Deep Learning APIs

Scenarios

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY HYDERABAD OUTREACH DIVISION

talent sprint

Emotion detection at retail displays

Facial identification to find missing children

Sentiment analysis to learn how customers feel

Facial detection to calculate the male/female ratio at a nightclub

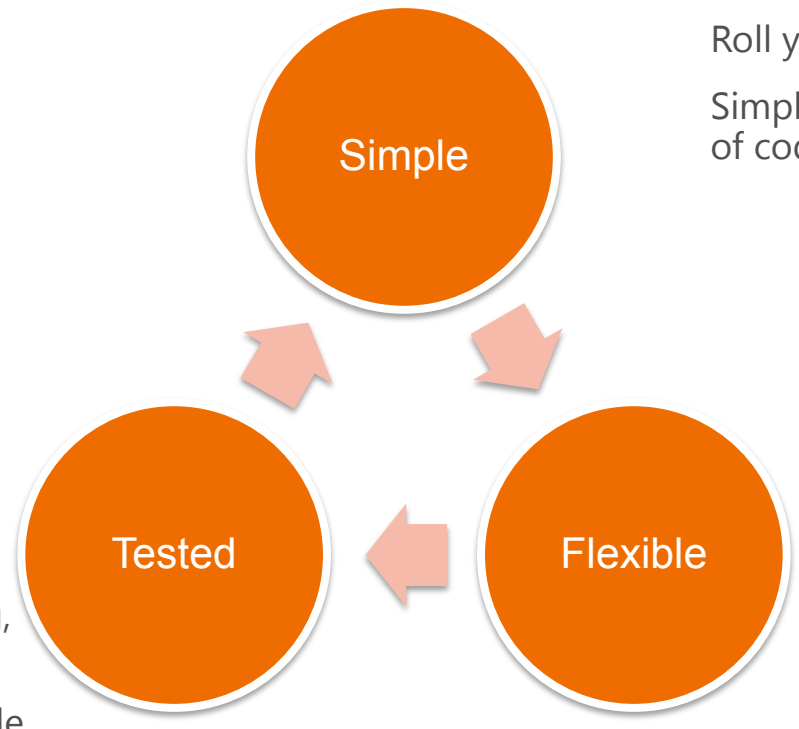
Language understanding to allow automated support bots to understand natural language

Object recognition to enable a blind person to read a menu

Approach

1. Prepare the data
2. Build a Model
3. Operationalize the model
4. Consume the Model

Cognitive Services :Key Tenets



Roll your own with REST APIs

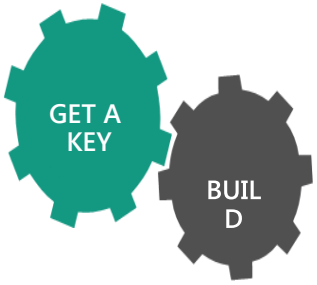
Simple to add: just a few lines of code required

Integrate into the language and platform of your choice with C#, Java, Python etc.

Breadth of offerings helps you find the right API for your app

Built by experts in their field from Microsoft Research, Bing, and Azure Machine Learning

Quality documentation, sample code, and community support

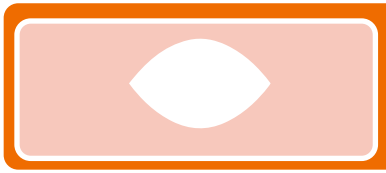


python

GitHub stackoverflow

msdn uservoice

Microsoft Cognitive Services



Vision

- From faces to feelings, allow s apps to understand images and video

Speech

- Hears and speaks to users by filtering noise, identifying speakers, and understanding intent

Language

- Processes text and learn how to recognize what users want

Knowledge

- Tap into rich knowledge amassed from the web, academia, or your own data

Search

- Access billions of web pages, images, videos, and news with the power of Bing APIs

How do I consume the APIs

Narrow the choices

- Based on the type of input Data – e.g. for text data, use service that take text as input
- In case you need to train a model with data you provide to improve the performance and accuracy

Getting Started

- Create an Azure account
- Under Market Place, click AI + Machine Learning
- Choose the service
- Start with the service providing necessary details for your account
- In the end use the Endpoint URL in the Overview section and keys in the Keys section to start making API calls in your applications.

References

- https://docs.microsoft.com/en-us/azure/cognitive-services/cognitive-services-apis-create-account
- https://docs.microsoft.com/en-us/azure/cognitive-services/

Example of an API call

- Get a subscription Key
- We take the **Face - Detect** method from the Vision API to detect faces in an image and return face attributes

Code

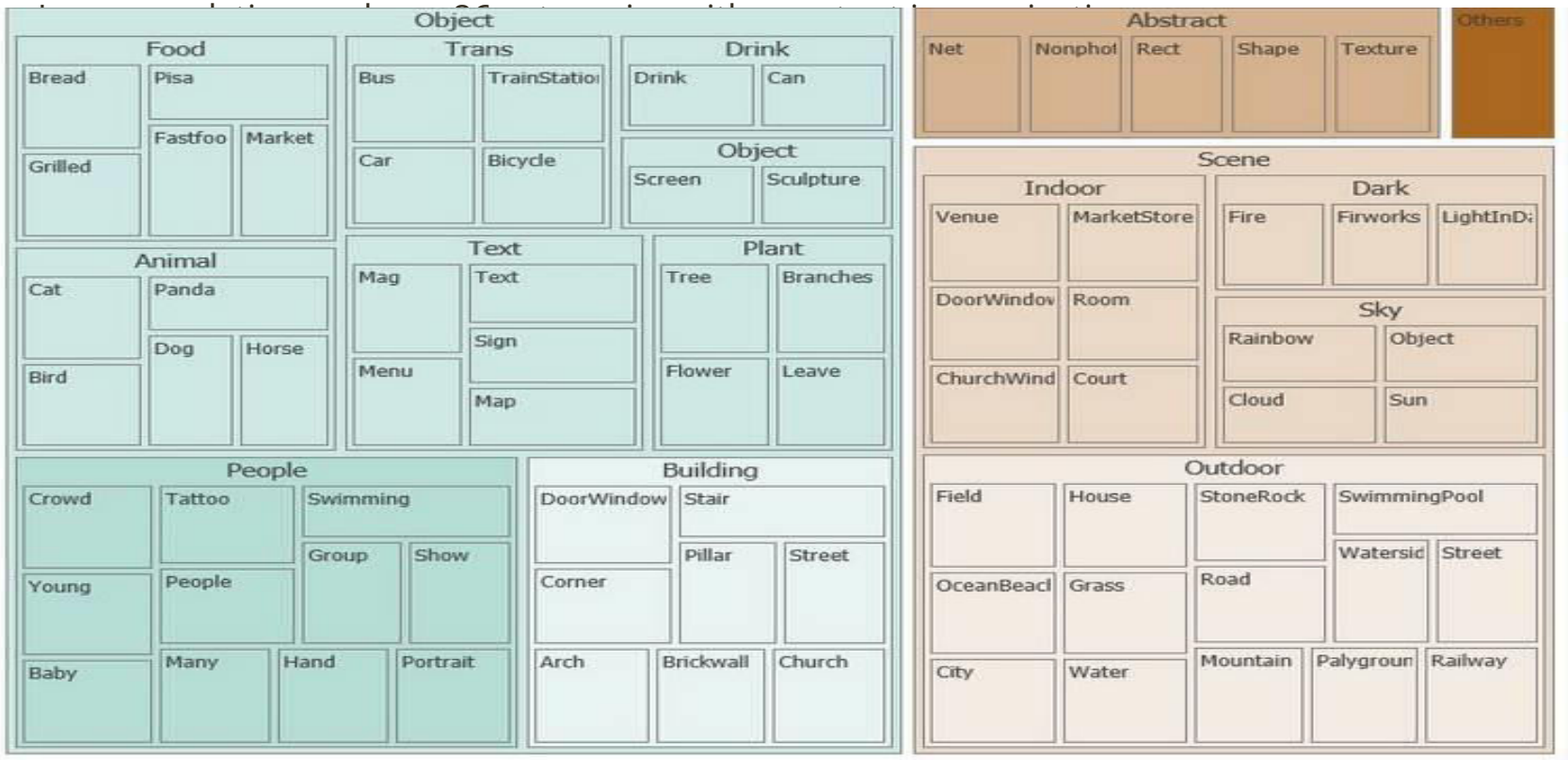
```
import requests
from PIL import Image
from io import BytesIO

subscription_key = "<Subscription Key>"
assert subscription_key
face_api_url =
'https://westcentralus.api.cognitive.microsoft.com/face/v1.0/dete
ct'

image_url = 'https://how-old.net/Images/faces2/main007.jpg'

headers = {'Ocp-Apim-Subscription-Key': subscription_key}
params = {
    'returnFaceId': 'true',
    'returnFaceLandmarks': 'false',
    'returnFaceAttributes':
'age,gender,headPose,smile,facialHair,glasses,' +

'emotion,hair,makeup,occlusion,accessories,blur,exposure,nois
e'
}
data = {'url': image_url}
response = requests.post(face_api_url, params=params,
headers=headers, json=data)
faces = response.json()
```

Output

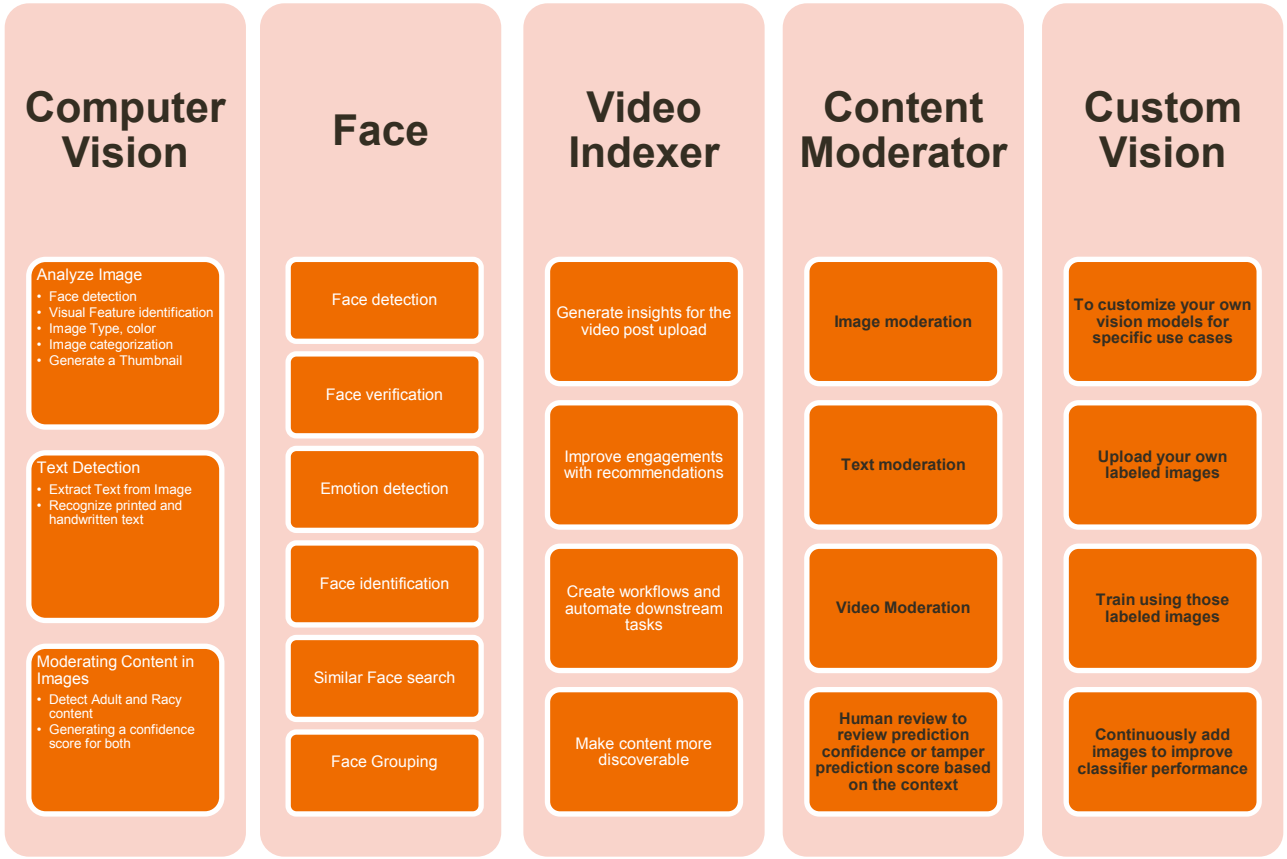
Response is in json as attached:

Microsoft Office Word Document

Vision APIs

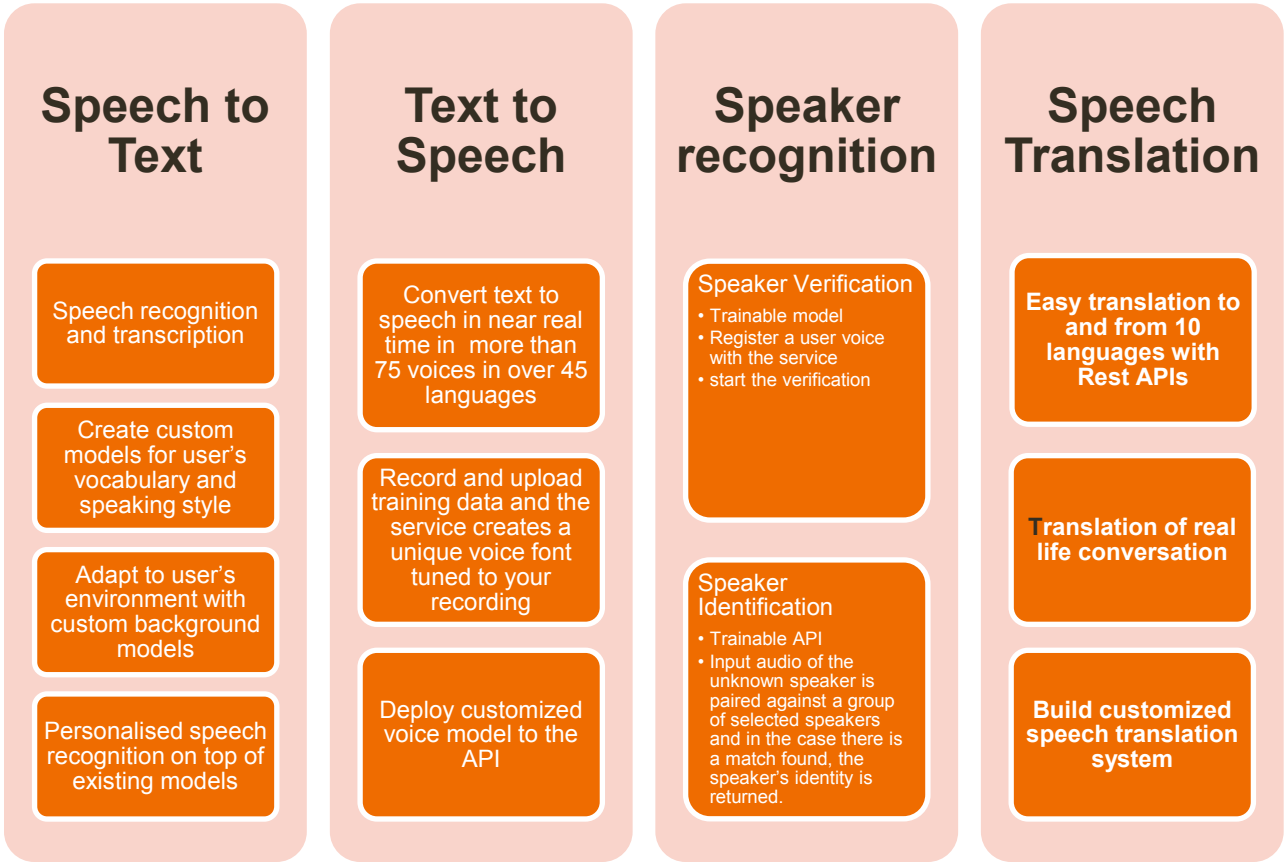
- The cloud-based Computer Vision API provides developers with access to advanced algorithms for processing images and returning information. By uploading an image or specifying an image URL, Microsoft Computer Vision algorithms can analyze visual content in different ways based on inputs and user choices. SDKs are available for : Windows, Node.js, Python, Go, Android, Swift
- 

Vision API Capabilities



Speech API capabilities

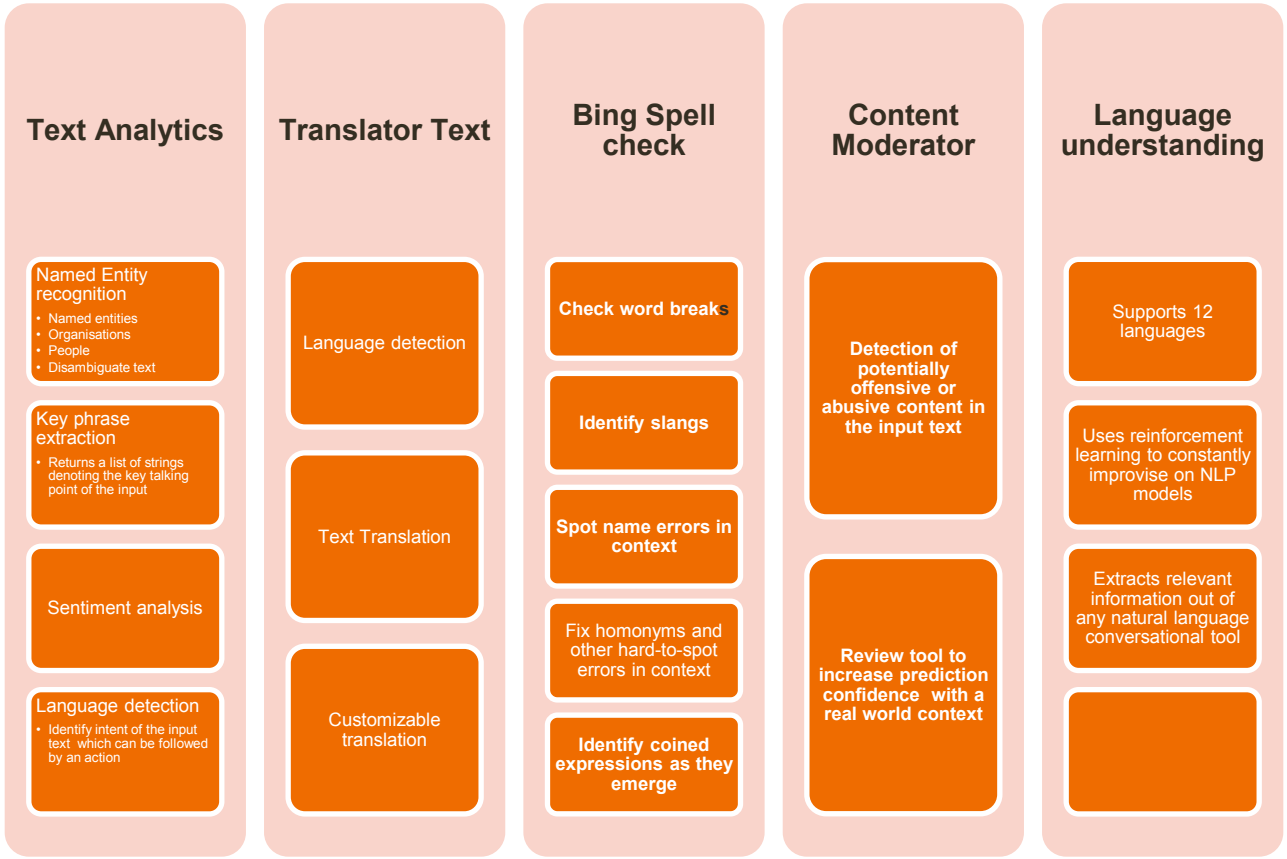
- Enables the integration of speech processing capabilities into any app or service.
- Convert spoken language into text or produce natural sounding speech from text using standard (or customizable) voice fonts



Knowledge API capabilities

- Leverage or create rich knowledge resources that can be integrated into apps and services with Knowledge services
  - Make intelligent recommendations
  - Semantic Search
- QnA maker
  - QnA extraction from unstructured text
  - Knowledge base creation from collections of Q&As
  - Semantic matching for knowledge bases

Language API capabilities



Search API capabilities

WebSearch And AutoSuggest

- Returns location aware search results
- Choose intent of Search – Adult or no
- Spelling corrected search with recommendations for similar queries
- Azure add in for statistics

Visual Search

- Search Visually similar content, extract barcodes or textual information and generate insights
- Create product recommendations for visually similar products for the domain in context
- Identify image content to recognize celebrities, monuments, search similar content etc.

Custom Search

- Define parts of the web that you would want to draw search results from
- Tailored search to drive the outcome needed with autosuggest
- Ad free searches with integration with statistics ad on

Entity Search

- **Customize search for named entity viz a viz celebrities, products, monuments, local businesses etc.**
- **Knowledge acquisition for named entities**

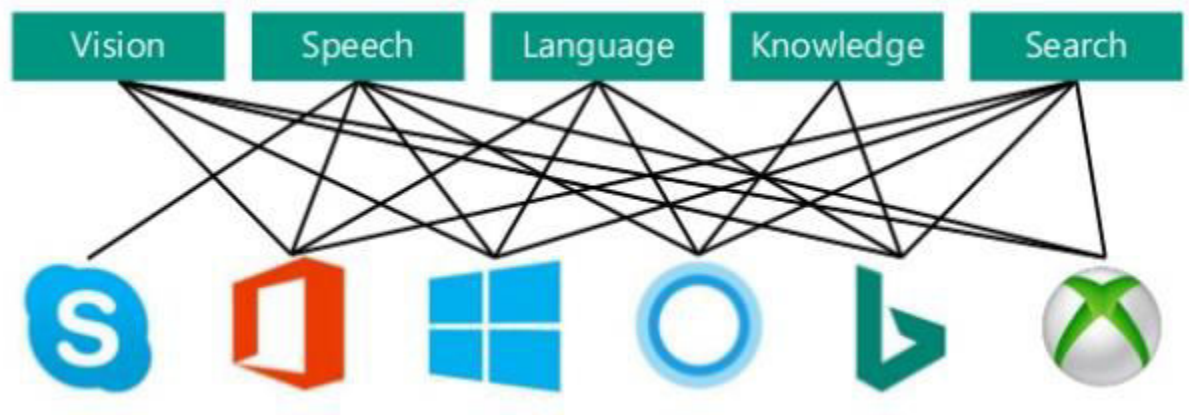
Video and Image Search

- Retrieve video and image search results harnessing the metadata on the content. The metadata includes machine generated insights(related content, visually similar etc.)
- Thumbnail preview with feed for trending content

News Search

- **Search fro news articles with** details like authoritative image of the news article, related news and categories, provider info, article URL and date added
- **Categorize searches by topics e.g. Sports, politics and customize the trending news from around the world by region and category**

API Implementations



TechnologyOne built a bot translation system to help international students talk to universities using LUIS
Detailed case study on implementation:
https://microsoft.github.io/techcasestudies/bot%20framework/cognitive%20services/2017/06/13/TechnologyOneEducationBot.html

Black Radley and the Shrewsbury Museum & Art Gallery uses Cognitive Services Face API to create a solution that would react to museum patrons as they interact with exhibits based upon their approximate age, gender, and emotional state. Additionally, the solution would provide detailed insight on how patrons traverse the museum, which exhibits they linger at, and for how long
https://microsoft.github.io/techcasestudies/cognitive%20services/2017/08/04/BlackRadley.html

Equadex uses Cognitige Services capabilities to help people with language and speech disorders. Their system, Helpicto provides a solution to help children with autism to communicate more easily with their environment, based on pictograms and associated keywords.
https://microsoft.github.io/techcasestudies/cognitive%20services/2017/08/04/equadexcognitives.html

Demo

References

https://azure.microsoft.com/en-us/services/cognitive-services

https://www.lynda.com/Azure-tutorials/Microsoft-Cognitive-Services-Developers/659280-2.html

https://www.slideshare.net/AmandaLange1/intro-to-microsoft-cognitive-services

https://sec.ch9.ms/sessions/build/2016/B855.pptx

Natural Language Processing Service

- Amazon Comprehend NLP
- Google Cloud Natural Language
- IBM Watson NLU
- Microsoft Azure Text Analytics

Feature Comparison

Features	Amazon Comprehend	Google Cloud Natural Language	MS Azure Text Analytics	IBM Watson NLU
Entity Extraction	✓	✓	✓	✓
Key Pharse Extraction	✓	✓	✓	✓
Sentiment Analysis	✓	✓	✓	✓
Syntax Analysis		✓	✓	

Feature Comparison

Features	Amazon Comprehend	Google Cloud Natural Language	MS Azure Text Analytics	IBM Watson NLU
Topic Modeling	✓	✓		
Multiple Language Support	✓	✓	✓	✓
Parts of Speech	✓	✓	✓	

Performance Comparison

Criteria	Amazon Comprehend	Google Cloud Natural Language	MS Azure Text Analytics	IBM Watson NLU
Execution Time	0:00:03.453	0:00:10.969	0:00:12:062	0:00:02.063

Source: https://goo.gl/jVyoPR

Price Comparison

Vision APIs

- Amazon Rekognition
- Google Vision
- IBM Watson Visual
- Azure Computer Vision API

Feature Comparison

Features	Amazon Rekognition	Google Vision	Azure Computer Vision	IBM Watson Visual Recognition
Object Detection	✓	✓	✓	✓
Scene Detection	✓	✓	✓	✓
Facial Recognition	✓		✓	✓
Flag Inappropriate Content	✓	✓	✓	

Feature Comparison

Features	Amazon Rekognition	Google Vision	Azure Computer Vision	IBM Watson Visual Recognition
Facial Analysis	✓	✓	✓	✓
Celebrity Recognition	✓	✓	✓	
Logo Detection		✓		
Text Recognition	✓	✓	✓	✓

Feature Comparison

Features	Amazon Rekognition	Google Vision	Azure Computer Vision	IBM Watson Visual Recognition
Web Detection		✓		
Landmark Detection	✓	✓	✓	✓
Dominant Colors Detection		✓	✓	
Thumbnail Generation			✓	

Performance Comparison

Criteria	Amazon Rekognition	Google Vision	Azure Computer Vision
Execution Time	0:00:15.625	0:00:13.343	0:00:04:484

Source: https://goo.gl/jVyoPR

Price Comparison

Processing Text and Speech

- Amazon Polly
- Microsoft Bing Text-to-Speech
- IBM Watson Text-to-Speech

Feature Comparison

Features	Amazon polly	Microsoft Bing Text-to-Speech	IBM Watson Text-to-Speech
SSML	✓	✓	✓
Multiple Language	✓	✓	✓
Formats	Mp3, Vorbis, PCM	WAV	FLAC, MP3, MPEG, PCM, WAV

Processing Speech-to-Text

- Amazon Lex
- Google Cloud Speech API
- Microsoft Bing Speech-to-Text
- IBM Watson Speech-to-Text

Feature Comparison

Features	Amazon Lex	Google Cloud Speech API	Microsoft Bing Speech-to-Text	IBM Watson Speech-to-Text
Automatic Speech Recognition	✓	✓	✓	✓
Multiple Language		✓	✓	✓
Noisy Accuracy		✓		

Performance Comparison

Criteria	Google Cloud Speech API	Microsoft Bing Speech-to-Text	IBM Watson Speech-to-Text
Execution Time	0:00:15.422	0:00:14.549	0:03:30:234

Source: https://goo.gl/jVyoPR

Price Comparison