Building Smart Applications with Conversational Interfaces and Language Services

Alex Casalboni
About me

- Software Engineer & Web Developer
- Startupper for 4.5 years
- Serverless Lover & AI Enthusiast
- ServerlessDays Organizer
- AWS Customer since 2013
Our mission

Put machine learning in the hands of every developer
Our Approach

- Customer-focused roadmap (90%+)
- 200+ new launches during 2018
- Wide range of services in production

- Most popular frameworks (Tensorflow, MXNET)
- Security and analytics capabilities
- Embedded Research & Development
AI capabilities

Vision

Speech-to-Text (STT)

Text-to-Speech (TTS)

Natural Language Understanding

Language Translation

Uniform programming interface for text and voice apps
AI capabilities on AWS *

Vision
Speech-to-Text (STT)
Text-to-Speech (TTS)
Natural Language Understanding
Language Translation
Chatbot Interface

Amazon Rekognition
Amazon Transcribe
Amazon Polly
Amazon Comprehend
Amazon Translate
Amazon Lex

* and more!
End-to-end workloads (media)

Audio Input

Amazon S3

store

trigg er

AWS Lambda

call

Amazon Transcribe

in
cascade

Amazon Comprehend

aggregate

Amazon Athena

analyze

Amazon QuickSight
Amazon Rekognition

Add intelligent image and video analysis to your applications
Deep Learning-Based Image and Video Analysis

PERSON
99.3%

CREST
83.0%

MOUNTAIN BIKE
99.1%

OUTDOORS
83.1%

ROCK
82.8%
Optical Character Recognition (OCR)

Results

- IT'S
- MONDAY
- but
- keep

Response

```json
{
  "TextDetections": [
    {
      "DetectedText": "IT'S",
      "Type": "LINE",
      "Id": 0,
      "Confidence": 95.83811950683594,
      "Geometry": {
        "BoundingBox": [null, null, null, null]
      }
    },
    {
      "DetectedText": "MONDAY",
      "Type": "LINE",
      "Id": 1,
      "Confidence": 95.83811950683594,
      "Geometry": {
        "BoundingBox": [null, null, null, null]
      }
    },
    {
      "DetectedText": "but",
      "Type": "LINE",
      "Id": 2,
      "Confidence": 95.83811950683594,
      "Geometry": {
        "BoundingBox": [null, null, null, null]
      }
    },
    {
      "DetectedText": "keep",
      "Type": "LINE",
      "Id": 3,
      "Confidence": 95.83811950683594,
      "Geometry": {
        "BoundingBox": [null, null, null, null]
      }
    },
    {
      "DetectedText": "Smiling",
      "Type": "LINE",
      "Id": 4,
      "Confidence": 95.83811950683594,
      "Geometry": {
        "BoundingBox": [null, null, null, null]
      }
    }
  ]
}
```
Object & Scene Detection

<table>
<thead>
<tr>
<th>Object</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automobile</td>
<td>98.8 %</td>
</tr>
<tr>
<td>Transportation</td>
<td>98.8 %</td>
</tr>
<tr>
<td>Vehicle</td>
<td>98.8 %</td>
</tr>
<tr>
<td>Car</td>
<td>98.8 %</td>
</tr>
<tr>
<td>Human</td>
<td>98.3 %</td>
</tr>
<tr>
<td>Person</td>
<td>98.3 %</td>
</tr>
<tr>
<td>Pedestrian</td>
<td>97.1 %</td>
</tr>
<tr>
<td>Skateboard</td>
<td>94.3 %</td>
</tr>
<tr>
<td>Sports</td>
<td>94.3 %</td>
</tr>
<tr>
<td>Sport</td>
<td>94.3 %</td>
</tr>
<tr>
<td>Road</td>
<td>92.4 %</td>
</tr>
</tbody>
</table>
Facial Analysis

<table>
<thead>
<tr>
<th>Feature</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>looks like a face</td>
<td>99.9 %</td>
</tr>
<tr>
<td>appears to be female</td>
<td>99.8 %</td>
</tr>
<tr>
<td>age range</td>
<td>20 - 38 years old</td>
</tr>
<tr>
<td>smiling</td>
<td>98.6 %</td>
</tr>
<tr>
<td>appears to be happy</td>
<td>92.1 %</td>
</tr>
<tr>
<td>wearing glasses</td>
<td>99.9 %</td>
</tr>
<tr>
<td>wearing sunglasses</td>
<td>87.8 %</td>
</tr>
<tr>
<td>eyes are open</td>
<td>100 %</td>
</tr>
<tr>
<td>mouth is open</td>
<td>99.9 %</td>
</tr>
<tr>
<td>does not have a mustache</td>
<td>99.9 %</td>
</tr>
<tr>
<td>does not have a beard</td>
<td>99.9 %</td>
</tr>
</tbody>
</table>
Face Search/Comparison
Marinus Analytics uses facial recognition to stop human trafficking

“Now with Traffic Jam’s FaceSearch, powered by Amazon Rekognition, investigators are able to take effective action by searching through millions of records in seconds to find victims.”
Amazon Comprehend

Discover insights and relationships in text
NLU Functionalities

- Sentiment
- Entities
- Keyphrases
- Languages
- Topic modeling

Syntax
Supported Languages

- English
- Italian
- French
- German
- Portuguese
Syntax detection

$ aws comprehend detect-synta
--language-code 'en' --text 'I love cloud!' 
{
  "SyntaxTokens": [ 
  {
    "TokenId": 1,
    "Text": "I",
    "BeginOffset": 0,
    "EndOffset": 1,
    "PartOfSpeech": { 
      "Tag": "PRON",
      "Score": 0.9999802112579346
    } 
  },

  
}
Sentiment Analysis

$ aws comprehend detect-sentiment
    --language-code 'en' --text 'I love cloud!'

{
    "Sentiment": "POSITIVE",
    "SentimentScore": {
        "Mixed": 0.012617903761565685,
        "Positive": 0.9599817991256714,
        "Neutral": 0.021758323535323143,
        "Negative": 0.005641999188810587
    }
}
«Amazon Comprehend helps us analyze the key sentiments, objects, and geos in our 30 million plus reviews & testimonies. Now we are able to discover new insights into the unique experiences available at each property, so our customers can make the best decision possible for their travel.»

Matt Fryer, VP and Chief Data Science Officer, Hotels.com
Comprehend Medical

Extract health information from patient notes, clinical trial reports, and other electronic health records
Amazon Polly

Turn text into lifelike speech using deep learning
Amazon Polly

Use cases

• Content creation
• Mobile & desktop applications
• Internet of Things (IoT)
• Education & e-learning
• Telephony
• Game development

Key features

• 58 voices across 28 languages
• Lip-syncing & text highlighting
• Fine-grained voice control
• Custom vocabularies
• Available in 17 AWS Regions
“With Amazon Polly our users benefit from the most lifelike Text-to-Speech voices available on the market.”

Severin Hacker
CTO, Duolingo
Amazon Translate

Yes, natural language translation 😊
Supported languages *

Arabic
Simplified Chinese
French
German
Spanish
Portuguese

Japanese
Traditional Chinese
Italian
Russian
Turkish
Czech

Danish
Dutch
Finnish
Swedish
Polish
Hebrew

* 417 translation combinations
“We operate 90 localized websites in 41 languages. (…)

Having evaluated Amazon Translate and several other solutions, we believe that Amazon Translate presents a quick, efficient and most importantly, accurate solution.”

Matt Fryer, VP and Chief Data Science Officer, Hotels.com
Amazon Transcribe

Turn speech into text
Supported Languages

- English
- Italian
- French
- Spanish
- Portuguese
Use cases

- Call centers
- Subtitles for VOD
- Broadcast closed captions
- Transcribe meetings
*End-to-end communications platform for sales teams.*

*Hundreds of enterprise organizations use RingDNA to dramatically increase productivity, engage in smarter sales conversations, gain predictive sales insights and improve their win rate.*

**Speech to Text**

"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

[https://www.youtube.com/watch?v=1ZJ_f1bDdog](https://www.youtube.com/watch?v=1ZJ_f1bDdog)
Amazon Lex

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

**Dialog Management**
- Native support & maintains context

**Speech to Intent**
- ASR+NLU integrated into one API

**Text to Speech**
- Amazon Polly integrated into API

**Analytics**
- Monitor and improve

**Deployment**
- One click deployment

**Scale**
- Completely managed service

**Business Logic**
- Native integration with AWS Lambda

**Security**
- Encrypted data in transit & at rest

**End to End**
“Alexa, help me get started with skills.”
NEW (Preview)

Amazon Textract

OCR++ service to extract data from virtually any document
(No ML experience required)
Amazon Textract: How it works

<table>
<thead>
<tr>
<th>Method</th>
<th>Num. clusters</th>
<th>Rand index</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM-score</td>
<td>8</td>
<td>89.7%</td>
</tr>
<tr>
<td>FPFH</td>
<td>9</td>
<td>89.3%</td>
</tr>
<tr>
<td>3DSC</td>
<td>9</td>
<td>89.5%</td>
</tr>
<tr>
<td>RSD</td>
<td>7</td>
<td>92.0%</td>
</tr>
<tr>
<td>VFH</td>
<td>8</td>
<td>85.3%</td>
</tr>
<tr>
<td>Combined silhouette weights</td>
<td>7</td>
<td>92.2%</td>
</tr>
<tr>
<td>Combined equal weights</td>
<td>7</td>
<td>90.2%</td>
</tr>
</tbody>
</table>

**GENERAL CLAIM SUBMISSION FORM**

**NAME** Presley, Elvis Aaron

**ADDRESS** Graceland, Memphis, TN

**ID** 12-12-1234

**COMPANY** TCB Limited
NEW (Preview)

Amazon Forecast

Accurate time-series forecasting service
(No ML experience required)
The perils of poor predictions in forecasting.
Amazon Forecast: How it works

Historical data: Supply chain, inventory, etc.
Related “causal” data: Weather, special offers, product details

Load data → Inspect data → Identify features → Select from 8 algorithms → Select Hyperparameters → Train models → Optimize models → Host models → Customized forecasting API

Amazon Forecast

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Amazon Personalize

Real-time personalization and recommendation service
(No ML experience required)
Amazon Personalize: How it works

Activity stream from app
Views, signups, conversion, etc.

Inventory
Articles, products, videos, etc.

Demographics (optional)
Age, location, etc.

LOAD DATA (EMR Cluster)
INSPECT DATA
IDENTIFY FEATURES
SELECT ALGORITHMS
SELECT HYPERPARAMETERS
TRAIN MODELS
OPTIMIZE MODELS
HOST MODELS
BUILD FEATURE STORE
CREATE REAL-TIME CACHES

Amazon Personalize

Customized personalization & recommendation API

Fully managed by Amazon Personalize

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Please complete the session survey.