



**AWS
re:Invent**

Amazon AI

An introduction to the new services

Julien Simon, Principal Technical Evangelist, AWS
@julsimon

Amazon Polly

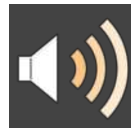
What is Amazon Polly

- A service that converts text into lifelike speech
- Offers 47 lifelike voices across 24 languages
- Low latency responses enable developers to build real-time systems
- Developers can store, replay and distribute generated speech

Amazon Polly: Quality

Natural sounding speech

A subjective measure of how close TTS output is to human speech.



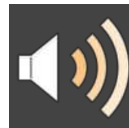
Accurate text processing

Ability of the system to interpret common text formats such as abbreviations, numerical sequences, homographs etc.



Today in Las Vegas, NV it's 54°F.

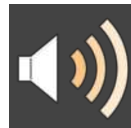
"We live for the music", live from the Madison Square Garden.



Highly intelligible

A measure of how comprehensible speech is.

"Peter Piper picked a peck of pickled peppers."



Amazon Polly: Language Portfolio

EMEA:

- British English
- Danish
- Dutch
- French
- German
- Icelandic
- Italian
- Norwegian
- Polish
- Portuguese
- Romanian
- Russian
- Spanish
- Swedish
- Turkish
- Welsh
- Welsh English

Americas:

- Brazilian Portuguese
- Canadian French
- English (US)
- Spanish (US)

A-PAC:

- Australian English
- Indian English
- Japanese

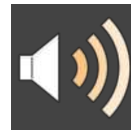


Amazon Polly features: SSML

Speech Synthesis Markup Language

is a W3C recommendation, an XML-based markup language for speech synthesis applications

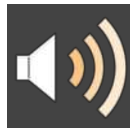
```
<speak>  
  My name is Kuklinski. It is spelled  
  <prosody rate='x-slow'>  
    <say-as interpret-as="characters">Kuklinski</say-as>  
  </prosody>  
</speak>
```



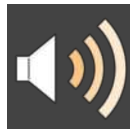
Amazon Polly features: Lexicons

Enables developers to customize the pronunciation of words or phrases

My daughter's name is Kaja.



```
<lexeme>  
  <grapheme>Kaja</grapheme>  
  <grapheme>kaja</grapheme>  
  <grapheme>KAJA</grapheme>  
  <phoneme>"kaI.ə</phoneme>  
</lexeme>
```



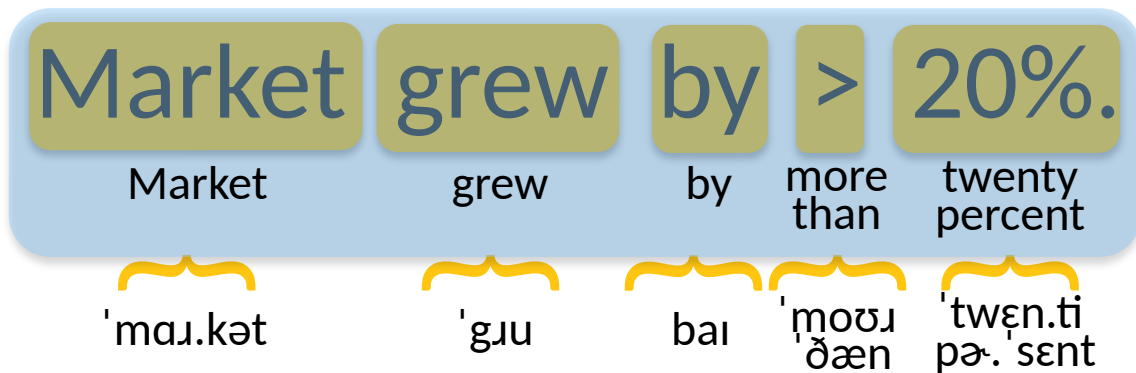
Main Challenges of Text-to-Speech

Goal: Convert text into intelligible, accurate, and natural speech

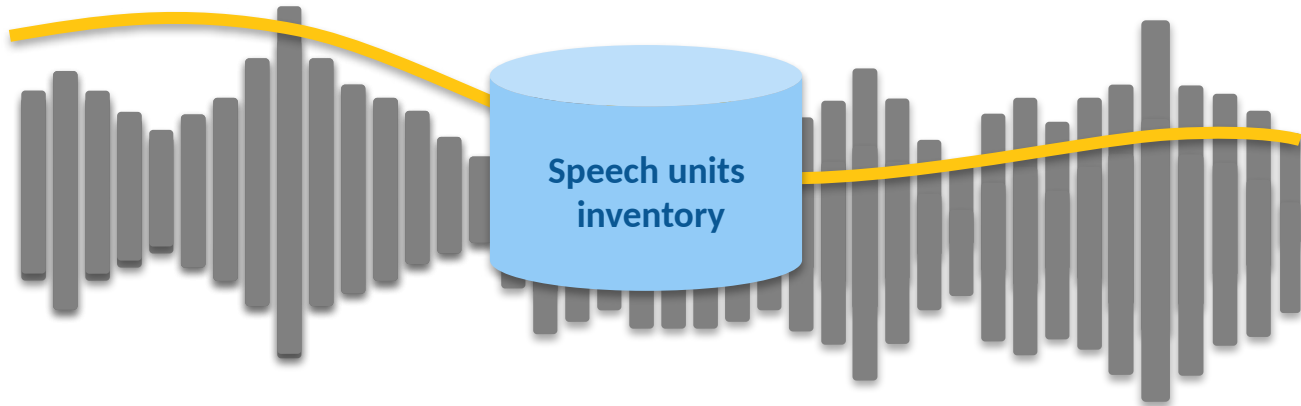
Challenges:

- Homographs: words written identically that have different pronunciation
*I **live** in Las Vegas vs This presentation broadcasts **live** from Las Vegas*
- Text normalization: disambiguation of abbreviations, acronyms, units
*'**St.**' expanded as '**street**' or '**saint**'*
- Conversion of text to phonemes (Grapheme-to-Phoneme) in languages with complex mapping such as English e.g. ***tough**, **through**, **though***
- Foreign words (***déjà vu***), proper names (***François Hollande***), slang (***ASAP**, **LOL***) etc.

TEXT PROCESSING



UNIT SPEECH RECOGNITION



Get started

Amazon Polly

Amazon Polly converts text to lifelike speech in the cloud. You can download the generated audio from the console, or stream it directly to your applications and services through the API.

[Get started](#)

[Getting started guide](#)

Customize pronunciation

Change the way Text-to-Speech pronounces unusual words such as industry-specific acronyms or uncommon words.

[Learn more](#)

Explore speech

Just type text, and then choose Listen to speech. You can experiment with any of the 24 available languages.

[Learn more](#)

Download or stream

Generate audio files with ease in the console, or integrate with the API to stream speech directly to your apps or services.

[Learn more](#)

Amazon Polly documentation and support

[Getting started guide](#) | [Amazon Polly documentation](#) | [Amazon Polly account](#) | [Feedback](#)

© 2009 - 2016, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy Policy](#) [Terms of Use](#)

audio (33).mp3 [Show all](#)

First app

```
from boto3 import Session
from contextlib import closing

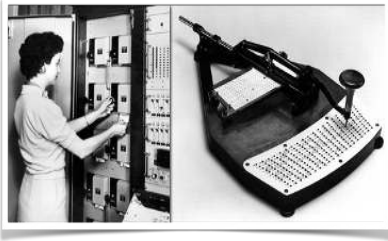
polly = Session().client("polly")

response = polly.synthesize_speech(
    Text="Hello world!",
    OutputFormat="mp3",
    VoiceId="Joanna")

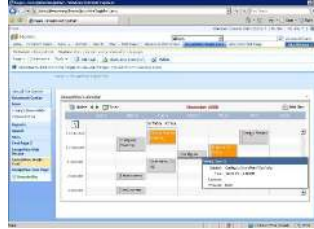
with closing(response["AudioStream"]) as stream:
    with open("speech.mp3", "wb") as file:
        file.write(stream.read())
```

Amazon Lex

Advent of Conversational Interactions



1st Gen:
Punch Cards & Memory Registers

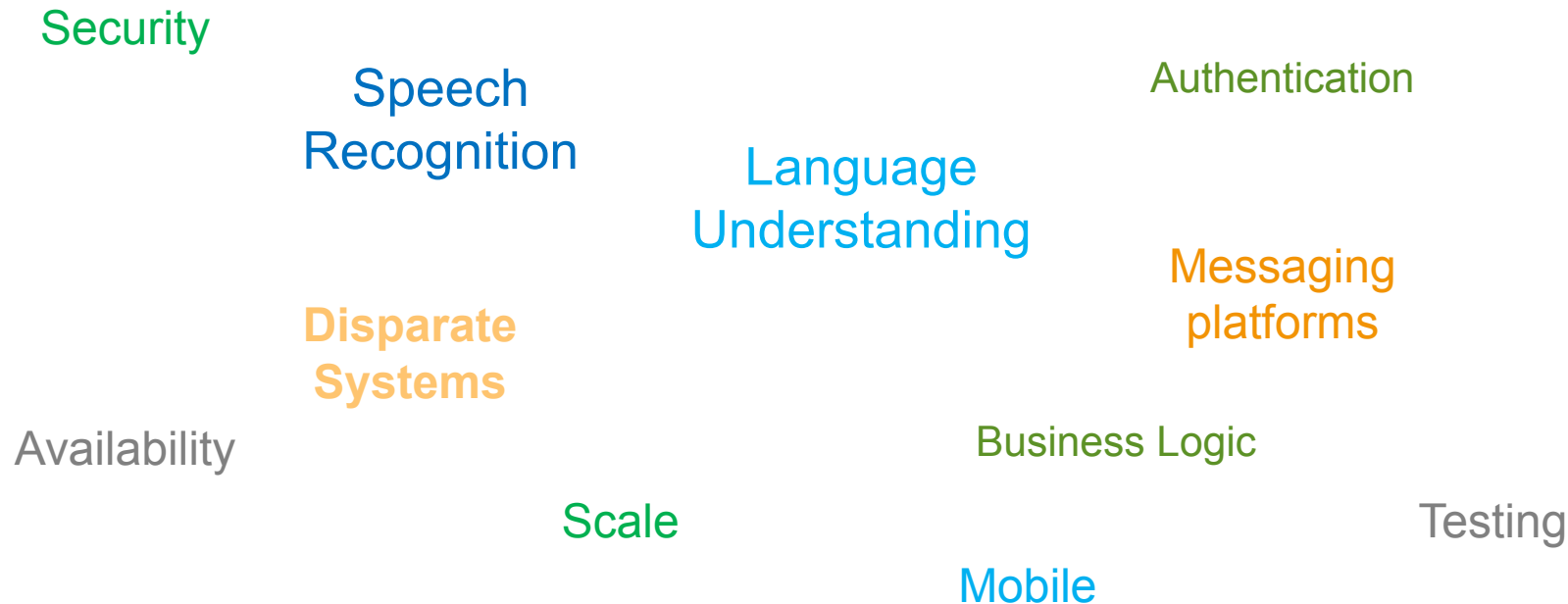


2nd Gen:
Pointers & Sliders



3rd Gen:
Conversational Interfaces

Developer Challenges



Conversational interfaces need to combine a large number of sophisticated algorithms and technologies

Text and Speech Language Understanding

Speech
Recognition

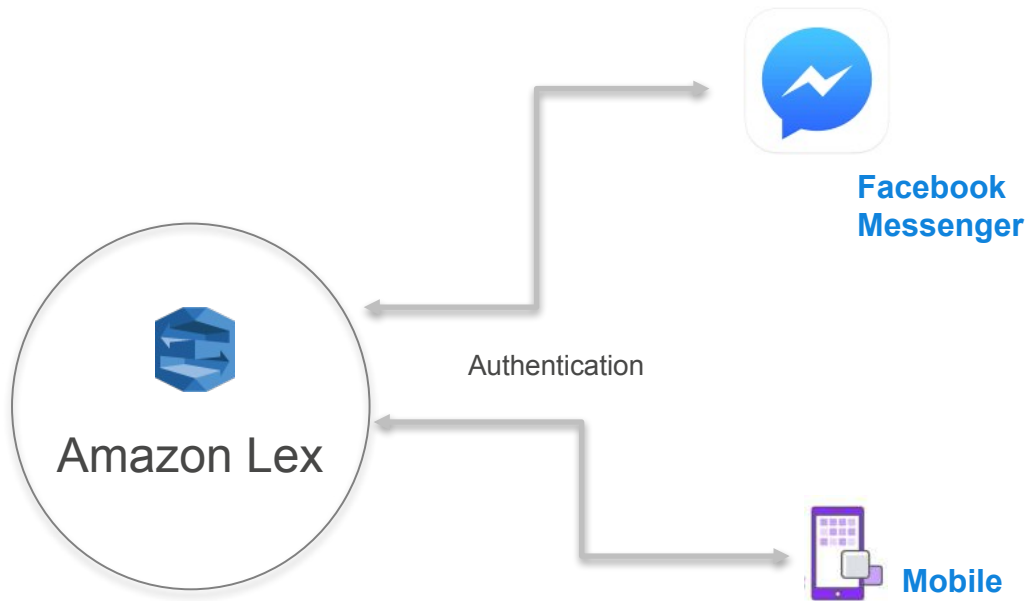


Natural Language
Understanding

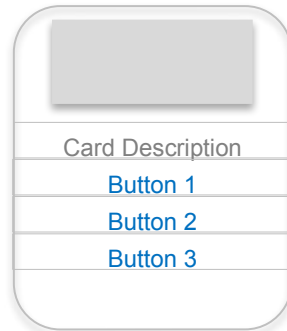
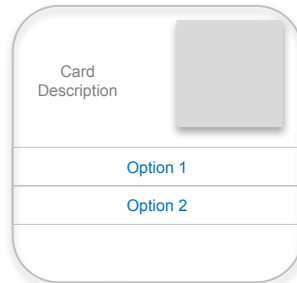
Powered by the same Deep Learning technology as Alexa

Deployment to Chat Services

One-Click Deployment



Rich Formatting



Designed for Builders

< BookTrip Latest +

Build Publish

Editor Settings

Intents

- BookCar
- BookHotel

Slot types

- @CarType
- @ReservationType
- @RoomTypes

Error Handling

BookHotel Latest +

Delete Save

Sample utterances

- e.g. I would like to book a flight.
- Book a hotel
- I want to make a hotel reservations
- I want to book a hotel in {Location}

Slots

| Required | Name | Slot type | Latest - | Prompt |
|-------------------------------------|---------------|------------------|----------|--|
| <input type="checkbox"/> | e.g. Location | e.g. AMAZON.CITY | Latest - | e.g. What city? |
| <input checked="" type="checkbox"/> | CheckInDate | AMAZON.DATE | Latest - | What day do you want to check in? |
| <input checked="" type="checkbox"/> | Nights | AMAZON.NUMBER | Latest - | How many nights will you be staying? |
| <input checked="" type="checkbox"/> | Location | AMAZON.US_CITY | Latest - | What city will you be staying in? |
| <input checked="" type="checkbox"/> | RoomType | RoomTypes | Latest - | What type of room would you like, queen, king or deluxe? |
| <input type="checkbox"/> | StreetAddress | Address | Latest - | What address will you be traveling to? |

Options

- ☐ Initialization and validation code hook
- ☒ Confirmation prompt

Confirm

The price of this {Nights} night stay in {Location} is {currentReservationPrice} dollars. Shall I book the reservation?

Cancel (if the user says "no")

Okay, I have cancelled your reservation in progress

Fulfillment

☒ AWS Lambda function ☐ Return parameters to client

BookStavFunction

☐ Goodbye message ☒ Follow-up message

I have placed your reservation, your confirmation code is {lastConfirmationCode}. Can I help you with anything else?

Cancel (if the user says "no")

Thanks for booking with us, come again soon!

Test App

I want to book a hotel

What city will you be staying in?

Seattle

What day do you want to check in?

30th November

How many nights will you be staying?

Time or speak to your application

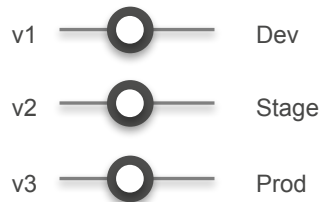
Efficient and intuitive tools to build conversations

Versioning and Alias Support



- Supported for Intents, Slots and Bots
- Enables multi-developer environment
- Rollback to previous versions

Versioning



- Deploy different aliases to different platforms
- Run different stacks for dev, stage and prod environments
- Target different user groups with different aliases

Alias

Amazon Lex – Use Cases



Informational Bots

Chatbots for everyday consumer requests

- News updates
- Weather information
- Game scores



Application Bots

Build powerful interfaces to mobile applications

- Book tickets
- Order food
- Manage bank accounts



Enterprise Productivity Bots

Streamline enterprise work activities and improve efficiencies

- Check sales numbers
- Marketing performance
- Inventory status



Internet of Things (IoT) Bots

Enable conversational interfaces for device interactions

- Wearables
- Appliances
- Auto

Amazon Lex - Benefits



Easy to use



High quality Text and Speech Language Understanding



Seamlessly deploy and scale



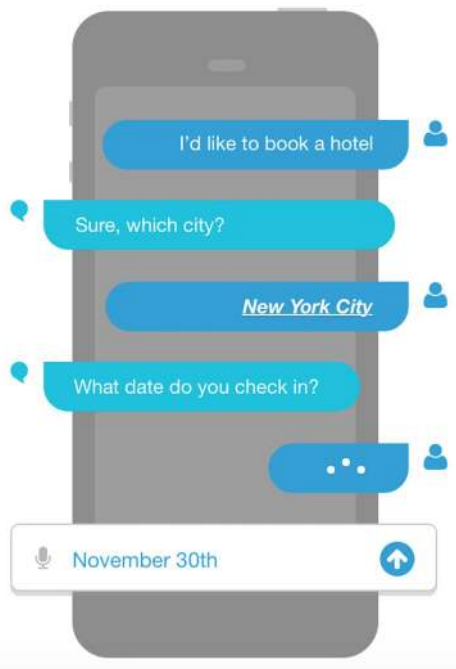
Built-in integration with the AWS platform



Cost effective

Lex Bot Structure

BookHotel



Intents

An Intent performs an action in response to natural language user input

Utterances

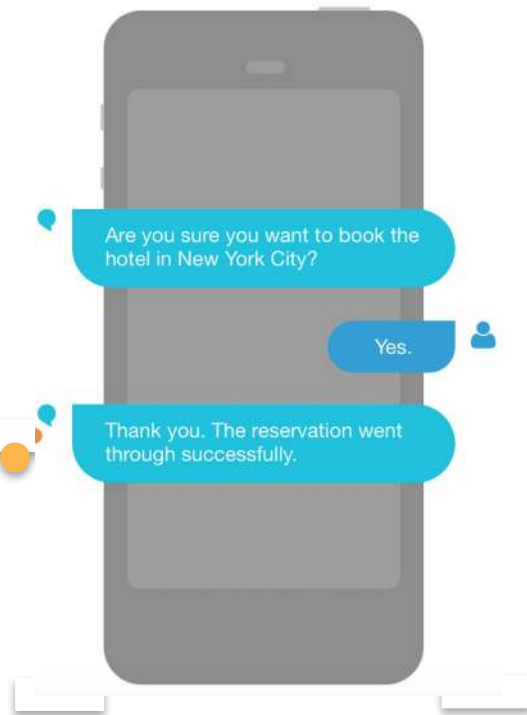
Spoken or typed phrases that invoke your intent

Slots

Slots are input data required to fulfill the intent

Fulfillment

Fulfillment mechanism for your intent



Utterances

I'd like to book a hotel

I want to make my hotel reservations

Can you help me book my hotel?

I want to book a hotel in New York City

Slots

| Slot | Type | Values |
|-------------|------|-------------------------------------|
| destination | City | New York City, Seattle, London, ... |
| Check In | Date | Valid dates |
| Check Out | Date | Valid dates |

Slot Elicitation

I'd like to book a hotel

Sure what city do you want to book?

New York City

City
New York City

What date do you check in?

Nov 30th

Check In
11/30/2016

Fulfillment



Intents and slots passed to
AWS Lambda function for
business logic
implementation.

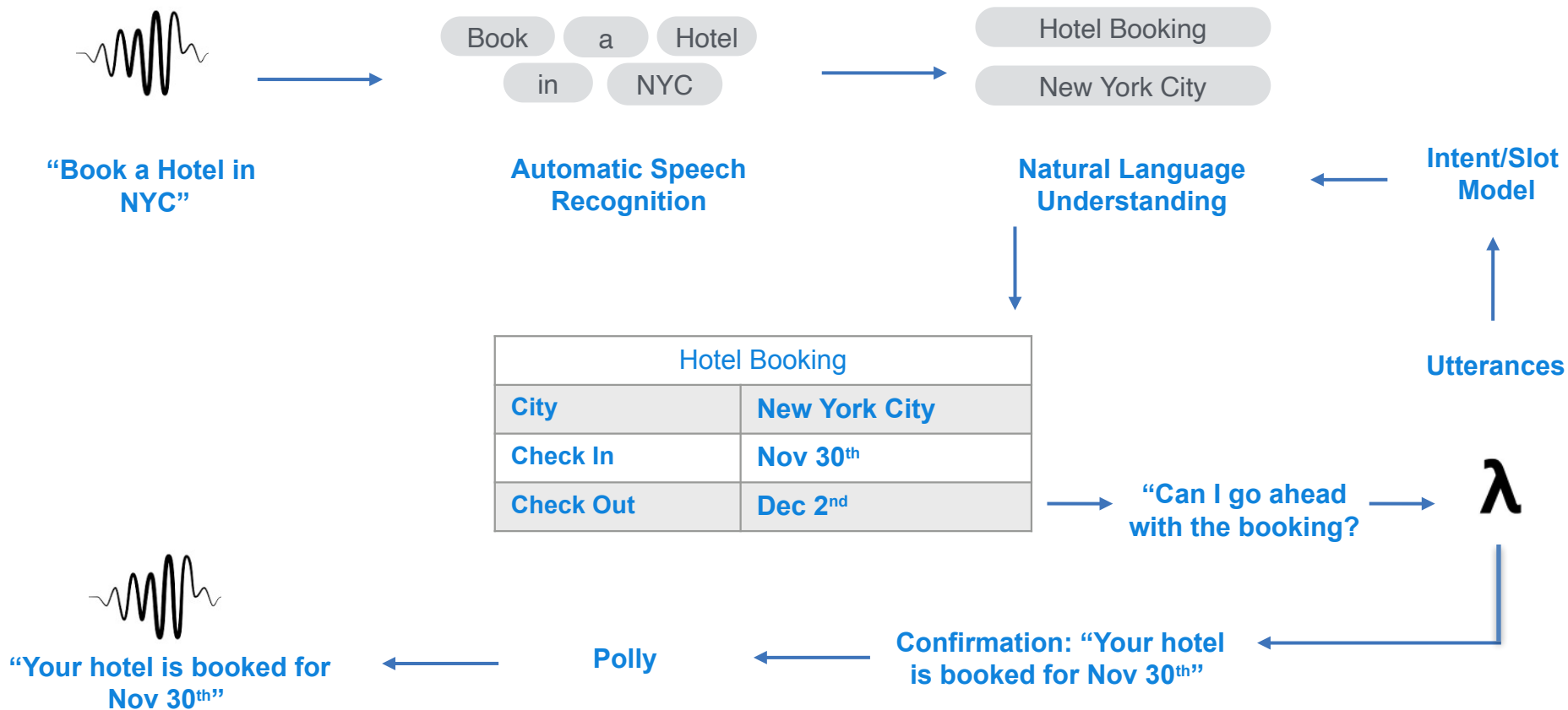
**AWS Lambda
Integration**



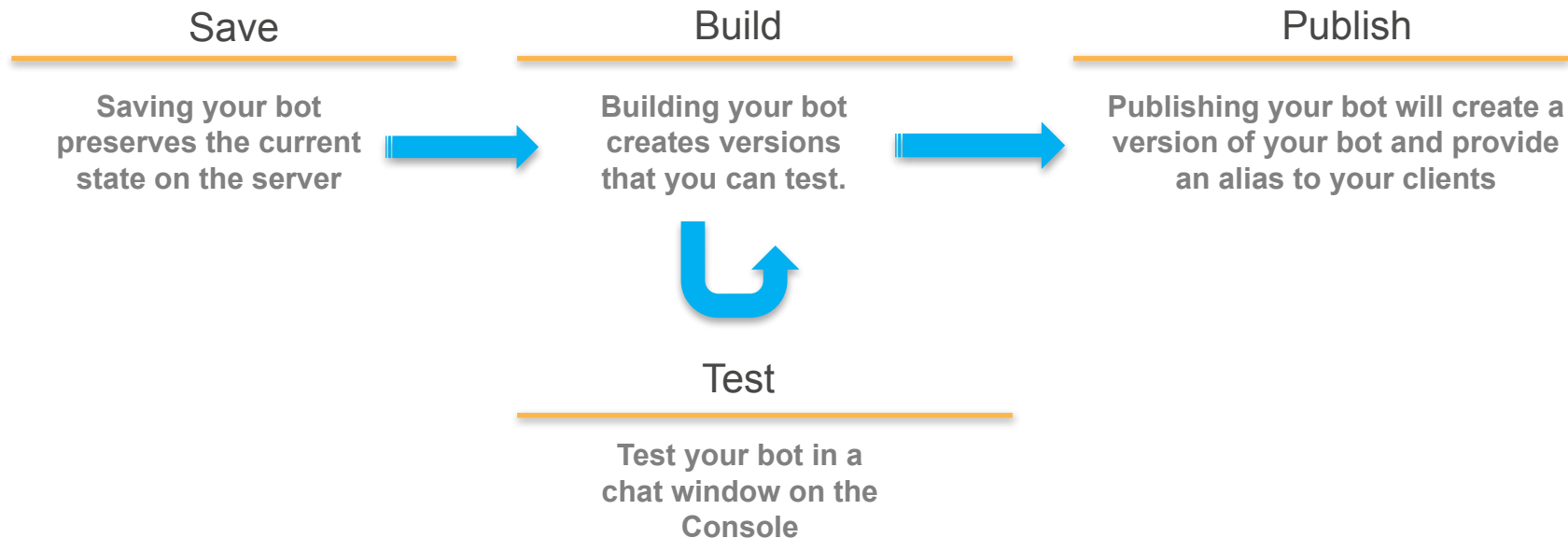
User input parsed to derive
intents and slot values.
Output returned to client for
further processing.

Return to Client

“Book a Hotel”



Save, Build and Publish



Monitoring



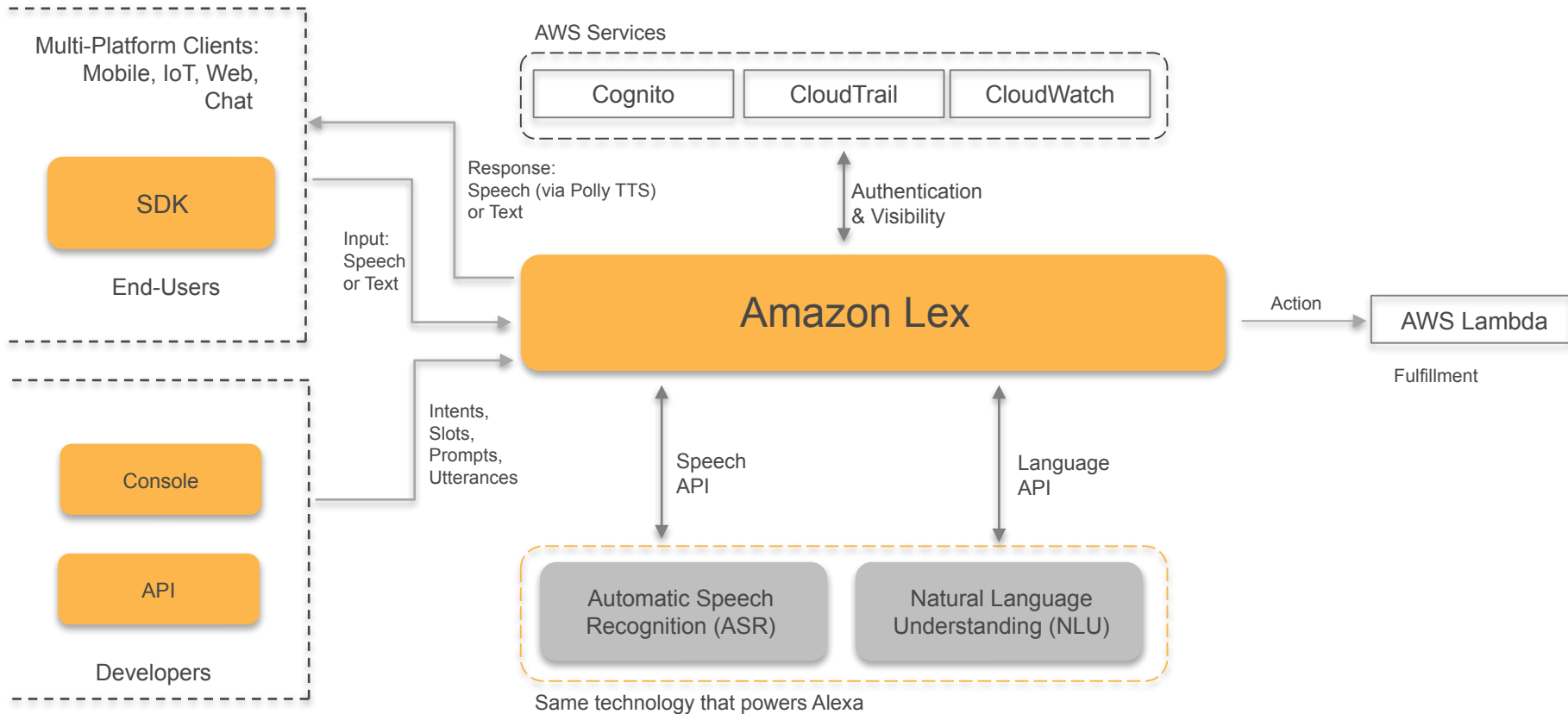
Missed Utterance Count



Request Latency

Track your bot

Amazon Lex - Technology



Amazon Rekognition

Amazon Rekognition

Deep learning-based image recognition service
Search, verify, and organize millions of images



Object and Scene
Detection



Facial
Analysis



Face
Comparison



Facial
Recognition



Amazon Rekognition API

Object and Scene Detection

Detect objects, scenes, and concepts in images



DetectLabels



Amazon Rekognition API



DetectLabels

```
{  
  "Confidence": 94.62968444824219,  
  "Name": "adventure"  
},  
{  
  "Confidence": 94.62968444824219,  
  "Name": "boat"  
},  
{  
  "Confidence": 94.62968444824219,  
  "Name": "rafting"  
},  
. . .
```



Object and Scene Detection



Flower

Indoors

Coffee Table

Chair

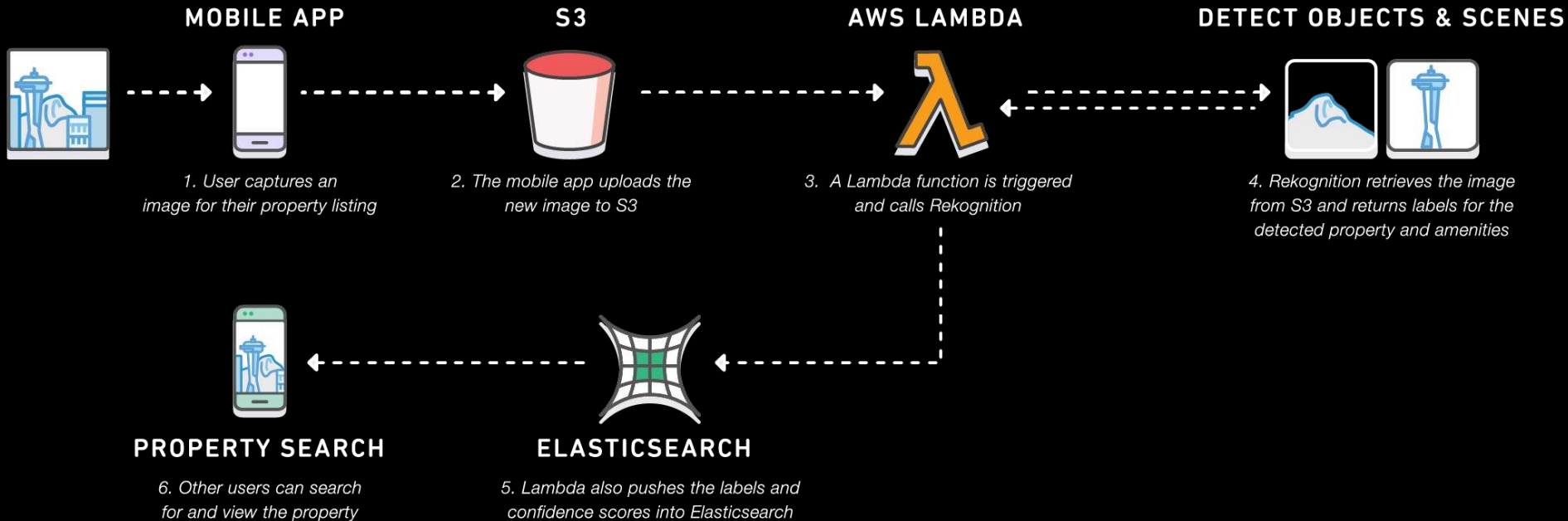
Living Room

Object and Scene Detection



Object and Scene Detection – Use Case

Dynamic Search Indexing



Amazon Rekognition API

Facial Analysis

Detect face and key facial characteristics



DetectFaces



Amazon Rekognition API



[

{

```
"BoundingBox": {  
  "Height": 0.3449999988079071,  
  "Left": 0.096666666388511658,  
  "Top": 0.27166667580604553,  
  "Width": 0.23000000417232513  
},  
"Confidence": 100,  
"Emotions": [  
  {"Confidence": 99.1335220336914,  
   "Type": "HAPPY"},  
  {"Confidence": 3.3275485038757324,  
   "Type": "CALM"},  
  {"Confidence": 0.31517744064331055,  
   "Type": "SAD"}  
],
```

```
"Eyeglasses": {"Confidence": 99.8050537109375,  
  "Value": false},  
"EyesOpen": {"Confidence": 99.99979400634766,  
  "Value": true},  
"Gender": {"Confidence": 100,  
  "Value": "Female"}
```



DetectFaces



Facial Analysis

Demographic Data

Sentiment Expressed

Facial Landmarks

Image Quality

General Attributes



Facial Analysis

Demographic Data

Male
97.4%

Facial Landmarks

Happy
96.0%

Smiling
96.3%

Beard
65.3%

Sentiment Expressed

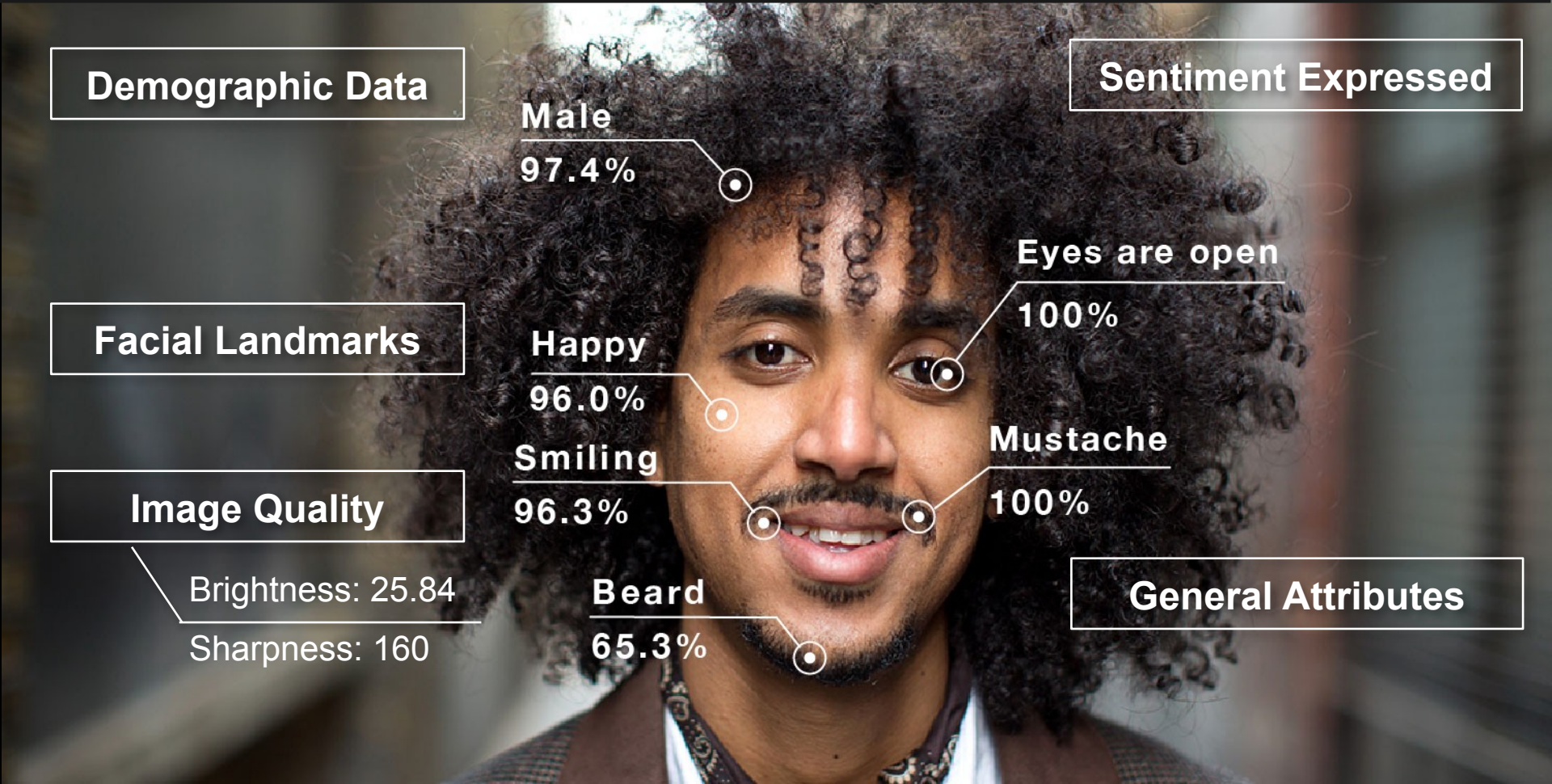
Eyes are open
100%

Mustache
100%

Image Quality

Brightness: 25.84
Sharpness: 160

General Attributes



Using Rekognition Facial Analysis

- Photo printing service can recommend the best photos to their users
- Online dating applications can improve their match recommendations using face attributes
- Retail businesses can understand the demographics and sentiment of in-store customers
- Ad-tech services can display dynamic and personalized content to customers



Amazon Rekognition API

Face Comparison

Face-based user verification



CompareFaces



Amazon Rekognition API

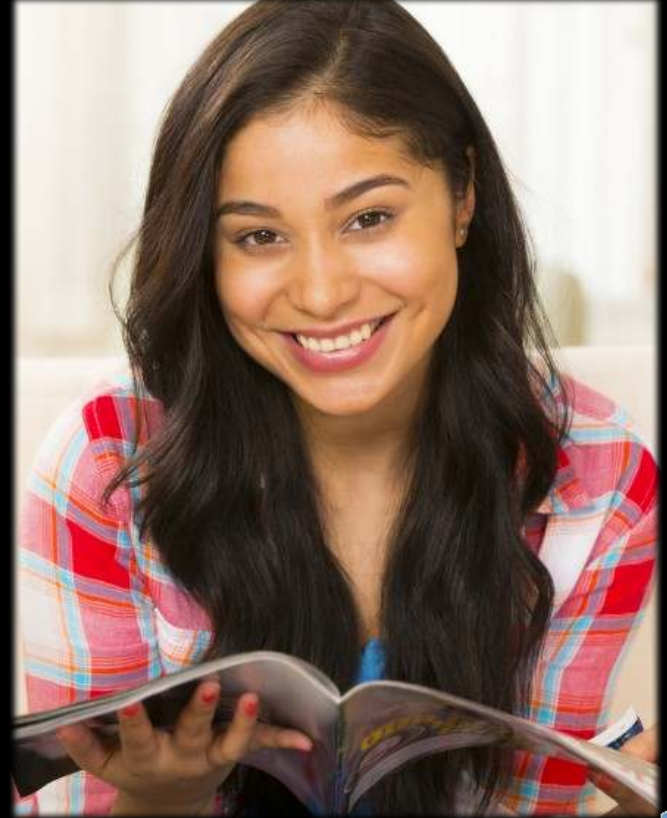
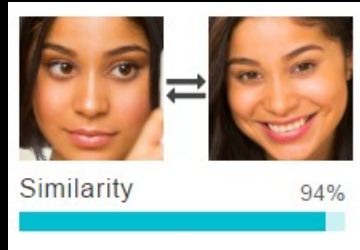


CompareFaces

```
{
  "FaceMatches": [
    {
      "Face": {
        "BoundingBox": {
          "Height": 0.2683333456516266,
          "Left": 0.5099999904632568,
          "Top": 0.1783333271741867,
          "Width": 0.17888888716697693,
          "Confidence": 99.99845123291016,
          "Similarity": 96
        },
        "SourceImageFace": {
          "BoundingBox": {
            "Height": 0.23983436822891235,
            "Left": 0.28333333134651184,
            "Top": 0.351423978805542,
            "Width": 0.1599999964237213,
            "Confidence": 99.99344635009766
          }
        }
      }
    },
    {
      "Face": {
        "BoundingBox": {
          "Height": 0.2383333295583725,
          "Left": 0.6233333349227905,
          "Top": 0.3016666769981384,
          "Width": 0.15888889133930206,
          "Confidence": 99.71249389648438,
          "Similarity": 0
        }
      }
    }
  ],
}
```



Face Comparison



Amazon Rekognition API

Face Recognition
Index and Search faces in a collection



IndexFaces

SearchFacesByImage

Index

Collection

Search



Amazon Rekognition API



IndexFace

```
{  
  f7a3a278-2a59-5102-a549-a12ab1a8cae8,  
  02e56305-1579-5b39-ba57-9afb0fd8782d,  
  4c55926e-69b3-5c80-8c9b-78ea01d30690  
}
```

Face



transformed

Face ID & vector<float>

f7a3a278-2a59-5102-a549-a12ab1a8cae8
&
v1

02e56305-1579-5b39-ba57-9afb0fd8782d
&
v2

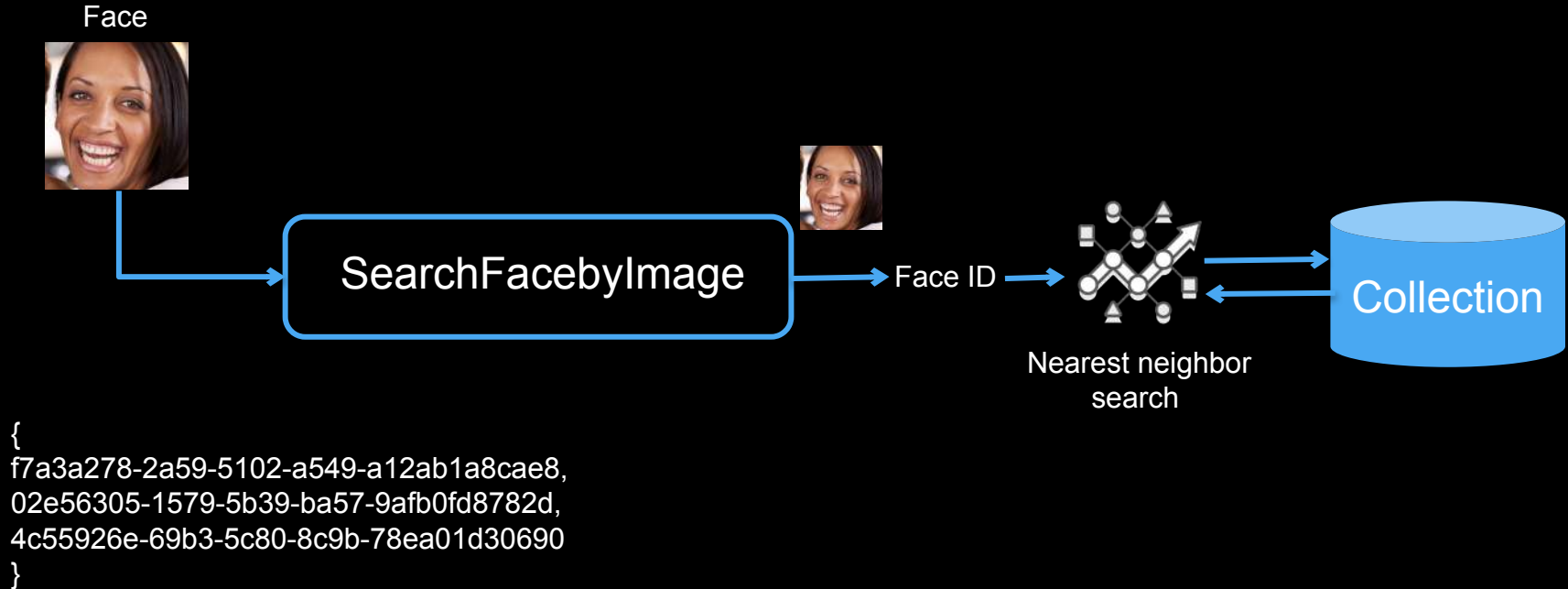
4c55926e-69b3-5c80-8c9b-78ea01d30690
&
v3

stored

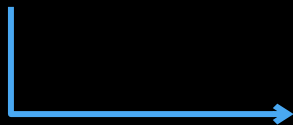
Collection



Amazon Rekognition API



Face Recognition





**AWS
re:Invent**

Thank you!

Julien Simon, Principal Technical Evangelist, AWS
@julsimon