Please

Ask questions through the app

Rate Session

Thank you!

@GOTOber

gotober.com

Breaking language barriers with AI

Boaz Ziniman
Principal Technical Evangelist - Amazon Web Services

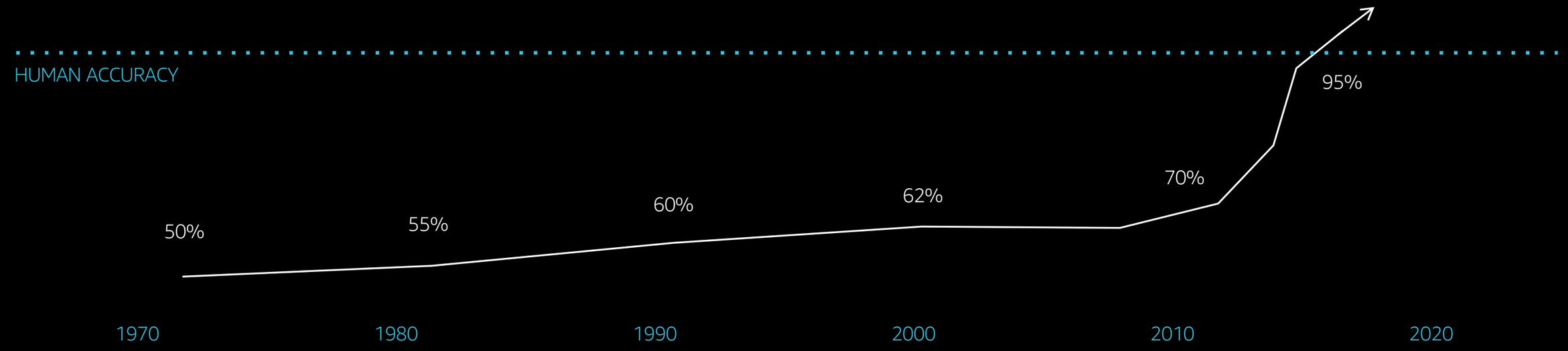
 @ziniman  ziniman

Natural language processing (NLP)

- Automatic speech recognition (ASR)
- Natural language understanding (NLU)
- Text to speech
- Translation

Natural language processing (NLP)

- Automatic speech recognition (ASR)
- Natural language understanding (NLU)
- Text to speech
- Translation



Source: MindMeld


amazon

UBER
Tide
AMEX
THIS DAY IN
HISTORY
NBA
epi
SONOS
ROYAL
OPERA
HOUSE
TED
WebMD
MR. ROBOT
DAILY FIVE / NINE
DUNKIRK
StubHub!
SHOWTIME
SESAME STREET
CHEF'D
BOSE
food
network
Polycom
geeni
Joule
PHILIPS
hue
wemo
CNBC
NASA
lyft
CNN
NASCAR
1-800-flowers
arlo
unicef
KID
POWER
DEAL OR NO DEAL
seamless
BMW
P&G
GARMIN
THE DAILY SHOW
WITH TREVOR NOAH
FOX NEWS
restaurants
The Economist
Espresso
works with
nest
STAR WARS
dish
npr
redbox
TECH CHECK
MTA
INFINITI CONNECTION
ZEEQ
SYFY
GLAD
LIE SWATTER
GoodRx
Allstate
TranslateIt
20 Q


100,000+
UBER
Tide
AMEX
MR. ROBOT
DAILY FIVE / NINE
TED
WebMD
StubHub!
SESAME STREET
HISTORY
NBA
DUNKIRK
SONOS
ROYAL OPERA HOUSE
DEAL OR NO DEAL
THE DAILY SHOW
unicef KID POWER
CHEF'D
BOSE
NASCAR
CNBC
SHOWTIME
Polycom
PHILIPS hue
Joule
geeni
lyft
NASA
CNN
BMW
P&G
GARMIN
MTA
epi
food network
Disney
redbox
npr
FOX NEWS
restaurants
seamless
1-800 flowers
INFINITI CONNECTION
TECH CHECK
BLACK FRIDAY
ZEEQ
SYFY
Allstate
LIE SWATTER
Jamie Oliver
Translated!

Use cases for common language

- Voice of customer applications
- Knowledge management
- Education
- Customer service/ Call centers
- Semantic search
- Accessibility
- Enterprise digital assistant
- Captioning workflows
- Information bots
- Personalization
- Localization

The Amazon ML stack: Broadest & deepest set of capabilities

AI SERVICES

Vision


AMAZON REKOGNITION IMAGE


AMAZON REKOGNITION VIDEO


AMAZON TEXTRACT

Speech


AMAZON POLLY


AMAZON TRANSCRIBE

Language


AMAZON TRANSLATE


AMAZON COMPREHEND & AMAZON COMPREHEND MEDICAL

Chatbots


AMAZON LEX

Forecasting


AMAZON FORECAST

Recommendations


AMAZON PERSONALIZE

ML SERVICES


AMAZON SAGEMAKER

BUILD

Pre-built algorithms & notebooks

Data labeling (GROUND TRUTH)

Algorithms & models (AWS MARKETPLACE FOR MACHINE LEARNING)

TRAIN

One-click model training & tuning

Optimization (NEO)

Reinforcement learning

DEPLOY

One-click deployment & hosting

ML FRAMEWORKS & INFRASTRUCTURE

Frameworks

TensorFlow

mxnet

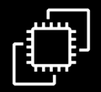
PYTORCH

Interfaces

GLUON

Keras

Infrastructure


EC2 P3 & P3dn


EC2 C5


FPGAs


AWS IoT GREENGRASS

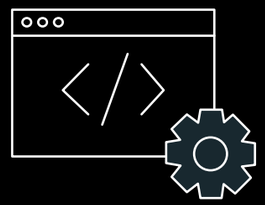

AMAZON ELASTIC INFERENCE

aws

AI services

	Vision			Speech		Language		Chatbots	Forecasting	Recommendations
AI SERVICES	 AMAZON REKOGNITION IMAGE	 AMAZON REKOGNITION VIDEO	 AMAZON TEXTRACT	 AMAZON POLLY	 AMAZON TRANSCRIBE	 AMAZON TRANSLATE	 AMAZON COMPREHEND & AMAZON COMPREHEND MEDICAL	 AMAZON LEX	 AMAZON FORECAST	 AMAZON PERSONALIZE

Pre-trained AI services that require no ML skills or training

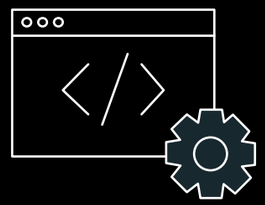


Ability to easily add intelligence to your existing apps and workflows

Quality and accuracy from continuously learning APIs

AI services

Pre-trained AI services that require no ML skills or training



Ability to easily add intelligence to your existing apps and workflows

Quality and accuracy from continuously learning APIs

Amazon Polly

Turn text into lifelike speech using deep learning

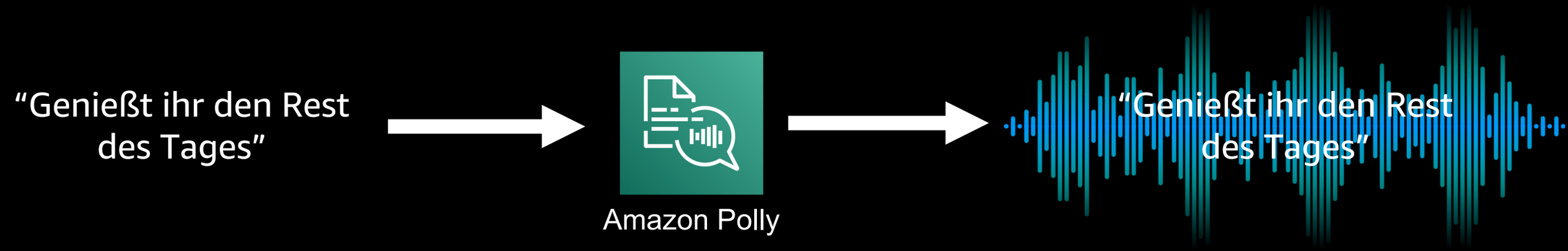
Amazon Polly: Text in, lifelike speech out

"Today in Seattle, WA, it's 11°F"

Amazon Polly

29 languages

Amazon Polly: Text in, lifelike speech out



29 languages

Add semantic meaning to text

https://www.w3.org/TR/speech-synthesis/

```
<speak>
The spelling of my name is
      <prosody rate='x-slow'>
            <say-as interpret-as="characters">Boaz</say-as>
      </prosody>
</speak>
```

aws

Add semantic meaning to text

https://www.w3.org/TR/speech-synthesis/

"Richard’s number is 2125550123"

Add semantic meaning to text

https://www.w3.org/TR/speech-synthesis/

```
<speak>
      Richard’s number is
      <say-as interpret-as="telephone">2125550123</say-as>
</speak>
```

Standard vs. neural TTS

Text

Sentence to synthesize.

ˈsɛntəns tə ˈsɪnθəˌsaɪz.

Standard vs. neural TTS

Text

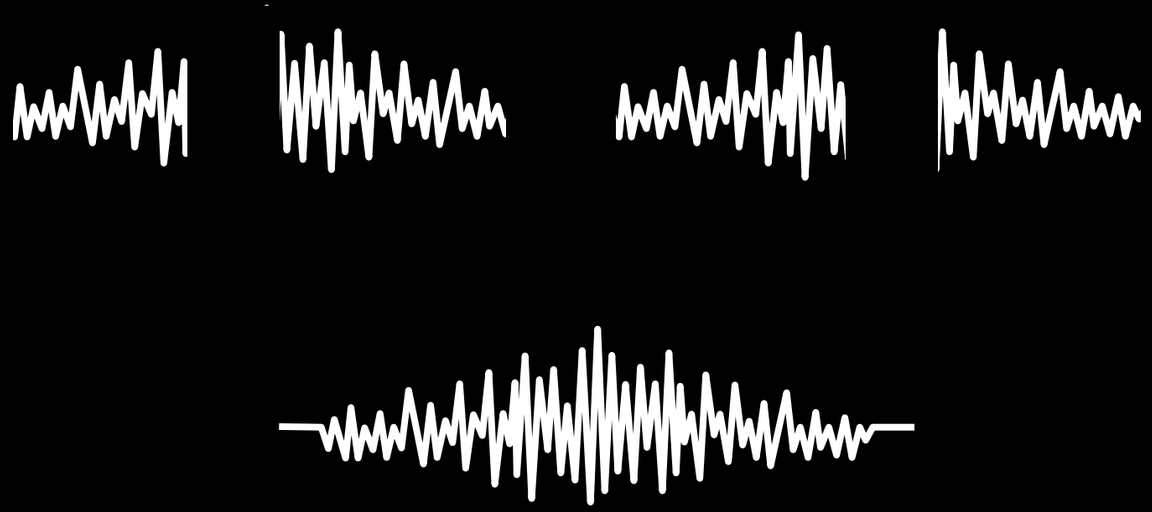
Sentence to synthesize.

ˈsɛntəns tə ˈsɪnθəˌsaɪz.

Phonetic transcription

Concatenative TTS

ˈsɛnt sɛntəns tə ˈsɪnθ əˌsaɪz.



Standard vs. neural TTS

Text

Sentence to synthesize.

ˈsɛntəns tə ˈsɪnθəˌsaɪz.

Phonetic transcription

Concatenative TTS

ˈsɛnt sɛntəns tə ˈsɪnθ əˌsaɪz.

Neural TTS

Neural TTS with Polly

Standard vs. neural

"Hi. My name is Boaz, and I will be your host today during this Innovate online session."

Standard

Neural TTS with Polly

Standard vs. neural

"Hi. My name is Boaz, and I will be your host today during this Innovate online session."

Standard

Neural TTS

Neural TTS with Polly

Standard vs. neural

"Hi. My name is Boaz, and I will be your host today during this Innovate online session."

Standard

Neural TTS

Styles with neural TTS

"Hi. My name is Boaz, and I will be your host today during this Innovate online session."

Newscaster NTTS

Add voice to your app

"This is a new technology that can give users more choice and better accessibility to our content, so we wanted to create an experiment to dive deeper into the user experience. After a month, we'll take what we've learned about how users engage with this feature to develop our first iteration of a product with **Amazon Polly**."

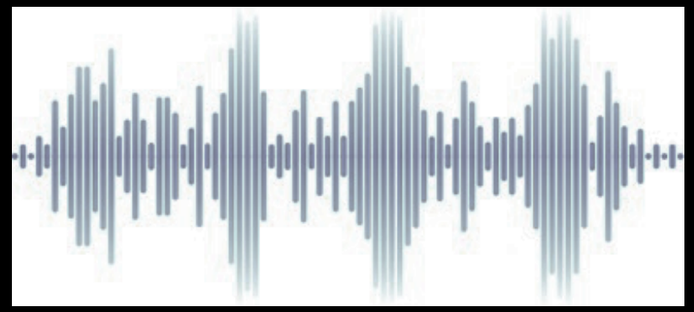
Joseph Price, Product Manager
The Washington Post

aws

Amazon Transcribe

Automatic speech recognition

Automatic speech recognition service



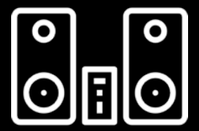
Amazon Transcribe

“Hello. This is Allan speaking.”

Amazon Transcribe: Key features

Custom vocabulary

Speaker identification



Channel identification

Punctuation and capitalization

Word-level time stamps

Word-level confidence scores

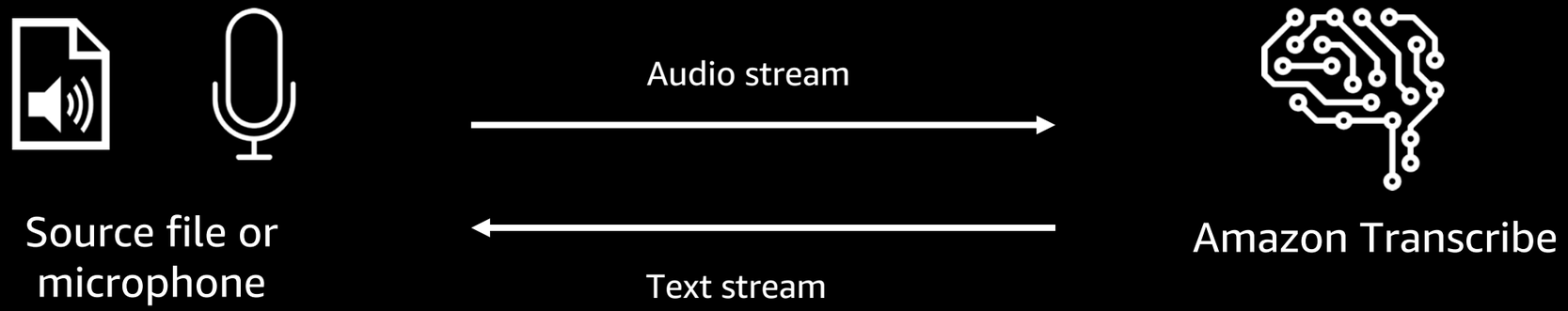
Amazon Transcribe: Streaming transcription

Bi-directional stream over:

HTTP/2 protocol

OR

WebSocket (WSS) protocol



Speech to text

RingDNA is an end-to-end communications platform for sales teams.
Hundreds of enterprise organizations use RingDNA to increase productivity, engage in smarter sales conversations, gain predictive sales insights, and improve their win rate.

"A critical component of RingDNA's Conversation AI requires best-of-breed speech-to-text to deliver transcriptions of every phone call. RingDNA is excited about Amazon Transcribe since it provides high-quality speech recognition at scale, helping us to better transcribe every call to text."

Howard Brown, CEO & Founder
RingDNA

Amazon Translate

Natural and accurate language translation

Key Features

32 Languages

987 Combinations

$15/1M characters

Or $0.000075 per word;
Pay as you go, 2M characters monthly free tier

Real-time

< 500ms / sentence on average
< 150ms / conversational / short form

< / > Tag Handling

XML tags placement maintains styling and formatting through translation

Data Security

Data ownership
Encryption
Access Management

Ease of Use

Simple API calls and partner solutions

HIPAA **Eligible**

Amazon Translate

Natural and fluent language translation

“Hello, what’s up? Do you want to go see a movie tonight?”

"Hallo, wat is er? Wil je vanavond naar de film gaan?"

Translate API example

```
boazz: ~$ aws translate translate-text \
--text "Hello, what’s up? Do you want to go see a movie tonight?" \
--source-language-code auto --target-language-code nl


{
    "TargetLanguageCode": "nl",
    "TranslatedText": "Hallo, wat is er? Wil je vanavond naar de film gaan?",
    "SourceLanguageCode": "en"
}
```

Translate API example

```
boazz: ~$ aws translate translate-text \

--text "Hello, what’s up? Do you want to go see a movie tonight?" \

--source-language-code auto --target-language-code nl


{

    "TargetLanguageCode": ”nl",

    "TranslatedText": "Hallo, wat is er? Wil je vanavond naar de film gaan?",

    "SourceLanguageCode": "en"

}
```

Translate API example

```
import boto3
translate = boto3.client("translate")
lang_flag_pairs = [("fr", "🇫🇷"), ("de", "🇩🇪"), ("es", "🇪🇸"),
                   ("pt", "🇵🇹"),("zh", "🇨🇳"), ("ja", "🇯🇵"),
                   ("ru", "🇷🇺"),("it", "🇮🇹"), ("zh-TW", "🇹🇼"),
                   ("tr", "🇹🇷"), ("cs", "🇨🇿"), (”he", "🇮🇱")]
for lang, flag in lang_flag_pairs:
    print(flag)
    print(translate.translate_text(
        Text="Hello, World",
        SourceLanguageCode="en",
        TargetLanguageCode=lang
    )['TranslatedText'])
```


aws

Translate API example

https://github.com/ziniman/aws-translate-demo

🇫🇷

Bonjour, Monde

🇩🇪

Hallo, Welt

🇪🇸

Hola, Mundo

🇵🇹

Olá, Mundo

🇨🇳

您好，世界

🇯🇵

ハローワールド

🇷🇺

Привет, Мир

🇮🇹

Ciao, Mondo

🇹🇼

大家好，世界

🇹🇷

Merhaba, Dünya.

🇨🇿

Ahoj, světe.

🇮🇱

.שלום, עולם

Scaling real-time translation

Using **Amazon Translate**, Lionbridge is able to scale machine translation in order to localize content faster and in more languages. Using Amazon Translate, Lionbridge was able to reduce translation costs by 20 percent.

Real-time audio transcription & translation

https://voice.boaz.cloud/

https://github.com/ziniman/amazon-transcribe-websocket-static

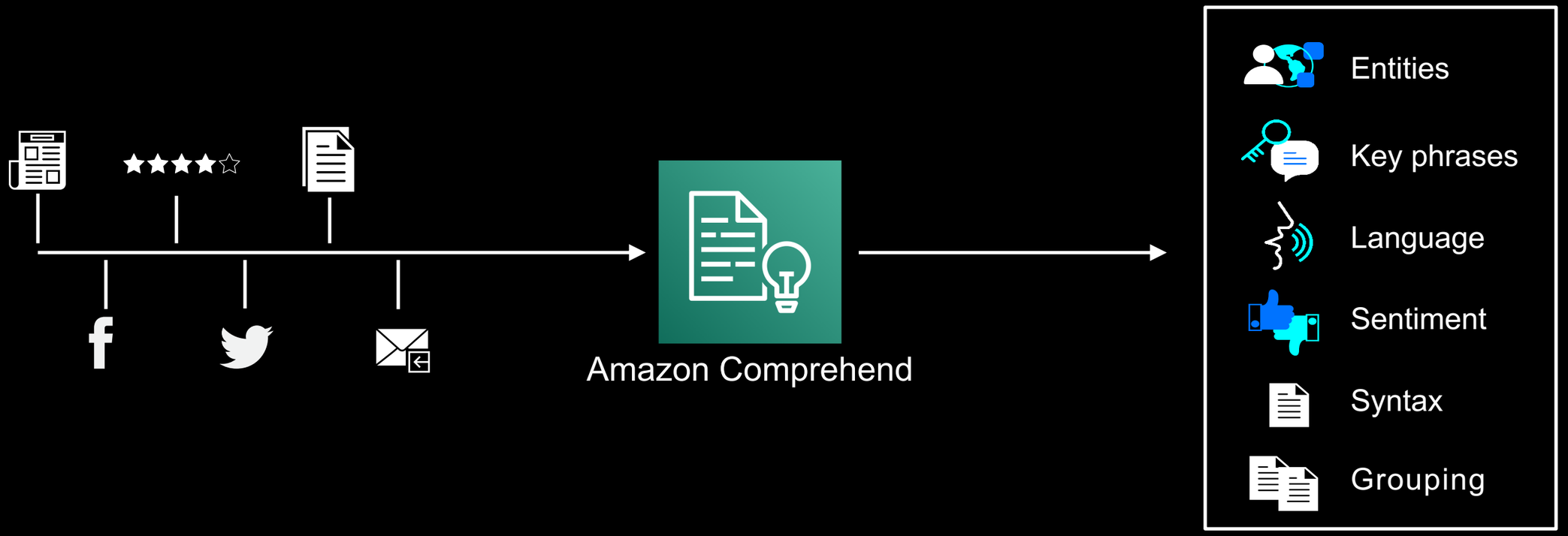
aws

Amazon Comprehend

Discover insights and relationships in text

Amazon Comprehend

Discover insights and relationships in text



English, Spanish, German, French, Italian, Portuguese

Run Amazon Comprehend on an S3 bucket

```
import boto3
import json

s3 = boto3.resource('s3')
bucket_name = 'my_bucket'
region_name = 'us-east-1'
bucket = s3.Bucket(bucket_name)

comprehend = boto3.client(service_name='comprehend', region_name=region)

for obj in bucket.objects.all():
    body = obj.get()['Body'].read()
    text = body

    sentiment_response = comprehend.detect_sentiment(Text=text, LanguageCode='en')
    print(json.dumps(sentiment_response, sort_keys=True, indent=4))
```


aws

Amazon Lex

Conversational interfaces for your applications powered by the same deep learning technologies as Alexa

Amazon Lex: Use cases

A service for building conversational interfaces into your applications using voice and text

CONTACT CENTER BOTS

Customer service IVR

Account inquiries
Bill payments
Service updates

APPLICATION BOTS

Conversational interfaces

Book tickets
Order food
Manage bank accounts

IoT BOTS

Device interactions

Kiosks
Appliances
Auto

INFORMATIONAL BOTS

Answer questions

News updates
Weather information
Game scores

PRODUCTIVITY BOTS

Enterprise efficiencies

Check sales numbers
Inventory status
Expense reports

Amazon Lex use case: Digital assistant to book a hotel

"Book a hotel in NYC"

Book | a | hotel | in | NYC

Automatic speech recognition

Hotel booking

New York City

Natural language understanding

Intent/slot model

Utterances

Hotel booking	
City	New York City
Check in	November 30
Check out	December 2

"Can I go ahead with the booking?"

λ

Confirmation: "Your hotel is booked for November 30."

Amazon Polly

"Your hotel is booked for November 30."

aws

Put AI to work for your business

Modernize your contact center to improve customer service

conversational chat bots | call transcription | intelligent routing | sentiment analysis | VoC analytics
text-to speech | multilingual omni-channel communication

AMAZON POLLY

AMAZON TRANSCRIBE

AMAZON TRANSLATE

AMAZON COMPREHEND

AMAZON LEX

Vonage

RedAwning

Liberty Mutual INSURANCE

American Heart Association.

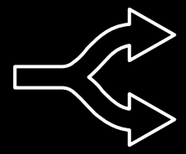
@ziniman

aws

Amazon Connect

Easy to use, cloud-based contact center solution
that scales to support businesses of any size

With tools that grow with your needs



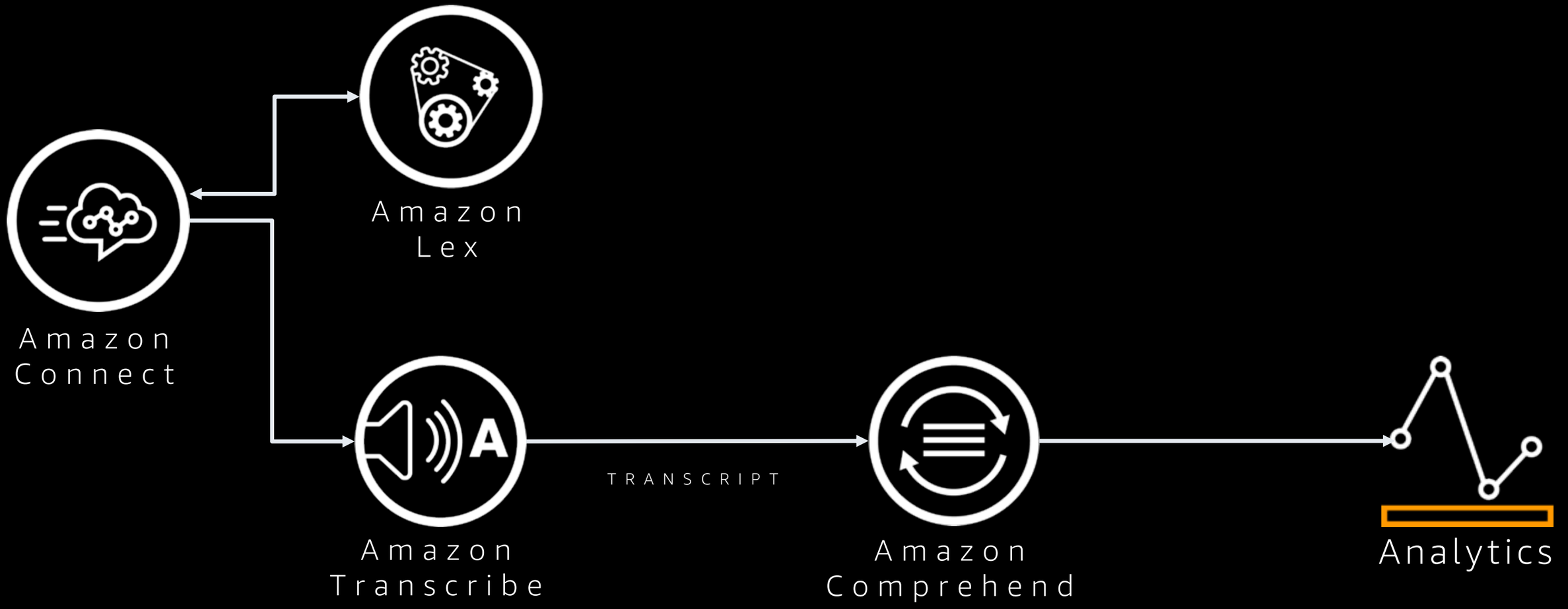
Skills-based routing
[Automatic Call Distribution (ACD)]

Call
recording

Real time and
historical analytics

High-quality
voice capability

Improving Contact Centers with Artificial Intelligence



Demo: Amazon Connect

+44 117 4565386
+31 800 0200802
+49 800 5052761
+1 979 3355593

Demo: Amazon Connect

+44 117 4565386
+31 800 0200802
+49 800 5052761
+1 979 3355593

Put AI to work for your business

Reduce localization costs & improve accuracy

Custom vocabulary | Timestamp generation | Secure real-time translation | Language identification

AMAZON POLLY

AMAZON TRANSCRIBE

AMAZON TRANSLATE

AMAZON COMPREHEND

isentia

Hotels.com

ringDNA

Lionbridge

aws

Babel fish

"The Babel fish is small, yellow, leech-like—and probably the oddest thing in the universe. It feeds on brain wave energy, absorbing all unconscious frequencies and then excreting telepathically a matrix formed from the conscious frequencies and nerve signals picked up from the speech centres of the brain, the practical upshot of which is that if you stick one in your ear, you can instantly understand anything said to you in any form of language: the speech you hear decodes the brain wave matrix."

The Hitchhiker's Guide to the Galaxy, Douglas Adams

Build your own Babel fish

Person speaks in
English

Build your own Babel fish

Amazon
Transcribe

Person speaks in
English

Transcribes voice to text
in **English**

Build your own Babel fish

Person speaks in **English**

Transcribes voice to text in **English**

Translates text to **German**

Build your own Babel fish

Amazon Transcribe

Amazon Translate

Amazon Polly

Person speaks in **English**

Transcribes voice to text in **English**

Translates text to **German**

Speaks in **German**

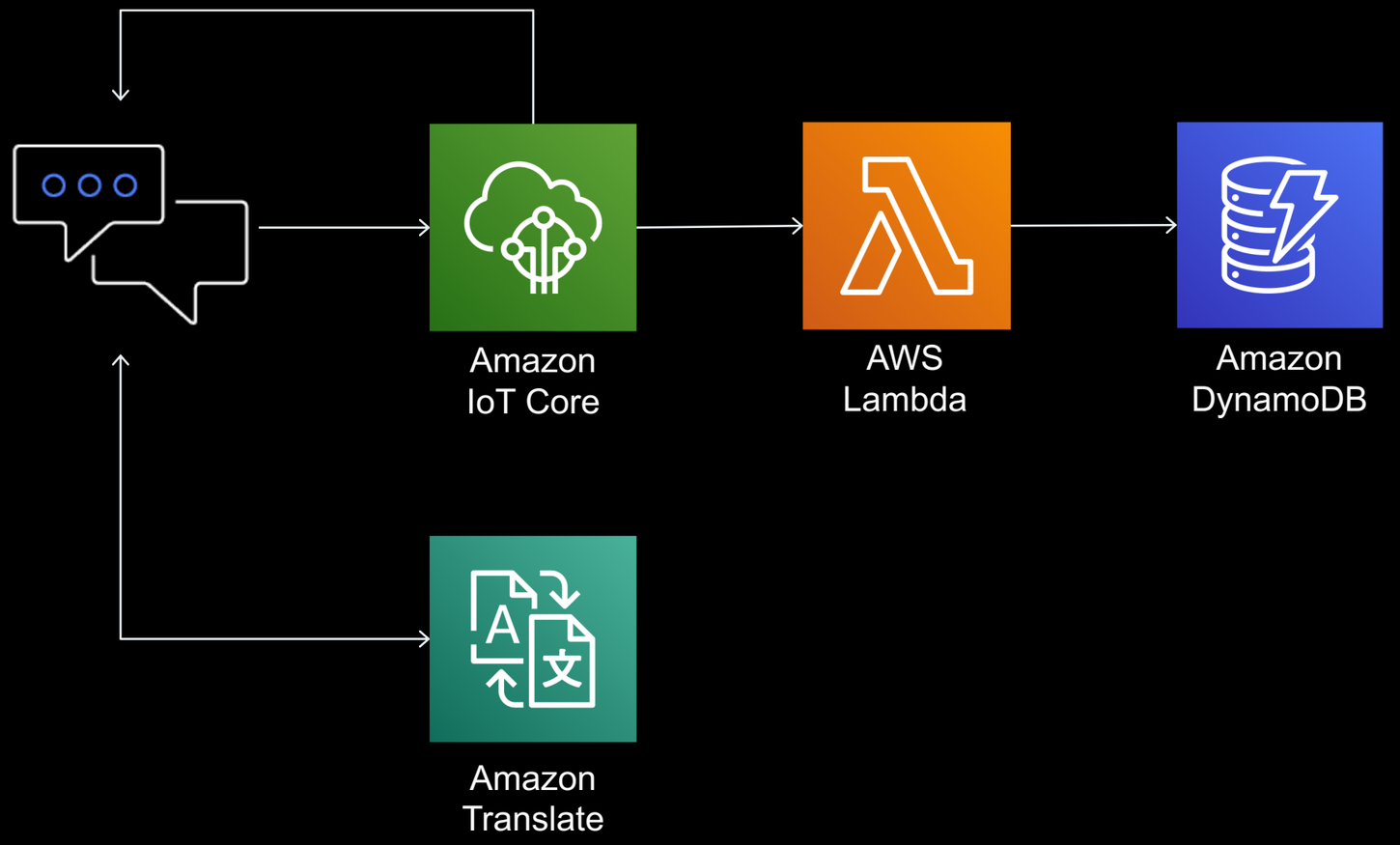
Babel Chat

https://chat.boaz.cloud

Babel Chat

https://chat.boaz.cloud


```
{
  "room":"/",
  "message":{
    "user":"Boaz",
    "text":"I'm doing great",
    "timestamp":1540978160578
  },
  "clientId":"07958771059731382"
}
```

The Amazon ML stack: Broadest & deepest set of capabilities

AI SERVICES

Vision


AMAZON REKOGNITION IMAGE


AMAZON REKOGNITION VIDEO


AMAZON TEXTRACT

Speech


AMAZON POLLY


AMAZON TRANSCRIBE

Language


AMAZON TRANSLATE


AMAZON COMPREHEND & AMAZON COMPREHEND MEDICAL

Chatbots


AMAZON LEX

Forecasting


AMAZON FORECAST

Recommendations


AMAZON PERSONALIZE

ML SERVICES


AMAZON SAGEMAKER

BUILD

Pre-built algorithms & notebooks

Data labeling (GROUND TRUTH)

Algorithms & models (AWS MARKETPLACE FOR MACHINE LEARNING)

TRAIN

One-click model training & tuning

Optimization (NEO)

Reinforcement learning

DEPLOY

One-click deployment & hosting

ML FRAMEWORKS & INFRASTRUCTURE

Frameworks

TensorFlow

mxnet

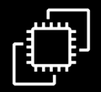
PYTORCH

Interfaces

GLUON

Keras

Infrastructure


EC2 P3 & P3dn


EC2 C5


FPGAs


AWS IoT GREENGRASS


AMAZON ELASTIC INFERENCE

aws

Thank you!

Boaz Ziniman
Principal Technical Evangelist - Amazon Web Services

@ziniman ziniman

Please

Remember to rate this session

Thank you!