Big Five: A comparison of the artificial intelligence technologies

Siddhi Khamitkar

Worcester Polytechnic Institute

Table of Contents

Abstract **4**

Introduction **5**

Devices **10**

Speech Recognition 10

Amazon Lex 10

Image Analysis 16

Amazon Rekognition 16

Cloud Vision 21

Visual Recognition Service 24

Computer Vision API 25

Custom Vision API 28

Text to live speech 30

Amazon Polly 30

Text to Speech 33

Custom Speech Service 34

Bing Speech API 36

Audio to text 38

Cloud Speech API 38

Speech to text API 39

Training and Evaluation 42

Amazon Machine Learning 42

Google Cloud Machine Learning 44

Watson Analytics 46

Trade off analytics 48

Job search 49

Cloud jobs API 49

Text Analysis 52

Natural Language API 52

Watson Explorer 53

Natural Language Classifier 55

Retrieve and Rank 56

Discovery....57
Alchemy Data News....59
Web language model API....60
Text Analytics API....61
Bing Spell Check....62
Translate....65
Cloud translation API....65
Document Conversion....66
Language Translator....67
Translator text API....68
Translator speech API....69
Relationship between the data....71
IBM Watson Knowledge Studio....71
Language Processing....73
Alchemy Language....73
Conversation Service....73
Dialog....75
Language Understanding Intelligent Service....76
Linguistic Analysis API....77
Personality Insights....78
Personality Insights Services....78
Tone analyzer....79
Speaker Recognition API....80
Conclusion....83
References....84
Appendix....87

Abstract

Artificial intelligence is taking developing at a lightening pace in the industry. Every machine is getting automated with minimal human intervention. IBM, Amazon, Google, Facebook and Microsoft in this scenario have joined hands to come together and develop advancements in the field of artificial intelligence. The main aim of the research team is to recommend best practices in the open areas of ethics, fairness, inclusivity, transparency, privacy and interoperability. They assure that this research will benefit the society. Considering the advancements in machine learning, personality analysis, tone analyzing, translation which the companies have been focusing on, this paper gives a comparison of the products and services proposed by the five companies and suggests what further advancements need to be carried out.

Keywords: Artificial Intelligence, Personality analysis, tone analyzing

Big Five: A comparison of the Artificial Intelligence Technologies

A "Big 5" research team has been formed by Amazon, Google, Microsoft, IBM and Facebook to **advance public understanding** of Artificial Intelligence. The team will not only involve members from the Big 5 but also representatives of academia and the nonprofit organizations which indicates that steps of already been taken when it comes to discussion with the Association for Advancement of Artificial Intelligence. Each of the five corporate members have a strong research team for AI, some of which include IBM's Watson and Amazon's Alexa.

Apple and Elon Musk are not involved in this search as Apple is advancing its own developments in the field of Artificial Intelligence. Microsoft's Eric Horvitz has mentioned that Microsoft has been in discussion with Apple to join the team and that Apple was excited. However, Apple did not respond to this comment. Elon Musk's Open AI is also absent from the team.

In the recent years, Artificial Intelligence is booming with its technologies and innovations. For example:

1) **Google's Deepmind Technology Alpha Go** beat the World Go Champion Lee Sedol at one of the most complex world games in history on March 9th 2016.
2) **Tesla's** auto pilot has the driving feature of speed adjusting, lane switching and automatic braking. Yet it dashed into 18-year-old crossing the path. However, auto pilot did take the injured to the hospital. Thus proving that auto driven vehicles would be safer than human driven vehicles.

3) **China`s** leading company **Baidu** is focusing on voice recognition and Natural Language Processing and has identified Machine Learning and artificial Intelligence as the next big areas of innovation

Let us trace back the growth of artificial intelligence from the past years to the current state. The history traces back to the 1950's. Alan Turing created the Turing test to determine if the computer has real intelligence. In this test, the computer had to fool a human and prove that it is also a human. In 1952 Arthur Samuel wrote the first computer learning program which was a game of checkers. In 1957 Frank Rosenblatt designed the ***first neural network*** for computers. In 1967, the ***nearest neighbor algorithm*** was written to formulate the travelling salesperson problem. In 1981, Gerald Dejong introduced the concept of ***Explanation Based Learning*** in which computer discards unimportant data. It was in the 1990's that machine learning had a data driven approach and not a knowledge driven approach. Large computer programs were written to learn from the results or draw conclusions from the data. These days the large code is replaced by scripts or test plans. The scripting language has to be known instead of the code or at times just the desired keyword, thus saving time and effort. In 2006 Geoffrey Hinton coined the term ***deep learning*** to explain algorithms that let computers distinguish between objects and texts in videos and images. In 2011 ***Google Brain*** was developed and its deep neural network can learn to discover and categorize objects much as a cat does. In 2014 Facebook developed ***Deepface***, a software algorithm to recognize individuals in photos just as humans do.

Not only all these, in 2015 over 3000 AI and robotic researchers endorsed by Stephan Hawking, Elon Musk and Steve Wozniak signed an open letter of warning of autonomous weapons which select and engage targets without human intervention. Then in March 2016,

Google's Deepmind technology AlphaGo defeated World champion Lee Sedol in the game Go which as mentioned above is one of the most challenging games in world games.

Given the growth of AI in the recent years, public relations are also going to be impacted to a great extent. The way we communicate today is different than the means of communication used ten years ago. Similarly, advancements in technology will impact the mode of communication in the next 10 years as well. For example, a telephone was replaced by a mobile phone which has now advanced to touch screen devices. With automation taking over human labor, focus should not be on reduction of jobs, but on the improvement of performance. How can AI not scare off employees and startups for loss of employment, and yet improve the company's performance is what will be investigated in this paper.

AI is now a blend of a number of technologies. Machine learning and Big Data are claimed to remake the world in upcoming years. By 2020, the market for machine learning will reach $40 billion and 60 % of those software's will run on the platform of four big companies- Amazon, Google IBM and Microsoft. This paper puts forth the advancements of each firm in the field of AI, what technology can the smaller firms adapt depending on their needs, and the use cases suggested.

The paper highlights the ***description, key features, benefits, use cases, and price*** of each service offered by the four firms. The smaller companies can then choose the respective service and firm according to their needs after comparison. According to the research till date, we only know that the four companies have come together for research and will be coming up with a product or service that would be most essential for people to use. However, the product or service is not disclosed. The paper puts up an idea as to what advancements in the field of

Artificial Intelligence can be carried out by IBM, Amazon Google, Microsoft and Facebook since they have joined hands together

Let us have a brief description of each device which involved a short description, features and benefits.

Criteria	Amazon	Google	IBM Watson API	IBM Watson Services	Microsoft
Speech Recognition	Amazon Lex				
Image Analysis	Amazon Rekognition	Cloud Vision		Visual Recognition Service	Computer vision API
					Custom vision API
Text to live speech	Amazon Polly(only text to audio)			Text to Speech	Custom Speech Service
					Bing Speech API
Audio to text		Cloud Speech API		Speech to Text	
Training and evaluation	Amazon machine learning(visualization)	Google Cloud Machine learning	Watson Analytics	Trade off analytics	
Job search		Cloud Jobs API			
Text Analysis		Natural Language API	Watson Explorer	Natural Language Classifier	Web language model API
				Retrieve and rank	Text Analytics API
				Discovery	Bing Spell Check
				Alchemy Data News	
Translate		Cloud Translation API		Document Conversion	Translator text API
				Language translator	Translator Speech API
Relationship between data			IBM Watson Knowledge Studio		
Language Processing				Alchemy Language	Language Understanding Intelligent Service
				Conversation Service	Linguistic Analysis API
				Dialog(Expired)	

Table 1. Summary of Technologies

DEVICES

A) Speech Recognition:

1) Amazon Lex:

Amazon Lex is used to build conversational interfaces in any application using text and voice. Automatic Speech Recognition and Natural Language Understanding for understanding the intent of the text are its special features. The above two mentioned concepts are the most difficult concepts of machine learning which are solved by Amazon Lex by integrating it with Amazon Alexa. Conversational interface enables new categories of products possible.

Key Features:

1) High quality speech recognition and natural language understanding

2) Multi turn conversations by identifying intent and prompting user for options

3) Utility prompts: Confirmation and error prompts

4) Integration with AWS Lambda for data retrieval , business logic execution

5) Connect to enterprise systems which includes applications as well as databases.

6) One click deployment to multiple platforms including intuitive hub for Facebook messenger

7) Lifecycle management which includes visioning to intents slot types, bots created. Multiple aliases are created and assigned to each model.

Benefits:

1) Highly scalable

2) No upfront costs

3) Pay for what you use

4) Built in integration with AWS

5) Cost effective

6) Easy to use

Use cases:

1) Informational bots:

Chatbots can be built for everyday consumer requests, such as accessing the latest news, updates, game score and weather. After building Amazon Lex bot it can be deployed on mobile devices, chat services and IOT devise with support for rich message formatting.

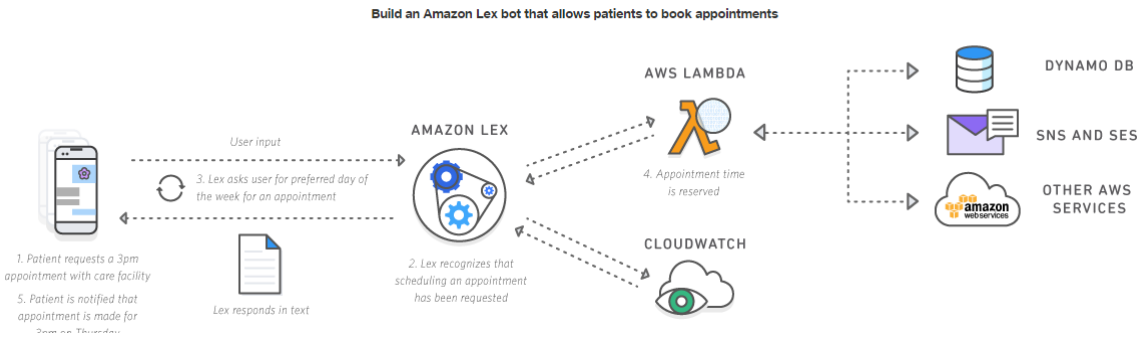

Fig 1. Amazon Lex Bot

Deploying firm:

OhioHealth: Amazon Lex helps OhioHealth to offer a better experience to the patients. While everything at OhioHealth is about providing the right care to patients, Amazon Lex with its next generation technology and the applications developed by OhioHealth with the help of this technology will help provide an improved customer experience.

2) Application bots:

High quality speech recognition and natural language understanding capabilities make it possible to build powerful interfaces for mobile applications. Voice or text chat interface can be

added to help customers access their bank account, book tickets, order food or call a cab.

Fig 2. Amazon Lex Chatbot

Deploying firm:

Capital One: Seamless integration of Amazon Lex w1ith AWS Lambda and Amazon Dynamo DB, its scalability is worth noting. The potential to speed time to market for a new generation of voice and text interactions have impressed Capital One.

3) Enterprise productivity bots:

Amazon Lex can be used to build enterprise chatbots that streamline common work activities and improve organizational efficiencies. Performance of the organization and the data can be checked from chatbots within minutes. A variety of productivity tools can be built through AWS Lambda functions.

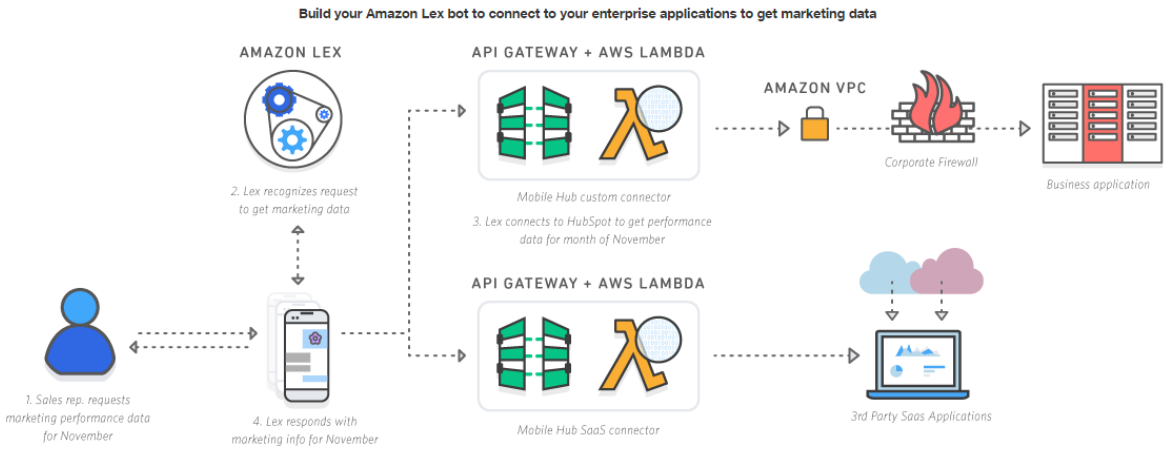

Fig 3. Amazon Lex Bot Enterprise Applications

Deploying firm:

Hubspot: Hubspot's Growthbot provides access to relevant data and services by using a conversational interface. Growthbot helps marketers to create content, research competitors and monitor their analytics. Amazon Lex has helped to add sophisticated natural language processing capabilities that help GrowthBot provide a more intuitive UI for users. There is no need to code the algorithm and advanced AI and machine learning.

4) Internet of things

highly interactive and conversational user experiences for connected devices in the rapidly growing segment of Internet of Things. This creates opportunities for new categories ranging

from cars and devices to wearables and appliances.

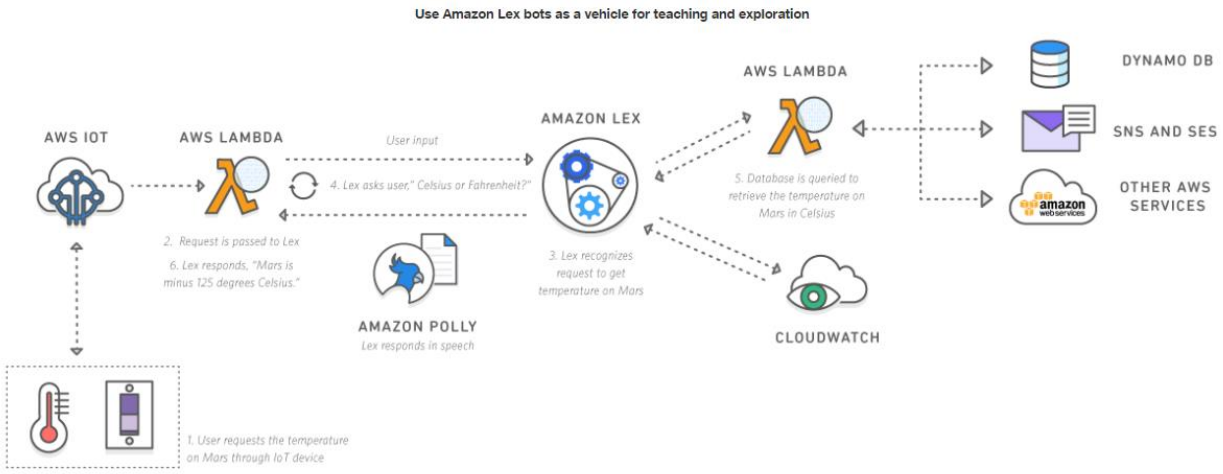

Fig 4. Amazon Lex teaching and exploration

Deploying firm:

1) NASA: NASA reaches out to students in schools, community organizations and public events. NASA uses Amazon Lex which helps it to navigate a star robotic ambassador “Rov-E”, a close replica of real NASA Mars rovers. NASA can now easily navigate Rover-E thanks to Amazon Lex. It can answer student’s questions about Mars in an engaging way. Integration allows Rov-E to connect and scale with various data sources to retrieve Mars data exploration information

Price:

Rate of speech request: 0.004$

Rate of text: 0.00075$

Consider you have 5000 speech requests and 1000 text requests. The total cost would be

(5000*0.004)+(1000*0.00075)=20.75$

Note that the speech and text requests are the request the client places to the device

There are no other products for speech recognition by IBM Watson, API, IBM

Watson Services and Google till now. So no comparison is made yet

B) Image Analysis:

1) Amazon Rekognition:

Amazon Rekognition makes it easy to add image analysis in your application. Detection of objects, scenes and faces becomes easy. Facial features are represented as vectors. These vectors represent prominent facial landmark points. Egs. Left eye, right eye, nose, left corner of the mouth and right corner of the mouth. Deep learning helps get a high level understanding of image content quickly and at scale. Engagement of the viewers can be identified based on visual attributes and emotions. Organization of images in a photo app, recognition of people present in the image can all be recognized based on Amazon Rekognition.

Key features:

1) Object and scene detection to identify vehicles pets, furniture, other objects and give a confidence score

2) Facial analysis to analyze facial expressions such as a smiling or straight face, detect if eyes are open or close and a smiling or a straight face.

3) Rekognition lets you compare faces of two or more different people. The user is verified against a reference and a score is given

4) Facial Recognition can be done which has a group of similar looking faces. When a single photo is specified, Rekognition searches visually similar images from the collection.

Benefits:

1) Easy integration of image recognition in the app

2) AI at core

3) Integration with Amazon AWS

Use cases:

1) Searchable image library:

Amazon Rekognition makes it possible to search for images so you can discover objects and the scenes that appear within them. AWS Lambda adds newly detected image labels directly into an elastic search index when an image is uploaded to S3.

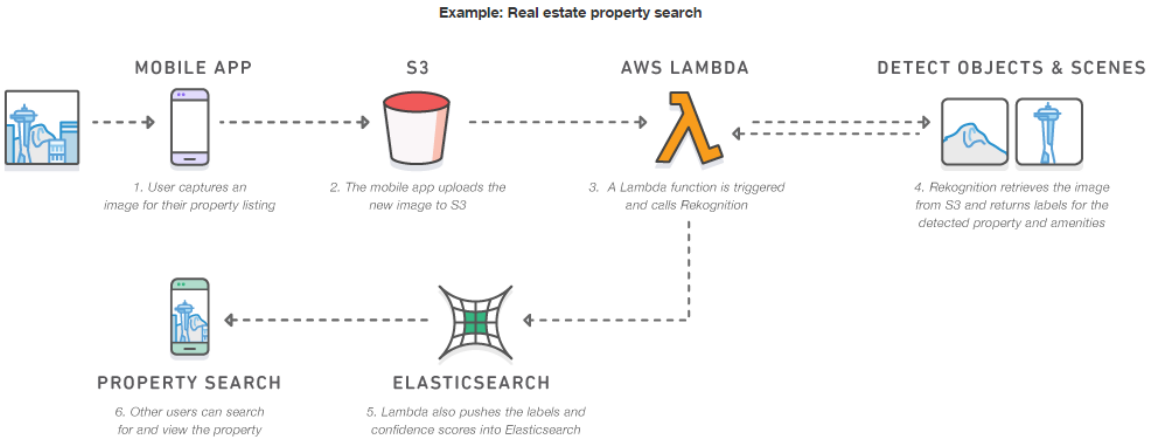

Fig 5. Real Estate Property Search

Deploying firm:

C-SPAN: C-SPAN is a public service that provides proceedings of the US house of representatives and the US senate where the public policy is discussed and decided. Rekognition helps C-SPAN with twice the indexing of content.

Go Girl: Happy Snap, from Go Girl apps is a mobile application designed for kids. It states that after testing various product, Amazon Rekognition is the most reliable.

2) Face based User verification

User identities are confirmed by comparing live image with a reference image.

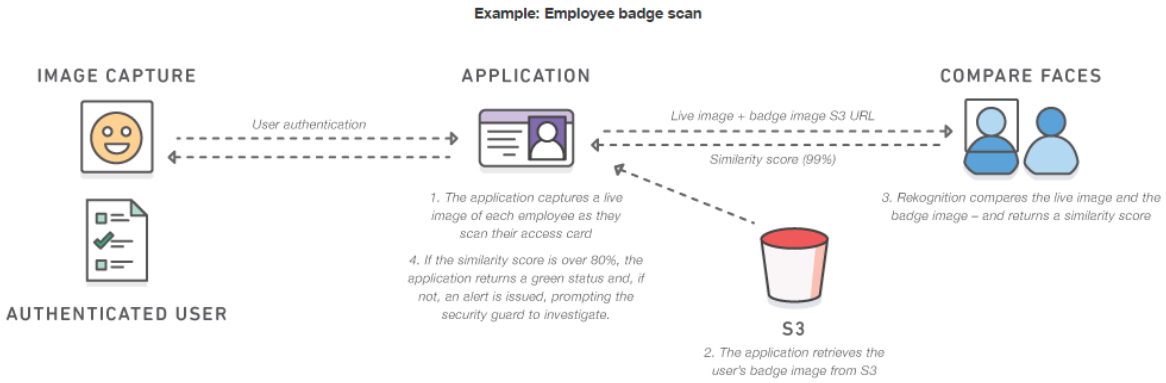

Fig 6. Employee Badge scan

4) Sentiment analysis

Amazon Rekognition can detect emotions such as happy, sad or surprised from facial images. Rekognition can analyze live images and send emotion attributes to RedShift for periodic reporting

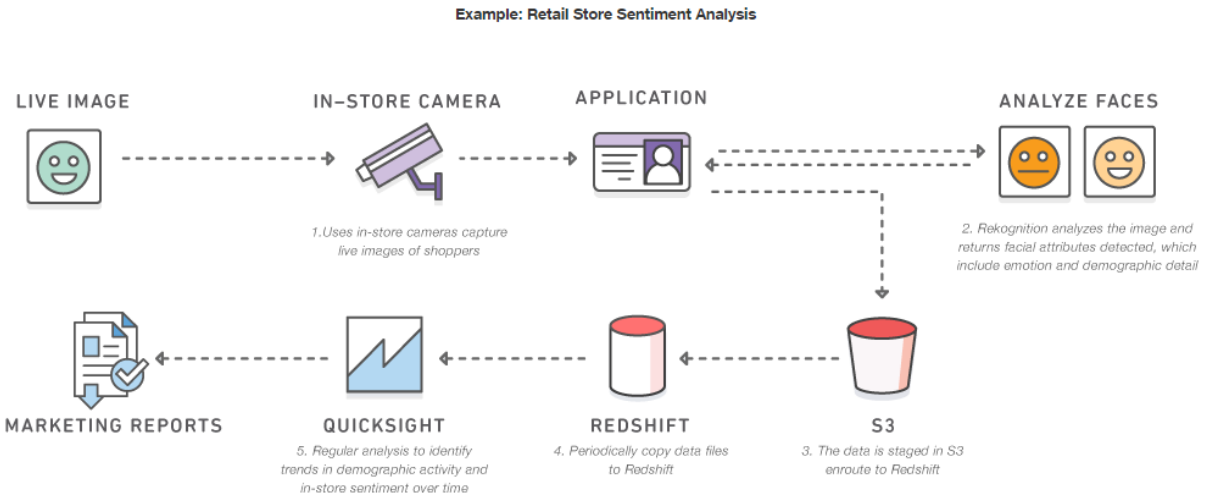

Fig 7. Retail store sentiment analysis

4) Facial Recognition:

Amazon Rekognition makes it easier to search for a collection of images for similar faces by storing face metadata using the IndexFaces API function. You can also use SearchFaces to return high confidence matches. A face collection is an index of faces that the user can own and manage.

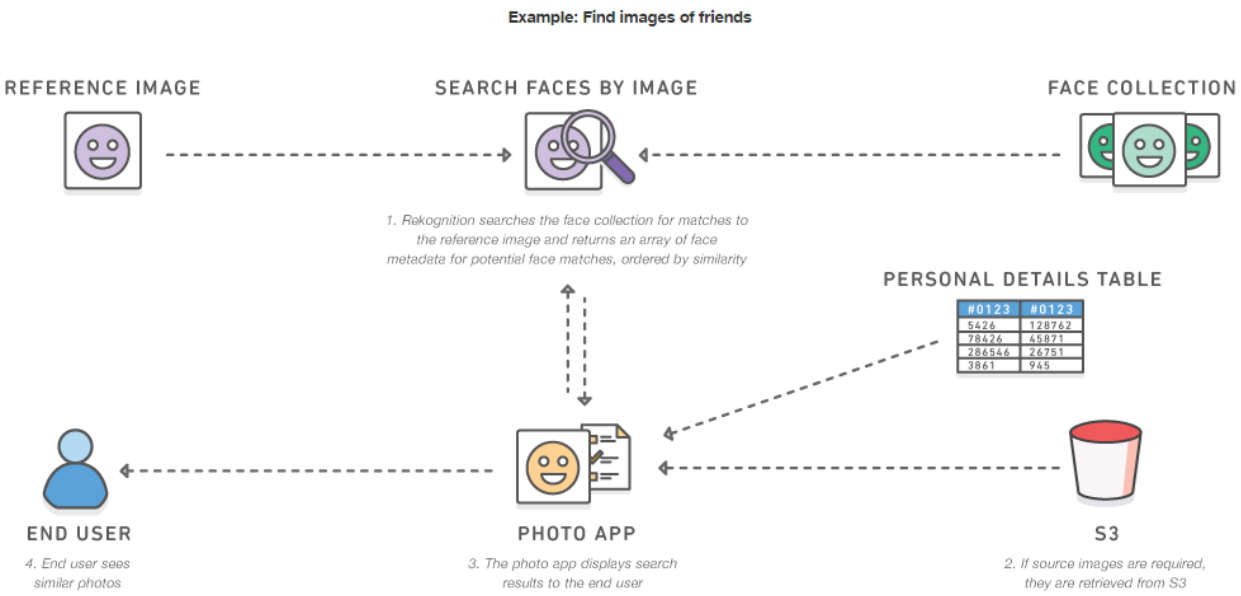

Fig 8. Find images of friends

Deploying firm:

SmugMug: Amazon helps SmugMug identify the content in customers photos, unlocking a host of features, thus allowing the visitors to have more time to celebrate, rather than sort and select photos.

5) Celebrity Recognition:

The RecognizeCelebrities API uses neural network based models to search photo libraries and automatically identify famous noteworthy individuals prominent in their field with high scale

and high accuracy. The celebrity's name, id and image id are integrated in Amazon elastic search index to make the images searchable for the celebrities.

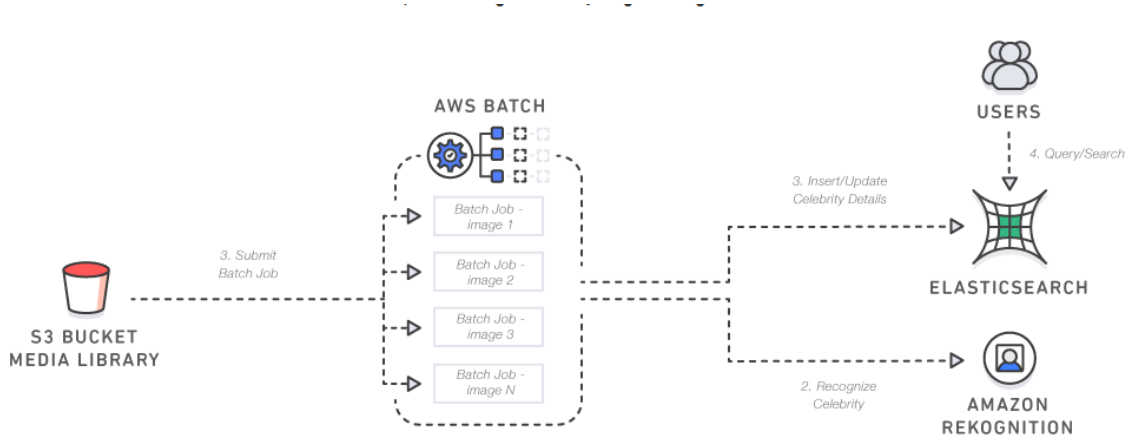

Fig 9. Celebrity Images

Price:

Image Analysis Tiers	**Price per 1000 images**
First 1 million images processed per month	$1.00
Next 9 million images processed per month	$0.80
Next 90 million images processed per month	$0.60
Next 100 million images processed per month	$0.40

2) Cloud Vision API

Cloud vision API enables content analysis of an image by encapsulating machine learning models in an easy to use REST API. It classifies images in a number of categories. It also reads contents of text within an image. Metadata can be built on the image catalog and offensive content can be blocked. Images can be integrated with the Google Cloud Storage and can also be uploaded for analysis.

Key features:

1) Insights of the image helps to detect a broad set of objects ranging from animals, transportations to thousands of other categories.

2) Detection of inappropriate content ranging from adult to violent content

3) Vision API uses the power of Google Image Search to find celebrities, logos and news events

4) Text detection along with automatic language identification is possible.

Benefits:

1) Explicit content detection

2) Logo detection

3) Landmark detection

4) Optical Character Recognition

5) Face detection

Use cases*:*

One of the most important use case of the cloud vision API that can be highlighted includes image detection. The various use cases that can be mentioned include:

5) **Tesla's** auto pilot has the driving feature of speed adjusting, lane switching and automatic braking. Yet it dashed into 18-year-old crossing the path. However, auto pilot did take the injured to the hospital. Thus proving that auto driven vehicles would be safer than human driven vehicles. (Parloff Roger, September 28, 2016)

This simple single activity sample shows how you can make a call to cloud vision API with an image picked from your device's gallery.

1) Image detection using ios device photos

The swift and objective C version of this app use cloud vision API for running the label and face detection on an image from the device's photo library. The resulting labels and face metadata from the API response are displayed in the UI

2) Get insight of thousands of images:

If there are thousands of images on a website(the website is a repository of the images), however no details of the image are given, this is where the Cloud vision API comes into play. The broad, unordered and disparate set of images is a great set to demonstrate vision API capabilities.

3) Processing invoices and receipts

A JSON code is written and a response is received for the request. The text pat is detected and processed. For any OCR application, the results depend on the quality of the scan. The edges of the image you take have to straight, the image should have a good contrast with little noise. Apart from text detection of the image, entity detection, layout analysis can also be performed

Price:

Feature	1 - 1000 units/month	1001- 1,000,000 units/month	1,000,001 to 5,000,000 units/month	5,000,001 - 20,000,000 units/month
Label Detection	Free	$1.50	$1.50	$1.00
OCR	Free	$1.50	$1.50	$0.60
Explicit Content Detection	Free		Now free with Label Detection*	
Facial Detection	Free	$1.50	$1.50	$0.60
Landmark Detection	Free	$1.50	$1.50	$0.60
Logo Detection	Free	$1.50	$1.50	$0.60
Image Properties	Free	$1.50	$1.50	$0.60
Web Detection	Free	$3.50	$3.50	Contact Google for more information
Document Text Detection	Free	$3.50	$3.50	Contact Google for more information

3) Visual Recognition Service:

Visual Recognition uses deep learning algorithms to analyze images for scenes, objects, face and other content. This includes keywords that provide information about the content.

Built in classes provide accurate results without training. Custom collections of the image can be created and upload an image to search the collection for similar images.

Key features:

1) General classification by generating class keywords to describe an image. Using own images or URL`s from publicly accessible web pages for analysis.

2) Detect human faces in the image and provide an age range and a gender

3) Create custom, unique visual classifiers. Recognize visual concepts that are not available with general classification.

4) Upload and search through image collections to find visually similar images

Benefits:

1) Online shopping

2) Insurance claims

3) Resource management

Use cases:

1) Manufacturing: Make use of image from a manufacturing setting to ensure products are being positioned currently on an assembly line

2) Visual auditing: Look for visual compliance or deterioration in a fleet of trucks, planes or windmills out in the field, train custom classifiers to understand what defects look like.
3) Insurance: Rapidly process claims by using images to classify claims into different categories.
4) Social listening: Use images from your logo or product line to track buzz about your company on social media.
5) Social commerce: Use an image of a plated dish to find which restaurant serves it and find reviews. Use a travel photo to find locations based on similar experiences.
6) Retail: Find stores with your kind of an outfit by taking a photo of your favorite outfit. Use travel image to find similar outfits in your area
7) Education: Image based applications to educate about taxonomies.

Price:

Image Analysis Tiers	**Price per 1000 images**
Image tagging events pay per use	$0.002/event
Face detection events pay per use	$0.004/event
Training events pay per use	$0.10/event
Custom tagging events pay per use	$0.004/event

4) Computer Vision API

Computer Vision SPI is used to extract rich information from images to categorize and process visual data and machine assisted moderation of images to help curate services.

Key features:

1) Analyze an image: This feature returns information about the visual content found in an image. Automated restriction of adult content is possible. Image types and colour schemes can be identified
2) Read text in images: OCR detects text in an image and extract recognized words into machine readable character scheme
3) Read handwritten text from images: Detection of hand written text from notes, letters, essays, white boards
4) Recognize celebrities and landmarks: Celebrity and landmark models are examples of domain specific models. Azure celebrity recognition model recognizes around 200 k celebrities and landmark recognition recognizes around 900 natural and man made landmarks
5) Analyze video in real time: Use any computer vision API's with your video files by extracting frames of the video from your device and sending those frames to API calls of your choice.
6) Generate a thumbnail: Generate high quality storage efficient thumb nail based on an input image

Benefits:

1) Analyze image
2) Use domain specific model

3) Intelligently generate a thumb nail

4) Detect and extract printed text from an image

Use cases:

Putting old and new people in a different category. If a person was previously tagged, the API will start to identify them in images. Besides that, it will also detect emotions

Price:

Tier	**Features**	**Price**
Computer vision API- Free		5000 transactions free per month
$1- upto 10 transactions per second for these features	Tag, face, Get Thumbnail color and Image Type	0-1 M transactions-$1/1000 transactions 1-5 M transactions- 0.80/1000 transactions 5M+ transactions- 0.65$/1000 transactions
	OCR(Printed), Adult, Celebrity and Landmark	0-1 M transactions- $1.5/1000 transactions 1-5 M transactions- $0.80/1000 transactions 5M+ transactions- $0.65/1000 transactions

	Describe and OCR(handwriting)	$2.5/1000 transactions

5) Custom Vision API

Custom vision is a tool for building custom image classifiers. It makes it easier and fast to deploy and improve an image classifier. REST API and web interface is present to upload the images and to train them.

Key features:

1) Upload images: Upload own labelled images to teach Custom Vision Service the concepts you want to learn
2) Train: Teach Custom Vision Services the concepts you want to learn
3) Evaluate: Use simple REST API to tag images with your new custom computer vision model
4) Active learning: Images evaluated through your custom vision model become a part of the loop to improve your classifier.

Benefits:

1) Build a classifier
2) Improve a classifier
3) Test the model
4) Use prediction API
5) Export the model to mobile

Use cases:

Unlabeled images can be tagged and trained to understand the concepts you care about. Use REST API calls to quickly tag images with your custom model

Price:

Contact the company for pricing details and its use has no additional cost

If you compare the three devices we realise the benefits of each are varied depending on the use case. Amazon's recognition does recognize celebrities/ Google has not come up with a product that does so but Microsoft has. If you consider the price, Google's Visualization is worth the extra money for the additional features it is offering.

C) Text to live Speech:

1) Amazon Polly (Text to audio)

Amazon Polly helps to convert text to live speech. Polly includes 47 life like voices and supports 27 different languages. Send the text you want to convert to speech and it will be saved in the mp4 format.

Key features:

1) Easy to use API for quick integration of speech in applications

2) Wide selection of voices and languages

3) Stream all kinds of video through application to users in real time and optimize audio quality and bandwidth

4)Use Speech Synthesis Markup Language(SSML) to adjust speech rate, pitch and loudness to attract attention of the audience.

5) AWS SDK and AWS mobile SDK support HTTP API.

6) Speech synthesis via API console and command line

7) Custom lexicons help to modify the pronunciations of company names and acronyms

Benefits:

1) Natural sounding voice

2) Fast response

3) Store and redistribute speech

4) Low cost

5) Easy integration

Use cases:

1) Content creation

Amazon Polly helps to add speech to video, presentation or online training course. Amazon Polly can generate speech in around 24 languages making it easier to add voice to applications with a global audience . RSS Feed, news or email can be read and synthesized speech can be stored in the form of audio files.

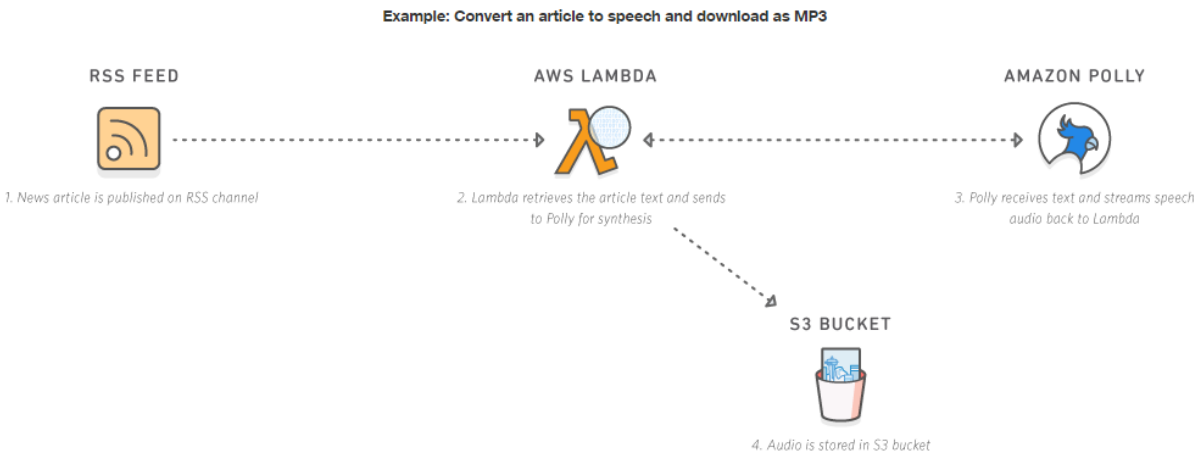

Fig 10. Convert an article to speech and download as mp3

Deploying firm:

GoAnimate: GoAnimate users immediately give voice to the characters that are animated by GoAnimate. This voice is given by Amazon Polly. Amazon Polly is extremely helpful when live voice over is either resource or time prohibitive such as developing a video in a number of languages or if the approval process has to be fastened. The Speech re-enforces GoAnimate's ease of use and affords both efficiency and speed to market to the customers.

2)Education/ E-learning

Amazon Polly enables developers to provide an enhanced visual experience to the

customers such as speech synchronized facial animation or karaoke style word highlighting. It is easier to request an additional stream of metadata with information that contains when the particular words, sentences and sounds are being pronounced with Amazon Poll. Use of this metadata stream alongside the synthesized speech audio stream helps customers to animate avatars and highlight the text as it is spoken in their app.

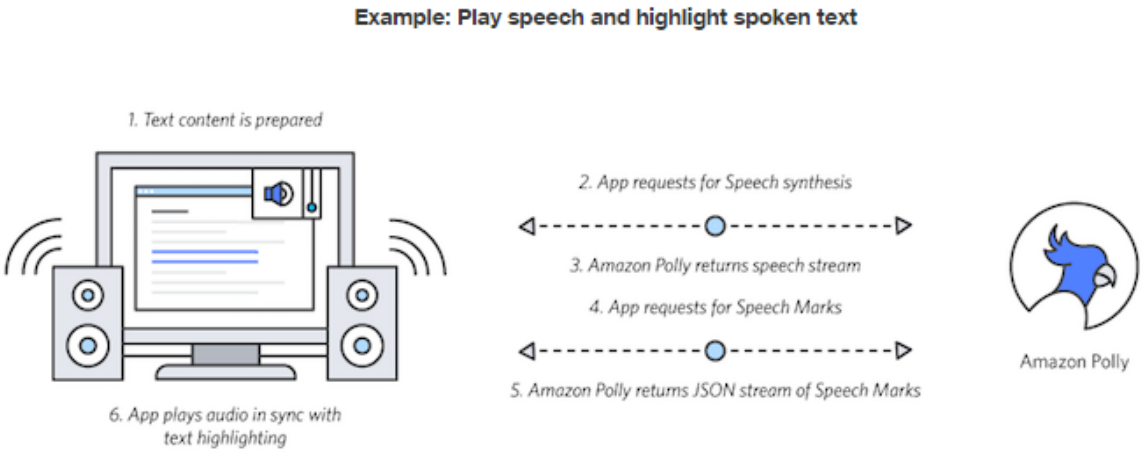

Fig 11. Play speech and highlight spoken text

Deploying firm:

WizKids: While WizKids is a cloud learning platform, Amazon Polly enhances their platform by integrating high quality text to speech voices with AppWriter products. Stephan Pal, the COO of WizKids feels that it is absolutely essential for the users to see real-time highlighting of text when it is being read aloud.

3) Mobile and Desktop Apps:

Adding voice to mobile apps and games is easier because of Amazon Polly. Standard speech responses can be stored and also dynamic, real time responses such as in game

character dialog, leaderboard rankings and game invitations can be enabled.

Fig 12. Mobile desktop apps

Deploying firm:

The Washington Post: Amazon Polly is cost effective for the Text to Speech Solution it offers says Joseph Prince, the senior product manager of Washington Post.

Price:

Free tier	Free for first 12 months(5 million characters/month)
Policy pricing	$4.00/1 million characters

2) Text to Speech

Coverts text to audio in a variety of natural sounding languages. Allows seamless voice interaction catering to the audience. It is used to automate call center interactions

Key features:

1) HTTP and web socket interface supports speech synthesis

2) OGG Opus audio format with Opus Codec waveform audio file format

3) Supports English, German, French, Italian, Japanese, Spanish

4) Accepts plan text/ text in SSML

5) Uses expressiveness to indicate good news, apology or uncertainty

6) Voice transformation controls voice pitch, rate timbre. It has two voice transformations: young and soft

7) Customization helps in pronunciation and phonetic representation

Benefits:

1) Language translation

2) Consistent voice across all compatible platforms

3) API used

Use cases:

1) Voice driven and screenless interfaces and interfaces for the disabled
2) Home automation solutions
3) Assistance tools for the visually impaired
4) Reading texts and emails out loud to the drivers
5) Video script narration and voice over

Price:

Standard service	First 1 million characters free per month. Additional characters are 0.02/1000
Premium	Lets talk

3) Custom Speech service

Custom Speech service is used to overcome speech recognition barriers suc as speaking style, vocabulary and background noise.

Key features:

1) Create custom language models: Customize language models of speech recognizer by tailoring it to the vocabulary of the application and speaking style of the users
2) Create custom acoustic models: Customize acoustic model of the speech recognizer to better match the environment and user population of the application
3) Deploy custom models: Deploy models to create speech recognition customized to your application
4) Access end point from any device: Send request to custom end point using RESTful API or cognitive services speech client library

Benefits:

1) Tune speech recognition to specific topics, actions and speeches
2) Can work with accents of non-native speakers
3) It gives the number of users an application is expected to accommodate

Use cases:

Recognizing speech of a specific set of users optionally using a custom vocabulary. Advanced options are for companies who train their own speech models

Pricing:

	Free	**S1**
Model Deployments	1 Model	$40.01/month/model

Model Adaption	3 hrs/month	Unlimited
Accuracy tests	2 hrs/month	2 hrs free then 1.401/hr
Scale out	NA	$6.452/unit/day where each unit allows you to send 5 concurrent requests
No trace	NA	$30.01/model/month
Request Pricing	2 hrs/month	2 hours free then $1.401/hr

4) Bing Speech API

Converts audio to text, understands intent and converts text back to speech for natural responsiveness.

Key features:

1) Speech Recognition: Convert spoken audio to text. The API can be directed to turn on and recognize audio coming from microphone in real time. As audio is being set to the server, partial recognition results are also returned. Smart apps that are voice triggered can be built
2) Text to speech: Convert text to spoken audio. When applications need to talk back to the users, this API can be used to convert text generated by the app into audio that can be played back to the user.

Benefits:

1) Voice activated commands
2) Smart apps speak back

Use cases:

Bing speech enables text to speech for 20 different languages with multiple voice options. The voices are significantly more natural and relatable than other computerized voices. It is useful for hands free mobile applications. It is often combined with other services

Price:

Tier	**Features**	**Unit**	**Price**
Bing speech API		Tansactions	5000 transactions free per month
Bing Speech to text API	Utterances upto 15 per seconds long	Transactions	$4/1000 transactions
Bing text to speech API		Transactions	$4/1000 transactions

While I have to compare IBM Watson Text to Speech and Amazon Polly, reviews state that IBM Watson has not set its standards too high and can be surpassed easily. Amazon has offered a wide range of services via Amazon Polly

D) Audio to text

1) Cloud speech API

Cloud speech API converts audio to text with the help of neural networks. It recognizes over 80 languages to support a global user base. Audio uploaded in the request can be recognized and can be integrated on Google Cloud Storage, by using the same technology as Google uses to power its products

Key features:

1) Google has the most advanced deep learning neural network algorithm for speech recognition with unparalleled accuracy powered by **Machine Learning.** This accuracy improves over time.

2) It supports **110 languages** and even filters inappropriate content in the text results

3) Returns text in real time. The results are immediately available while speaking. Alternatively, Speech API returns recognized text stored in the file

4) No need of advanced signal processing or noise cancellation as the API is accurate even in a noisy environment

5) Speech Recognition can be tailored to context by providing separate set of word hints with each API. Thus, it has **context aware recognition**

Benefits:

1) Automatic speech recognition

2) Global vocabulary

3) Streaming recognition

4) word hints

5) real time/ prerecorded audio support

6) Noise robustness

7) Integrated API

8) Inappropriate content filtering

Use case:

1) Transcribe text of users dictating to an application's microphone
2) Enable command and control through voice
3) Transcribe audio files
4) Recognize audio uploaded in the request and integrate with your audio storage on Google cloud storage.

Price:

Feature	**0-60 minutes**	**61+- 1million minutes**
Speech Recognition	Free	0.006$/15 seconds

Each request is rounded off to nearest increment of 15 seconds. Each separate requests containing 7 seconds audio would be rounded to 45 seconds

2) Speech to text

Audio voice is converted to written text. Transcribe calls at a contact center to know what is being discussed, when to escalate calls and understand the content being discussed. Accuracy of the language can be improved and emphasized. E.g Product names, sensitive subjects or names of individuals.

Key Features

Input features:

1) Wide range of languages including Brazilian, Portuguese, French Chinese, Japanese, Spanish, UK English, US English.

2) Broadband and narrowband models for most languages

3) Audio formats ranging from PCM, WAV. Ogg, mu-law audio data or basic audio

4)100 mb of audio to the service as a one shot delivery.

Output features:

1) Recognizes different speakers from narrowband audio in US English, Spanish or Japanese. Each speaker's contribution is sent out to multi participant conversation.

2) Focuses on individual words which can be more important than the entire conversation.

3) Gives alternative words that can be used similar to the word used previously and timestamps at the start and end of each word.

4) Maximum alternatives and interim results providing different possible hypotheses. Final results only having the maximum confidence are included

5) Censors profanity from English US transcripts.

6) Smart formatting converts dates, time's numbers and phone numbers into conventional forms.

Benefits:

1) Transcribing conversations

2) Direct interaction with customers

3) Communication in multiple languages.

Price:

Standard service	First 1 million characters: free Additional characters: $0.02/thousand
Premium service	For security and sensitive data requirements. Contact representative for the pricing.
Telephony add-on	First thousand minutes per month are free. Additional $0.02 in addition to using the standard service

On comparing Google and IBM's speech recognition services, we realize that Google recognizes over 40 languages in the Beta version while IBM Speech to text recognizes only nine. Google also has a better structured documentation and easily deployable SDK's. IBM Watson Engine is faster than Google

E) Training and Evaluation

1) Amazon machine learning

Amazon helps to use machine learning technology. Visualization tools and wizards are used to guide in machine learning. There is no need to go through complex algorithms or to implement code. Amazon Machine Learning helps to predict training and evaluation, find patterns in data and deploy machine learning models.

Key features:

1) Integration of AWS for ease of data access and working with the data. Databases in CSV, MySQL or Redshift can be used to store data.

2) Data visualization and exploration provides interactive charts to spot the missing attributes easily

3) Model evaluation and interpretation tools to predict and deliver optimum results

4) Creating, automating and modelling new API`s. Previous models can be examined for tracking and repeatability.

5) Machine learning algorithms for binary, numeric and categorical prediction.

Examples of each are as below:

Binary: If website is spam or not

Regression: No of days before next interaction

Categorical: Where to route customer request

6) Data transformations to improve the quality of data during training time

7) Predictions from machine learning models are divided as batch production and real time

production. This gives high throughput, low latency and is real time

8) Fully managed infrastructure and workflows managed b Amazon. A number of models can be created and the volume and throughput can be scaled.

9) It is easy, cost effective and scalable models can be built.

Benefits:

1) Low cost

2) Efficient

3) Scalable

4) Easy Machine Learning modelling

Use Cases:

While working with machine learning, the two things of utmost importance include:

1) Making sure what to classify/predict

2) Collecting as much data as possible about the context without making too many assumptions

A few use cases from low level to high level include:

1) Predicting whether a particular user will become a paid customer based on his activities in the first week/month or day

2) Detecting spammers, fake users or bots depending on website activity records

3) Classification of song genre based on signal level features

4) Character recognition from a plain image(OCR)

5) Detecting gyroscope signals

Thus, you can say prediction of completely unknown stuff.

Price:

Batch predictions	0.10 per 1000 predictions
Real time predictions	$0.0001 per prediction rounded to the nearest penny
Data analysis and model building fees	$0.42 per hr

2) Google cloud machine learning

Google has done an extensive research on Machine Learning in Artificial Intelligence. Google has a cloud machine learning platform to provide machine learning services with pre-trained models and a service to even generate your own model.

Key Features:

1) Prediction at scale using online and batch production and also load balancing.

2) Easy model building by tuning the model for better learning faster. Tensor flow and model graphs and train models to analyze their quality.

3) Fully managed service by proper provisioning of training and monitoring handled by serverless infrastructure for individual handling of models instead of clusters.

4) Deep learning techniques to build and use models on any type of data and any variety of scenarios.

Benefits:

1) Integrated

2) Scalable

3) Portable

4) Hypertune

5) Managed services

6) Scalable services

Use cases:

1) Get new customers and retain existing customers
2) Analyze tweet logs and messages
3) Insurance pricing optimization.

    7-10% of AXA's customers cause a car accident every year, most of them small but 1% large. The AXA'As R & D team in Japan has been researching the use of machine learning to predict the large loss case during the insurance period. After developing a machine learning model with Tensorflow, the team achieved 78% accuracy in its results, thus optimizing its insurance cost.

4) Professional services and educational programs

    Price:

Training clusters	**Price**	**Price**
Basic Tier	$0.49/hour	$0.54/hour
Standard Tier	$4.90/hour	$5.40/hour
Premium Tier	$36.75/hour	$40.50/hour
Custom Cluster Configuration	$0.49/hour per ML training unit	$0.54/hour per ML training unit
Basic GPU Tier	$1.47/hour	$1.62/hour
Prediction Requests		

Up to 100M per Month	$0.10 / 1K +$0.40/Node Hour	$0.11 / 1K +$0.44/Node Hour
Requests over 100M per month	$0.05 / 1K +$0.40/Node Hour	$0.05 / 1K +$0.44/Node Hour

3) Watson Analytics

Analytics is a smart data analysis and visualization service to quickly discover patterns and meaning in the data by your own self. With guided data discovery, automated predictive analytics and cognitive capabilities you can quickly spot a trend and report data in a dashboard.

Key features:

1) Add or connect to data and instantly get a list of starting points on your desktop or ipad

2) Discover new patterns or trends as well as factors that are most likely to influence your business outcomes and math behind the findings.

3) Add visualizations that you want to show your audience

4) Analyze the data that comes in various forms. You can explore, predict or assemble the data.

Benefits:

1) Natural language dialogue

2) Automatic predictive analytics

3) One click analysis

4) Simplified analysis

5) Self-service dashboard

6) Accessible advanced analytics

Use cases:

1) Analyze supplier delays
2) Analyze sales effectiveness
3) Demystifying Incentive Plan effectiveness
4) Exploring banking loss event data
5) Analyzing customer campaigns
6) Using customer behavior data to improve detention
7) Watson analytics use case for marketing

Price:

PRICE	FEATURES
$0 USD for a month	Upload spreadsheets, get visualizations, discover insights, build dashboards
Starting $30 USD per month	All features of free, more storage, more data including databases and twitter.
Starting $80 USD per month	Multi user tenant to collaborate, more storage, more data

From the various bogs read and comparisons made, I have come to the conclusion that the scope of Amazon and Google when it comes to machine leaning is a lot narrower or you

may say focused as compared to IBM. Thus, depending on your requirements, if you want a wide range of features, I would recommend IBM

Trade off analytics

Watson trade off analytics helps people to make decisions while balancing multiple objectives. With analytics, a list of endless options can be avoided and right option can be determined by considering multiple objectives

F) Job Search

1) Google Cloud jobs API

Google Cloud jobs can be used for intuitive job search. To identify what the job seekers are looking for and send targeted recommendations to improve opportunities and preferences. The API uses machine learning to understand job titles and the skills related to them as well as the location and content according to the preference of the job seeker

Key features:

1) Synonym and acronym expansion. For example if an engineer is searching the term BA , Google Cloud jobs API will interpret it as Business Analyst according to the options selected by the user.

2) Job enrichment with relevant additional information such as street address, employment type and benefits

3) Advanced location mapping to filter jobs by distance and commute times

4) Jobs API understands seniority of positions and returns only relevant results.

5) When a job search returns a low set of results, search is expanded to the elected location, similar job roles etc

6) Users can mark their choice of jobs so the results are returned accordingly.

Benefits:

1) Relevant results

2) Dynamic job discovery

3) Simple integration

Use case:

Cloud job discovery is a part of Google for Jobs and helps people find jobs easily. Thus enabling the betterment of the entire recruiting system- jb boards, applicant tracking systems and staffing agencies to improve job site engagement and candidate conversion.

Working:

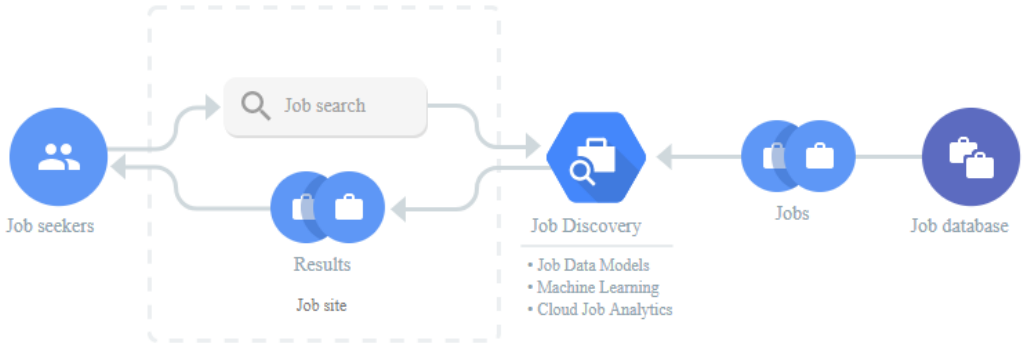

Fig 13. Job discovery working

Deploying firm:

Career Builder:

We are excited about our collaboration with the Google team combining our shared expertise to empower employment. Cloud job discovery enables us to provide an engaging experience to our job seekers, while out sourcing a time intensive product area to experts in the space. The API was simple to plug in and continuously improve performance over time with built in machine learning capabilities. This allows our

product, technology and data science experts to focus on solving deeper and more critical challenges for job seekers to employers (Leigh Margeret Stull, VP Product development at Career Builder)

Price:

Contact Google Cloud Jobs directly

Since Cloud jobs API is the only application offering such a service and no such services are offered by Amazon and IBM, there is no comparison

---

G) Text Analysis

1) Natural language API

Natural language API helps in powerful text analysis by deploying machine learning techniques. The structure and meaning of the text are revealed by offering powerful machine learning models. It helps in text analysis and parsing intent from customer's conversations. It helps to extract information about people, places and even understand sentiments.

Key Features:

1) Syntax analysis which includes extracting tokens and sentences, identifying parts of speech and creating dependency parse trees for each sentence.

2) Entity analysis to identify entity and labels by types such as person, organization, location, events, products and media

3) Sentiment analysis which includes understanding the overall sentiment expressed in a block of text

4) Entity sentiment analysis for all the entities included in a block of text

5) Multilanguage analysis ranging from English, Spanish, Japanese, Chinese, French, German, Italian, Korean and Portuguese

6) Integrated REST API

***Benefits*:**

1) Insights from customers

2) Multi-media and multi-lingual support

3) Content classification and relationship graphs

4) Google Deep Learning models

Use cases:

1) Content classification by comparing content of a document against a preset list of classification categories.
2) Brands interested in capturing public opinion related to products or physical locations.
3) Tag content with common topics from a list of preset categories

Price:

(Per 1000 units by monthly usage)

Feature	0-5k	5K-1M	1M-5M	5M-20M
Entity analysis	Free	$1.00	$0.50	$0.25
Sentiment analysis	Free	$1.00	$0.50	$0.25
Syntax analysis	Free	$0.50	$0.25	$0.125
Entity sentiment analysis(beta)	Free	$2.00	$1.00	$0.50

2) Watson Explorer

Watson explorer is a cloud based enterprise search platform for content analysis as well as cognitive solutions to connect and view data insights across silos for better outcome, decision making and ROI. It helps to gain a 360 degrees view of the

customers to deliver superior experiences.

Key features:

1) Create unified information applications that bring data, analytics and cognitive insights together in one place to improve employee performance and customer engagement.

2) Automatic index rebalancing across clusters and remove barriers of scale, support redundancy and reliability.

3) Access on premise, cloud based, external and internal applications and extract data for indexing, analysis and interpretation to discover new insights.

4) Keep data wherever you want to make business flexible and future proof.

Benefits:

1) Improve customer engagement.

2) Achieve increased savings.

3) Reduce response time.

Use cases:

1) Identify actionable insights for preventive actions for connectivity, performance, bandwidth, utilization/ capacity and recommendation for optimization upgrades across globally dispersed manufacturing sites.
2) Improve client satisfaction across different desk operations beyond obvious self-service such as password reset and id services
3) Analysis of patient treatment and clinical data in healthcare.
4) Analysis and machine learning of security threats for IT security
5) Analysis and prediction of weather using big data and IoT

6) Analysis of consumer behavior who is shopping

Pricing:

Contact explorer representative for information on pricing.

3) Natural Language Classifier

Natural Language Classifier uses machine learning algorithms to return top matched pre-defined classes for short text inputs. It understands the intent behind the text and runs a corresponding classification with a confidence score. For example, what is the weather like today?

Key features:

1) Understand language of short texts.

2) Make predictions on how to handle them.

3) Routing users questions to the correct users and clarifying questions by severity

4) Route voice questions to a specific department.

Benefits:

1) Recombination of phrases

2) Optimization of grammar

3) Acoustic modelling.

Use cases:

1) Natural Language Classifier to interact with an IOT device
2) Answering questions with natural language classifier

Price:

Type	Description

First Classifier	First is Free. Each additional is $20.00 per month
API calls	First thousand per month are free. Additional $0.035 per call
Classifier training events	First 4 events free. Additional event $3.00 per training event

4) Retrieve and rank.

Key Features:

1) Customer Support: Find quick answers for customers from growing set of documents.

2) Field technicians: Resolve technical issues onsite

3) Professional Services: Finding the right set of people with the right skillset

Benefits:

1) Improved information retrieval

2) Ranker models take advantage of rich data in documents to provide relevant answers to queries

3) Benefits of both, open source community and advanced retrieval techniques build by Watson algorithm teams.

Use cases:

1) Train machine learning model based on relevant results
2) Leverage model to provide improved results to the end users
3) Load data into service

Pricing:

TYPE	**DESCRIPTION**
Retrieve	One Soir cluster with upto 50 Mb of storage is free. Highly available cluster costs $0.30 per instance hour for 4 Gb memory and 32 Gb storage. You can create a cluster upto 7X size
Ranker models	One model free per month. Additional are $10 per ranker model
Ranker API calls	First 1000 free. Additional $0.0009 per call
Rank training events	First 4 free. Additional $2.00 per training event

5) Discovery

Discovery helps to convert, normalize and enrich data. Then use the simplified query language to embed discovery into existing applications. You can include propriety data as well as third party data

Key Features:

1) Crawl, convert, enrich and normalize data

2) Explore propriety content and free license content

3) Concepts, relations and sentiments through Natural Language Processing

4) Simplify development while providing direct access to API`s

Benefits:

1) Extract business value from data

2) Solutions to have this information

3) Solutions at infrastructure back and front level.

Use cases:

1) Capturing insights from data
2) Add customer service representatives to deliver answers to complex questions
3) Help researchers understand the content from a vast number of research documents
4) Extend Watson conversation service based apps to find answers when the intents modelled are insufficient to find questions

Price:

1 GB RAM+ 2GB storage+unlimited enrichments+1000 new queries+custom model	**Free** for one month
Environment size 1: 2 GB RAM+48 GB Storage+4000 enrichments	**$960.00** USD/env/month
Environment size 2: 8 GB RAM+192 GB storage+ 16000 enrichments	**$3460.00** USD/env/month
Environment size 3: 16 GB RAM+384 GB storage+32000 enrichments	**$6530.00** USD/env/month
Enrichment overage	**$0.01** USD/ enrichment
News queries	**$0.10** USD/ query

6) Alchemy Data News

(Won't be available beyond March 7th 2018)) Has been replaced by Watson Discovery News Service

Alchemy Data News gives access to searchable dataset of blogs and news that are regularly updated. The data is indexed by Natural Language Processing. The historical search has data available since 60 days. You can go beyond the database with no need to acquire, enrich and store the data yourself enabling you to go beyond keyword searches. It is enhanced during indexing by applying natural language processing techniques and leveraging metadata extractions and creation functions provided by Alchemy Language API.

Key features:

1) News alerting

2) Event detection

3) Trending topics in the news

Benefits:

1) Combined with Alchemy Language(6)

2) Excellent guest service(5)

Use cases:

1) News alerting: Create powerful news alerts by taking advantage of API support for entities, cncepts, keywords, taxonomies and sentiment analysis as well to understand both, the news and how it is perceived

2) Event detection: Innovative event detection by leveraging API support for subject, action or object relationship extraction and checking for terms like acquisition.

3) Trending topics in the news: Identify the trending topics in the news and identify how frequently they are being mentioned.

Price:

Start free nothing else mentioned.

7) Web language model API

Automate a variety of standard natural language processing tasks using state of art language modeling API's

Key features:

1) Word breaking: Insert spaces in a list of words lacking spaces, like hashtag or part of the URL. Input a string of words with no spaces in between by entering only alpha numeric characters.

Benefits:

1) Assist customers with intelligent relevant suggestions and assistance
2) Assist customers with intelligent content completion
3) Enable the understanding of what is being said about your business

Use cases:

Web Language Model API provides probabilistic algorithms for natural language processing. Examples include probabilities for words appearing together and completing sentences. There are multiple use cases when processing text, including parsing OCR-ed content to determine word breaks.

Price:

Tier	Features	Unit	Price
Web language model API-Free		Transactions	100,000 transactions free per month
Web language model API-Standard	Upto 1000 transactions per seconds	Transactions	$0.05 per 1000 transactions

8) Text analytics API

Detects sentiments, key phrases, topics and language from your text

Key features:

1) Sentiment analysis: The API returns a numeric score between 0 and 1. A score close to 0, is a negative sentiment and close to 1 is a positive sentiment. The score is generated using classification techniques. The input features include n- grams, features generated from part of speech tags and word embeddings
2) Key phrase extraction: The API returns a list of strings denoting the key talking points in the input text. We employ techniques from Microsoft Office's sophisticated Natural Language Processing tool kit. English, German, Spanis and Japanese texts are supported.

3) Language detection: The API returns the detected language and a numeric score between 0 and 1. Score close to 1 indicates 100% the language is true. A total of 120 languages is supported

Benefits:

Use cases:

Txt analytics API extracts information and sentiment from the uploaded text. It returns the interpreted language, extracts key phrases and determines a sentiment score ranging from 0 to 100%. It is used to analyze reviews and social media comments.

Price:

	Free	**Standard S0**	**Standard S1**	**Standard S2**	**Standard S3**	**Standard S4**
Price	$0	$74.71	$249.86	$999.75	$2499.84	$4999.99
Max transactions	5000	25,000	100,000	500,000	2,500,000	10,000,000
Overage rate(per 1000 transactions)	N/A	$3	$2.50	$2	$1	$0.50

9) Bing Spell Check API:

Bing Spell check API helps users correct spelling errors, recognize the difference among names, brand names, slang as well as understand homophones as they are typing

Key features:

1) Easily recognize slang and informal language
2) Spot common name errors in context.
3) Spot common name errors in context
4) Fix homonyms and hard to spot errors
5) Spot new brands and other names as they emerge

Benefits:

1) Fast
2) Precise
3) Easy
4) Inviting

Use cases

A chatbot for proofing and sending emails can auto correct misspellings and add capitalization or punctuation as appropriate to a draft message

Price:

Features	**Unit**	**Price**
Up to 100 transactions/ sec	Transactions	$5/10,000 transactions

As mentioned in the use cases, the service best useful in a scenario can be used

H) Translate

1) Cloud translation API

Cloud translation API helps in fast and dynamic translation. It is responsive, It shows constant seamless improvement and introduces new languages and language pairs. Language detection is also available in case the source of the language is unknown. The underlying technology pushes the boundary of machine translation and is updated constantly.

Key features:

1) Accessible via a single Google REST API including Python, Objective C and Ruby.

2) More than 100 languages are supported in HTML with a translated text back. No need to extract source text or re- assemble translated content.

3) Log analysis and human translation examples being learnt. Existing language pairs being imporved and new language pairs being learnt at no additional cost.

4) Increase quota from 2M characters to 50M characters per day or receive a higher quota.

5) Languages being successfully detected and translated.

Benefits:

1) Simple integration

2) Highly scalable

3) Language detection

4) Language translation

5) Affordable pricing

Use cases:

1) Language translation
2) Language detection
3) HTML support
4) Translation API limit

Price:

Feature	Price
Text translation	**$20** per million characters
Language detection	**$20** per million characters

2) Document Conversion:

Converts documents into formats that have to be well understood by Watson services such as retrieve and rank services. The input is transformed to normalized HTML, plaintext or a set of JSON- formatted answer units that can be used with Watson Retrieve and Rank services.

Key features:

1) Custom query builder optimized (7)

2) Set of algorithms to score semantic relationship between a query and Solr document which get fed in the rank position of the R&R. (7)

Benefits:

1) Dictionary based segmentation with high recall benefits of non-dictionary based segmentation.

2) Contextually relevant insights

3) Flexible architecture

Price:

Service	**Price**
Standard service	First 100 Megabytes per month are free and then **$0.05** per mega byte
Premium(High level of security and isolation)	Contact sales.

3) Language translator

Text is translated from one language to another. The domain specific models can be customized based on unique terminology and language. News from across the globe can be taken and translated to your language, customers can be communicated in their language.

Key features:

1) Identify the language of the text

2) Translate the text from one language to another

3) Perform domain specific translations

4) Add your own terms of glossary and phrases

5) Upload a large body of text to serve as a language sample.

Benefits:

1) Targeted as news articles and transcripts

2) Targeted conversational colloquialisms

3) Patents targeted at technical and legal terminology.

Use cases:

1) Create speakers captions in a video
2) Take news from across the globe and present it in the readers language
3) Help desk assistant chat bot that communicates with the clients/customers in their own language

Price:

Service type	**Description**	**Price**
Standard plan	Includes news, conversational and patent models	**$0.02**/thousand characters after every first 250000
Advanced plan	Custom model maintenance	**$0.02**/thousand for standard translations **$0.10**/ thousand characters for custom model translations **$15 .00**/model/month, pro-rated daily
Premium	High level of security	Contact sales advisor

4) Translator text API

Microsoft translator text API is a cloud based machine learning translation service supporting multiple languages that reach more than 95% of the world's gross domestic product. Translator can be used t build applications, websites, tols or any solution requiring Multilanguage support

Key features:

1) Extend the reach of your applicatons: Across mobile, desktop and web applications easily power translation to and from 60+ supported languages fot text through open REST interface of translator API
2) Automatically detect languages: Easily and automatically detect the languages of any text string, simplifying development process allowing you to quickly send for translation and serve for localized content
3) Build customized translation systems with hub using previous translations to build a system reflecting your needs or starting with a dictionary of specific words.
4) Crowd source translation improvement by creating specific user groups that provides suggestions and designated approvers who verify inclusion using Collaborative Translation Framework

Benefits:

1) Fast
2) Cost effective
3) Easy
4) High capacity

Use cases:

Automatic translation of training and sales materials

Price:

	Free	**S1**	S2	S3	S4
Price					$45,000

Max characters in tire	2,000,000	N/A	250,000,000	1,000,000,000	10,000,000,000
Overage rate	N/A	N/A	$8.22	$6	$4.50

Google supports a higher number of characters i.e. upto 1.5 billion. Both of those support 100+ languages. Google translation is best for words and short sentences translations. IBM Watson is best for video translations, emails and articles and website services.

I) Relation Between The Data

1) IBM Watson Knowledge Studio

Enables developers and domain experts to collaborate and create custom annotator components. These annotators can identify mentions and relationships in unstructured data and can be administered throughout the lifecycle using one common tool. These annotator components can be deployed directly to IBM Watson Explorer, Watson Natural Language Understanding and Watson Discovery Service.

Key Features:

1) Accelerated training process by using dictionaries and pre-annotated documents

2) Shared knowledge artifacts to expedite model training process

3) Continuous performance improvement of annotator components

4) Integrated development environment: Deploy text annotators directly to Watson Explorer, Natural Language Understanding and Watson Discovery Service

5) Intuitive knowledge transfer: Single point and click to teach watson about domain entities and relations

6) Work with SME`s and resolve conflicts between domain experts by using a built-in interface

7) Improve annotation accuracy. Leverage quantitative measurements such as Inter Annotator Agreement (IAA) and precision and recall scores to produce highest quality ground truth.

Benefits:

1) Train Watson without coding

2) Engage your experts

3) Save time and money

Use cases:

1) Identify safety defects using traffic incident reports
2) Train and deploy custom text annotator components
3) Used by data scientists, natural language experts and subject matter experts to create custom annotator.

Price:

Tier	**Description**	**Price**
Free	-	**$0.00 USD**
Standard	Infuse knowledge into cognitive applications	**$150.00 USD** per user per month
Premium	Collaborate with more users and have access to isolated data storage with a premium offer designed for enterprises	Contact IBM

J) Language Processing:

1) Alchemy Language:

Will remain supported till March 2018

Helps you understand sentiments, keywords, entities, high level concepts and more. To understand how social media followers feel about your products to automatically classify the contents of a web page or topics trending in the news.

Key Features:

1) Text extraction which includes normalization of HTML content, removing ads, navigation links and other unimportant contents

2) Title extraction from web pages

3) Combined call for semantic information.

Benefits:

1) Natural Language Understanding

2) Image Recognition

3) Sentiment analysis.

Use cases:

1) Understand the cause behind negative and positive reviews
2) Track important industry events like acquisitions and new competitors
3) Tailor recommended content and advertisement to the content on the page

***Price*:**

Free

2) Conversation Service:

Creates an application that understands natural language input. It uses machine

learning to simulate conversation between individuals. It allows you to quickly build, test and deploy a bot or virtual agent across mobile devices or messaging platforms. It creates natural conversations between the application and the user, without any coding experience required.

Key Features

1) Control mobile app using natural language

2) Build twitter, slack and other social network platform bots.

3) Chatbot that responds automatically to customers most frequently asked questions

Benefits:

1) General availability

2) Natural conversation flows.

3) Easy tooling

4) Free of cost

Pricing:

Type	Description	Price
Free	Upto 1000 API queries/month Upto 3 workspaces Upto 25 intents Shared public cloud	
Standard	Unlimited API queries/month Up to 20 workspaces	$0.0025 USD per API call

	Upto 2000 intents Shared public cloud	
Premium	High level of security and isolation	Talk to staff

3) Language Understanding Intelligent Service

The aim to understand the language contextually, so the app communicates with people the way they speak

Key features:

Build custom language models:

One of the key problems in understanding human-computer interactions is understanding what exactly the person wants and finding the person pieces of information relevant to his/her intentions. The Language Understanding Intelligent Service provides simple tools that enable to build one's own model which allow any application bot to understand your commands and act accordingly.

Benefits:

1) Fast and easy
2) Learns and adapts
3) Offers pre-build applications
4) It is a power developer tool

Use cases:

LUIS integrates seamlessly with multiple services and works well with Bing Speech API to integrate speech, Bot framework to create chatbots, IoT devices, mobile applications and more

Fig 14. LUIS working

Price:

Tiers	**Units**	**Price**
Free	Transactions	10,000 transactions free per month
Basic	Transactions	0.75$/1000 transactions

4) Linguistic Analysis API

Simplify complex language concepts and parse text with Linguistic Analysis API.

Key features:

1) Explore the structure of your text:

   The Linguistic API uses advanced linguistic analysis tools for natural language processing, giving you access to part of speech tagging and parsing. These tools allow to hone in on important concepts and actions. The API can find phrases and concepts using natural language parsers. Whether you are mining customer feedback, interpreting user commands or consuming web text , understanding the structure of the web text is really important.

   Benefits:

   1) Sentence separation and tokenization
   2) Part of speech tagging
   3) Constituency parsing

   Price:

   No additional costs

Both services, Amazon as well as those related to IBM can be used in specific cases

---

K) Personality Insights:

1) Personality insights services

Extracts personality characteristics based on how a person writes. Services can be used to match individuals to other individuals, opportunities, products w.r.t personalized messaging and recommendation

Key Features:

1) Only those preferences necessary to analyze consumer behavior are selected

2) Specify only the necessary dominant habits.

3) Customer understanding at a deeper level

4) Tailor products, services, campaigns and communications.

Benefits:

1) Higher customer satisfaction

2) Acquisition retention

3) Strengthen relations.

Use cases:

1) Targeted marketing
2) Customer care
3) Customer service
4) Personal connections
5) Resume Writing

Price:

Tier	Description	Price
Free tier	First 100 API calls per month. Big 5, values and Needs are the models included.	**-----**
Tier 2	1 - 100,000 API calls	**$0.02 USD per API call**
Tier 3	100,001 - 250,000 API calls	**$0.01 USD per API call**
Tier 4	250,001 and greater API calls	**$0.005 USD per API call**
Premium		**Talk to staff**

2) Tone Analyzer:

Uses linguistic analysis to detect emotions, social tendencies and writing style. It understands conversations and communications and responds to customers at an appropriate scale.

Key Features:

1) Can be integrated with other Watson services

2) Calculates the tone of the text

3) Focuses only on the speech of the author

Benefits:

1) Infers emotions

2) Infers social tones

3) Response is confident, organized and thoughtful

Use cases:

1) Social listening and audience monitoring: Monitor social media to understand customer reviews of your brand in real time
2) Personalized marketing: Determine whom to target with personal messaging
3) Chatbots: Respond to customers after analyzing their mood and tones to craft responses
4) Customer engagement monitoring and quality assurance: Monitor tone of agent

Price:

Standard service		
	First 1000 API calls per month	Free
	1-250,000 calls	$0.0088
	250,001-5,000,000	$0.0013
	5.000.000+ calls	$0.0008
Premium service	Talk to the IBM Representative	

3) Speaker Recognition API

The Microsoft speech recognition API has an HHTP 1.1 definition for building simple applications that perform speech recognition. It is useful where continuous feedback is not required from the customers.

Features:

1) Utterances limited to a maximum of 15 seconds
2) Partial results are not returned. Only the final phase result is returned
3) Service end of speech detection is not supported. Clients must determine end of speech.
4) Continuous recognition is not supported
5) A single recognition phrase result is returned to the client only after the client stops writing to the request stream and not beforehand

Benefits:

1) User can control the application by speaking.
2) No need to create own engine
3) Capture information from events generated by speech recognition
4) Cost efficient

Use Cases:

1) Create custom language models
2) Deploy models
3) Access end point from any device

Price:

	Standard S1	**Standard S2**	**Standard S3**
Price	5	50	450

Maximum transactions included in a tier	10000	100000	1000000
Average rate per 100000 transactions	N/A	50	50

The IBM Tone Analyser has a chart accompanied by the various emotional, social and writing tones of the language. It has a better review than Microsoft Speaker Recognition API

Conclusion

This paper compares the artificial intelligence API's and soft wares pertaining to their features and use cases that can be referred to when a solution is needed for a particular product. It can be used by consultants to suggest which software is better than the rest depending on the price of the product, the usability in various scenarios

REFERENCES:

Ackerson Chris, (March 16, 2016), Developing with IBM Watson Retrieve and Rank: Part 1 Solr Configuration, Retrieved from https://medium.com/machine-learning-with-ibm-watson/developing-with-ibm-watson-retrieve-and-rank-part-1-solr-configuration-29c18e52966f

Casalboni Alex (July 20, 2017), What is Amazon Machine Learning and How does it work Retrieved from https://cloudacademy.com/blog/aws-machine-learning/

Hern Alex (September 28, 2016), Partnership on AI formed by Google, Facebook, Amazon, IBM and Microsoft. Retrieved from https://www.theguardian.com/technology/2016/sep/28/google-facebook-amazon-ibm-microsoft-partnership-on-ai-tech-firms

Ganesan Neerja (2016, December 19), Visual Recognition puts data to work to improve decision making. Retrieved from https://mobilebusinessinsights.com/2016/12/visual-recognition-puts-data-to-work-to-improve-decision-making/

Hope Reese (December 14 2016), The 6 most exciting AI advance of 2016. Retrieved from http://www.techrepublic.com/article/the-6-most-exciting-ai-advances-of-2016/

IBM, (October 6, 2016), IBM Watson is now fluent in nine languages (and counting). Retrieved from http://www.wired.co.uk/article/connecting-the-cognitive-world IBM Bluemix Docs, Retrieved from https://console.bluemix.net/docs/services/visual-recognition/index.html#about

IBM Bluemix Docs, Retrieved from https://console.bluemix.net/docs/services/visual-recognition/index.html#about

Knight Will (January 4, 2017), Five Big Predictions for Artificial Intelligence in 2017, Retrieved from https://www.technologyreview.com/s/603216/5-big-predictions-for-artificial-intelligence-in-2017/

Mannes John(2017, September 19), 'Google Cloud's Natural Language API gets content classification and more granular sentiment analysis' Retrieved from https://techcrunch.com/2017/09/19/google-clouds-natural-language-api-gets-content-classification-and-more-granular-sentiment-analysis/

McAssey Jim, (October 25, 2016), Building a business case for cognitive computing (tangible benefits). Retrieved form https://www.linkedin.com/pulse/oct-2016-update-part-3-building-business-case-tangible-jim-mcassey/

Microsoft Azure, Retrieved from https://docs.microsoft.com/en-gb/azure/cognitive-services/speech/concepts

Pkahanal, Retrieved from https://github.com/pkhanal/openwhisk-nlc-action

Parloff Roger (September 28, 2016), AI Partnership launched by Amazon, Facebook, Google, IBM, and Microsoft. Retrieved from http://fortune.com/2016/09/28/ai-partnership-facebook-google-amazon/

Shen Hsieh Angela B, (July 11, 2016), Building better bots with Watson Conversation. Retrieved from https://www.ibm.com/blogs/watson/2016/07/building-better-bots-watson-conversation/

Thanh Lam Muhammad, Valerie Lampkin ,Zain Mustafa Lak Sr, May 2017, Retrieved from Building Cognitive Appications with IBM Watson Services: Volume 5, Language Translator, Retrieved from http://www.redbooks.ibm.com/redbooks/pdfs/sg248392.pdf

Underwood Jenn, (March 28, 2017) IBM Watson Cognitive Computing. Retrieved from http://www.jenunderwood.com/2017/03/28/ibm-watson-cognitive-computing/

Yegulalp Serder (February 29, 2016), How IBM, Google, Microsoft and Amazon do machine learning in the cloud, Retrieved from https://www.infoworld.com/article/3039052/cloud-computing/how-ibm-google-microsoft-and-amazon-do-machine-learning-in-the-cloud.html

Zachary Chase Lipton, Cloud Machine Learning Wars: Amazon vs IBM Watson vs Microsoft Azure. Retrieved from http://www.kdnuggets.com/2015/04/cloud-machine-learning-amazon-ibm-watson-microsoft-azure.html

(2016, November 11),' Real World Use Cases of IBM Watson Explorer' , Retrieved from http://bitnine.net/blog-computing/real-world-use-cases-of-ibm-watson-explorer/?ckattempt=1

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

APPENDIX

List of figures

Fig 1. Amazon Lex Bot **11**
Fig 2. Amazon Lex Chatbot **12**
Fig 3. Amazon Lex Bot Enterprise Applications **13**
Fig 4. Amazon Lex teaching and exploration **14**
Fig 5. Real Estate Property Search **17**
Fig 6. Employee Badge Scan **18**
Fig 7. Retail Store Sentiment Analysis **18**
Fig 8. Find images of friends **19**
Fig 9. Celebrity Images **20**
Fig 10. Convert an article to speech and download as mp3 ..... **31**
Fig 11. Play speech and highlight spoken text **32**
Fig 12. Mobile desktop applications **33**
Fig 13. Job discovery working **50**
Fig 14. LUIS Working **76**

List of tables

1. **Summary of technologies** **9**