

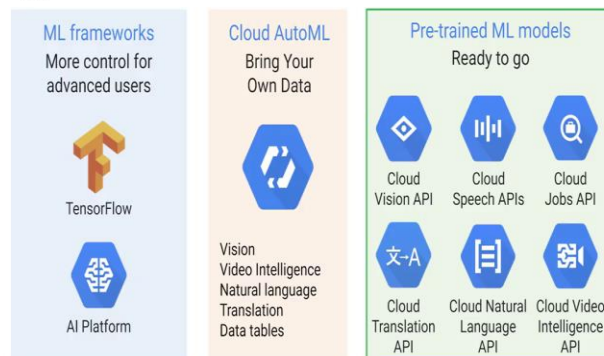
**Google Cloud Computing Foundation Course**  
**Evan Jones**  
**Technical Curriculum Developer**  
**Google Cloud**

**Lecture-81**  
**Googles Pre-trained ML APIs**

In the previous topic you learned how you can build custom ML models with minimal effort or ML expertise by leveraging the suite of ML products offered through cloud auto ML. When using cloud auto ML you define a domain-specific label the training dataset that is used to create the custom ML model you require if you do not need a domain specific data set however Google suite of pre trained ML api's might meet your needs. In this topic you will explore some of those api's and apply them through a series of labs.

**(Refer Slide Time: 00:35)**

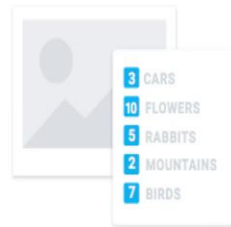
Access pre-trained ML APIs for common applications



API is like the vision API, natural language API or translation API or are already trained for common ML use cases like image classification they save you time and the effort of building curating and training a new data set so you can just jump ahead right to your predictions. For pre-trained models Google has already figured out a lot of those hard problems for you.

**(Refer Slide Time: 01:03)**

## Understand image content With Cloud Vision API



Detect and label



Let us explore some of these pretrained machine learning api's. let us start with the cloud vision API there are three major components that all roll up into this restful api and behind the scenes each of these are powered by many ML models and years of research. The first is detecting what an image is and then classifying it. The vision API picks out the dominant entity for example a car or a cat within an image from a broad set of object categories this allows you to easily detect broad sets of objects within your images.

Facial detection can detect when a face appears in photos along with the associated facial features such as the eyes the nose and the mouth placement and the likelihood of over eight attributes like joy and sorrow. Facial recognition however is not supported and Google does not store facial detection information on any Google server. You can use the API to easily build metadata on your image catalog enabling new scenarios like image based searches or recommendations.

**(Refer Slide Time: 02:09)**

Understand image content  
With Cloud Vision API



Extract text



Next are images with text like scanned documents or a sign out there on the vision API uses optical character recognition or OCR to extract the text of a wide range of languages into a selectable and searchable format.

**(Refer Slide Time: 02:26)**

Understand image content  
With Cloud Vision API



Identify entities

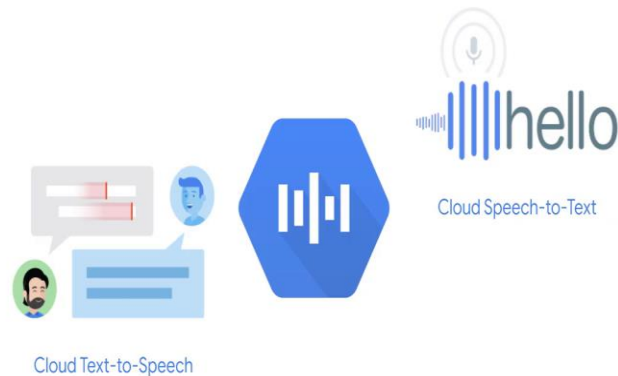


Lastly is a bit of intuition from the web and it uses the power of Google Image Search does the image contain entities that we know like the Eiffel Tower or a famous person and then landmark detection allows you to identify a popular natural and man-made structures along with the associated latitude and longitude of the landmark. And the logo detection allows you to identify product logos within an image.

You can build metadata on your image catalog extract text moderate offensive content or enable new marketing scenarios through image sentiment analysis. You can also analyze images uploaded in the request or integrate with image storage on Google Cloud Storage.

**(Refer Slide Time: 03:11)**

Convert speech to text and vice versa

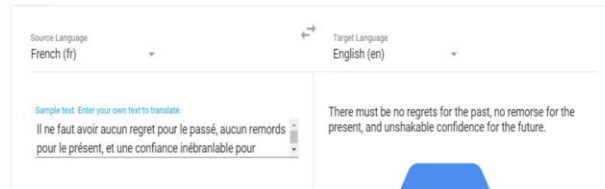


There are two api's that apply to speech the cloud text-to-speech API converts text into human-like speech in more than 180 voices across more than 30 languages and variants. It applies research in speech synthesis and Google's powerful neural networks to deliver high fidelity audio with this API you can create lifelike interactions with users that transform customer service device interaction and other applications.

The cloud speech-to-text API enables you to convert real-time streaming or pre-recorded audio into text the API recognizes a 120 languages in variants to support a global user base. You can enable voice command and control transcribe audio from call centers and so on.

**(Refer Slide Time: 04:04)**

## Dynamically translate between languages using the Cloud Translation API



The cloud translation API provides a simple programmatic interface for translating an arbitrary string into any supported language. The API is highly responsive so web sites and applications can integrate with it for fast dynamic translation of source text from the source language to a target language for example from French to English. Language detection is also available in the cases where the source language is unknown.

Let us look at a short video that shows how Bloomberg a global leader in business and financial data news and insight applies the cloud translation API to reach all of their customers regardless of language.

**(Refer Slide Time: 04:44)**

## Derive insights from unstructured text with the Cloud Natural Language API

The powerful pre-trained models of the Natural Language API let developers work with natural language understanding features including sentiment analysis, entity analysis, entity sentiment analysis, content classification, and syntax analysis. RESET

[See supported languages](#)

Entities    Sentiment    **Syntax**    Categories

Dependency     Parse label     Part of speech     Lemma     Morphology

det	amod	amod	nsubj	prep	det	nn	nn	potbj
The	powerful	pre-trained	models	of	the	Natural	Language	API
DET	ADJ	ADJ	model NOUN tense=PAST number=PLURAL	ADP	DET	NOUN number=SINGULAR proper=PROPER	NOUN number=SINGULAR proper=PROPER	NOUN number=SINGULAR proper=PROPER

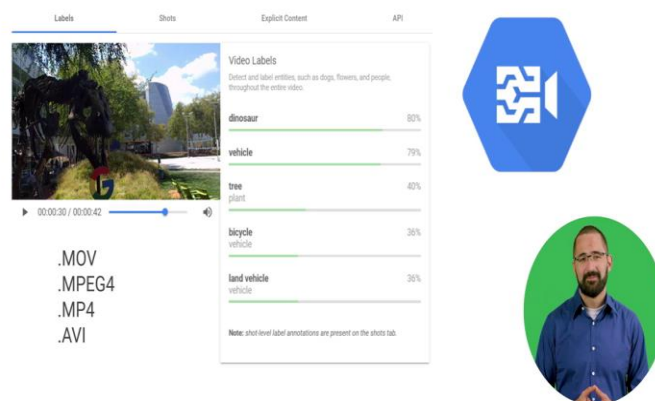


The cloud natural language API offers a variety of natural language understanding technologies it can do syntax analysis breaking down sentences into tokens identify the nouns verbs and adjectives and other parts of speech. And then also figure out the relationships among the words. It can also do entity recognition in other words it can parse text and flag mentions of people organizations locations events products and media.

Sentiment analysis allows you to understand customer opinions to find actionable product and UX insights.

**(Refer Slide Time: 05:22)**

Make your media more discoverable with the  
Cloud Video Intelligence API



The screenshot displays the Cloud Video Intelligence API interface. On the left, there is a video player showing a scene with a dinosaur. Below the player, supported video formats are listed: .MOV, .MPEG4, .MP4, and .AVI. On the right, the 'Video Labels' section lists detected entities with their confidence scores: dinosaur (80%), vehicle (79%), tree/plant (40%), bicycle/vehicle (36%), and land vehicle/vehicle (36%). A note at the bottom states: 'Note: shot-level label annotations are present on the shots tab.' To the right of the interface is a blue hexagonal logo with a white camera icon, and below it is a circular portrait of a man with a beard wearing a blue shirt.

The cloud video intelligence API supports the annotation of common video formats and allows you to use Google video analysis technology as part of your applications. This REST API enables you to annotate videos stored in Google Cloud storage with video and one frame per second contextual information. It helps you identify key entities that are the nouns that are within your video and when they occur. You can also use it to make your content more accessible searchable and discoverable.