

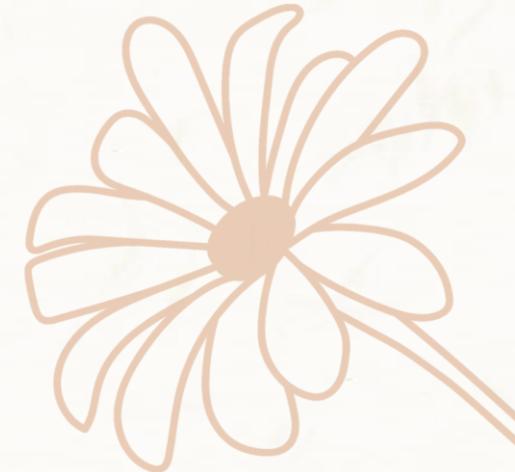
Coding Angels



Microsoft Technology Center
Bobo Lu



Cognitive Service



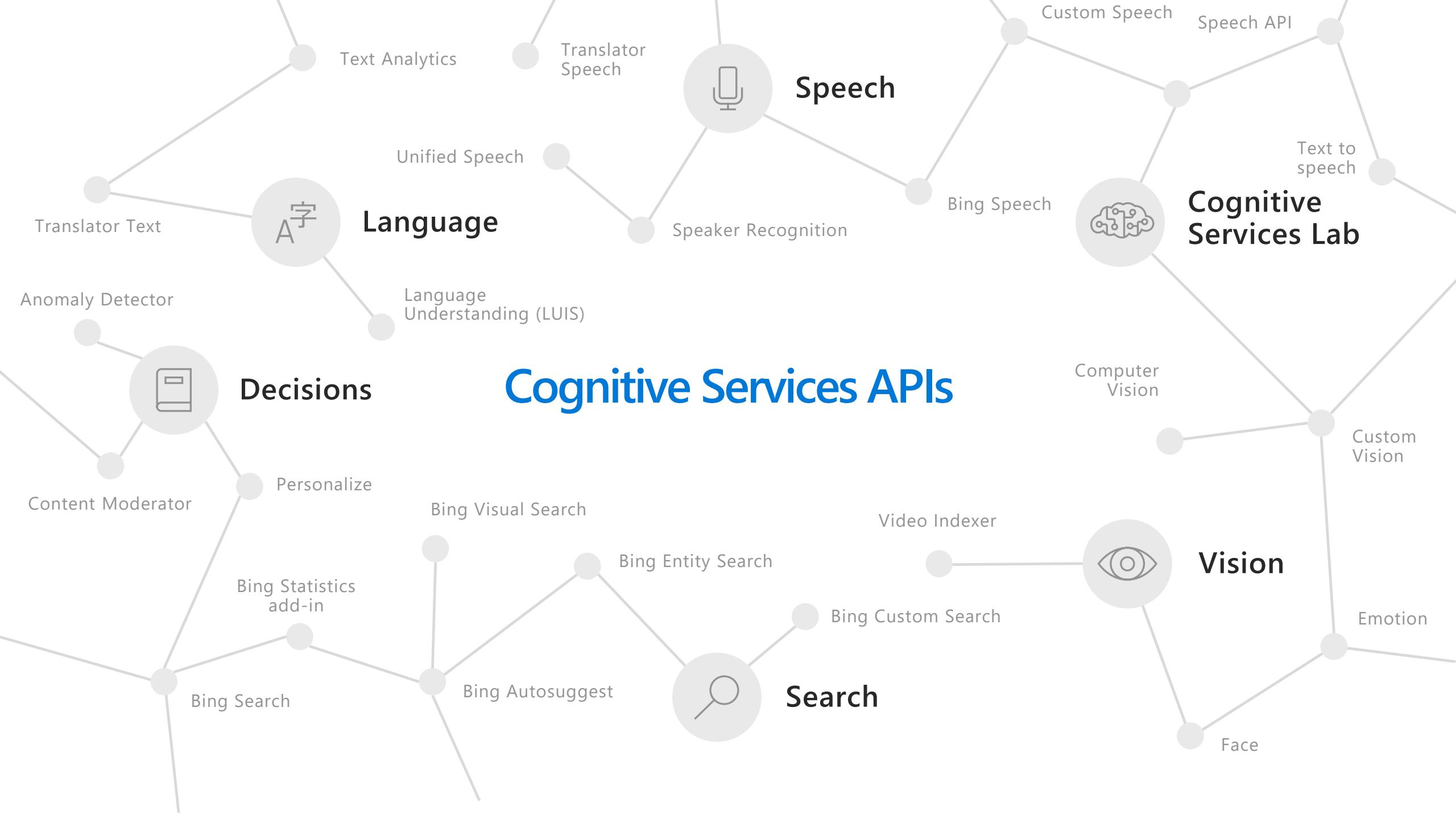
AI

Vision

Speech

Language

Customized language **Text-to-speech**
understanding Content moderation **Spell
Speech translation check**
Custom image classification
Speaker recognition Entity linking
Sentiment analysis, & augmentation
key phrase extraction **Image tagging**
Custom Object detection Text translation **Intend
voice** **OCR handwriting** analysis
Emotion detection Video insights **recognition** Custom translation
Face Custom speech Assisted text moderation
identification Speech transcription

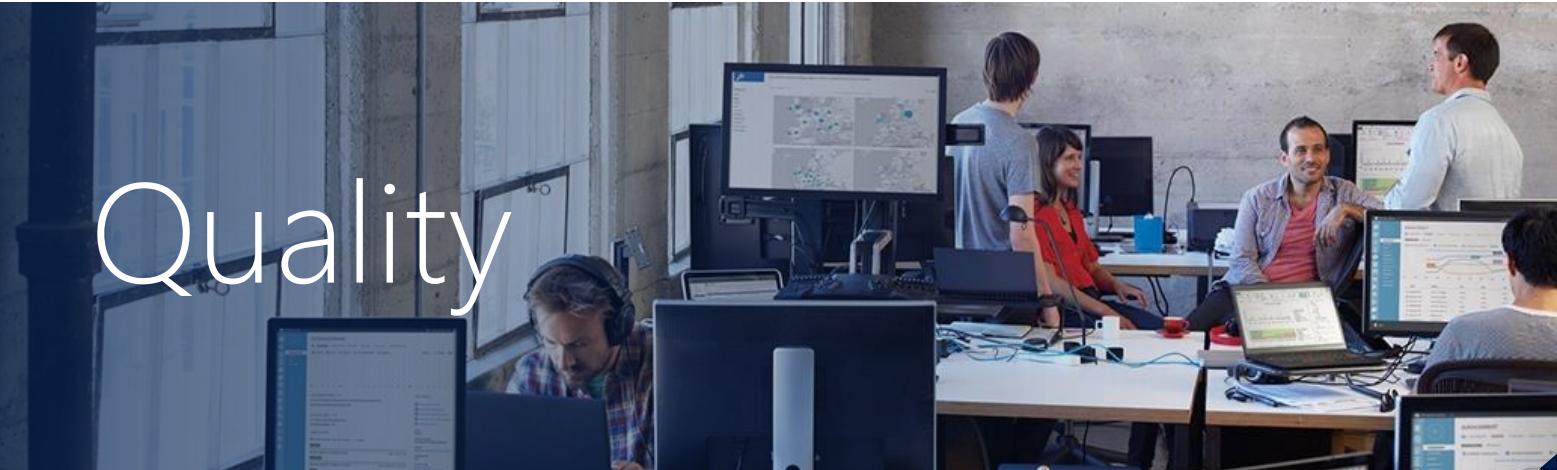


Why Microsoft Cognitive Services?

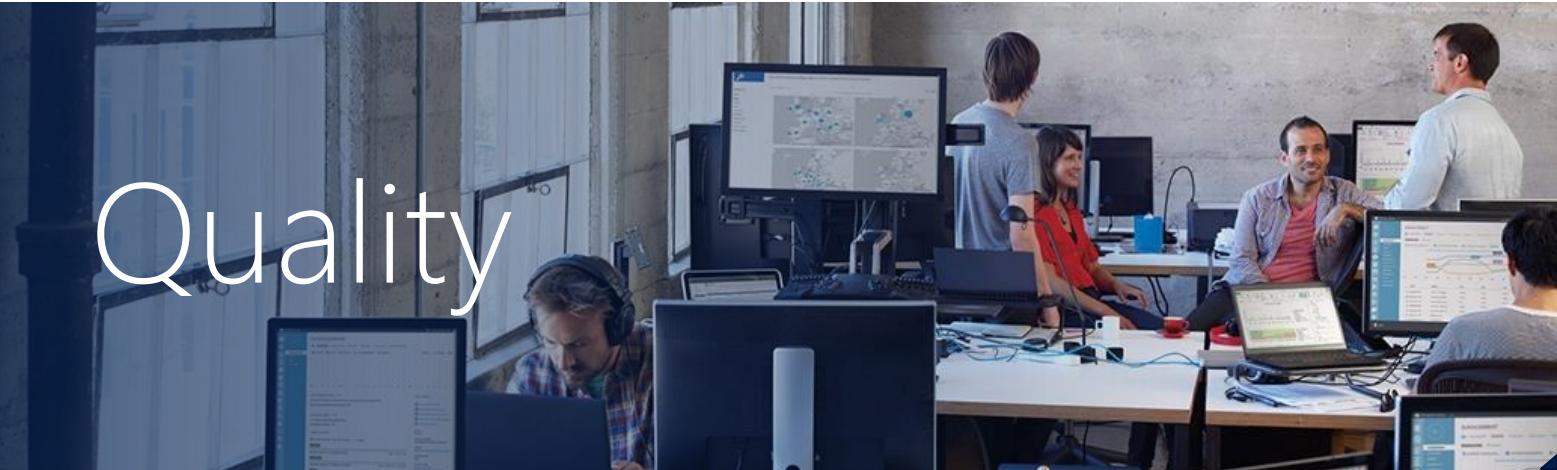
Easy



Flexible



Quality



Microsoft Cognitive Services

Give your apps a human side



Vision

From faces to feelings, allow your apps to understand images and video



Speech

Hear and speak to your users by filtering noise, identifying speakers, and understanding intent



Language

Process text and learn how to recognize what users want



Knowledge

Map complex information and data in order to solve specific tasks



Search

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



Labs

An early look at emerging Cognitive Services technologies: discover, try, and give feedback on new technologies before general availability

Microsoft Cognitive Services updates



Vision

Video Indexer
Computer Vision
Face
Emotion
Content Moderator
Custom Vision



Speech

Speaker Recognition
Bing Speech
Custom Speech
Translator Speech
Unified Speech
Speech to Text w. Custom Speech
Text to Speech w. Custom Voice
Speech Translation w. Custom Translator



Language

Text Analytics
Bing Spell Check
Translator Text
Language Understanding (LUIS)



Knowledge

QnA Maker
Custom Decision



Search

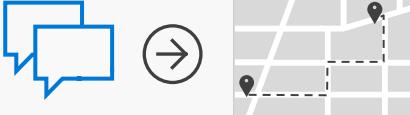
Bing Entity Search
Bing Autosuggest
Bing Search
Web Search
Image Search
News Search
Video Search
Bing Statistics add-in
Bing Visual Search
Bing Custom Search



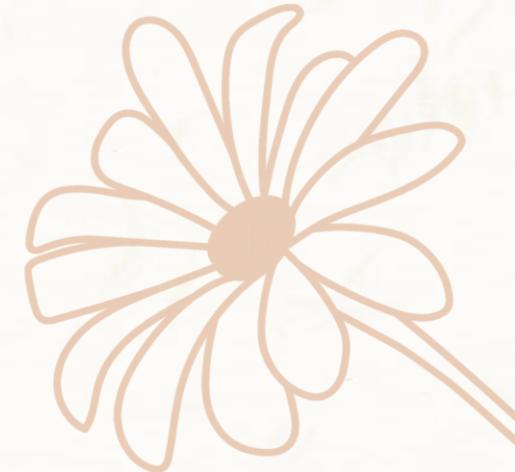
Labs

Project Gesture
Project Local Insights
Project Academic Knowledge
Project Entity Linking
Project Knowledge Exploration
Project Event Tracking
Project Answer Search
Project URL Preview
Project Anomaly Finder
Project Conversation Learner
Project Personality Chat

A variety of real-world applications

Vision	Speech	Language	Knowledge	Search														
 What is in the image or video? Intelligent Image insights  <table border="1"><tr><td>Category</td><td>People; 5 faces</td></tr><tr><td>Adult/Racy?</td><td>False/False</td></tr><tr><td>Dominant colors</td><td></td></tr><tr><td>Accent color</td><td></td></tr></table> Computer Vision	Category	People; 5 faces	Adult/Racy?	False/False	Dominant colors		Accent color		 Give me directions to the nearest local branch Speech to text  <table border="1"><tr><td>Convert spoken audio to text</td></tr><tr><td>Convert text to spoken audio</td></tr><tr><td>Extract intent of user</td></tr></table> Speech Service	Convert spoken audio to text	Convert text to spoken audio	Extract intent of user	 Play today's customer call recording Natural Language Processing  Intent: PlayCall Content: Customer# Date/Time.date: today Now Playing 11/29/2016 Customer Call Language Understanding	 QnA Pair of this site? Automatic extraction of questions and answers  QnA Maker	 Search for 'fraud prevention' Intelligent web search <table border="1"><tr><td> Information Communications Media Market News It also investigates the top three expected Fraud Detection and Prevention programs, in terms of demand in key markets...</td></tr><tr><td> The Big Question: In-House or Outsourced Fraud Protection? First, let's point out that there is not one absolute answer—there are "pros" and "cons" to each. Those who favor in-house...</td></tr><tr><td> How to Protect Your Business from Online Fraud this Holiday Season Michael heads fraud prevention tool. Online and mobile shopping are expected to continue growing apace...</td></tr></table> Bing News Search	 Information Communications Media Market News It also investigates the top three expected Fraud Detection and Prevention programs, in terms of demand in key markets...	 The Big Question: In-House or Outsourced Fraud Protection? First, let's point out that there is not one absolute answer—there are "pros" and "cons" to each. Those who favor in-house...	 How to Protect Your Business from Online Fraud this Holiday Season Michael heads fraud prevention tool. Online and mobile shopping are expected to continue growing apace...
Category	People; 5 faces																	
Adult/Racy?	False/False																	
Dominant colors																		
Accent color																		
Convert spoken audio to text																		
Convert text to spoken audio																		
Extract intent of user																		
 Information Communications Media Market News It also investigates the top three expected Fraud Detection and Prevention programs, in terms of demand in key markets...																		
 The Big Question: In-House or Outsourced Fraud Protection? First, let's point out that there is not one absolute answer—there are "pros" and "cons" to each. Those who favor in-house...																		
 How to Protect Your Business from Online Fraud this Holiday Season Michael heads fraud prevention tool. Online and mobile shopping are expected to continue growing apace...																		

Computer Vision





Vision

Image-processing algorithms to smartly identify, caption, and moderate your pictures





Vision

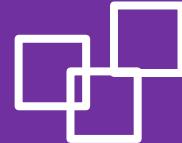
IMAGE CLASSIFICATION

What are my images about?



OBJECT DETECTION

Locate rectangular areas containing known objects in an image



SEMANTIC SEGMENTATION

Locate known objects in an image, at pixel level





Vision



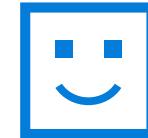
Computer Vision

Distill actionable information from images



Video Indexer

Process and extract smart insights from videos



Face

Detect, identify, analyze, organize, tag faces in photos, and even recognize emotions



Content Moderator

Machine-assisted moderation of text and images, augmented with human review tools



Custom Vision

Customizable web service that learns to recognize specific content in imagery



Vision

From faces to feelings, allow your apps to understand images and video

Computer Vision | Video Indexer | Custom Vision |
Face | Content Moderator



Face



Detection

```
"faceRectangle": {"width": 193, "height": 193,  
"left": 326, "top": 204}  
...
```

Feature attributes

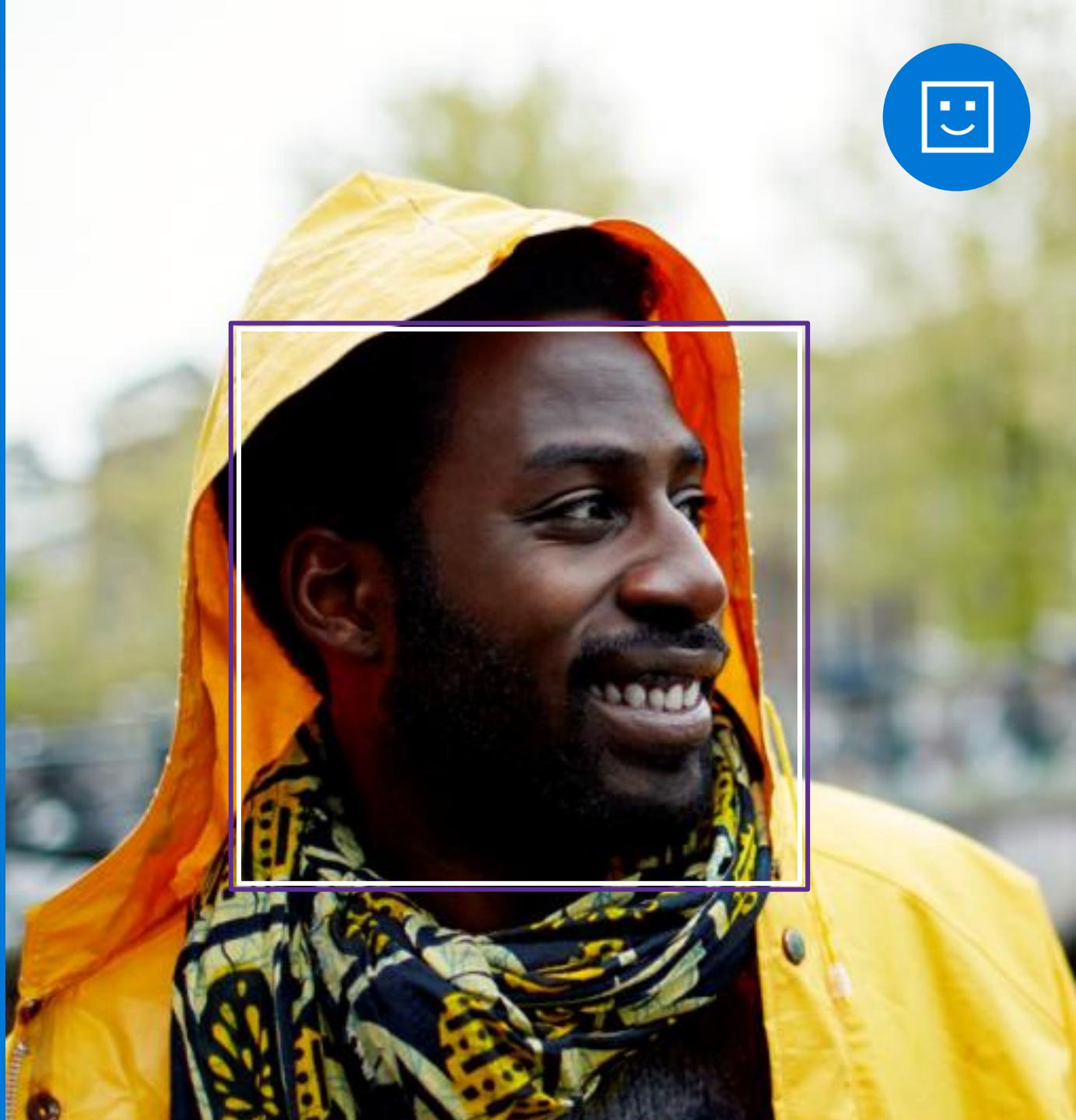
```
"attributes": { "age": 42, "gender": "male",  
"headPose": { "roll": "8.2", "yaw": "-37.8",  
"pitch": "0.0" }}
```

Grouping



Identification

Jasper Williams



Face



Face detection

Detect faces and their attributes within an image

Face verification

Check if two faces belong to the same person

Similar face searching

Find similar faces within a set of images

Face grouping

Organize many faces into groups

Face identification

Search which person a face belongs to



Emotion

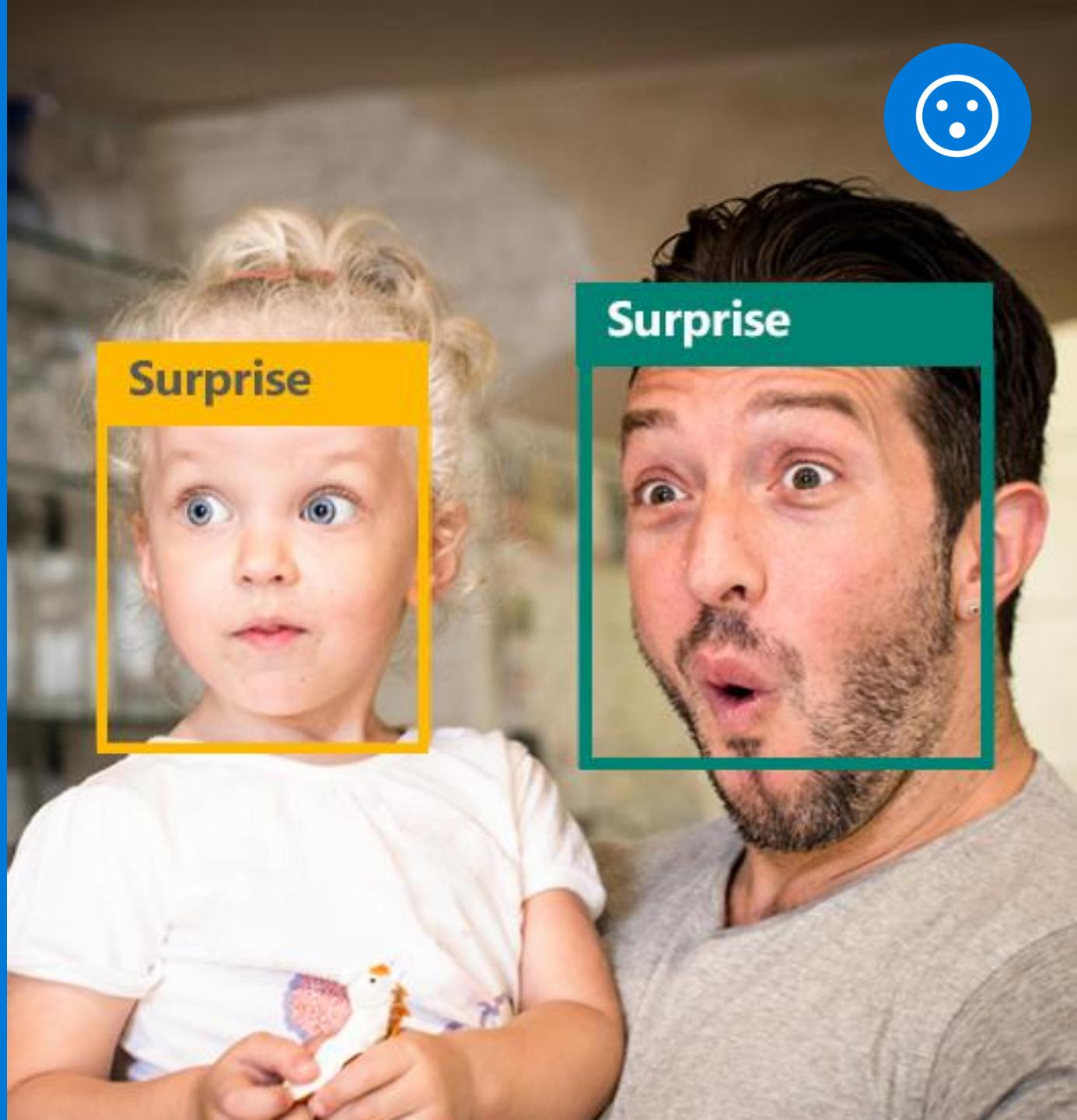


Face detection

```
"faceRectangle": {"width": 193,  
                 "height": 193,  
                 "left": 326,  
                 "top": 204} ...
```

Emotion scores

```
"scores": { "anger": 5.182241e-8,  
            "contempt": 0.0000242813,  
            "disgust": 5.621025e-7,  
            "fear": 0.00115027453,  
            "happiness": 1.06114619e-8,  
            "neutral": 0.003540177,  
            "sadness": 9.30888746e-7,  
            "surprise": 0.9952837}
```



Computer Vision



FEATURE	VALUE
NAME:	
Description	{ "tags": ["train", "platform", "station", "building", "indoor", "subway", "track", "walking", "waiting", "pulling", "board", "people", "man", "luggage", "standing", "holding", "large", "woman", "yellow", "suitcase"], "captions": [{ "text": "people waiting at a train station", "confidence": 0.833099365 }] }
Tags	[{ "name": "train", "confidence": 0.9975446 }, { "name": "platform", "confidence": 0.995543063 }, { "name": "station", "confidence": 0.9798007 }, { "name": "indoor", "confidence": 0.927719653 }, { "name": "subway", "confidence": 0.838939846 }, { "name": "pulling", "confidence": 0.431715637 }]
Image format	"Jpeg"

Analyze image



Type of image

Clip Art Type	0 Non-clipart
Line Drawing Type	0 Non-Line Drawing
Black & White Image	False

Content of image

Categories	[{"name": "people_swimming", "score": 0.099609375}]
Adult Content	False
Adult Score	0.18533889949321747
Faces	[{"age": 27, "gender": "Male", "faceRectangle": {"left": 472, "top": 258, "width": 199, "height": 199}}]

Image colors

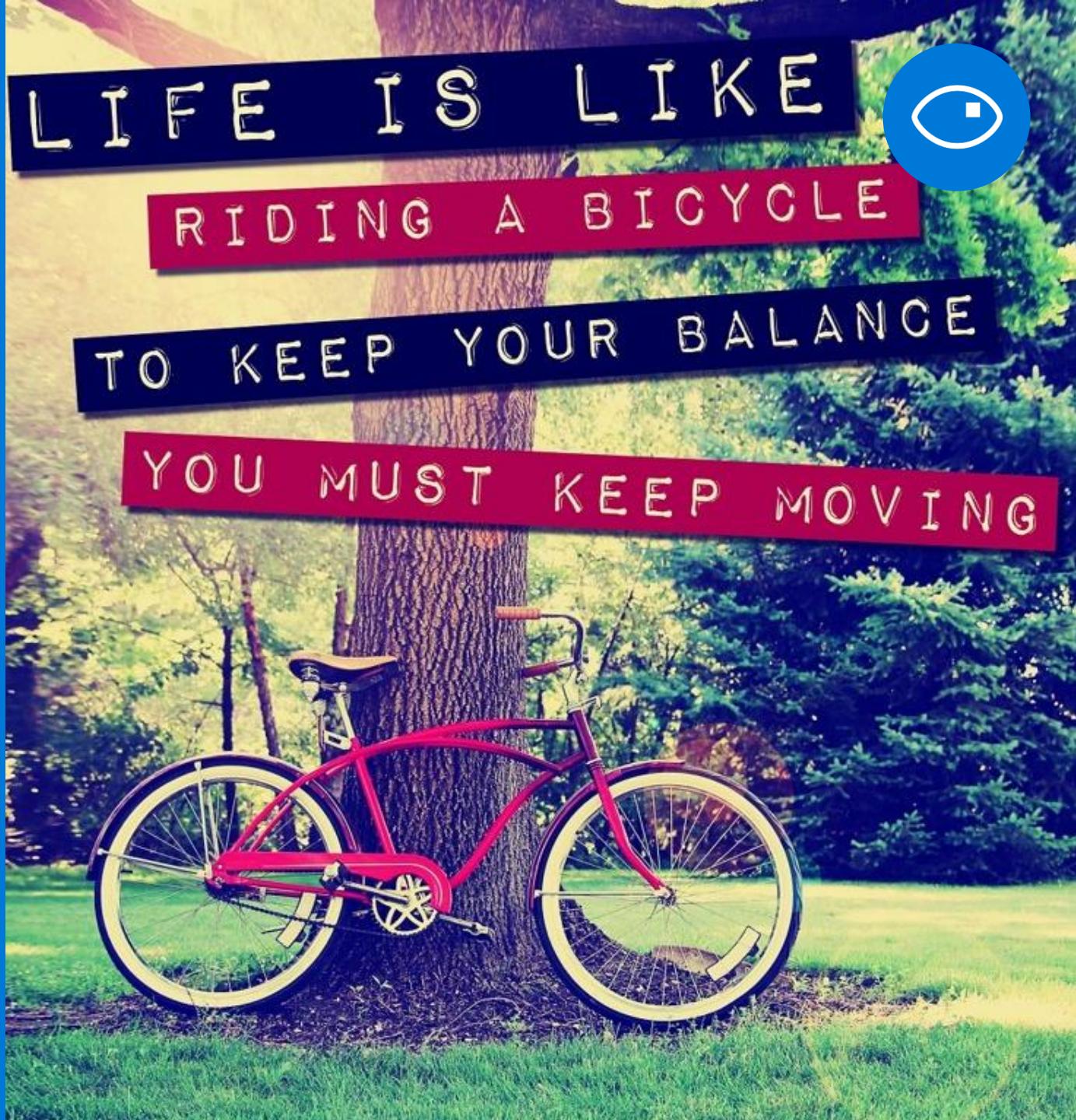
Dominant Color Background	White
Dominant Color Foreground	Grey
Dominant Colors	White
Accent Color	



OCR

JSON:

```
{  
  "language": "en",  
  "orientation": "Up",  
  "regions": [  
    {  
      "boundingBox": "41,77,918,440",  
      "lines": [  
        {  
          "boundingBox": "41,77,723,89",  
          "words": [  
            {  
              "boundingBox": "41,102,225,64",  
              "text": "LIFE"  
            },  
            {  
              "boundingBox": "356,89,94,62",  
              "text": "IS"  
            },  
            {  
              "boundingBox": "539,77,225,64",  
              "text": "LIKE"  
            }  
          ]  
        ]  
      ]  
    }  
  ]  
}
```



Smart thumbnail

Smart cropping off



Custom Vision

Customize

Design your own state-of-the-art models for unique use cases

Upload

Use labeled images to quickly train and update your models

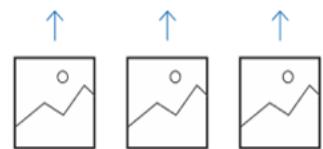
Export

Run models on a device or as a Docker container with just one click



Results

Tag	Probability
daisy	99.9%
trillium	3.1%
lily of the valley	0.1%
dogwood	0.0%



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.







Microsoft



#WholeNewGame

1

Microsoft
ULTRA
CourtSide



Powered by Microsoft Teams Microsoft Teams Microsoft Teams

#WholeNewGame

Microsoft
ULTRA
CourtSide



Microsoft Teams Microsoft Teams

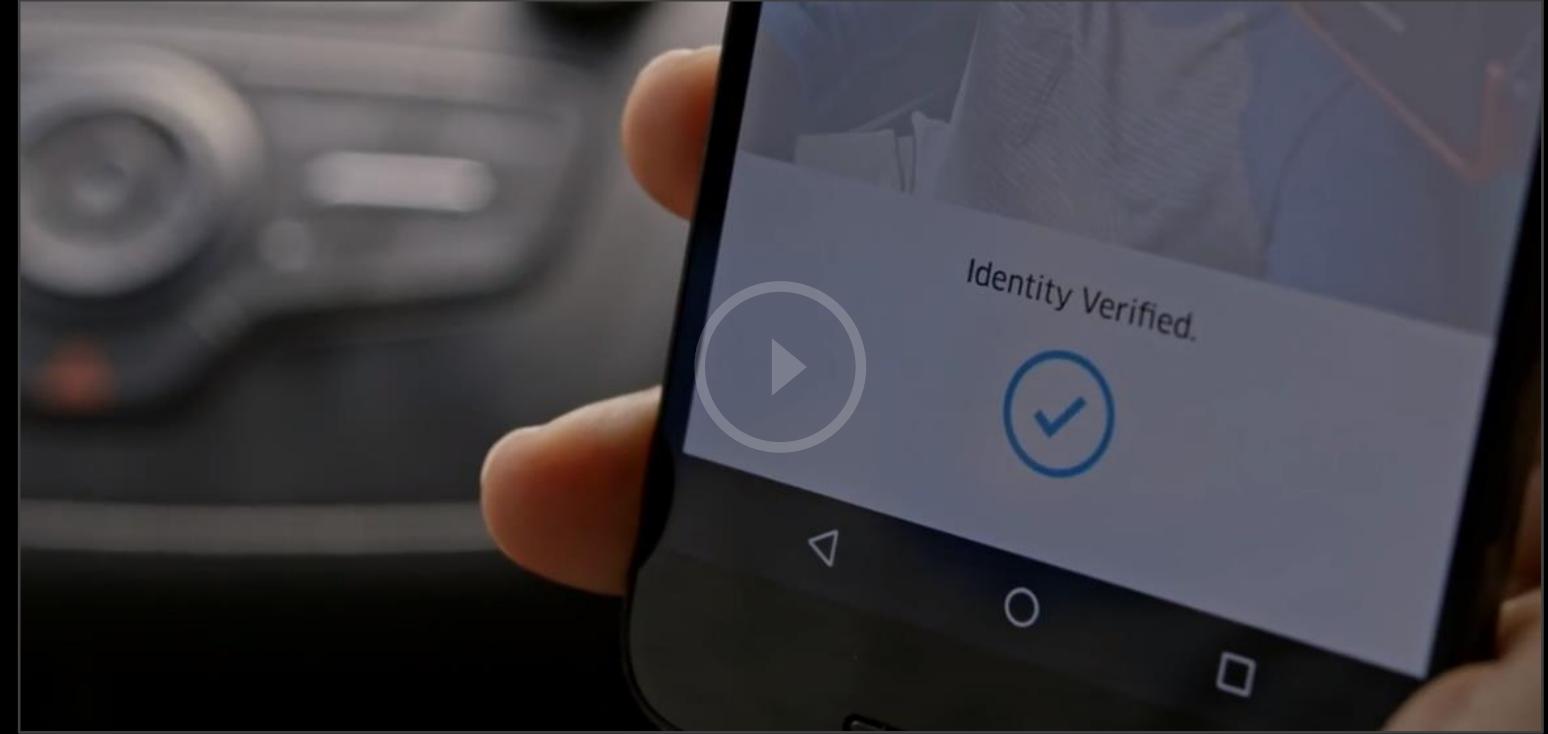
NBA WHOLE NEW GAME



BLACK LIVES MATTER



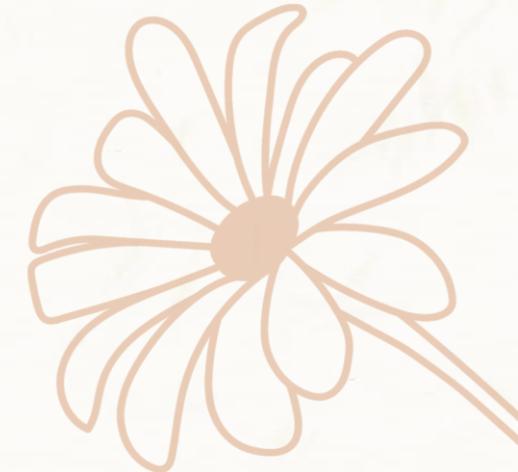
Uber



Uber boosts platform security through selfies

"Thousands of partners sign into our platform every hour. The response time from the Face API is incredible, enabling us to verify our drivers without slowing them down." *Dima Kovalez, Product Manager*

Custom Vision



Custom Vision Workflow

Collect Data

Upload & Tag

Train

Test

Custom Vision Workflow

Collect Data

Upload & Tag

Train

Test

1. 越真實越好，太乾淨反而會脫離現實
2. **每類至少15張**
3. **各種類別數量要均衡**
4. 數據多樣性，一個類別要有各種資料
5. 可以先在電腦上完成分類，再一次上傳
6. 可以先把data分成 train, validate, test三類

Custom Vision Workflow

Collect Data

Upload & Tag

Train

Test

BBQ

Training Images

Performance

Predictions

Train

Quick Test

Filter

Add images

Delete

Select all

Iteration

Workspace

Tags

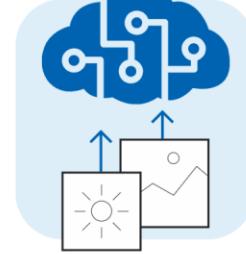
Tagged Untagged

Showing: all untagged images

Suggested Tags

Quickly label your untagged images with suggested objects. [Learn more.](#)

Get suggested objects



Looks like you don't have any images here!

Go ahead and browse for images to upload to your project, tag them, and they will be ready to be trained.

Add images

JPG, .PNG, .BMP format, up to 6 MB per image

Get started

Custom Vision Workflow

Image Detail Undo Changes Regions Shown Suggested objects off

Collect Data

Upload & Tag

Train

Test

My Objects

Wear mask

The image shows a user interface for a computer vision application. On the left, there is a vertical stack of four buttons labeled 'Collect Data', 'Upload & Tag', 'Train', and 'Test' from top to bottom. The 'Upload & Tag' button is highlighted with a dark blue background. To the right of this stack is a large image of a man wearing a face mask. A white rectangular box highlights the area where the man's face is located, indicating a detected object. Above this box, the text 'Wear mask' is displayed. At the top of the interface, there are several control buttons: 'Image Detail', 'Undo Changes', a toggle switch for 'Regions Shown' (which is turned on), a toggle switch for 'Suggested objects off' (which is turned off), and a close button ('X') in the top right corner. The main image features a blue background with large white Chinese characters and some English text, including 'emission command' and 'Ministry of'. In the foreground, a nameplate on a table reads '中央流行疫情指揮中心' and '陳時中 指揮官'.

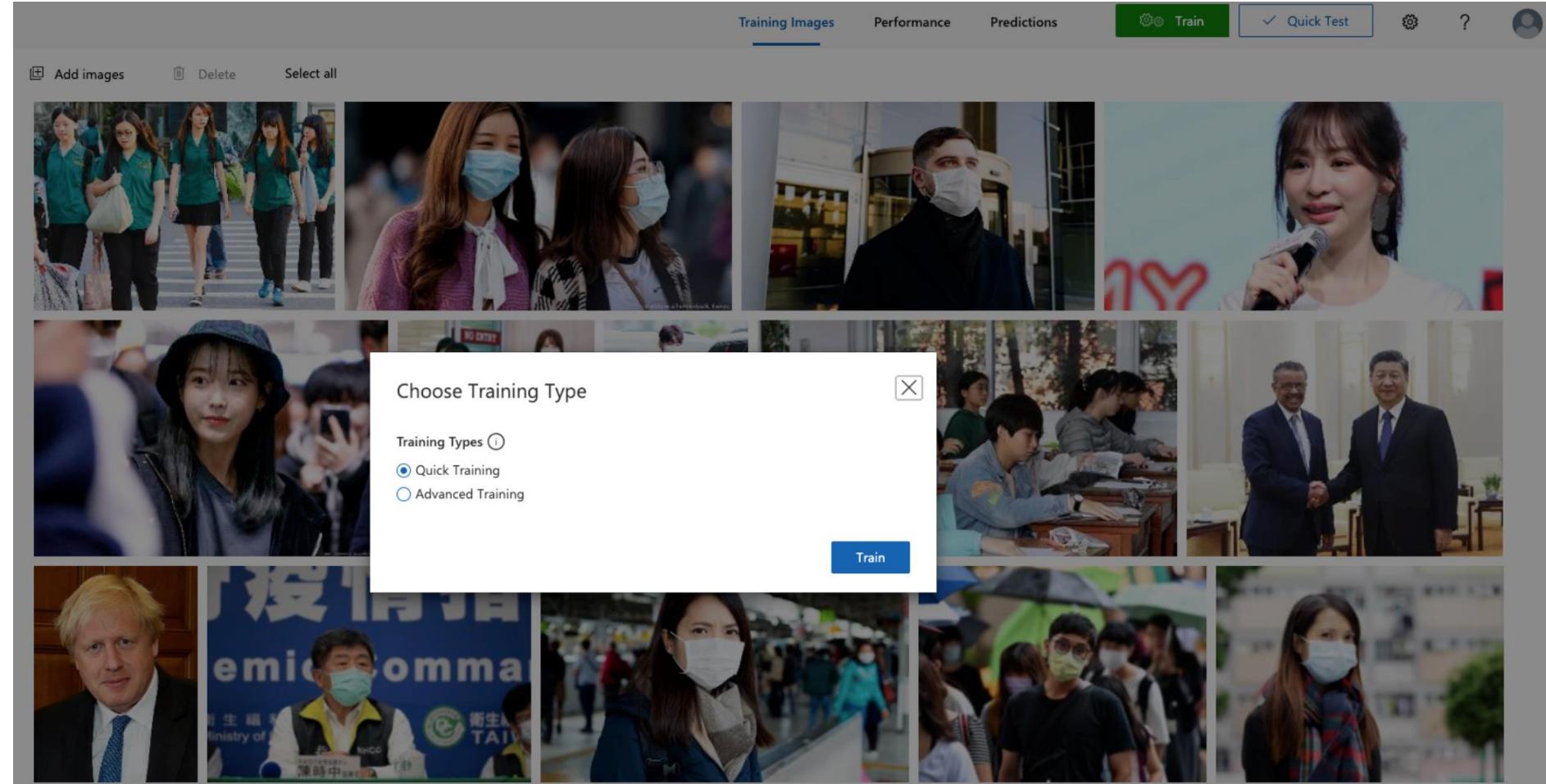
Custom Vision Workflow

Collect Data

Upload & Tag

Train

Test



Custom Vision Workflow

Collect Data

Upload & Tag

Train

Test

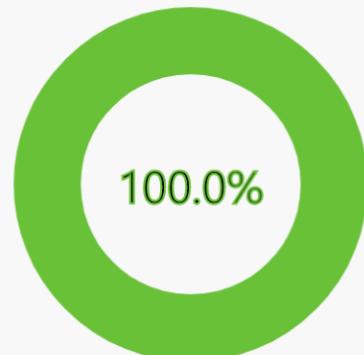
Precision ⓘ



Recall ⓘ



AP ⓘ



Precision

模型的精確度

Recall

真實狀況的吻合程度

AP

系統綜合評比資訊

Custom Vision Workflow

Collect Data

Upload & Tag

Train

Test

Threshold 越高，precision 越高，但recall可能下降

Iterations

Probability Threshold: 50% 

Overlap Threshold: 30% 

Custom Vision Workflow

Quick Test Regions Shown

Collect Data

Upload & Tag

Train

Test



The image shows a woman with dark hair and glasses, wearing a light blue face mask. She is standing behind a wooden podium with a microphone. A red rectangular box highlights her face area.

Image URL

or

Browse local files

File formats accepted: jpg, png, bmp
File size should not exceed: 4mb

Using model trained in

Iteration

Predicted Object Threshold
Only show suggested objects if the probability is above the selected threshold.
Threshold Value: 15%

Predictions
Predictions are shown in red

Tag	Probability
Wear mask	93%

IT'S YOUR TURN