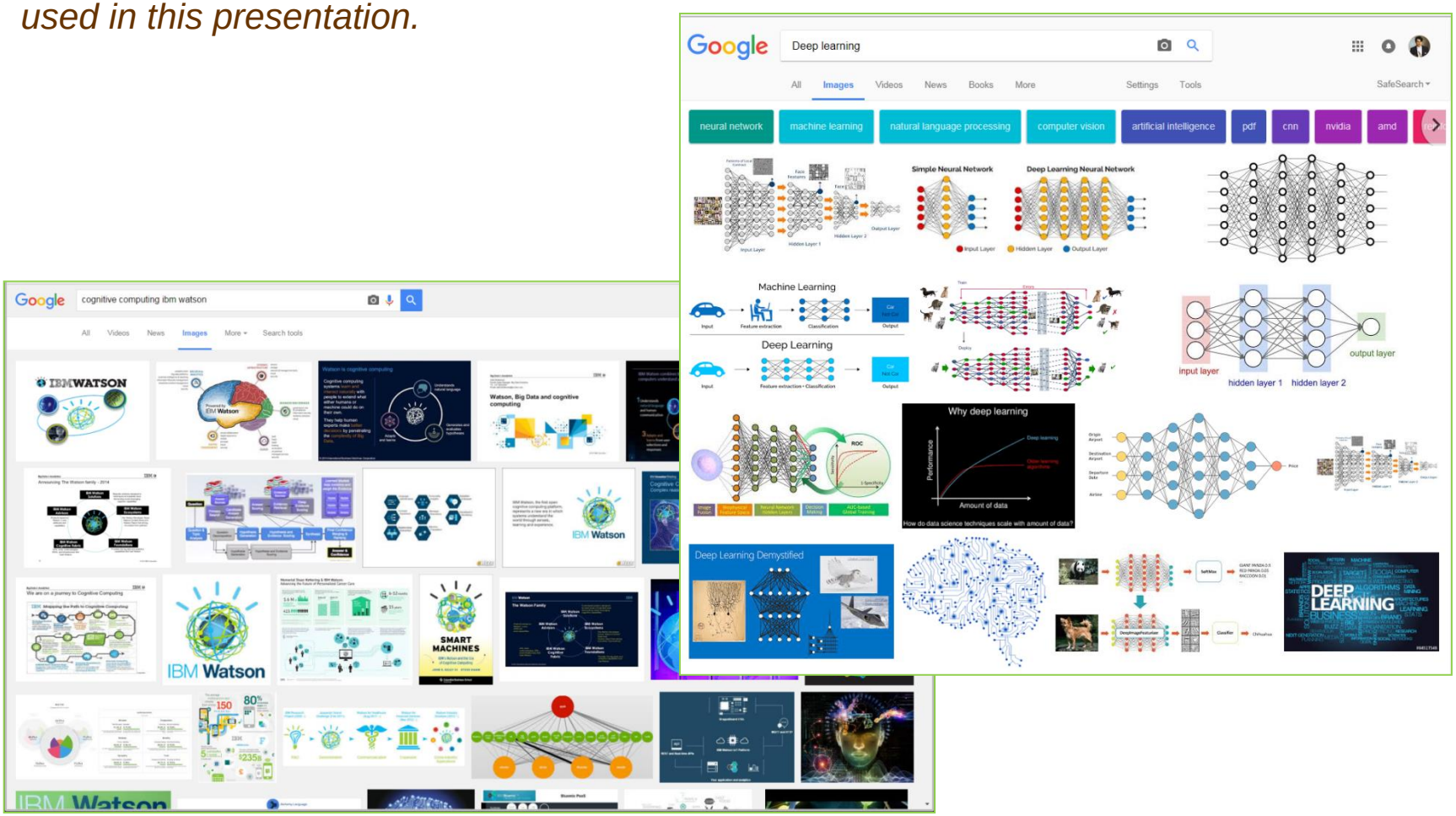
Lecture 8: Cognitive Computing Services

TIES4911 Deep-Learning for Cognitive Computing for Developers
Spring 2024

by:
Dr. Oleksiy Khriyenko
IT Faculty
University of Jyväskylä

Acknowledgement

I am grateful to all the creators/owners of the images that I found from Google and have used in this presentation.



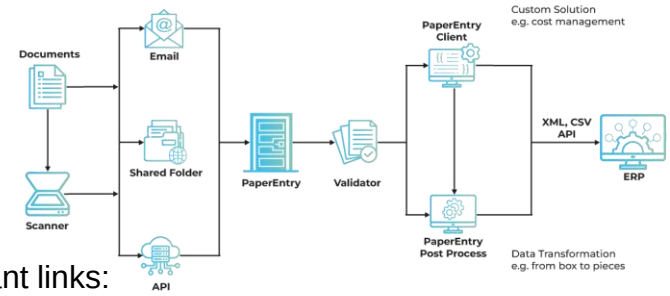
"Visual" Deep Learning

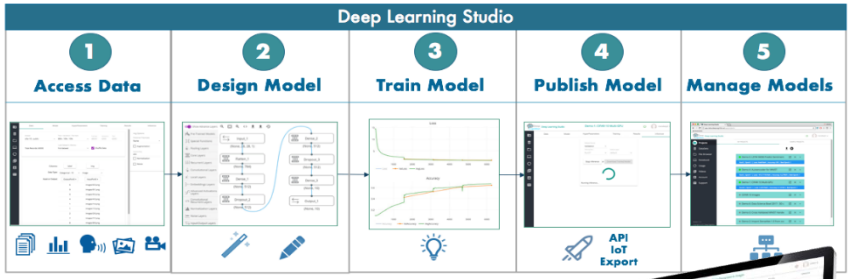
DeepCognition.ai

Deep Learning Studio *is a platform founded by* ***Deep Cognition*** *to "democratize AI" through visual API to create and deploy Deep Learning solutions with the click of a button. The simple drag & drop interface helps to design deep learning models with ease. Deep Cognition provides free access to cloud and desktop software with single sign on.* https://deeplearningstudio.com , https://docs.deepcognition.ai

- *Cloud Solution is a single-user solution for creating and deploying AI. The simple drag & drop interface helps you design deep learning models with ease. Pre-trained models as well as use built-in assistive features simplify and accelerate the model development process. You can import model code and edit the model with the visual interface.*
- *Desktop Solution is a single user solution that runs locally on your hardware. Desktop version allows you to train models on your GPU(s) without uploading data to the cloud.*
- *Enterprise Solution is a multi-user solution that supports both the cloud and on-premise hardware. The solution includes a full suite of features designed to accelerate enterprise AI including Big Data ingestion, model development, training, & deployment via code export, REST API, or IoT integration.*

PaperEntry *technology is designed to work with different types of documents. The documents can be extracted from email, Shared Folders, and can be integrated via APIs.* ***PaperEntry's*** *core technology is based on Artificial Intelligence. The technology enables relevant data extraction from documents. The extracted document can be analyzed via a Validator, and the validated data can then be routed to a client or a post processing engine for further digital transformation. Finally, the extracted, validated, transformed (optional) data can be integrated into ERP (Enterprise Resource Planning) or TMS (Transport Management System) or AP (Accounts Payable) systems. The diagram below illustrates the overall flow.* ***PaperEntry's*** *user-friendly interface, custom developed labeler, validation software, and integration allows the data extraction from any type of document.*



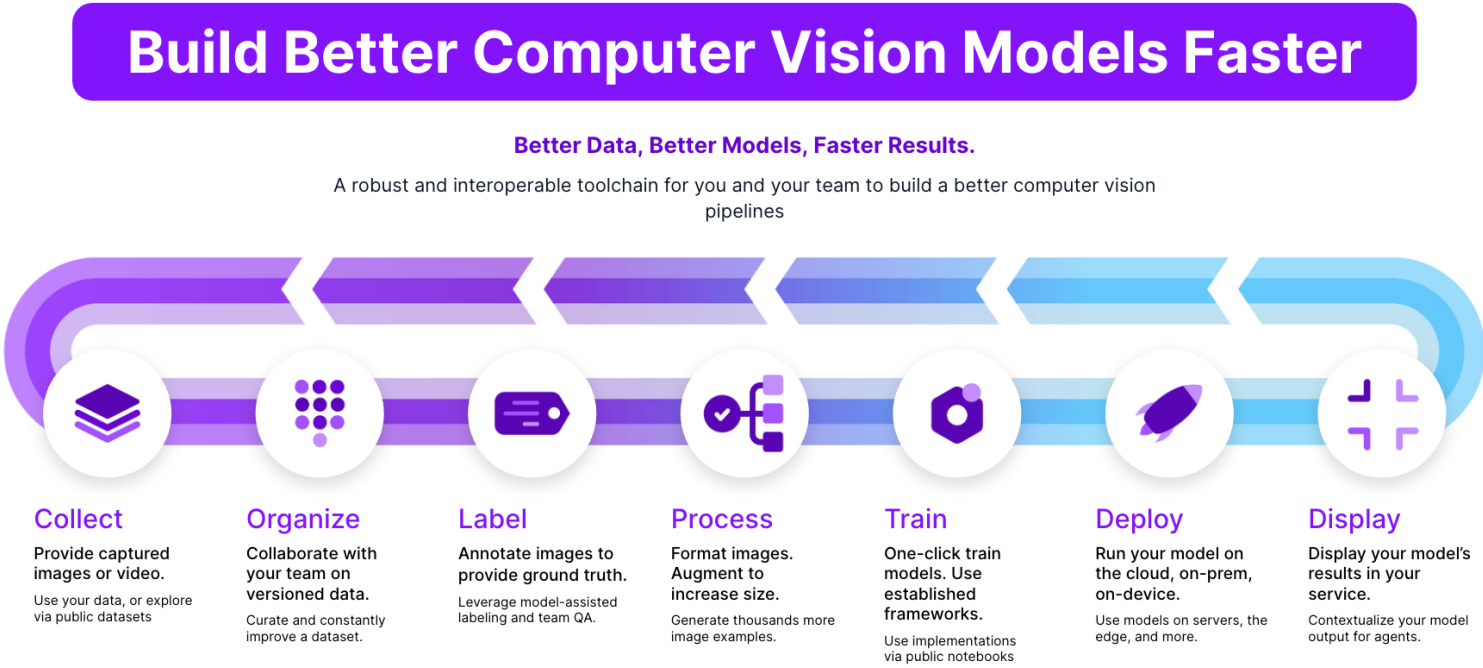


Relevant links:

https://towardsdatascience.com/a-video-walkthrough-of-deep-cognition-fd0ca59d2f76
https://towardsdatascience.com/towards-an-easier-deep-learning-life-with-deep-cognition-8fa8f41ca672
https://becominghuman.ai/deep-learning-made-easy-with-deep-cognition-403fbe445351

Roboflow

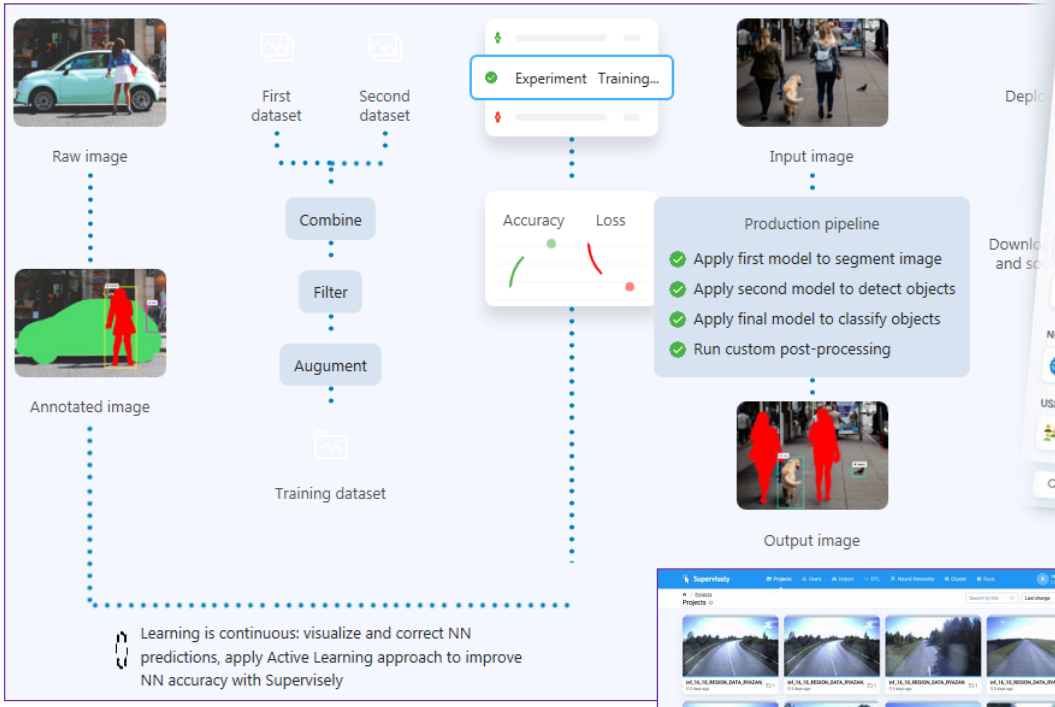
Roboflow *empowers developers to build their own computer vision applications, no matter their skillset or experience. It provides all of the tools needed to convert raw images into a custom trained computer vision model and deploy it for use in applications. Today, Roboflow supports object detection and classification models.*
https://roboflow.com/ , https://docs.roboflow.com/

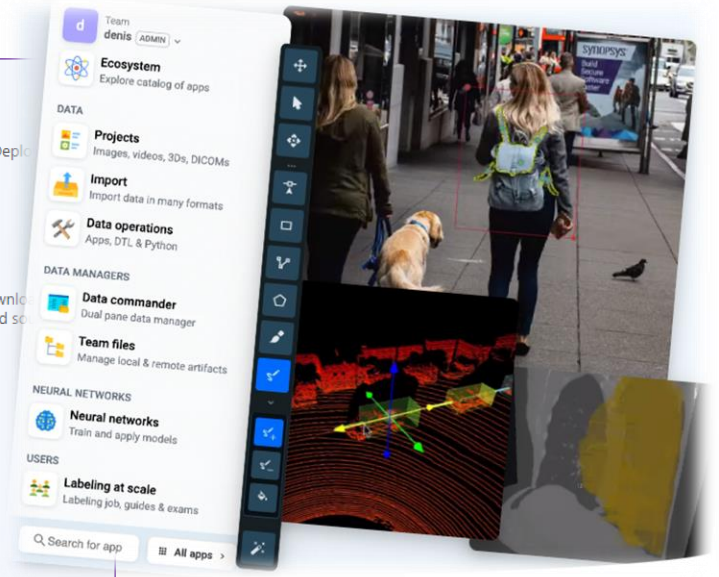


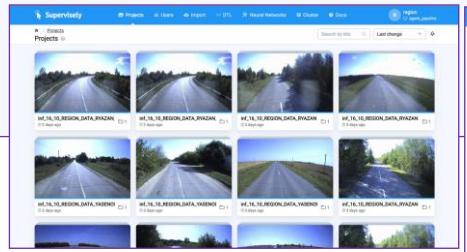
Relevant links:
https://blog.roboflow.com/how-to-train-yolov8-on-a-custom-dataset/

Supervisely

Supervisely *the leading platform for entire computer vision lifecycle. Image annotation and data management tools transform your images / videos / 3d point cloud into high-quality training data. Train your models, track experiments, visualize and continuously improve model predictions, build custom solution within the single environment.* https://supervise.ly





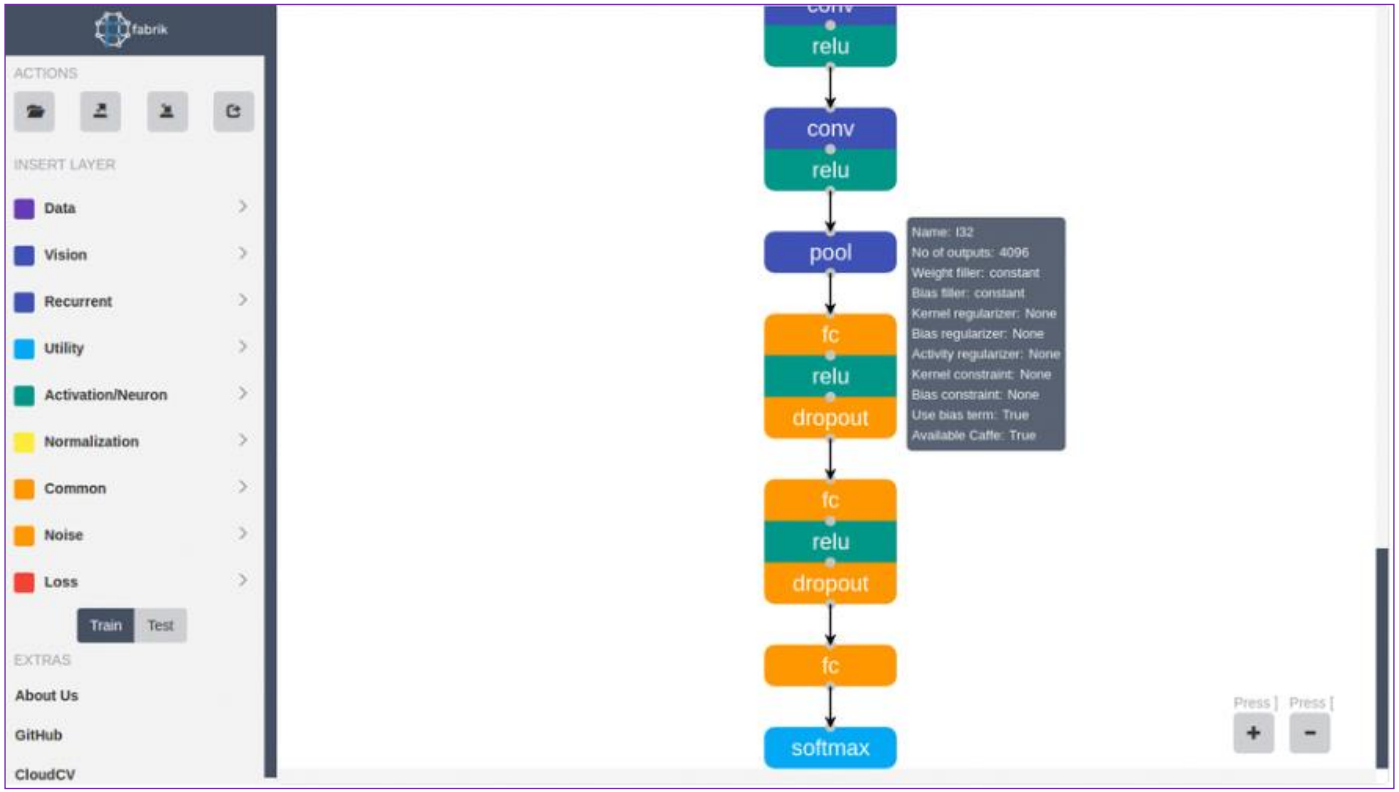


Relevant links:
https://www.youtube.com/watch?v=7eOLm_dsuz8

fabrik

“Visual” Deep Learning

Fabrik *is an online collaborative platform to build, visualize and train deep learning models via a simple drag-and-drop interface. It allows researchers to collaboratively develop and debug models using a web GUI that supports importing, editing and exporting networks written in widely popular frameworks like* ***Caffe****,* ***Keras****, and* ***TensorFlow****.* *https://github.com/Cloud-CV/Fabrik*



Cloud AutoML

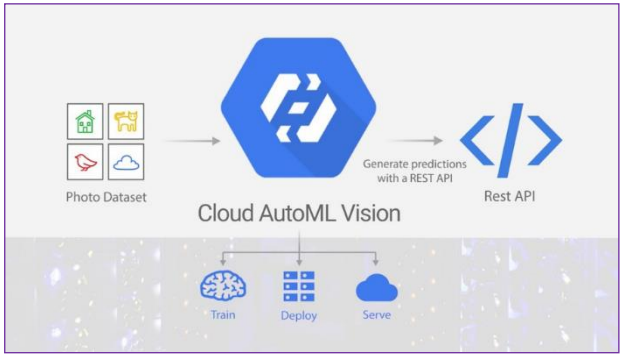
Cloud AutoML *allows training high-quality custom machine learning models with minimum effort and machine learning expertise.* https://cloud.google.com/automl/

- ***Vertex AI Platform.*** *Fully managed, end-to-end platform for data science and machine learning. Build and deploy custom models for image, video, text, and tabular data with a robust toolset in one complete, cohesive platform.* https://cloud.google.com/vertex-ai
- ***AutoML Vision****. Derive insights from object detection and image classification, in the cloud or at the edge. AutoML Vision offers higher model accuracy and faster time to create a production-ready model.* https://cloud.google.com/vertex-ai/docs/training-overview#image_data
- ***AutoML Video Intelligence*** *Enable powerful content discovery and engaging video experiences.* https://cloud.google.com/vertex-ai/docs/training-overview#video_data
- ***AutoML Natural Language****. Reveal the structure and meaning of text through machine learning.* https://cloud.google.com/vertex-ai/docs/training-overview#text
- ***AutoML Translation****. Dynamically detect and translate between languages. AutoML Translation enables you to create your own, custom translation models so that translation queries return results specific to your domain.* https://cloud.google.com/translate/automl/docs/
- ***AutoML Tables (beta)*** *Automatically build and deploy state-of-the-art machine learning models on structured data.* https://cloud.google.com/vertex-ai/docs/tabular-data/overview

Vertex AI *is a machine learning (ML) platform that lets you train and deploy ML models and AI applications and customize large language models (LLMs) for use in your AI-powered applications. Vertex AI brings* ***AutoML*** *and* ***AI Platform*** *together into a unified API, client library, and user interface.*

https://cloud.google.com/vertex-ai

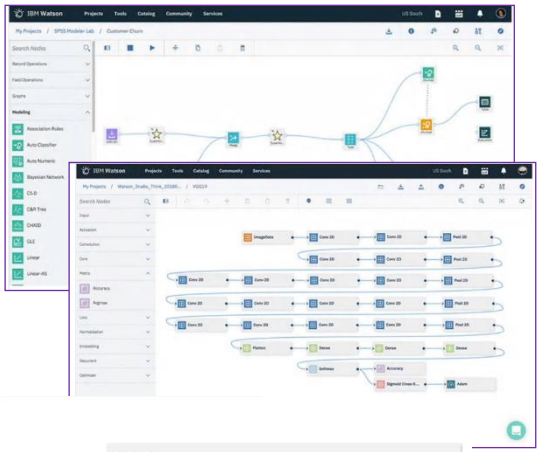
https://cloud.google.com/vertex-ai/docs/start/introduction-unified-platform

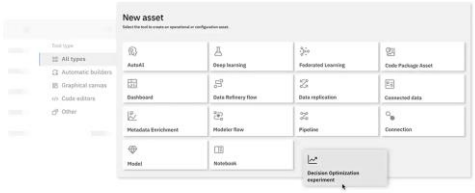


IBM Watson Studio

IBM Watson Studio *is one of the core services in Cloud Pak for* ***Data as a Service*** *that provides tools for data scientists, application developers and subject matter experts to collaboratively and easily work with data to build and train models at scale. It gives you the flexibility to build models where your data resides and deploy anywhere in a hybrid environment so you can operationalize data science faster.*

https://www.ibm.com/products/watson-studio/, https://dataplatform.cloud.ibm.com/docs/content/svc-welcome/wsl.html





Bring together open-source frameworks like PyTorch, TensorFlow and scikit-learn with IBM and its ecosystem tools for code-based and visual data science. Work with Jupyter notebooks, JupyterLab and CLIs — or in languages such as Python, R and Scala.

IBM Watsonx *In 2023, IBM announced the watsonx platform, which allows partners to train, tune and distribute models with generative AI and machine learning capabilities. Under development for three years, IBM designed watsonx to manage the life cycle of foundation models that are the basis of generative AI capabilities and for creating and tuning machine learning models.*

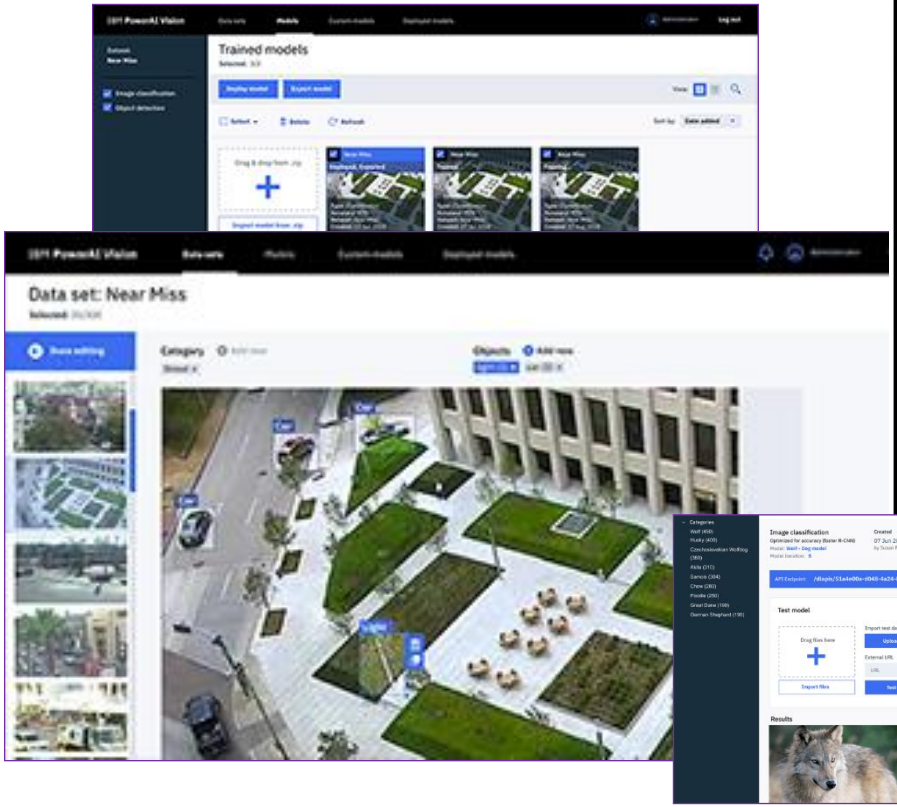
https://www.ibm.com/watsonx

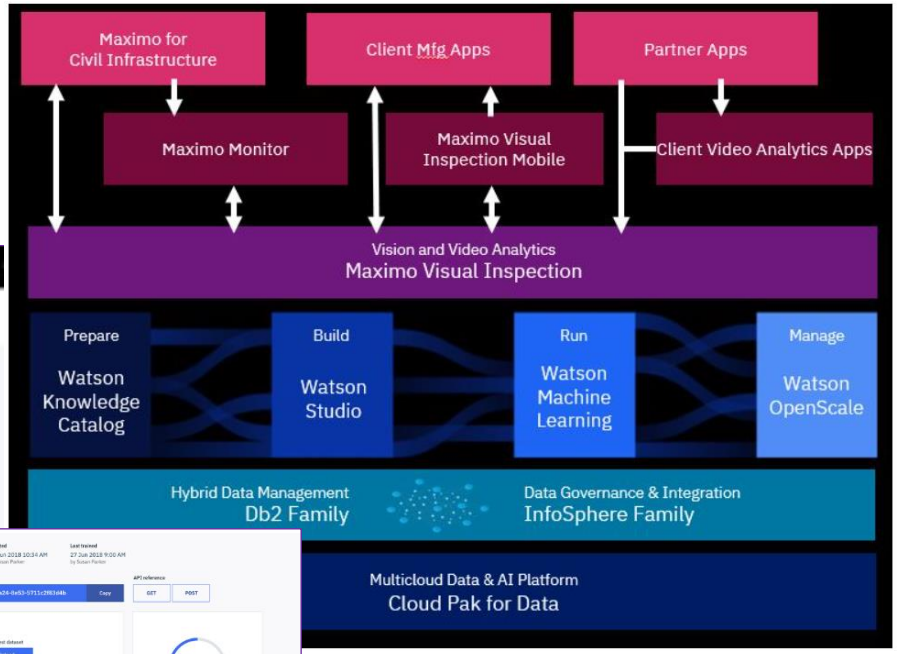
IBM PowerAI Vision

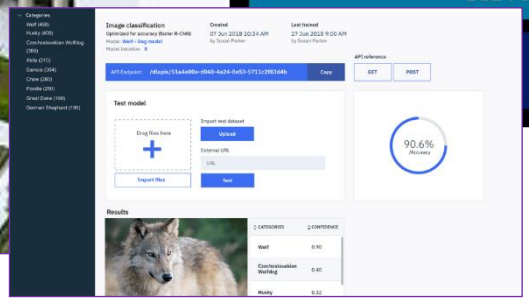
IBM PowerAI Vision

*The **IBM® Maximo® Visual Inspection** platform, built on cognitive infrastructure, is a new generation of video and image analysis platforms. The platform offers built-in deep learning models that learn to analyze images and video streams for classification, object detection, and anomaly detection.* https://www.ibm.com/docs/en/mas-cd/maximo-vi

https://www.ibm.com/docs/en/mas-cd/maximo-vi/continuous-delivery?topic=overview

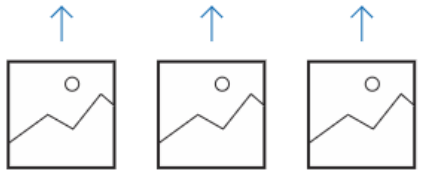






Microsoft Custom Vision

Custom Vision *Easily customize your own state-of-the-art computer vision models that fit perfectly with your unique use case. Just bring a few examples of labeled images and let Custom Vision do the hard work.* https://www.customvision.ai/

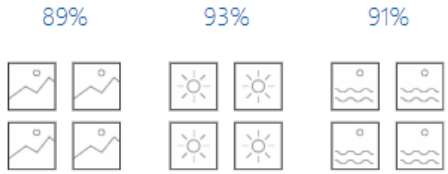


Upload Images

Bring your own labeled images, or use Custom Vision to quickly add tags to any unlabeled images.

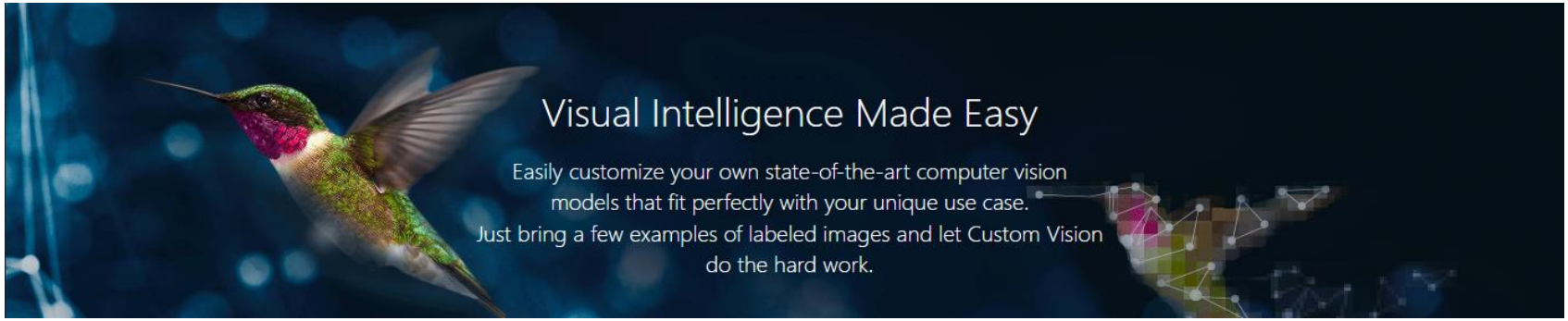
Train

Use your labeled images to teach Custom Vision the concepts you care about.



Evaluate

Use simple REST API calls to quickly tag images with your new custom computer vision model.



Cognitive Computing Services

Google (DeepMind)
https://cloud.google.com/vision/
https://cloud.google.com/video-intelligence
https://cloud.google.com/speech-to-text/
…

IBM Watson
https://www.ibm.com/cloud/watson-text-to-speech
https://www.ibm.com/cloud/watson-natural-language-understanding
https://www.ibm.com/cloud/watson-language-translator
…

Microsoft Azure
https://azure.microsoft.com/en-us/services/cognitive-services/computer-vision/
https://azure.microsoft.com/en-gb/services/cognitive-services/speech-services/
https://azure.microsoft.com/en-us/services/cognitive-services/face/
https://azure.microsoft.com/en-us/services/cognitive-services/text-analytics/
…
https://www.captionbot.ai/
https://www.how-old.net/
https://www.what-dog.net/

Amazon
https://aws.amazon.com/rekognition/
https://aws.amazon.com/comprehend/
…

Clarifai
https://www.clarifai.com/

MetaMind
https://einstein.ai/
https://metamind.readme.io/

Cognitec
http://www.cognitec.com/

Kairos
https://www.kairos.com/

Relevant links:
https://blog.filestack.com/thoughts-and-knowledge/comparing-google-vision-microsoft-cognitive-amazon-rekognition-clarifai/
https://www.predictiveanalyticstoday.com/what-is-cognitive-computing/

IBM Watson

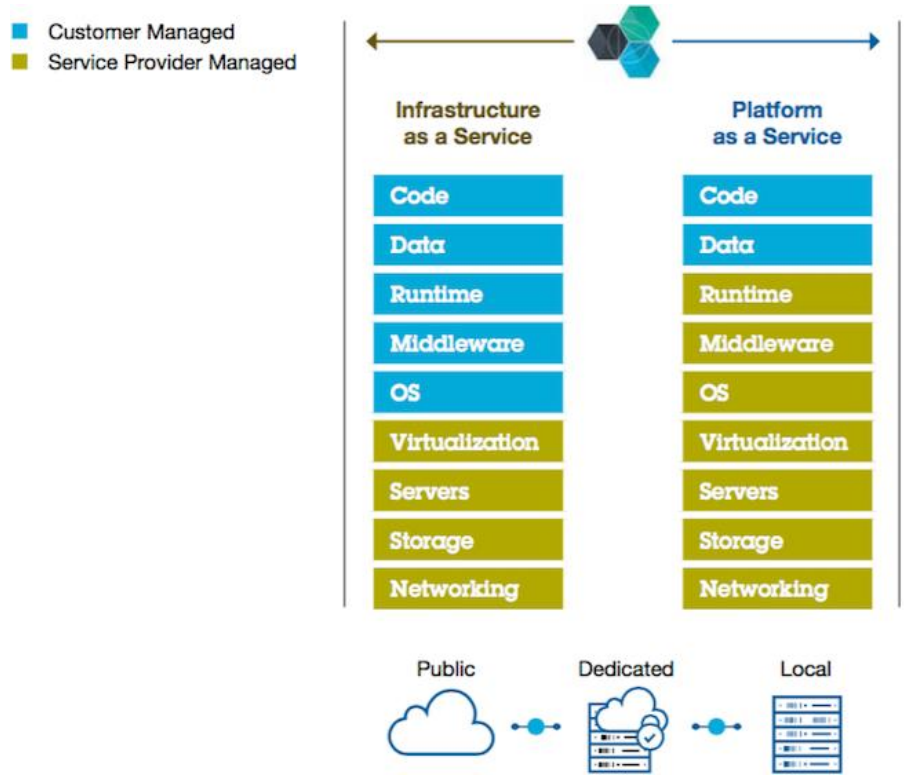
IBM Cloud Platform

IBM Cloud *is a* ***cloud platform*** *for rapidly building, running, and managing applications. The goal is:*

- *to simplify the delivery of an application by providing hosting capabilities and services that are ready for immediate use.*
- *helping to test and adopt a broad range of cloud services and capabilities from IBM, open-source communities, and third-party developers.*

Link: https://www.ibm.com/cloud/

IBM **Cloud**



Build on open technologies:

IBM Watson

IBM Cloud Platform

***Access to the Watson services**:*

- *Sign up to the platform by creating your personal IBMid. In addition to IBM ID and password, you suppose to specify organization name and space name that will be used in app deployment… (https://cloud.ibm.com/registration)*
  - *no time restrictions with Lite plan services.*
  - *Get access to Infrastructure and Platform services from the catalog (https://cloud.ibm.com/catalog/).*

- *Access a detailed service information (documentation, videos, demos, etc.) directly from **Watson Development Cloud**. Choose a service from the "Services" menu on the top of the page.*

https://www.ibm.com/watson/developer/

IBM Watson

IBM Watson

Now Watson is available as a set of open APIs and SaaS products.

SaaS products

- *Watson Explorer*
- *Watson Analytics*
- *Watson Virtual Agent*
- *Watson Knowledge Studio*
- *Watson Knowledge Catalog*
- *…*

Services at IBM Cloud

Language	*Language Translator*
	Natural Language Classifier
Speech	*Speech to Text*
	Text to Speech
Vision	*Visual Recognition*
Empathy	*Personality Insights*
	Tone Analyzer
Knowledge	*Natural Language Understanding*
	Discovery
Language	*Discovery News*
Conversation	*Watsonx Assistant*
Coding	*Watsonx Code Assistant*

IBM Cloud

IBM Watson

Speech Speech to Text

Speech to Text *service converts audio and voice into written text…*

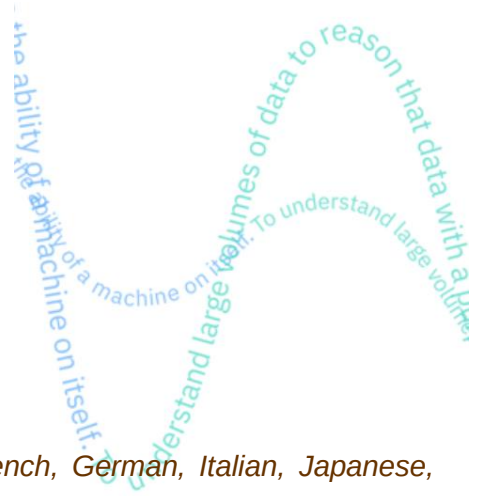
- *This easy-to-use service uses machine intelligence to combine information about grammar and language structure with knowledge of the composition of an audio signal to generate an accurate transcription.*
- *It uses IBM's speech recognition capabilities to convert speech in multiple languages into text.*
- *The transcription of incoming audio is continuously sent back to the client with minimal delay, and it is corrected as more speech is heard.*
- *Additionally, the service now includes the ability to detect one or more keywords in the audio stream.*
- *The service is accessed via REST API, a WebSocket connection and Asynchronous HTTP.*

Link: https://www.ibm.com/products/speech-to-text

https://cloud.ibm.com/docs/speech-to-text/index.html

https://cloud.ibm.com/apidocs/speech-to-text

Demo: https://www.ibm.com/demos/live/speech-to-text/self-service



Available Languages *(to be checked for updates…):*

Arabic, Portuguese (Brazil), Chinese (Mandarin), Dutch, English (Australian, US and UK), French, German, Italian, Japanese, Korean, Spanish (Castilian, Latin and North American)…

IBM Watson

Speech Text to Speech

Text to Speech *designed for streaming low-latency synthesis of audio from written text. The service synthesizes natural-sounding speech from input text in a variety of languages and voices that speak with appropriate cadence and intonation.*

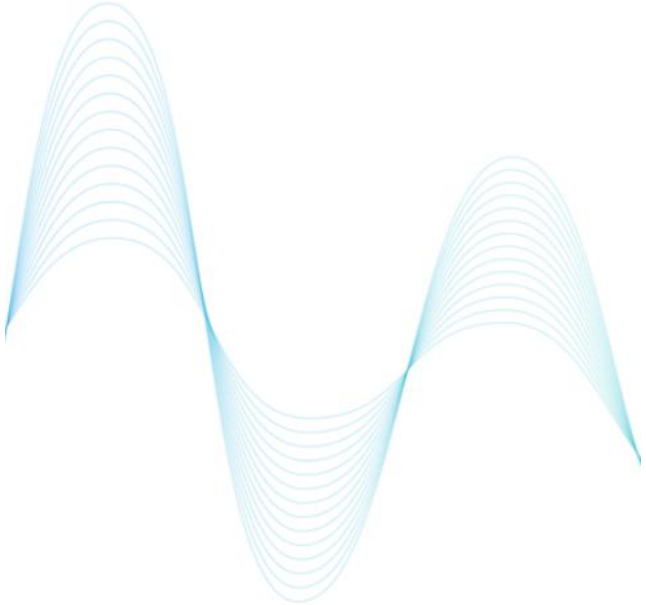
- *Watson Text to Speech provides REST and WebSocket interfaces to synthesize speech audio from an input of plain text.*
- *Multiple voices, both male and female, are available.*
- *The Text to Speech service now enables developers to control the pronunciation of specific words.*

Link: https://www.ibm.com/products/speech-to-text

https://cloud.ibm.com/docs/text-to-speech/index.html

https://cloud.ibm.com/apidocs/text-to-speech

Demo: https://www.ibm.com/demos/live/tts-demo/self-service



***Available Languages** (to be checked for updates…):*

Arabic, Portuguese (Brazil), Chinese (Mandarin), Dutch, English (Australian, US and UK), French, German, Italian, Japanese, Korean, Spanish (Castilian, Latin and North American)…

IBM Watson

Deprecated

Vision Visual Recognition

Visual Recognition *finds meaning in visual content. Allows to:*

- *Analyze images for scenes, objects, faces, colors, food, text, explicit content, and other subjects that can give you insights into your visual content.*
- *Choose a default model off the shelf, or create your own custom classifier.*
- *Find similar images within a collection.*
- *Detect faces, find all faces present in the image and receive their age and gender estimates*
- *Develop smart applications that analyze the visual content of images or video frames to understand what is happening in a scene.*

Link: https://www.ibm.com/fi-en/cloud/watson-visual-recognition

https://cloud.ibm.com/catalog/services/visual-recognition

https://cloud.ibm.com/docs/visual-recognition/getting-started.html

https://cloud.ibm.com/apidocs/visual-recognition/visual-recognition-v3

Take a look: https://www.ibm.com/products/maximo/remote-monitoring

Demo: https://visual-recognition-code-pattern.ng.bluemix.net/

***Features**:*

- ***General Classification*** *Generate class keywords that describe the image. Use your own images, or extract relevant image URLs from publicly accessible webpages for analysis.*
- ***Visual Training*** *Create custom, unique visual classifiers. Use the service to recognize custom visual concepts that are not available with general classification.*
- ***Face Detection*** *Detect human faces in the image. This service also provides a general indication of age range and gender of faces.*
- ***Similar Image Search (BETA)*** *Upload and search through image collections to find visually similar images.*

IBM Watson

Deprecated

Language Language Translator

Language Translator *translates text from one language to another…*

With Watson Language Translator you can:

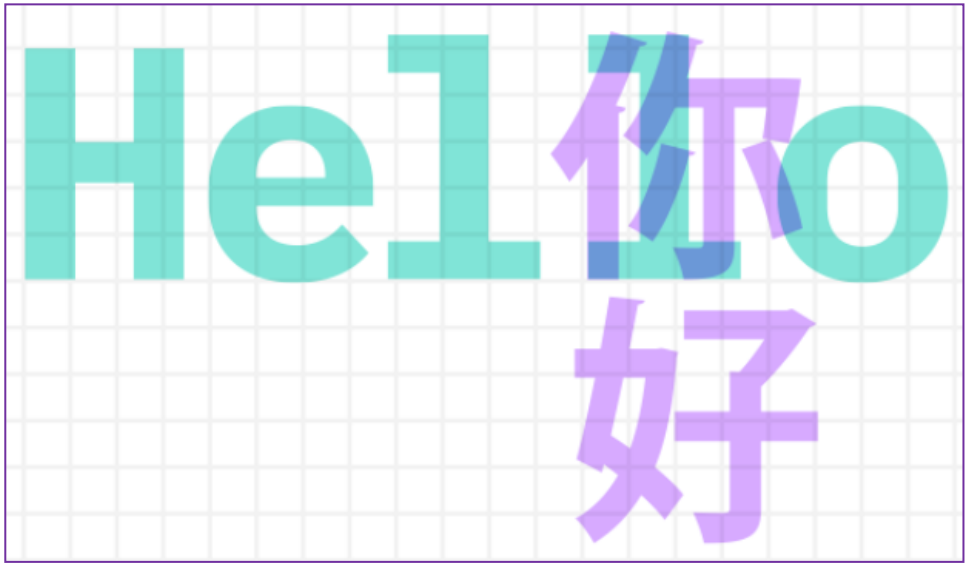
- *dynamically translate news, patents, or conversational documents;*
- *instantly publish content in multiple languages;*
- *allow your, for example, French-speaking staff to instantly send emails in English.*

Link: https://www.ibm.com/cloud/watson-language-translator

https://cloud.ibm.com/docs/language-translator/getting-started.html

https://cloud.ibm.com/apidocs/language-translator

Demo: https://www.ibm.com/demos/live/watson-language-translator/self-service



IBM Watson

Is a part of NLU service at the moment

Language Natural Language Classifier

Natural Language Classifier *interprets and classifies natural language with confidence. Service applies cognitive computing techniques to return the best matching classes for a sentence or phrase.*

For example:

- *you submit a question and the service returns keys to the best matching answers or next actions for your application.*
- *you create a classifier instance by providing a set of representative strings and a set of one or more correct classes for each training.*
- *after training, the new classifier can accept new questions or phrases and return the top matches with a probability value for each match.*

Link: https://www.ibm.com/cloud/watson-natural-language-classifier

https://cloud.ibm.com/docs/natural-language-classifier/getting-started.html#natural-language-classifier

https://cloud.ibm.com/apidocs/natural-language-classifier

Demo: https://natural-language-classifier-demo.ng.bluemix.net/

Intended uses:

- *Tackle common questions from your users that are typically handled by a live agent.*
- *Classify SMS texts as personal, work, or promotional*
- *Classify tweets into a set of classes, such as events, news, or opinions.*
- *Based on the response from the service, an application can control the outcome to the user.*

IBM Watson

Is a part of NLU service at the moment

Empathy Tone Analyzer

Tone Analyzer *leverages cognitive linguistic analysis to identify a variety of tones at both the sentence and document level. This insight can then used to refine and improve communications.*

It detects three types of tones, including:

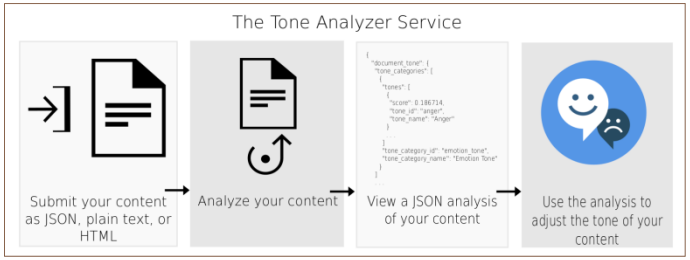
- *emotion (anger, disgust, fear, joy and sadness),*
- *social propensities (openness, conscientiousness, extroversion, agreeableness, and emotional range),*
- *and language styles (analytical, confident and tentative) from text.*

Link: https://www.ibm.com/watson/services/tone-analyzer/

https://cloud.ibm.com/docs/tone-analyzer/index.html

https://cloud.ibm.com/apidocs/tone-analyzer

Demo: https://tone-analyzer-demo.ng.bluemix.net/



Intended uses:

- *Personal and business communications - Anyone could use the Tone Analyzer service to get feedback about their communications, which could improve the effectiveness of the messages and how they are received.*
- *Message resonance - optimize the tones in communication to increase the impact on your audience*
- *Digital Virtual Agent for customer care - If a human client is interacting with an automated digital agent, and the client is agitated or angry, it is likely reflected in the choice of words they use to explain their problem. An automated agent could use the Tone Analyzer Service to detect those tones, and be programmed to respond appropriately to them.*
- *Self-branding - Bloggers and journalists could use the Tone Analyzer Service to get feedback on their tone and fine-tune their writing to reflect a specific personality or style.*

Is a part of NLU service at the moment

IBM Watson

Empathy Personality Insights

Personality Insights *uncover a deeper understanding of people's personality characteristics, needs, and values to drive personalization.*

- *Extracts and analyzes a spectrum of personality attributes to help discover actionable insights about people and entities, and in turn guides end users to highly personalized interactions.*
- *Processes linguistic input such a text messages, emails, posts, tweets to provide more customized answers and predict social behavior of the customers.*

Link: https://www.ibm.com/fi-en/cloud/watson-personality-insights

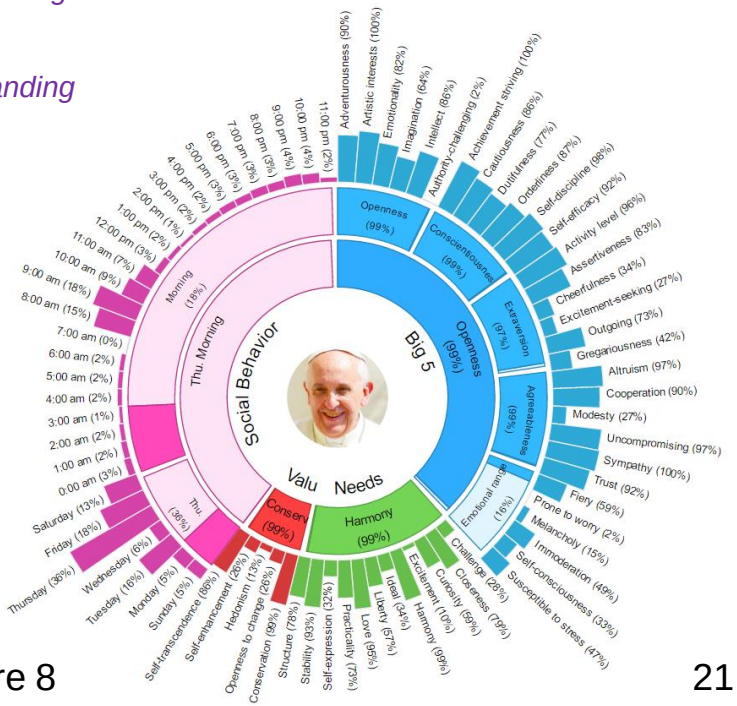
https://cloud.ibm.com/docs/personality-insights/getting-started.html#getting-started-tutorial

https://cloud.ibm.com/apidocs/personality-insights

Take a look: https://www.ibm.com/cloud/watson-natural-language-understanding

The service is based on psychology of language in combination with data analytics algorithms. The algorithm is trying to extract personality characteristics from social media activity, providing three models of personality:

- **Big Five** *(Agreeableness, Conscientiousness, Extraversion, Emotional Range, Openness)*
- **Needs** *(Excitement, Harmony, Curiosity, Ideal, Closeness, Self-expression, Liberty, Love, Practicality, Stability, Challenge, Structure)*
- **Values** *(Self-transcendence / Helping others, Conservation / Tradition, Hedonism / Taking pleasure in life, Self-enhancement / Achieving success, Open to change / Excitement)*



IBM Watson

Language Knowledge Natural Language Understanding

Natural Language Understanding *is a collection of APIs that offer text analysis through natural language processing. The APIs can analyze text and help to understand its sentiment, keywords, entities, high-level concepts and more.*

- *Clients can train their own custom model in a specific domain using Watson Knowledge Studio.*
- *Business can use Watson abilities to understand the content and context of text in webpages, news articles and blogs.*

Link: https://www.ibm.com/products/natural-language-understanding

https://cloud.ibm.com/docs/natural-language-understanding/getting-started.html

https://cloud.ibm.com/apidocs/natural-language-understanding/

Demo: https://www.ibm.com/demos/live/natural-language-understanding/self-service

***Available functions**:*

- *Entity Extraction*
- *Sentiment Analysis*
- *Emotion Analysis*
- *Keyword and Metadata Extraction*
- *Concept Tagging*
- *Relations and Semantic Roles Extraction*
- *Taxonomy (Category) Classification*

***Available Languages**: (currently only English language is supported across all features, others have limitations)*
Arabic, Chinese (Simplified), Czech, Danish, Dutch, English, Finnish, French, German, Hebrew, Italian, Japanese, Korean, Norwegian, Polish, Portuguese, Russian, Slovak, Spanish, Swedish…

IBM Watson

Knowledge Discovery

Discovery *helps to rapidly build a cognitive search and content analytics engine. Service unlocks hidden value in data to find answers, monitor trends and surface patterns with the world's most advanced cloud-native insight engine.*

- *Extract value from unstructured data by converting, normalizing, enriching it.*
- *Securely explore your proprietary content as well as free and licensed public content.*
- *Apply additional enrichments such as concepts, relations, and sentiment through Natural Language Understanding (NLU).*

Link: https://www.ibm.com/products/watson-discovery

https://cloud.ibm.com/docs/discovery-data?topic=discovery-data-getting-started

https://cloud.ibm.com/apidocs/discovery-data

Demo: https://www.ibm.com/demos/live/discovery-expert-assist/self-service

https://cloud.ibm.com/developer/watson/dashboard

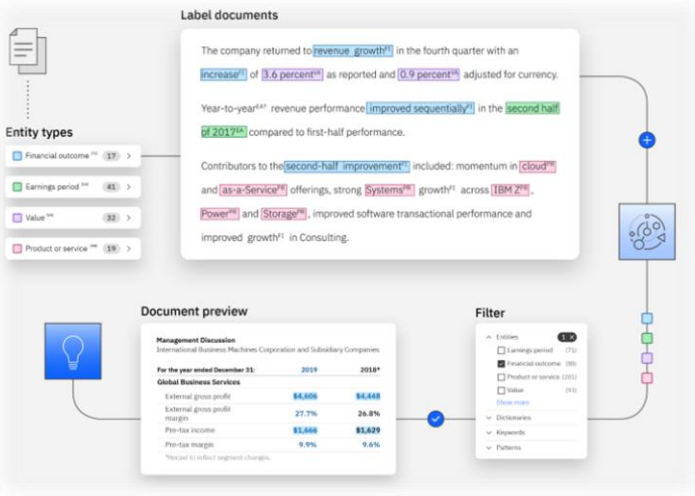
Intended uses:

Help a customer service team decrease support times, find hidden insights on customer pain points, behaviors and needs, and improve agent productivity

Enable research teams to extract the latest studies and insights from journals, newsletters, press releases and blogs to provide consolidated industry and domain-specific insights

Empower field service teams to locate hard-to-find answers in manuals, FAQ documents, collaboration sites and more

Augment a chat bot or other application created with Watson Conversation to provide possible responses to complex questions without modeling intents



IBM Watson

Deprecated

Knowledge Discovery News

Discovery News *(included with Discovery) is an indexed dataset that is pre-enriched with the following cognitive insights:* ***Keyword Extraction****,* ***Entity Extraction****,* ***Semantic Role Extraction****,* ***Sentiment Analysis****,* ***Relation Extraction****, and* ***Category Classification****.*

- *Watson also identifies important meta-information – like authors, publication and crawl dates*
- *Historical search is available for the past 60 days of news data*

Watson Discovery News is updated continuously with new articles (the news sources vary by language, so the query results for each collection will not be identical):

- *Discovery News English is updated with approximately 300,000 new articles daily;*
- *Discovery News Spanish is updated with approximately 60,000 new articles daily;*
- *Discovery News Korean with 10,000 new articles daily.*

Link: https://www.ibm.com/cloud/watson-discovery/

https://www.ibm.com/watson/services/discovery-news/

https://cloud.ibm.com/docs/discovery?topic=discovery-watson-discovery-news

Intended uses:

- ***News alerting*** *- Create news alerts by taking advantage of the support for entities, keywords, categories, and sentiment analysis to watch for both news and how it is perceived.*
- ***Event detection*** *- The subject/action/object semantic role extraction checks for terms/actions such as "acquisition", "election results", or "IPO".*
- ***Trending topics in the news*** *- Identify popular topics and monitor increases and decreases in how frequently they (or related topics) are mentioned.*

IBM Watson

Knowledge Watson Knowledge Studio

To become a subject matter expert in a given industry or domain, Watson must be trained. Use Watson Knowledge Studio to:

- *Teach IBM Watson® to extract meaningful information from unstructured text. Teach it the language of your domain with custom models that identify entities and relationships unique to your industry in unstructured text.*
- *Create a machine-learning model that understands the linguistic nuances, meaning, and relationships specific to your industry or to create a rule-based model that finds entities in documents based on rules that you define. Build your models in a collaborative environment designed for both developers and domain experts, without needing to write code.*
- *Use the models in IBM Watson Discovery, IBM Watson Natural Language Understanding and IBM Watson Explorer.*

***Watson Knowledge Studio** provides easy-to-use tools for annotating unstructured domain literature and uses those annotations to create a custom machine-learning model that understands the language of the domain.*

- *The accuracy of the model improves through iterative testing, ultimately resulting in an algorithm that can learn from the patterns that it sees and recognize those patterns in large collections of new documents.*
- *You can deploy the finished machine-learning model to other Watson cloud-based offerings and cognitive solutions to find and extract mentions of relations and entities, including entity co-references.*

Link: https://cloud.ibm.com/catalog/services/knowledge-studio
https://cloud.ibm.com/docs/watson-knowledge-studio/index.html?topic=watson-knowledge-studio-wks_tutintro
https://www.ibm.com/demos/live/watson-knowledge-studio/self-service/home

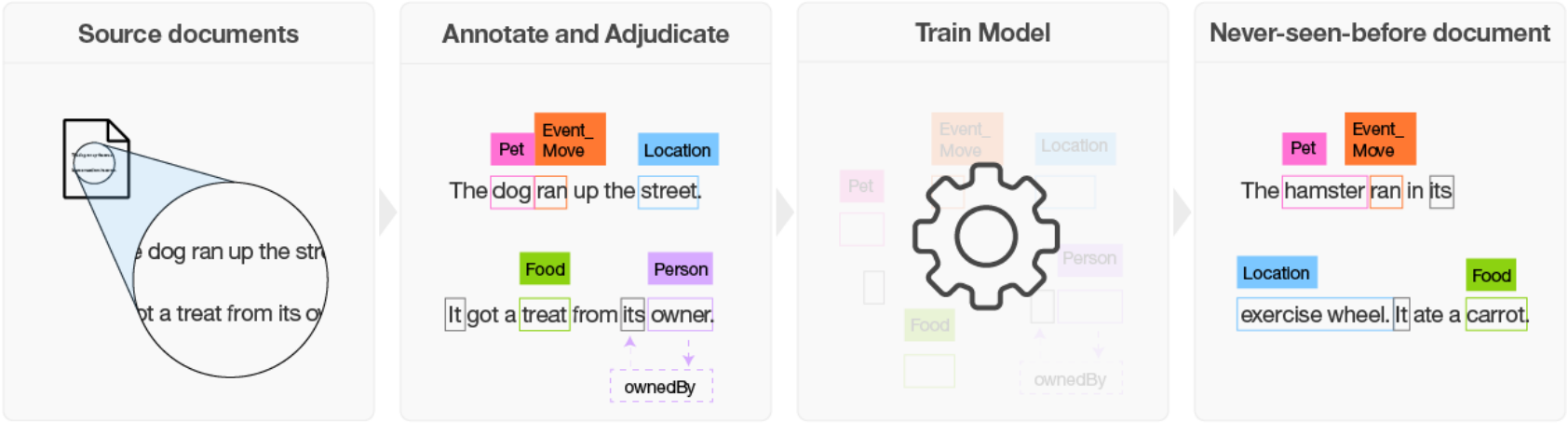
***Watson Knowledge Studio** provides a **Rules editor** that simplifies the process of finding and capturing common patterns in your documents as rules. You can then create a model that recognizes the rule patterns and deploy it for use in other services.*

Relevant links:
https://www.youtube.com/watch?v=r2xYHW0iyZM&list=PLZDyxLlNKRY9yM_Deg6GhTU1WS0-JLwyX
https://www.youtube.com/watch?v=byqpojcfDZM&list=PLZDyxLlNKRY9yM_Deg6GhTU1WS0-JLwyX&

IBM Watson

Knowledge Watson Knowledge Studio

Workflow:



1) *Based on a set of domain-specific source documents, the team creates a type system that defines entity types and relation types for the information of interest to the application that will use the model.*
2) *A group of two or more human annotators annotates a small set of source documents to label words that represent entity types, to identify relation types where the text identifies relationships between entity mentions, and to define co-references, which identify different mentions that refer to the same thing, that is, the same entity. Any inconsistencies in annotation are resolved, and one set of optimally annotated documents is built, which forms the ground truth.*
3) *The ground truth is used to train a model.*
4) *The trained model is used to find entities, relations, and co-references in new, never-seen-before documents.*

IBM Watson

Coding
Conversation

Watsonx Assistant

Watsonx Assistant *former (**Watson Assistant** and **Conversation**) service adds a natural language interface to the application to automate interactions with end users.*

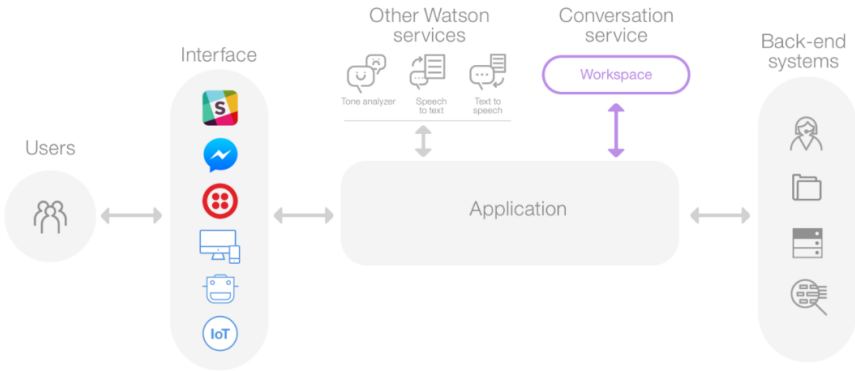
- *Quickly build and deploy chatbots and virtual agents across a variety of channels, including mobile devices, messaging platforms, and even robots.*
- *Train Watson Conversation service through an easy-to-use web application, designed so you can quickly build natural conversation flows between your apps and users, and deploy scalable, cost-effective solutions.*

Link: https://cloud.ibm.com/docs/assistant/getting-started.html#getting-started

https://www.ibm.com/products/watsonx-assistant

https://cloud.ibm.com/catalog/services/watsonx-assistant

- *Powered by large language models (LLMs) you can trust*
- *Intuitive user interface*
- *Build AI-powered voice agents and chatbots*
- *Deliver automated self-service support across all channels and touch-points*
- *Seamless integration to the tools that power your business*



*Also check the **IBM watsonx Code Assistant** - cloud service that leverages generative AI to accelerate code generation and increase developer productivity. It uses generative AI that is purpose-built for targeted use cases like Application Modernization and IT Automation.*

Link: https://cloud.ibm.com/catalog/services/ibm-watsonx-code-assistant#about

IBM Watson services

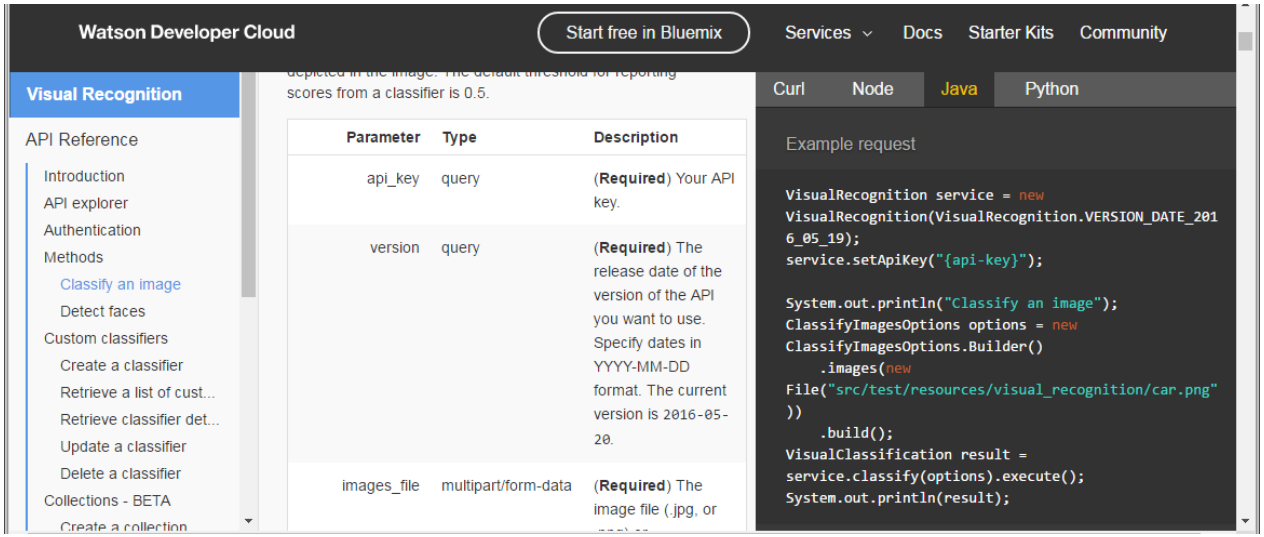
Getting started with Watson

Tutorial to begin with Watson services:

https://cloud.ibm.com/docs/watson/index.html

Service documentation

Check documentation of particular Watson service and find API reference with code examples in several programming languages…



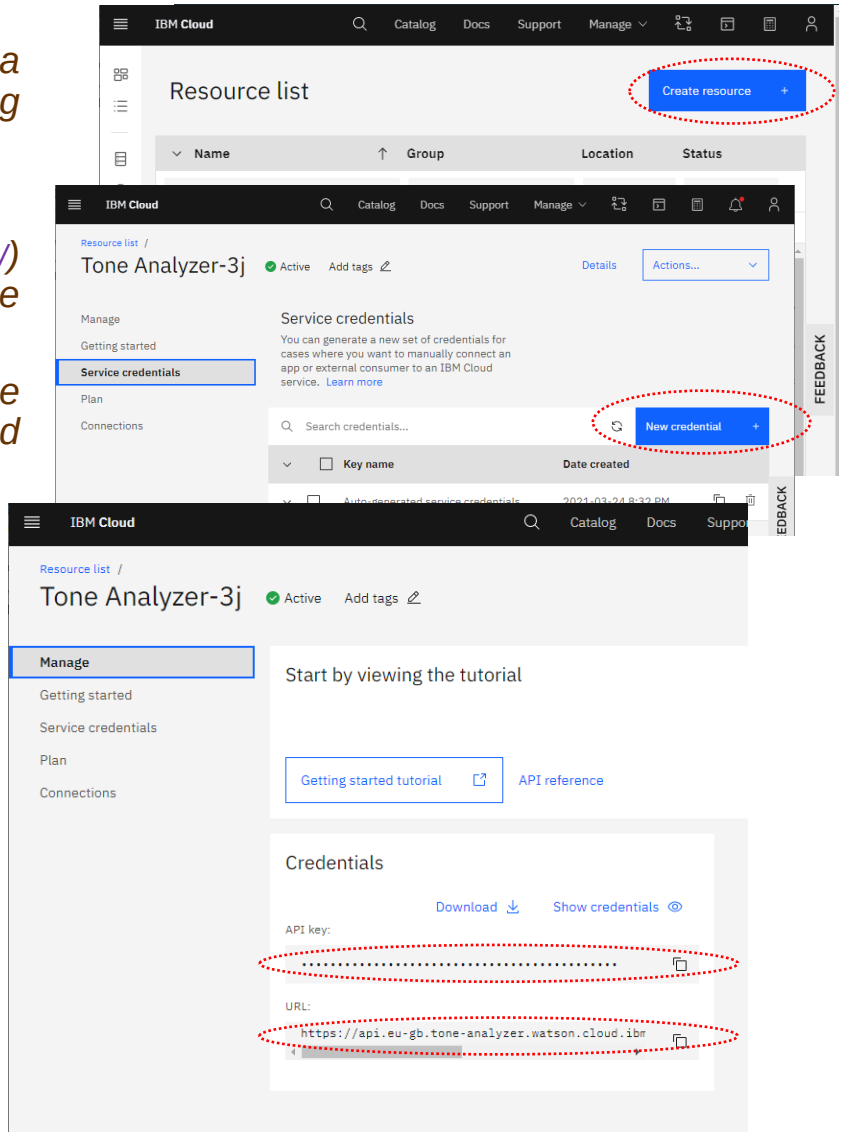
Starter Kits

Take a look at some examples of the apps that use Watson services:

https://cloud.ibm.com/developer/watson/dashboard

Service Use…

- *Create IBM Cloud Account, sign-in and select a service (e.g. Discovery) from Catalog (https://cloud.ibm.com/catalog/).*
- *CREATE the service instance from the service page.*
- *Go to the Dashboard (https://cloud.ibm.com/resources/) and select service from the list of created service instances.*
- *Select "Service credentials" from the menu on the left, create New credentials (or use auto-generated one) and view them…*



Service credentials

Get service url and apikey...

Access services from App

- *Use SDK for one of the supported languages*
- *Directly access service REST API via http client*

https://cloud.ibm.com/developer/watson/sdks-and-tools

https://cloud.ibm.com/docs?tab=api-docs&category=ai

IBM Watson

Service Use…

IBM Watson Natural Language Understanding *via Http Request…*

- *Pass services credentials (username = "apikey" and password = "<apikey value>") with **Basic Auth** in the request header*
- *Do **HTTP GET** request with a header: Content-Type: application/json*

```
<url>/v1/analyze?version=2022-04-07&text=I%20still%20have%20a%20dream%2C%20a%20dream%20deeply%20rooted%20in%20the%20American%20dream%20%E2%80%93%20one%20day%20this%20nation%20will%20rise%20up%20and%20live%20up%20to%20its%20creed%2C%20%22We%20hold%20these%20truths%20to%20be%20self%20evident%3A%20that%20all%20men%20are%20created%20equal.&features=keywords,entities&entities.emotion=true&entities.sentiment=true&keywords.emotion=true&keywords.sentiment=true
```

```
<url>/v1/analyze?version=2022-04-07&text=...&features=sentiment,keywords&keywords.sentiment=true
```

- *Do **HTTP POST** request with a header: Content-Type: application/json and Accept: application/json*

```
<url>/v1/analyze?version=2022-04-07
```

with corresponding text for analysis and request configuration in the request body:

```
{   "text": "IBM is an American multinational technology company headquartered in Armonk, New York, United States, with operations in over 170 countries.",
    "features": {   "entities": {      "emotion": true,
                                       "sentiment": true,
                                       "limit": 2   },
                    "keywords": {   "emotion": true,
                                    "sentiment": true,
                                    "limit": 2   }   }
}
```

IBM Watson

Deprecated

Service Use…

IBM Watson Tone Analyzer *via Http Request…*

- *Pass services credentials (username = "apikey" and password = "<apikey value>") with Basic Auth in the request header*
- *Do **HTTP GET** request with a header: Content-Type: application/json*

```
<url>/v3/tone?version=2017-09-21&text=Team,%20I%20know%20that%20times%20are%20tough!%20Product%20sales%20have%20been%20disappointing%20for%20the%20past%20three%20quarters.%20We%20have%20a%20competitive%20product,%20but%20we%20need%20to%20do%20a%20better%20job%20of%20selling%20it!
```

- *Do **HTTP POST** request with a header: Content-Type: application/json and Accept: application/json*

```
<url>/v3/tone?version=2017-09-21
```

with corresponding text for analysis in the request body:

```
{
  "text": "Team, I know that times are tough! Product sales have been disappointing for the past three quarters. We have a competitive product, but we need to do a better job of selling it!"
}
```

IBM Watson

Deprecated

Service Use…

IBM Watson Visual Recognition *via Http Request…*

- *Pass services credentials (username = "apikey" and password = "<apikey value>") with Basic Auth in the request header*
- *Do HTTP GET request with a header: Content-Type: application/json*

```
<url>/v3/classify? url=https://watson-developer-cloud.github.io/doc-tutorial-downloads/visual-recognition/fruitbowl.jpg&version=2018-03-19
```

Do HTTP POST request with a header: Content-Type: application/json

```
<url>/v3/classify?api_key={api-key}&version=2018-03-19
```

with corresponding image url and other configurations in the request body as a form-data:

```
url - https://watson-developer-cloud.github.io/doc-tutorial-downloads/visual-recognition/fruitbowl.jpg
threshold - 0.6
```

Google Cloud Platform

Google Cloud AI

*Google Cognitive Services are available as a set of **APIs** and **AI Platform**.* https://cloud.google.com/products#section-3

Pre-trained ML models

Speech-to-Text
Speech recognition and transcription supporting 125 languages.

Cloud Natural Language
Sentiment analysis and classification of unstructured text.

Recommendations AI
Deliver highly personalized product recommendations at scale.

Vision AI
Custom and pre-trained models to detect emotion, text, more.

Text-to-Speech
Speech synthesis in 220+ voices and 40+ languages.

Dialogflow
Conversation applications and systems development suite for virtual agents.

Cloud Translation
Language detection, translation, and glossary support.

Video AI
Video classification and recognition using machine learning.

Media Translation
Add dynamic audio translation directly to your content and applications.

Custom ML models

AI Platform
Platform for training, hosting, and managing ML models.

AutoML
Custom machine learning mod
and development.

TensorFlow

Generative AI on Vertex AI
Build, tune, and deploy foundation models on Vertex AI.

Vertex AI Search and Conversation
Generative AI apps for enterprise search and conversational AI.

Google Cloud Platform

Natural Language Processing

Natural Language API *provides natural language understanding technologies to developers revealing the structure and meaning of text by offering powerful machine learning models in an easy-to-use REST API.*

- *Sentiment Analysis and Entity Sentiment Analysis. Understand the overall sentiment expressed in a block of text and separate entities.*
- *Syntax Analysis. Extract tokens and sentences, identify parts of speech (PoS) and create dependency parse trees for each sentence.*
- *Entity Recognition. Identify entities and label by types such as person, organization, location, events, products and media.*
- *Multi-Language. Enables you to easily analyze text in multiple languages including English, Spanish, Japanese, Chinese (Simplified and Traditional), French, German, Italian, Korean and Portuguese.*
- *Content Classification. Classify documents in predefined 700+ categories.*

Link: https://cloud.google.com/natural-language/

https://cloud.google.com/natural-language/docs/

Google Cloud Platform

Translation

Translation API *can dynamically translate text between thousands of language pairs. The Cloud Translation API lets websites and programs integrate with the translation service programmatically.*

- *Programmatic Access. Translation API is accessible via a standard Google REST API.*
- *Text Translation. Translation API supports more than 100 languages and thousands of language pairs. You can send in HTML and receive HTML with translated text back.*
- *Language Detection. Detect a document's language and translate it using RESTful API.*
- *Continuous Updates. Behind the scenes, Translation API is learning from logs analysis and human translation examples. Existing language pairs improve and new language pairs come online at no additional cost.*

Link: https://cloud.google.com/translate/

https://cloud.google.com/translate/docs/

Google Cloud Platform

Deprecated

Will not be available after July 1, 2024

Translation

Media Translation (BETA) API *delivers real-time speech translation to your content and applications directly from your audio data. Leveraging Google's machine learning technologies, the API offers enhanced accuracy and simplified integration while equipping you with a comprehensive set of features to further refine your translation results. Improve user experience with low-latency streaming translation and scale quickly with straightforward internationalization.*

- ***Proven record of quality.*** *Bringing translation and speech recognition technologies together, Media Translation API provides state-of-the-art audio translation along with the features of the popular Translation API and Speech-to-Text API.*
- ***Seamless content translation.*** *Translate content directly from audio data. Media Translation API enhances the accuracy of interpretation by optimizing model integrations from audio to text and abstracts potential frictions that could be faced initiating multiple API calls. Simply make one API call, and Media Translation takes care of the rest.*
- ***Streaming translation at speed.*** *Stream translation output as you supply audio from a microphone or prerecorded audio file. Media Translation API minimizes the latency between input and translation results — enhancing user experience and enabling real-time engagement across languages and/or geographies.*

Link: https://cloud.google.com/media-translation

https://cloud.google.com/translate/media/docs

Google Cloud Platform

Speech

Speech-to-Text API *enables developers to convert audio to text by applying powerful neural network models in an easy-to-use API.*

- *Automatic Speech Recognition. Automatic Speech Recognition (ASR) powered by deep learning neural networking to power your applications like voice search or speech transcription.*
- *Global Vocabulary. Recognizes over 110 languages and variants with an extensive vocabulary.*
- *Streaming Recognition. Returns recognition results while the user is still speaking.*
- *Noise Robustness. Handles noisy audio from many environments without requiring additional noise cancellation.*

Text-to-Speech API *converts text into natural-sounding speech using an API powered by Google's AI technologies.*

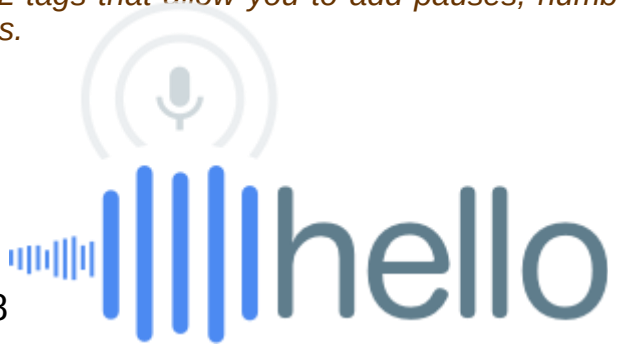
- *Custom Voice (beta). Train a custom speech synthesis model using your own audio recordings to create a unique and more natural-sounding voice for your organization.*
- *WaveNet voices. Take advantage of 90+ WaveNet voices built based on DeepMind's groundbreaking research to generate speech that significantly closes the gap with human performance.*
- *Voice tuning. Personalize the pitch of your selected voice, up to 20 semitones more or less from the default. Adjust your speaking rate to be 4x faster or slower than the normal rate..*
- *Text and SSML support. Customize your speech with SSML tags that allow you to add pauses, numbers, date and time formatting, and other pronunciation instructions.*

Link: https://cloud.google.com/speech-to-text

https://cloud.google.com/speech-to-text/docs

https://cloud.google.com/text-to-speech

https://cloud.google.com/text-to-speech/docs/



Google Cloud Platform

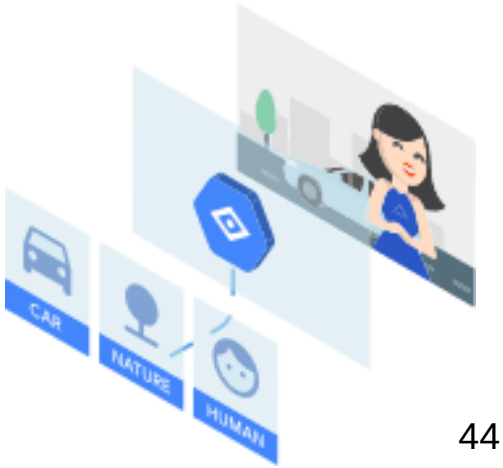
Vision

Vision API *enables developers to understand the content of an image by encapsulating powerful machine learning models in an easy-to-use REST API.*

- ***Label Detection***. *Detect broad sets of categories within an image, ranging from modes of transportation to animals.*
- ***Explicit Content Detection***. *Detect explicit content like adult content or violent content within an image.*
- ***Logo Detection***. *Detect popular product logos within an image.*
- ***Landmark Detection***. *Detect popular natural and man-made structures within an image.*
- ***Optical Character Recognition (OCR)***. *Detect and extract text within an image, with support for a broad range of languages, along with support for automatic language identification.*
- ***Face Detection***. *Detect multiple faces within an image, along with the associated key facial attributes like emotional state or wearing headwear.* **Facial Recognition is not supported.**
- ***Image Attributes***. *Detect general attributes of the image, such as dominant colors and appropriate crop hints.*
- ***Object Detection***.
- ***Web Detection***. *Search the Internet for similar images.*

Link: https://cloud.google.com/vision/

https://cloud.google.com/vision/docs/



Google Cloud Platform

Video Intelligence

Video Intelligence API *allows developers to use Google video analysis technology as part of their applications. The REST API enables users to annotate videos stored locally or in Google Cloud Storage with contextual information at the level of the entire video, per segment, per shot, and per frame.*

The API supports common video formats, including .MOV, .MPEG4, .MP4, and .AVI.

- *Label Detection. Detect entities within the video, such as "dog", "flower" or "car".*
- *Explicit Content Detection. Detect adult content within a video.*
- *Shot Change Detection. Detect scene changes within the video.*
- *Video Transcription. Automatically transcribes video content in English. More languages will be added in future releases.*
- *Object detection and tracking*
- *Logo and Text detection, extraction*
- *Face detection*
- *Pose estimation*

Link: https://cloud.google.com/video-intelligence/

https://cloud.google.com/video-intelligence/docs/

Google Cloud Platform

Dialogflow

Dialogflow API *is the lifelike conversational AI with state-of-the-art virtual agents. Dialogflow is a natural language understanding platform that makes it easy to design and integrate a conversational user interface into your mobile app, web application, device, bot, interactive voice response system, and so on.*

- *Visual flow builder. Reduce development time with interactive flow visualizations that allow builders to quickly see, understand, edit, and share their work. It also allows for easy collaboration across teams.*
- *Omnichannel implementation. Build once, deploy everywhere - in your contact centers and digital channels. Seamlessly integrate your agents across platforms including web, mobile, and messenger and with telephony partners such as Genesys, Avaya, and Cisco.*
- *Advanced AI. Improve your call/chat containment rate with the latest BERT-based natural language understanding (NLU) models that are capable of recognizing intent and context accurately and efficiently in more complex use cases..*
- *State-based data models. Reuse intents, intuitively define transitions and data conditions, and handle supplemental questions - allowing customers to deviate from the main topic, then gracefully return to the main flow.*
- *End-to-end management. Take care of all your agent management needs including CI/CD, analytics, experiments, and bot evaluation inside Dialogflow - you don't need any other custom software.*

Link: https://cloud.google.com/dialogflow/

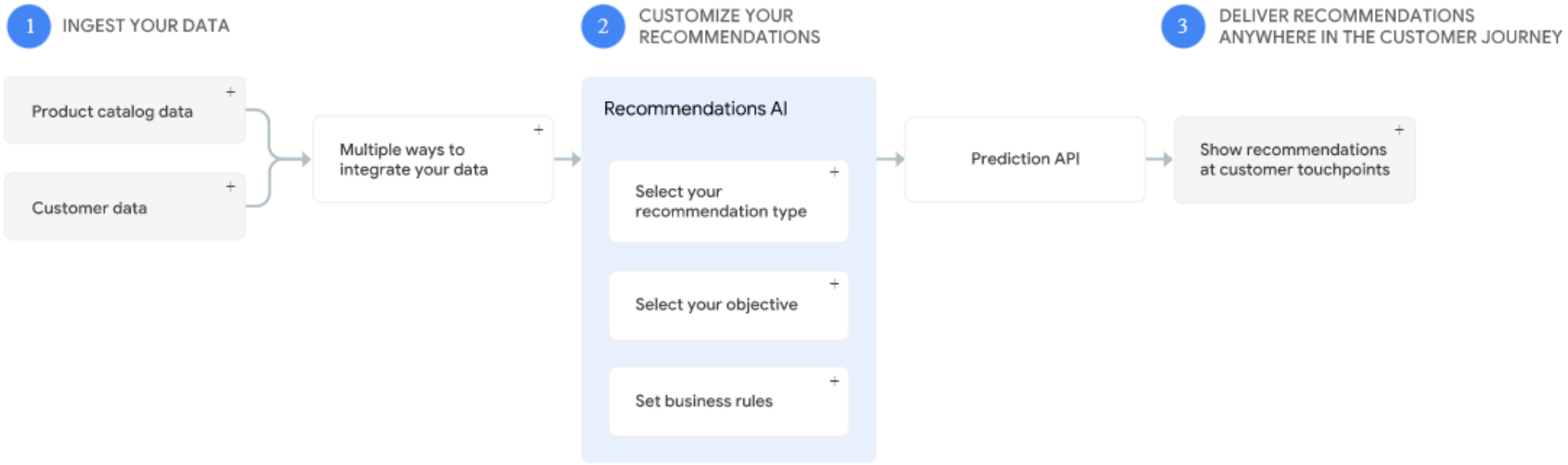
https://cloud.google.com/dialogflow/docs/

Google Cloud Platform

Recommendations

Recommendations API *draws on the experience (gained for years by Google delivering recommended content across flagship properties such as Google Ads, Google Search, and YouTube) and expertise in machine learning to deliver personalized recommendations that suit each customer's tastes and preferences across all your touchpoints.*

With Recommendations AI you can deliver high-performing recommendations to any customer touchpoint including your website, mobile experience, email, contact center, and more. Seamlessly deliver curated recommendations with low latency to customers anywhere in the world.



Link: https://cloud.google.com/recommendations

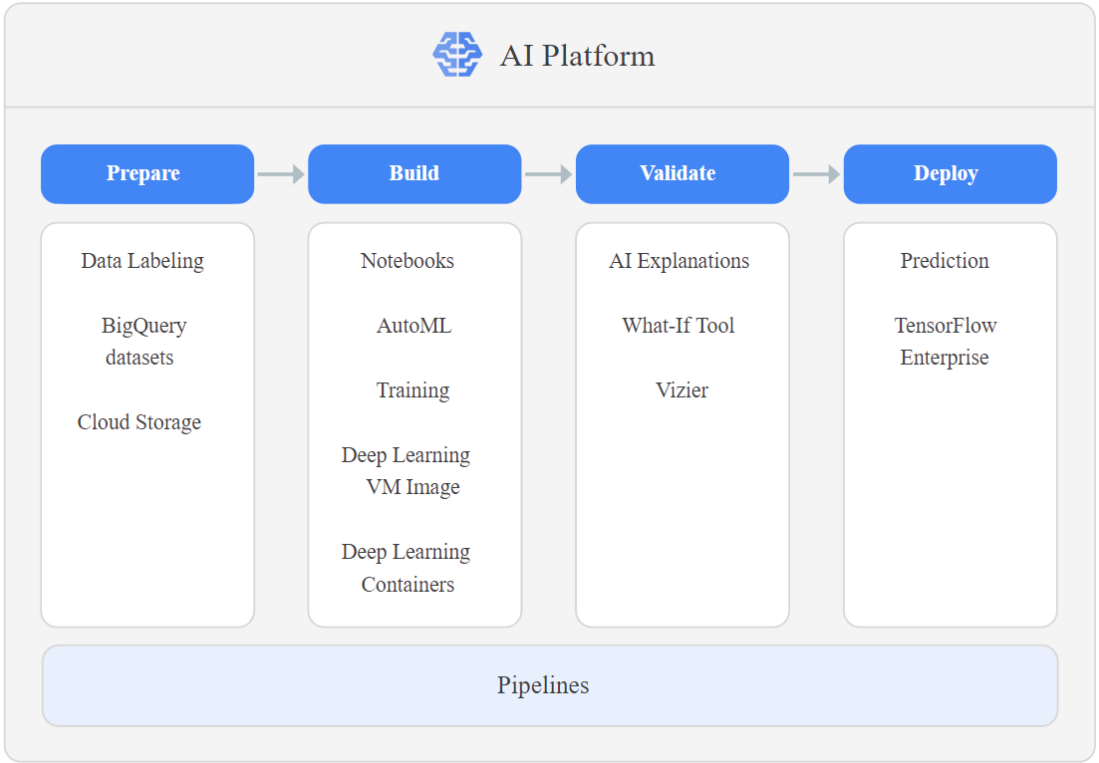
https://cloud.google.com/retail/docs

Google Cloud AI Platform

AI Platform *is a fully managed, end-to-end platform for data science and machine learning. AI Platform makes it easy for machine learning developers, data scientists, and data engineers to take their ML projects from ideation to production and deployment, quickly and cost-effectively. From data engineering to "no lock-in" flexibility, AI Platform's integrated tool chain helps you build and run your own machine learning applications.*

Link: https://cloud.google.com/ai-platform

https://cloud.google.com/ai-platform/docs



Google Cloud Platform

Vertex AI - Generative AI

Vertex AI *offers everything you need to build and use generative AI—from AI solutions, to Search and Conversation, to 130+ foundation models, to a unified AI platform.* https://cloud.google.com/vertex-ai

- *Vertex AI offers access to Gemini multimodal models from Google. In addition to Gemini, you also have access to Gemma, a family of lightweight, state-of-the-art open models built from the same research and technology used to create the Gemini models.*
- *Choose from the widest variety of models with first-party (PaLM API, Imagen, Codey), third-party (Anthropic's Claude 3), and open models (Gemma, Llama 2) in Model Garden. Use extensions to enable models to retrieve real-time information and trigger actions. Customize models to your use case with a variety of tuning options for Google's text, image, or code models.*
- *Open and integrated AI platform*
- *MLOps for predictive and generative AI*
- *Search and Conversation*

Customize and deploy generative models with ***Vertex AI Studio****. Vertex AI provides APIs for leading foundation models, and tools to rapidly prototype, easily tune models with your own data, and seamlessly deploy to applications.*

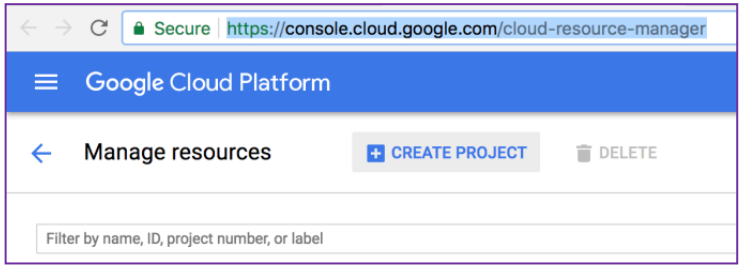
https://cloud.google.com/generative-ai-studio

Build generative AI search and conversations for customers and employees with ***Vertex AI Search and Conversation*** *(formerly Gen App Builder)— now generally available.* https://cloud.google.com/vertex-ai-search-and-conversation

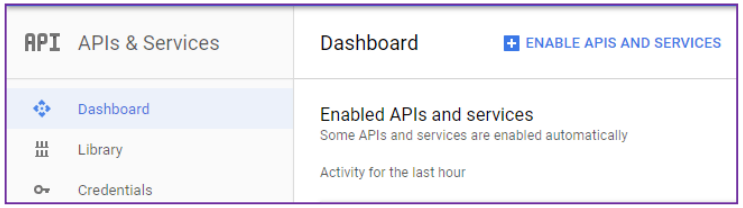
Google Cloud Platform

Service Use…

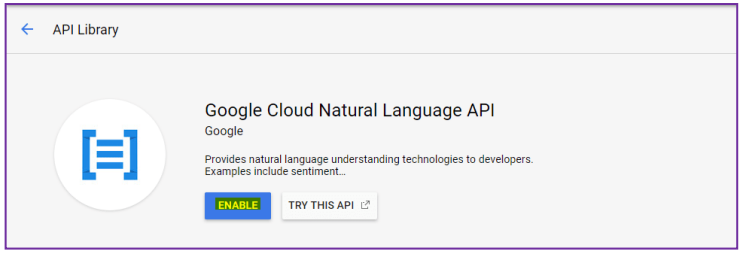
- *Create Google Account and sign-in to Google Cloud Platform Console (http://console.cloud.google.com/).*
- *Go to Manage Resources Page (https://console.cloud.google.com/cloud-resource-manager) and create a new project by click on **CREATE PROJECT**.*



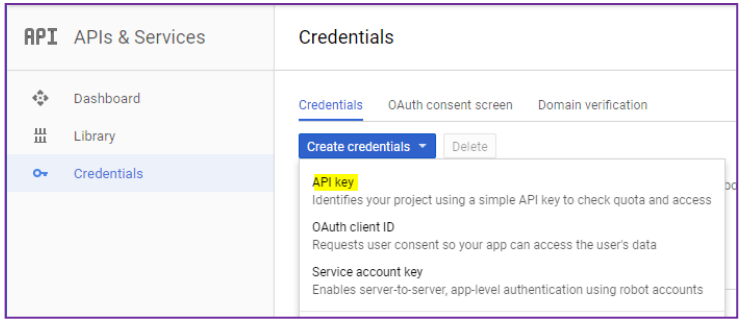
- *Go to Dashboard menu and click on **APIs & Services** and click on **Enable APIS AND SERVICES**.*



- *Search for and select corresponding service (e.g. Google Cloud Natural Language API) and click **ENABLE** to enable it.*



- *Generate a service API Key (to be used for service request). Navigate to the **Credentials section** of APIs & Services in the Cloud console, click **Create credentials** dropdown and choose **API Key**.*



Google Cloud Platform

Service Use…

Google Cloud Natural Language API *via Http Request…*

- *Service Endpoint:*

```
https://language.googleapis.com/v1/documents:analyzeSentiment?key=${API_KEY}
```

```
https://language.googleapis.com/v1/documents:analyzeEntities?key=${API_KEY}
```

…

- *Do **HTTP POST** request with a header: Content-Type: application/json and Accept: application/json*
- *Provide a text for analysis in the request body:*

```
{
  "document":{
    "type":"PLAIN_TEXT",
    "content":"The food was really bad at this restaurant. However, really enjoyed the drinks."
  },
  "encodingType": "UTF8"
}
```

Try also alternative text: Inception is one of the best movies of all time. I think everyone should watch it.

Google Cloud Platform

Service Use…

Google Cloud Vision API *via Http Request…*

- *Service Endpoint:*

```
https://vision.googleapis.com/v1/images:annotate?key=${API KEY}
```

Alternatively, API KEY could be sent as OAuth token via bearer token in an Authorization header (check documentation)…

- *Do **HTTP POST** request with a header: Content-Type: application/json and Accept: application/json*
- *Provide an image source for analysis in the request body:*

```
{
  "requests": [
    {
      "image": {
        "source": {
          "imageUri": "https://cloud.google.com/vision/docs/images/ferris-wheel.jpg"
        }
      },
      "features": [
        {
          "type": "LABEL_DETECTION"
        }
      ]
    }
  ]
}
```

LABEL_DETECTION
LOGO_DETECTION
IMAGE_PROPERTIES
FACE_DETECTION
LANDMARK_DETECTION
TEXT_DETECTION
DOCUMENT_TEXT_DETECTION
CROP_HINTS
WEB_DETECTION
SAFE_SEARCH_DETECTION

Microsoft Cognitive Services

Microsoft Cognitive Services

*Microsoft offers Artificial Intelligence productivity for every developer and every scenario with their **AI Platform (Azure AI)** including **AI Services**, **Infrastructure** and **Tools** https://azure.microsoft.com/en-us/solutions/ai/*

AI Services:

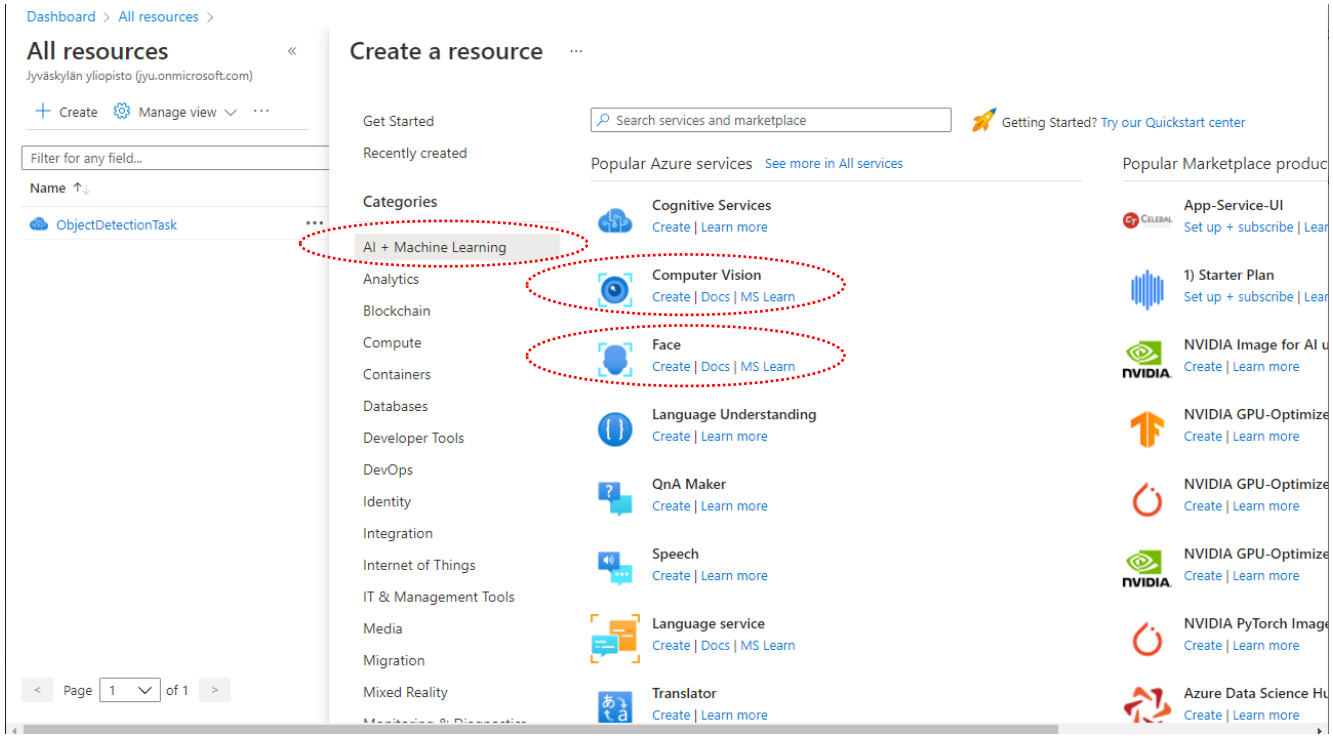
- **Cognitive Services** (https://azure.microsoft.com/en-us/services/cognitive-services/)
  - **Vision**. Image-processing algorithms to smartly identify, caption and moderate your pictures. (Computer Vision, Custom Vision, Face, Form Recognizer, Video Indexer)
  - **Speech**. Convert spoken audio into text, use voice for verification, or add speaker recognition to your app. (Speech to Text, Text to Speech, Speech Translation, Speaker Recognition)
  - **Language**. Allow your apps to process natural language with pre-built scripts, evaluate sentiment and learn how to recognize what users want. (Entity recognition, Language Understanding, QnA Maker, Text Analytics, Conversational language understanding, Translator)
  - **Decision**. Make smarter decisions faster. (Anomaly Detection, Content Moderator, Metrics Advisor, Personalizer)
  - **Azure OpenAI Service**. Apply large language models and generative AI to a variety of use cases. ChatGPT is now available in Azure OpenAI Service.

- **Machine Learning Services** (https://azure.microsoft.com/en-us/services/machine-learning/)
- **Azure Bot Service** (https://azure.microsoft.com/en-us/services/bot-services/)

Microsoft Cognitive Services

Service Use…

- *Create Microsoft Azure Account and sign-in to the Platform Console (https://portal.azure.com).*
- *Go to Dashboard menu and click All Resources.*
- *Create a resource, and select corresponding cognitive service from the list under "AI + Machine Learning" category (e.g. Face, Computer Vision, etc.).*
- *Check the service KEYs and Endpoint…*
- *Check API Console docs.*



Microsoft Cognitive Services

Service Use…

Microsoft Face API via Http Request…

- Do **HTTP POST** request with a header: Content-Type: application/json and Accept: application/json

```
<your resource Endpoint>/face/v1.0/detect?returnFaceId=true&returnFaceLandmarks=true&
returnFaceAttributes=age,gender,smile,facialHair,glasses,headPose
```

- Pass services **KEY** within the value of Ocp-Apim-Subscription-Key header
- Provide image url in the request body:

```
{
  "url": " https://people.com/thmb/O-xEgRssaPPDX-pp3-
2R_enUVdU=/750x0/filters:no_upscale():max_bytes(150000):strip_icc():focal(999x0:1001x2):format(webp)/elo
n-musk-secret-project-42-glass-house-071423-1-5505e11097414473a6b068b5a7d4fb6a.jpg"
}
```

Microsoft Computer Vision API via Http Request…

- Do **HTTP POST** request with a header: Content-Type: application/json and Accept: application/json

```
<your resource Endpoint>/computervision/imageanalysis:analyze?features=caption,read,objects&model-
version=latest&language=en&api-version=2023-02-01-preview
```

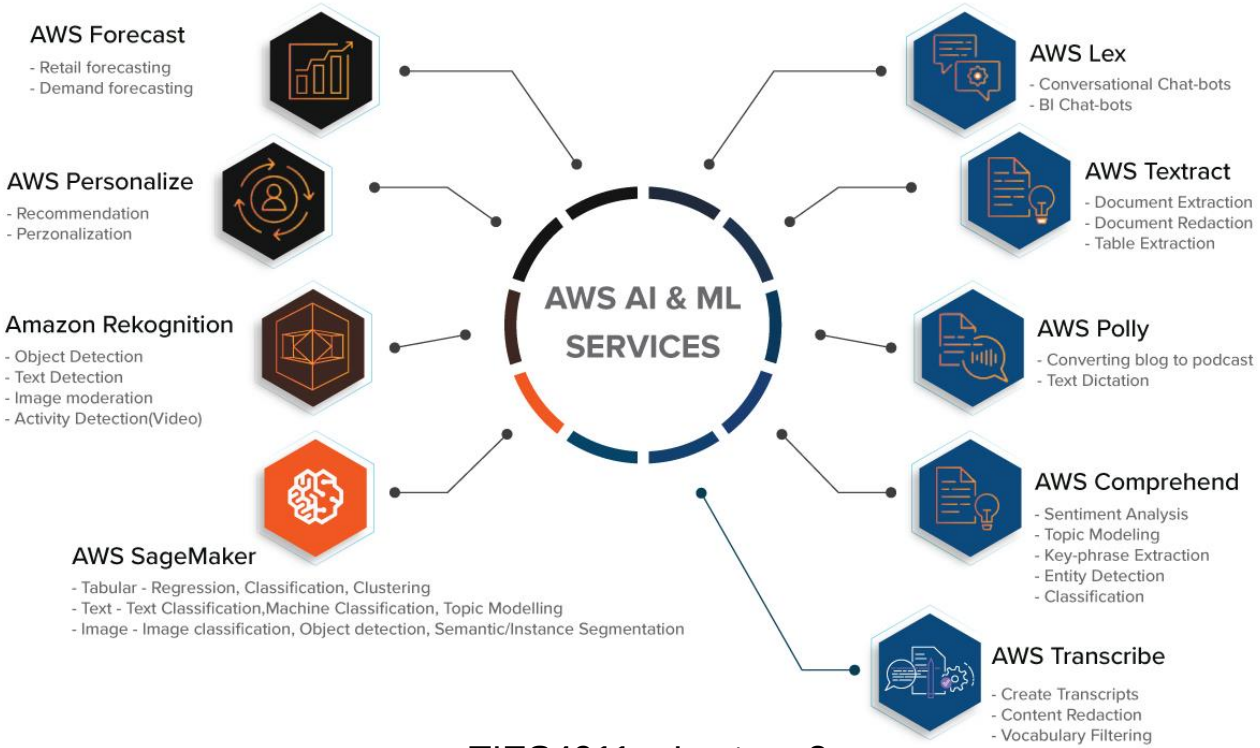
- Pass services **KEY** within the value of Ocp-Apim-Subscription-Key header
- Provide image url in the request body:

```
{
  "url": "https://image-tc.galaxy.tf/wijpeg-8sgtm7vii5u4pumz64ehpjtzy/zoo-negara-
4_wide.jpg?crop=0%2C46%2C894%2C503&width=1100"
}
```

AWS Cognitive Services

*Pre-trained **AWS AI Services** provide ready-made intelligence for your applications and workflows. AI Services easily integrate with your applications to address common use cases such as personalized recommendations, modernizing your contact center, improving safety and security, and increasing customer engagement. Due to use of the same deep learning technology that powers Amazon.com and AWS ML Services, you get quality and accuracy from continuously-learning APIs.*

https://aws.amazon.com/machine-learning/ai-services/



Computer Vision

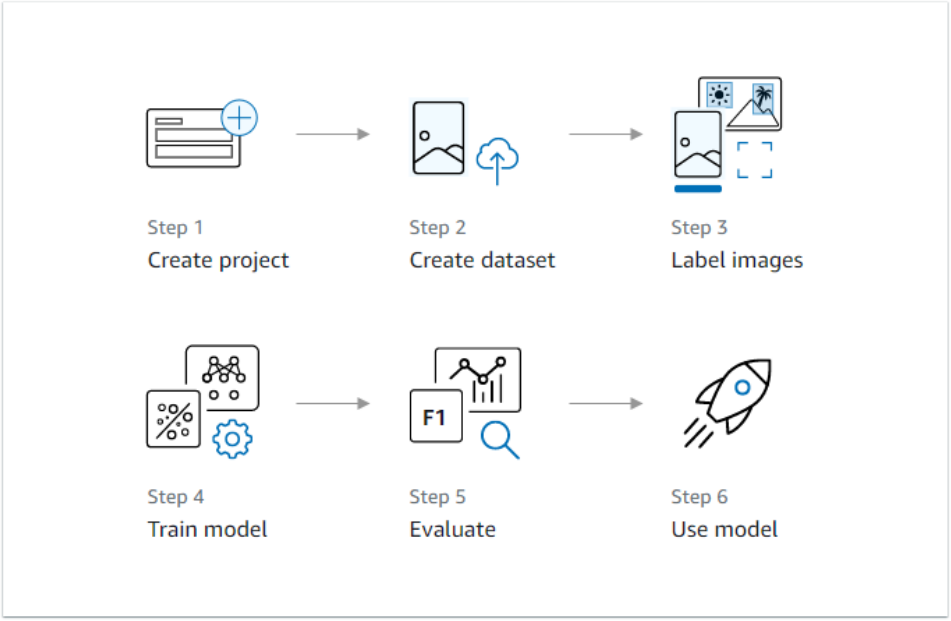
Amazon Rekognition *automates and lowers the cost of your image recognition and video analysis with machine learning.*

https://aws.amazon.com/rekognition/

Demo: https://us-east-2.console.aws.amazon.com/rekognition/home?region=us-east-2#/label-detection

With ***Amazon Rekognition Custom Labels*** *you can build own custom machine learning model to find objects and scenes unique to your business.*

https://us-east-2.console.aws.amazon.com/rekognition/custom-labels#/

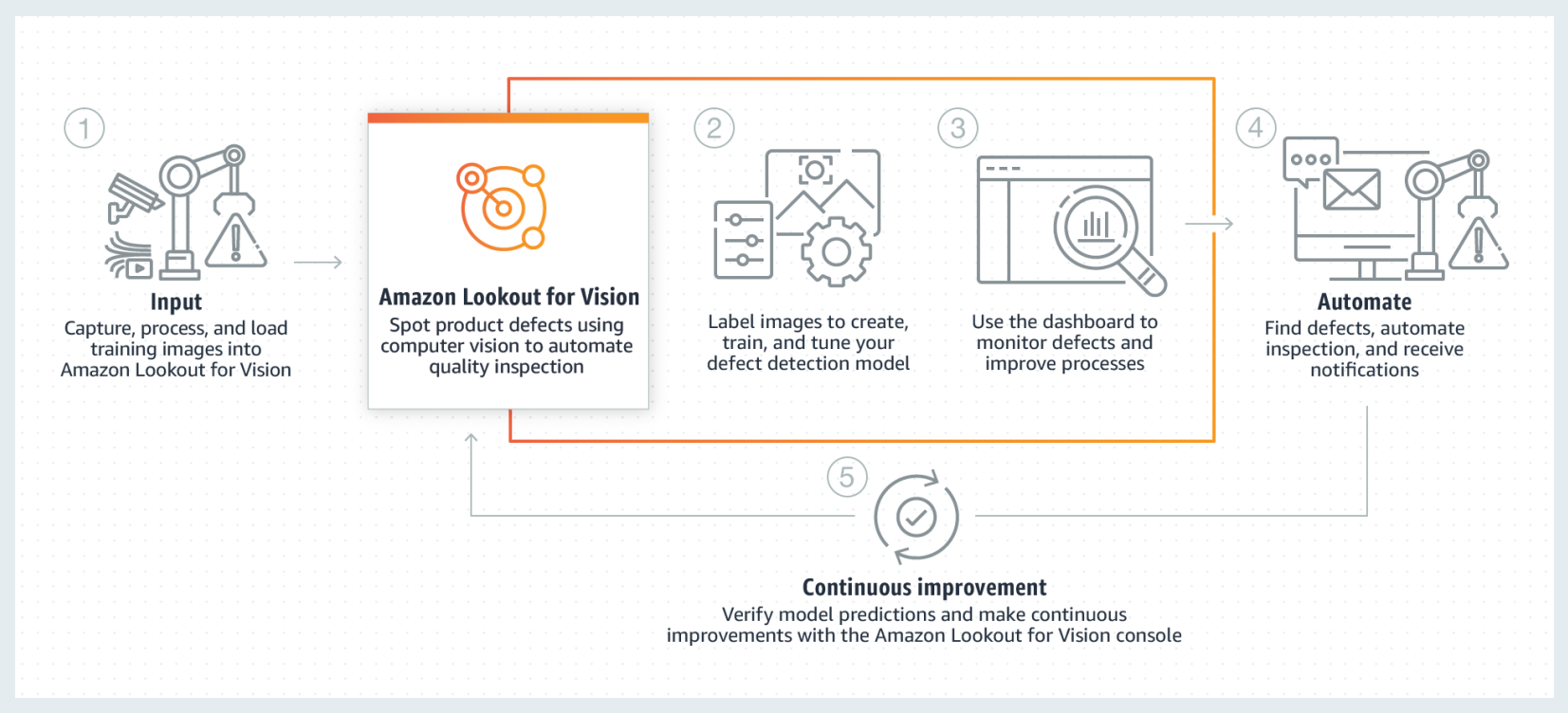


Computer Vision

Amazon Lookout for Vision *is a machine learning service that uses computer vision to automate visual inspection of product defects.*

https://aws.amazon.com/lookout-for-vision/

https://us-east-2.console.aws.amazon.com/lookoutvision#/

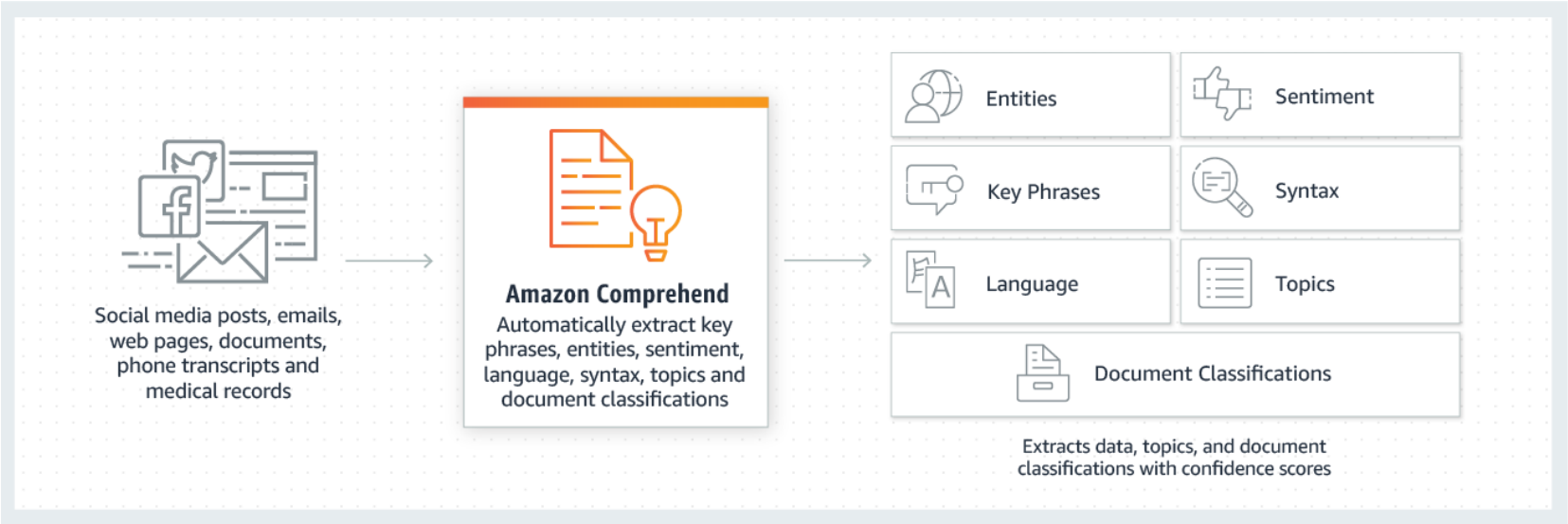


Text Processing

Amazon Comprehend *is a natural-language processing (NLP) service that uses machine learning to uncover valuable insights and connections in text.*
https://aws.amazon.com/comprehend/?nc2=h_ql_prod_ml_comp#

https://us-east-2.console.aws.amazon.com/comprehend/home?region=us-east-2#welcome

Demo: https://us-east-2.console.aws.amazon.com/comprehend/home?region=us-east-2#home



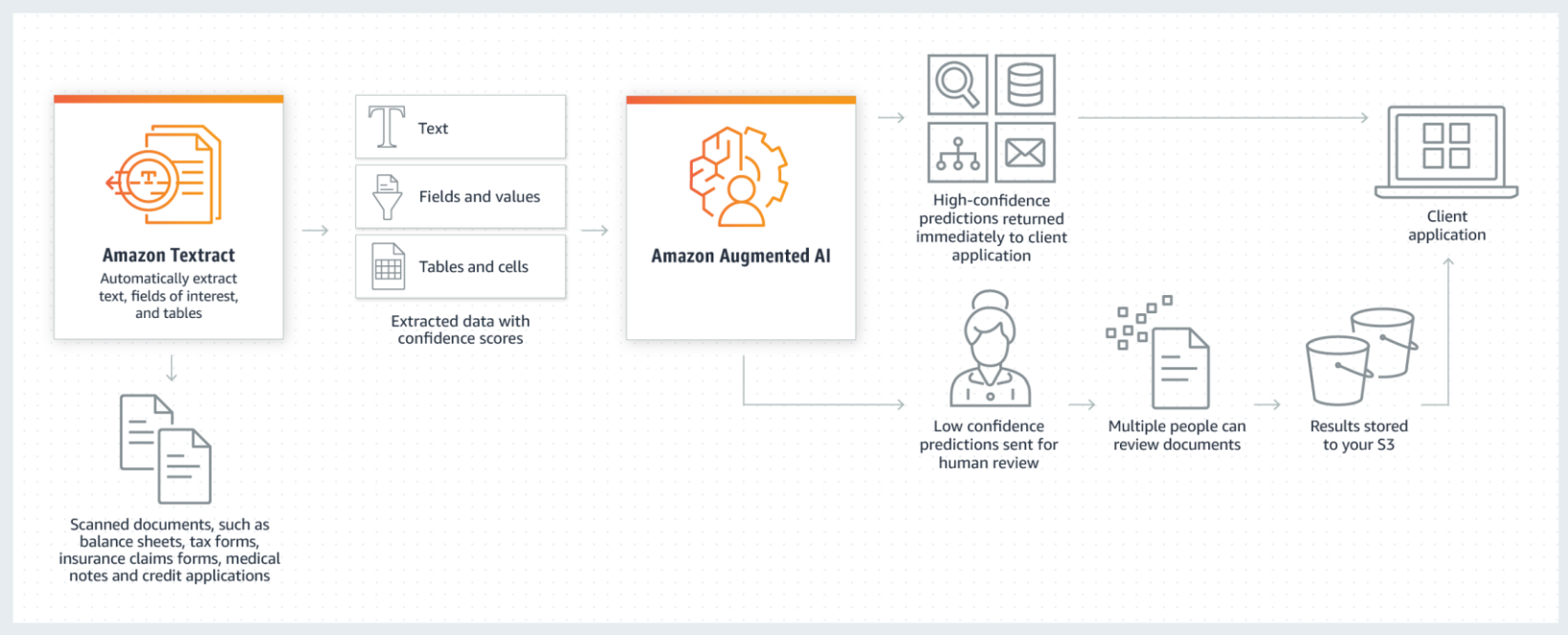
Text Processing

Amazon Textract *is a service that automatically detects and extracts text and data from scanned documents. It goes beyond simple optical character recognition (OCR) to also identify the contents of fields in forms and information stored in tables.*

https://aws.amazon.com/textract/

https://us-east-2.console.aws.amazon.com/textract/home?region=us-east-2#/

Demo: https://us-east-2.console.aws.amazon.com/textract/home?region=us-east-2#/demo



Speech

Amazon Polly *is a service that turns text into lifelike speech, enabling you to create applications that talk and build entirely new categories of speech-enabled products. Amazon Polly is a Text-to-Speech (TTS) service that uses advanced deep learning technologies to synthesize speech that sounds like a human voice.*

https://aws.amazon.com/polly/?nc2=h_ql_prod_ml_poll

https://us-east-2.console.aws.amazon.com/polly/home?region=us-east-2#

Demo: https://us-east-2.console.aws.amazon.com/polly/home/SynthesizeSpeech?region=us-east-2

Amazon Transcribe *provides high-quality and affordable speech-to-text transcription for a wide range of use cases.*

https://aws.amazon.com/transcribe/

https://us-east-2.console.aws.amazon.com/transcribe/home?region=us-east-2#welcome

Demo: https://us-east-2.console.aws.amazon.com/transcribe/home?region=us-east-2#realTimeTranscription

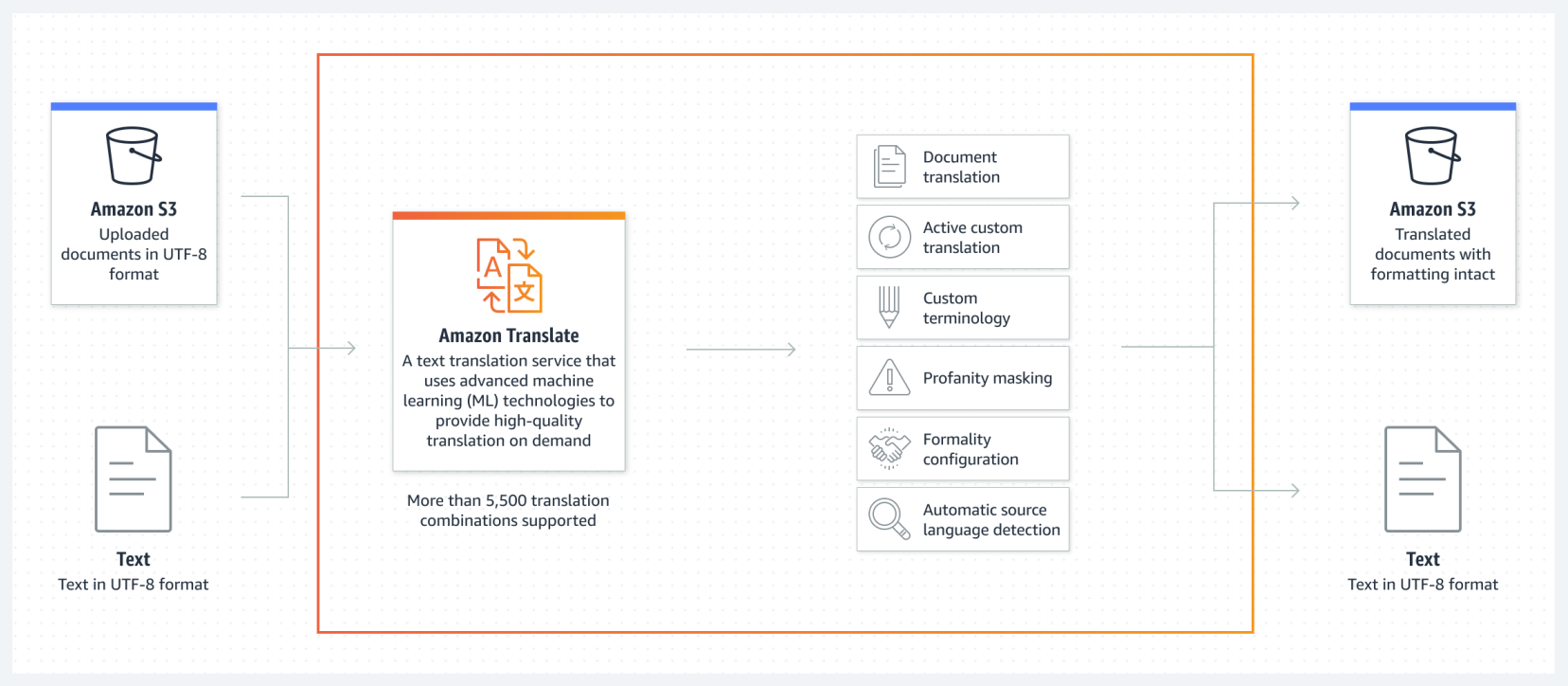
Translation

Amazon Translate *is a neural machine translation service that delivers fast, high-quality, affordable, and customizable language translation.*

https://aws.amazon.com/translate/

https://us-east-2.console.aws.amazon.com/translate/home?region=us-east-2#welcome

Demo: https://us-east-2.console.aws.amazon.com/translate/home?region=us-east-2#translation

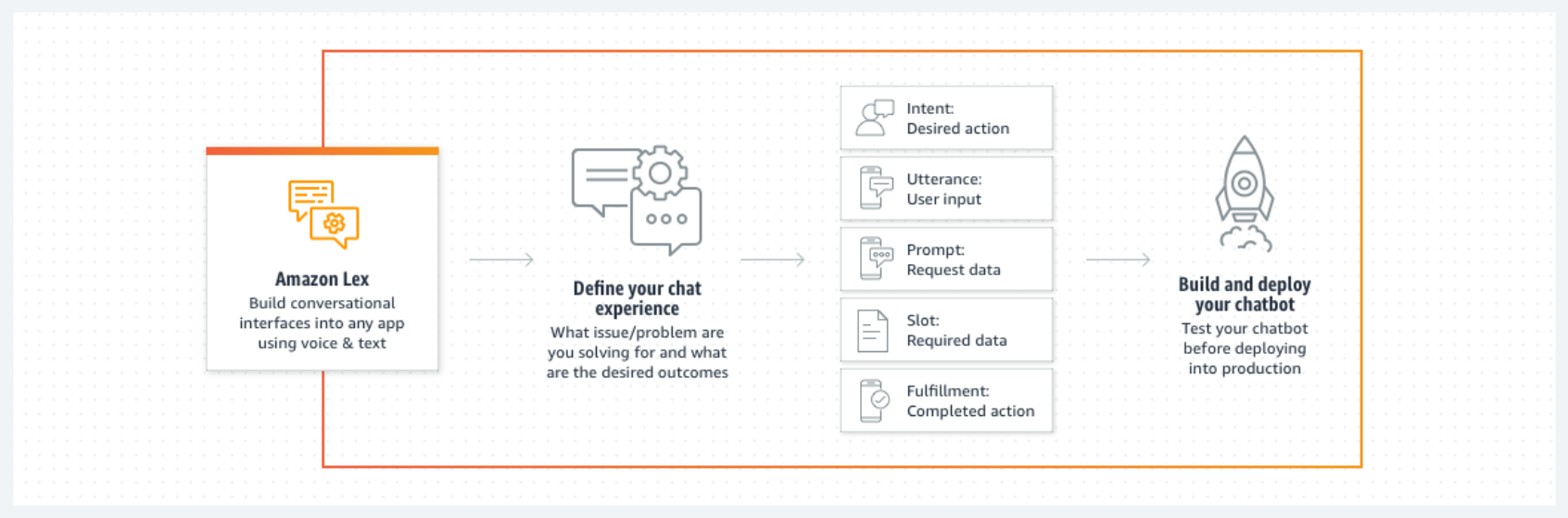


Conversation

Amazon Lex *is a fully managed artificial intelligence (AI) service with advanced natural language models to design, build, test, and deploy conversational interfaces in applications. It is a service for building conversational interfaces into any application using voice and text, enabling you to add sophisticated, natural language chatbots to your applications.*

https://aws.amazon.com/lex

https://us-west-2.console.aws.amazon.com/lexv2/home?region=us-west-2#welcome



Generation

AWS DeepComposer *gives developers a creative way to get started with machine learning. Get hands-on, literally, with a musical keyboard and the latest machine learning techniques, designed to expand your ML skills.*

https://aws.amazon.com/deepcomposer/

https://us-east-1.console.aws.amazon.com/deepcomposer/home?region=us-east-1#getStarted