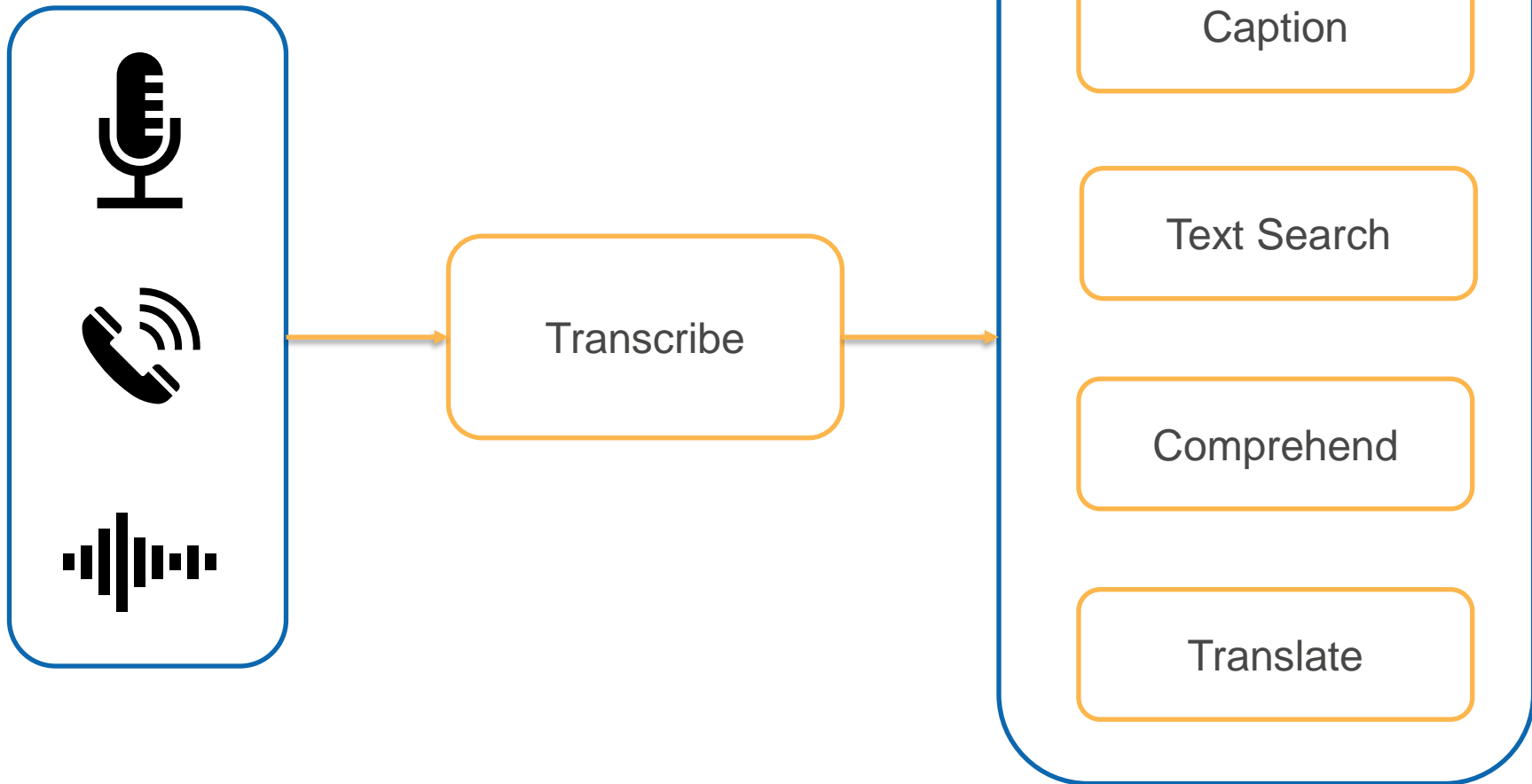
Service	Purpose	Use
Transcribe	Voice to Text	The first step in analyzing voice is to convert to text and then use text based algorithms
Translate	Text to Text - Language translation	Understand text/comment in other languages, Localization of your products to international markets
Comprehend	Analyze text to understand intent	Sentiment analysis (Positive, Negative, Neutral, Mixed), classification, data extraction, organize text by topics
Polly	Text to Voice	Speech enabled products
Lex	Conversation using voice and text	Alexa style interaction. Build your own self-service applications that can understand what customer is asking for using voice or text
Rekognition	Image and Video analysis	Detect objects, person, emotions, text, inappropriate content and track people
Textract	Extract text and data from any document	Convert Scanned documents, extract fields and forms,
DeepLens	Deep Learning enabled Video Camera for developers	Learning tool for building vision enabled applications

Amazon Transcribe

“Audio data is virtually impossible for computers to search and analyze. Therefore, recorded speech needs to be converted to text before it can be used in applications.”

Reference: <https://aws.amazon.com/transcribe/>

Amazon Transcribe



Transcribe Features

Automatically adds punctuation and formatting

Batch and Real-time transcription

Timestamp for each word – easily locate in original media

Custom Vocabulary Support

Detect and Tag multiple Speakers

Lab – Sample Audio File

“XGBoost routinely finds a top place in machine learning competitions. So, let's take a quick tour of XGBoost and compare it with a linear model and other tree-based methods. This spreadsheet contains the bike rental count prediction data set provided by Kaggle.”

Transcribe Issues

Actual

“XGBoost routinely finds a top place in machine learning competitions. So, let's take a quick tour of XGBoost and compare it with a linear model and other tree-based methods. This spreadsheet contains the bike rental count prediction data set provided by Kaggle.”

Transcribed

“**exhibition** is **vertically** finds the top place in machine learning competitions. So let's take a quick tour off **extra booze** and compare it with a **lady moral** and other **tree bees** methods. This spreadsheet contains the bike rental **contradiction**, **data said**, perturbed by **cargo**.”

Custom Words

“These [custom vocabulary] are generally domain-specific words and phrases, words that Amazon Transcribe isn’t recognizing, or proper nouns.”

Reference:

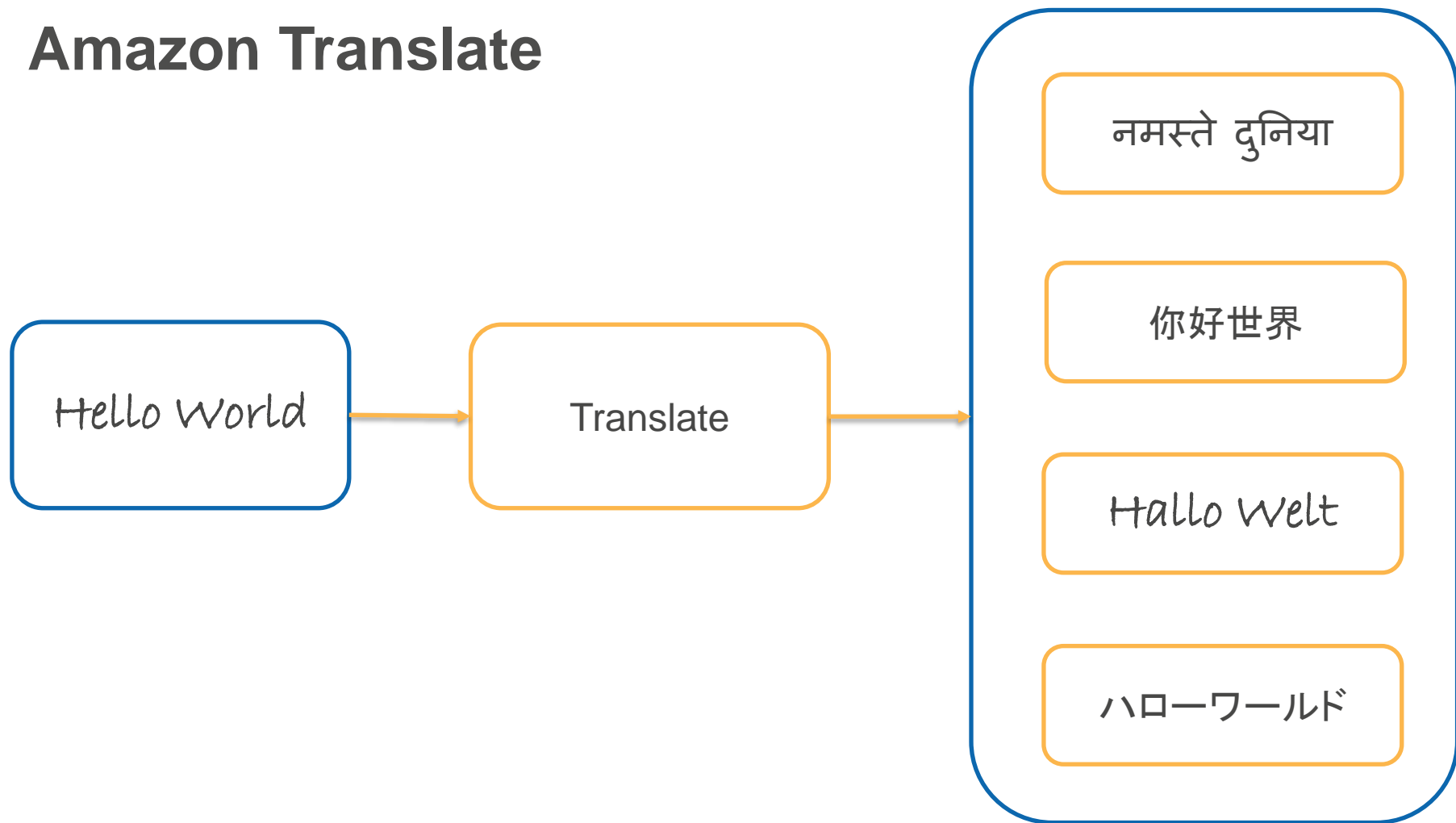
<https://docs.aws.amazon.com/transcribe/latest/dg/how-vocabulary.html?shortFooter=true>

Custom Words

List [of words]

Table [of words with phonetics, output hints]

Amazon Translate



Amazon Translate

Understand reviews, text, comment in other languages

Localization of your products and services

Custom Terminology Support

Real-time and Batch translation

Amazon Translate

“Neural network based - takes into account the entire context of the source sentence as well as translation it has generated so far, to create a more accurate and fluent translation”

Reference:

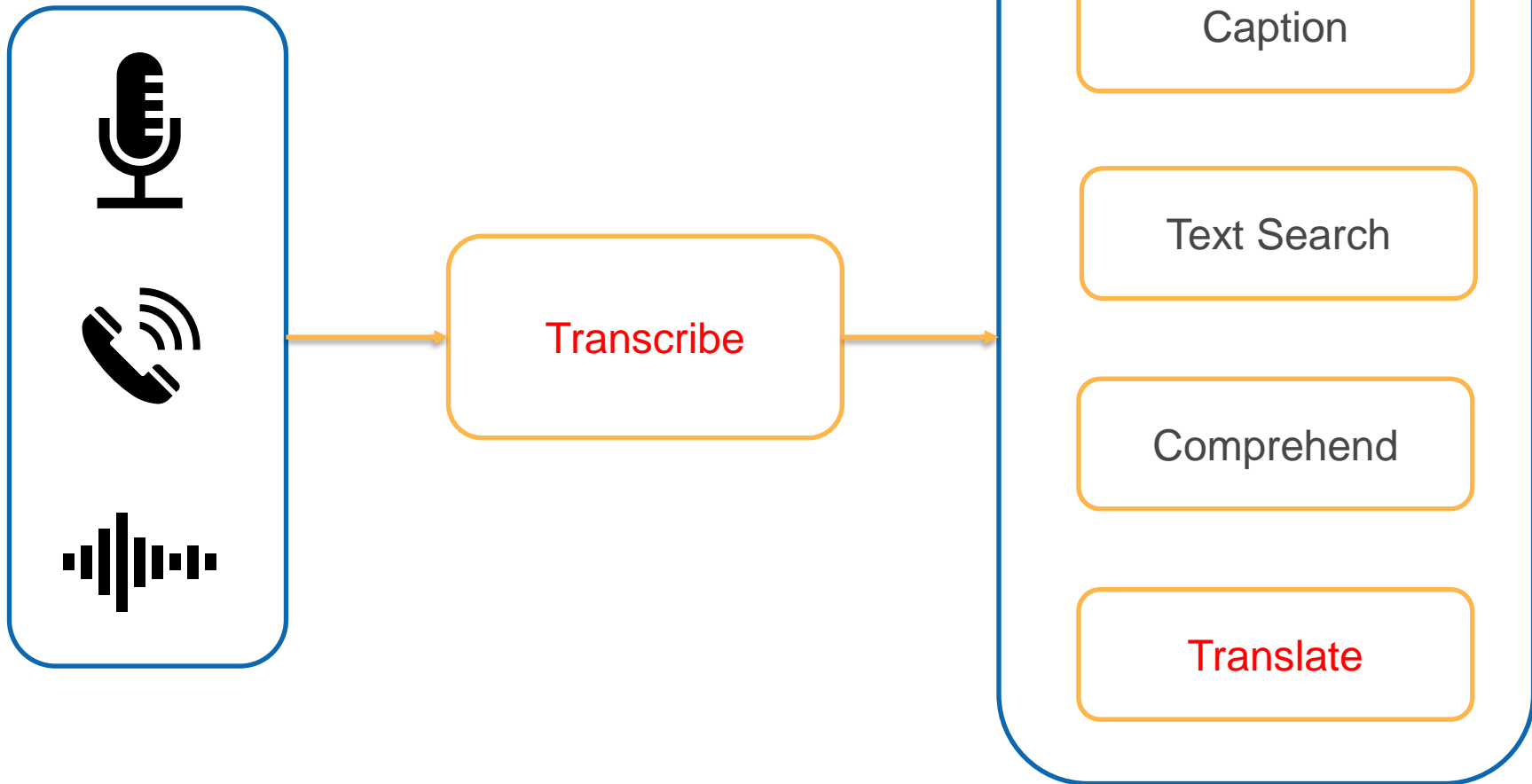
<https://aws.amazon.com/translate/details/>

Lab – Translate [English to Hindi, Spanish]

Source Text

“XGBoost routinely finds a top place in machine learning competitions. So, let's take a quick tour of XGBoost and compare it with a linear model and other tree-based methods. This spreadsheet contains the bike rental count prediction data set provided by Kaggle.”

Transcribe and Translate

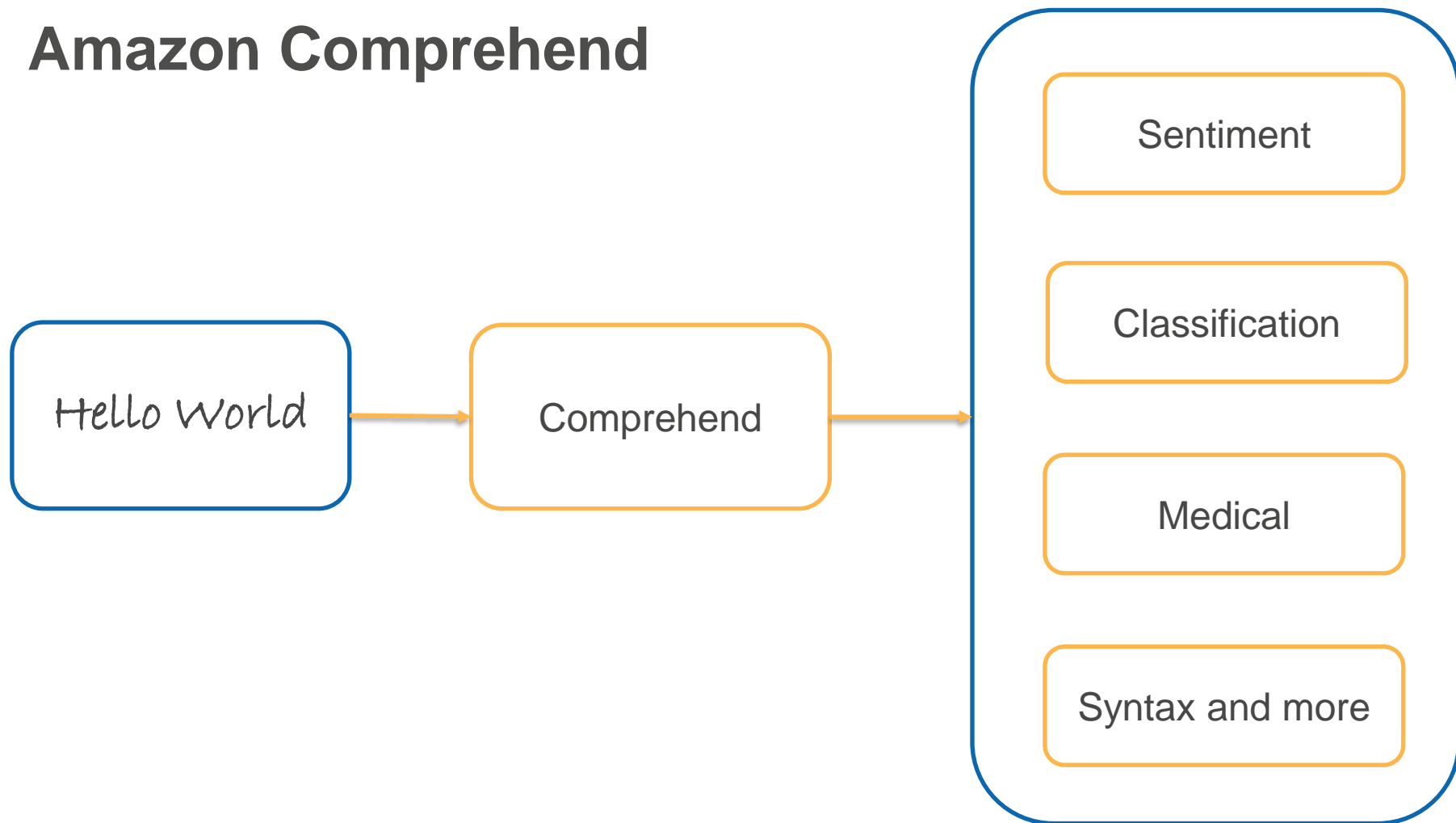


Lab 2 – Translation with Errors

Source Text

“exhibition is vertically finds the top place in machine learning competitions. So let's take a quick tour off extra booze and compare it with a lady moral and other tree bees methods. This spreadsheet contains the bike rental contradiction, data said, perturbed by cargo.”

Amazon Comprehend



Amazon Comprehend Features

Keyphrase Extraction

Sentiment Analysis

Syntax Analysis

Entity Recognition

Medical Information Extraction

Custom Entities

Classification

Topic Modeling

Comprehend Medical

“Amazon Comprehend Medical is a service that detects useful information in unstructured clinical text”

“As much as 75% of all health record data is found in unstructured text: Physician’s note, discharge summaries, test results, case notes and so on.”

Reference:

<https://docs.aws.amazon.com/comprehend/latest/dg/comprehend-med.html>

Classification Using Comprehend

Text based custom classification

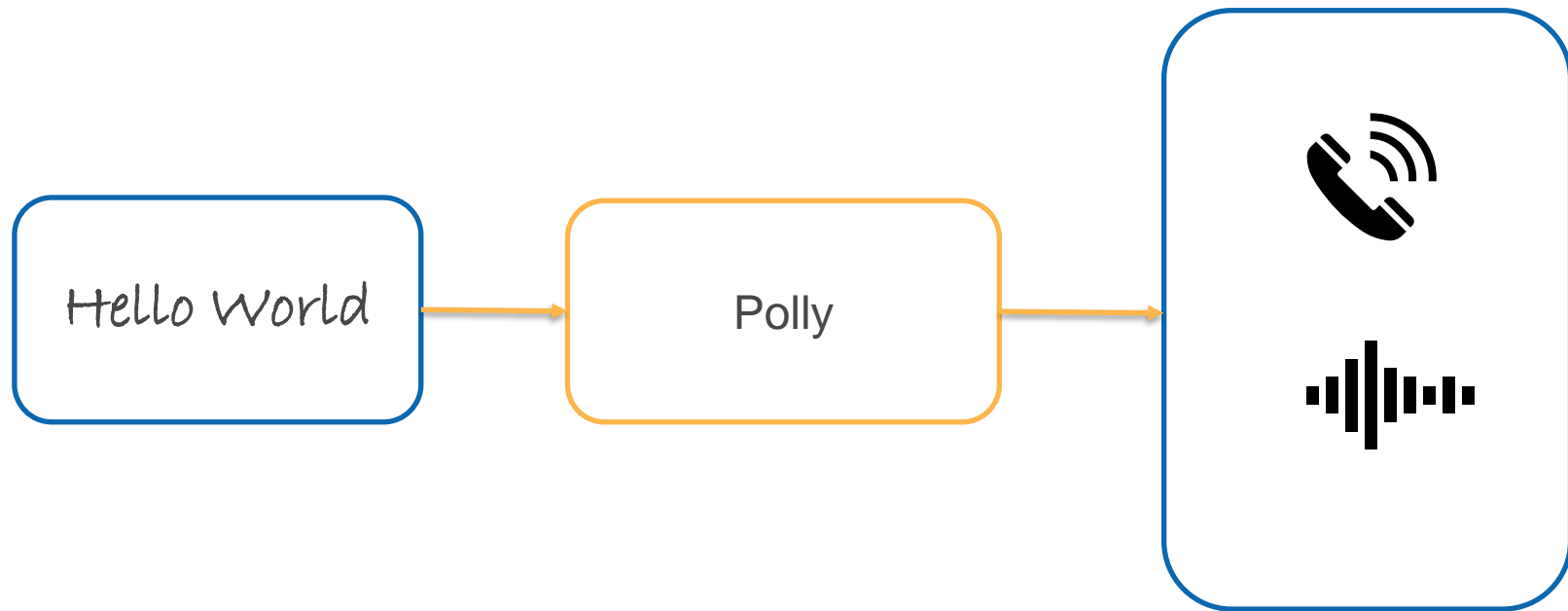
- Moderating Comments and Reviews
- Organizing Documents by Category

Lab – Classification using Comprehend

Objective: Identify tweets that require follow-up

Twitter Classification Dataset [AWS provided labeled dataset]

Amazon Polly



Amazon Polly - Text To Speech Service (TTS)

Speech Enabled Products

Real-time or batch mode

Wide selection of voices, languages

Customize with SSML (Speech Synthesis Markup Language)

Polly SSML

Control how Polly generates speech from text

Adjust pause, speech rate, pitch

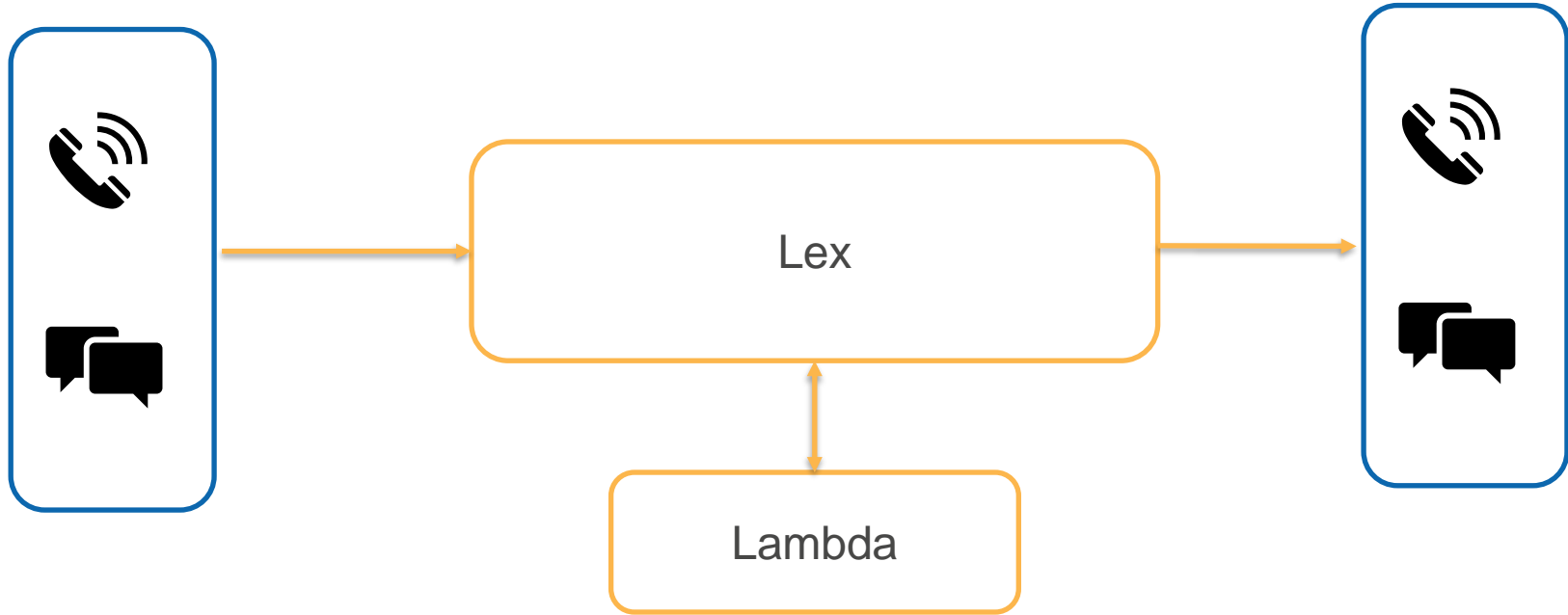
Emphasis of sounds, Phonetic pronunciation

Newscaster

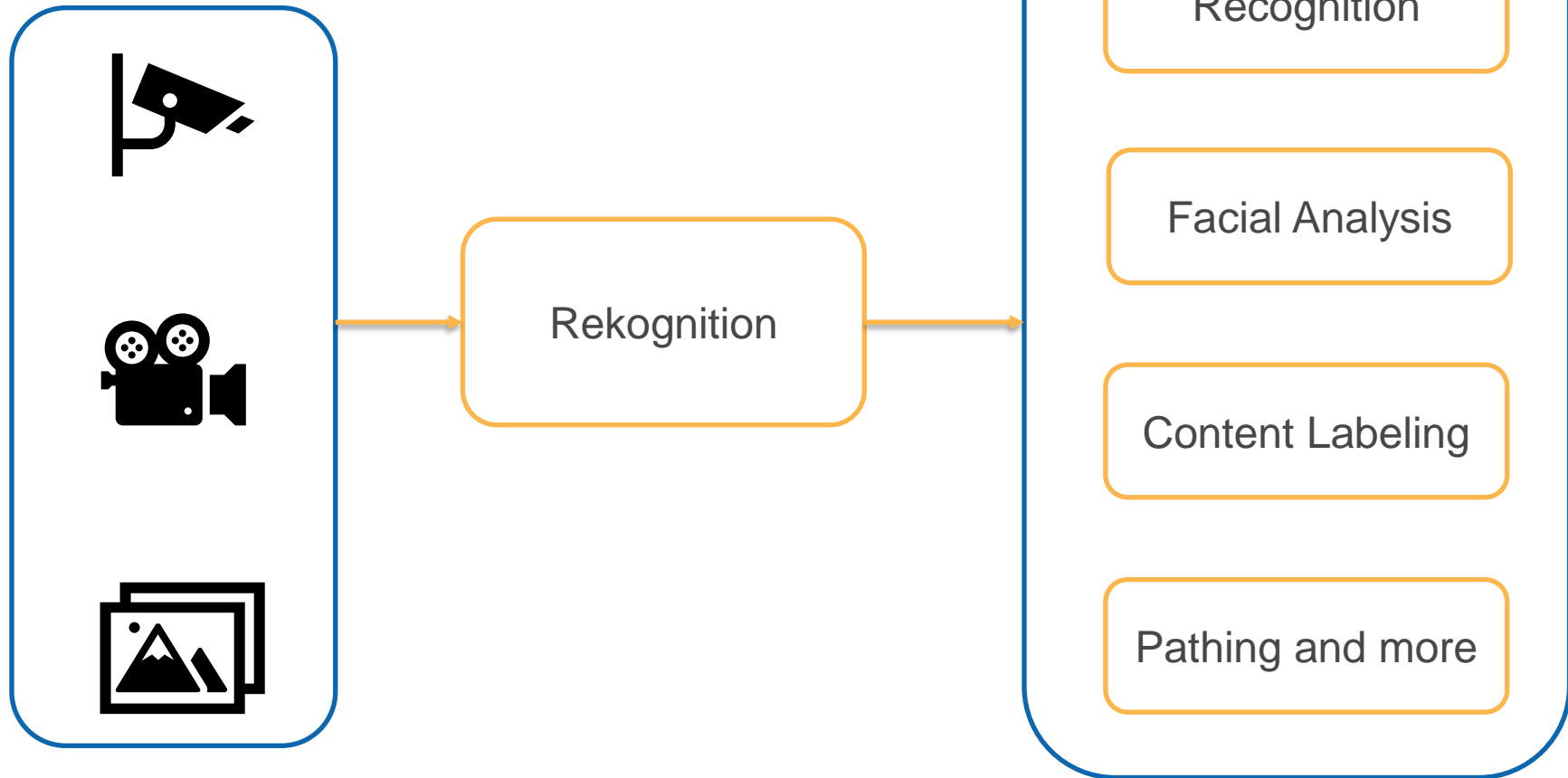
Amazon Lex - Motivation



Amazon Lex



Amazon Rekognition



Amazon Textract

