



Develop computer vision solutions with

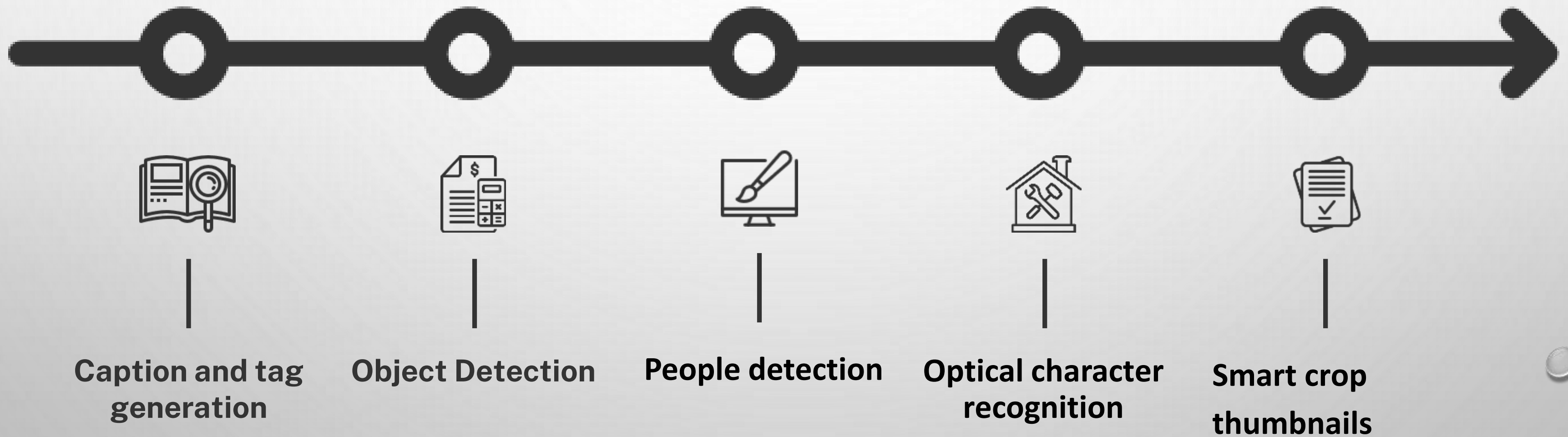
Azure AI Vision



Azure AI Vision – Image Analysis

- Computer Vision is a branch of artificial intelligence (AI) in which software interprets visual input, often from images or video feeds.
- In Microsoft Azure, you can use the Azure AI Vision service to implement multiple computer vision scenarios, including: Image Analysis, Face detection etc.
- Image analysis is the core of Azure AI Vision

Image Analysis



Captions

- The Caption feature generates a one-sentence description of all the image contents



A man pointing at a screen

confidence: 0.4891590476036072

Optical character recognition (OCR)

- It allow you to extract printed or handwritten text from images

Use one of your own files or choose from a sample below.

Drag and drop a file here
or
Browse for a file
or
Take a photo



Sample form #3



Detected attributes JSON

Nutrition Facts	Amount Per Serving
Serving size: 1 bar (40g)	
Serving Per Package: 4	
Total Fat 13g	
Saturated Fat 1.5g	
Amount Per Serving	
Calories 190	
Calories from Fat 110	
% Daily Values are based on	
calorie diet.	
Cholesterol 0mg	
Sodium 20mg	
Vitamin A 50%	

Object detection

- The bounding box coordinates of each detected person are returned.

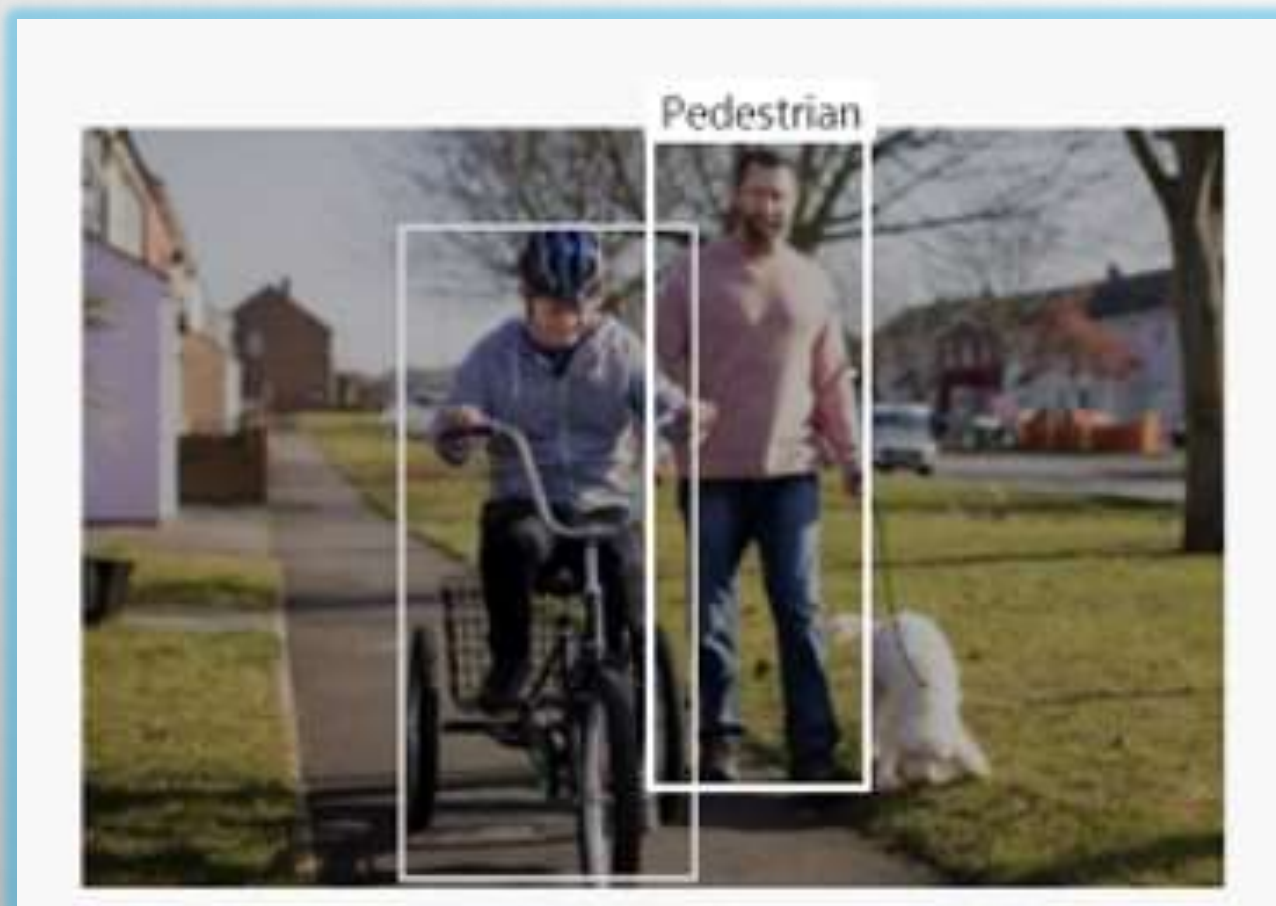


Image Analysis

Tags

- It gives it labels or keywords—like “House,” “grass,” or “plant”—to describe what’s in the image.



Deployment Options for Azure AI Vision Image Analysis

Can be used as:

- Standalone **Azure AI Vision** resource
- Multi-service **Azure AI Services** resource

Image Analysis 4.0 REST API

<https://<endpoint>/computervision/imageanalysis:analyze?features=caption,read&model-version=latest&language=en&api-version=2024-02-01>

Image Analysis – Regional API Endpoint (Old Version)

<https://eastus.api.cognitive.microsoft.com/computervision/imageanalysis:analyze?features=caption&model-version=latest&language=en&api-version=2024-02-01>

API

Regional API

*
cognitive.microsoft.com

<https://eastus.api.cognitive.microsoft.com>

Custom Subdomain API

*
.cognitiveservices.azure.com

<https://vision1.cognitiveservices.azure.com>

Question

You are developing an app that will use the Azure AI Vision API to analyze an image.

You need configure the request that will be used by the app to identify whether an image is clipart or a line drawing.

How should you complete the request? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

▼

GET
PATCH
POST

"https://*.cognitiveservices.azure.com/vision/v3.2/analyze?visualFeatures=

▼

description
imageType
objects
tags

&details={string}&language=e

Question

You have collection of pdf files. You need to extract text from files. Which service and feature should you use ?

- 1) Azure AI Vision and Image tags
- 2) Azure AI Vision and Optical character recognition (OCR)
- 3) Language service and Optical character recognition (OCR)

Image Analysis Options – C# vs Python

<https://microsoftlearning.github.io/AI-102-AIEngineer/Instructions/15-computer-vision.html>

Azure Custom Vision

Azure AI Custom Vision

Azure AI Vision

Microsoft managed pre trained training and models.

Azure AI Custom Vision

Trained on your images

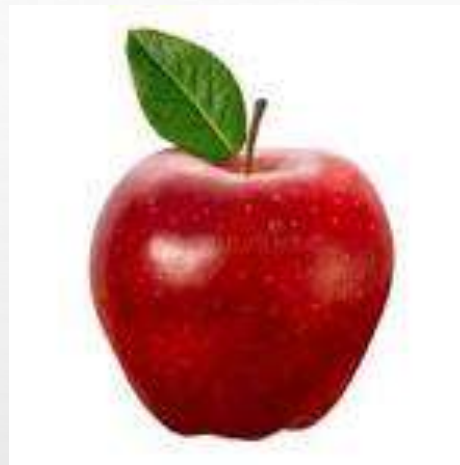
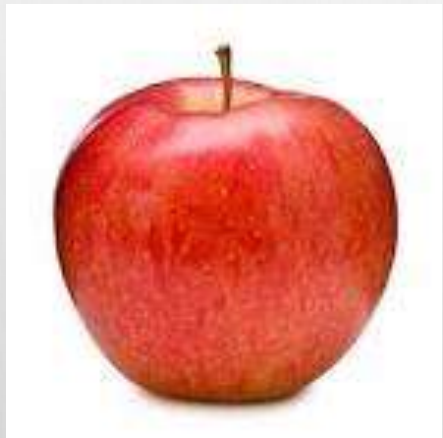
How Custom Vision Works

- Upload your images
- Label (Tag) your images
- Train your model
- Query your model with new images to predict labels

What is Image Classification?

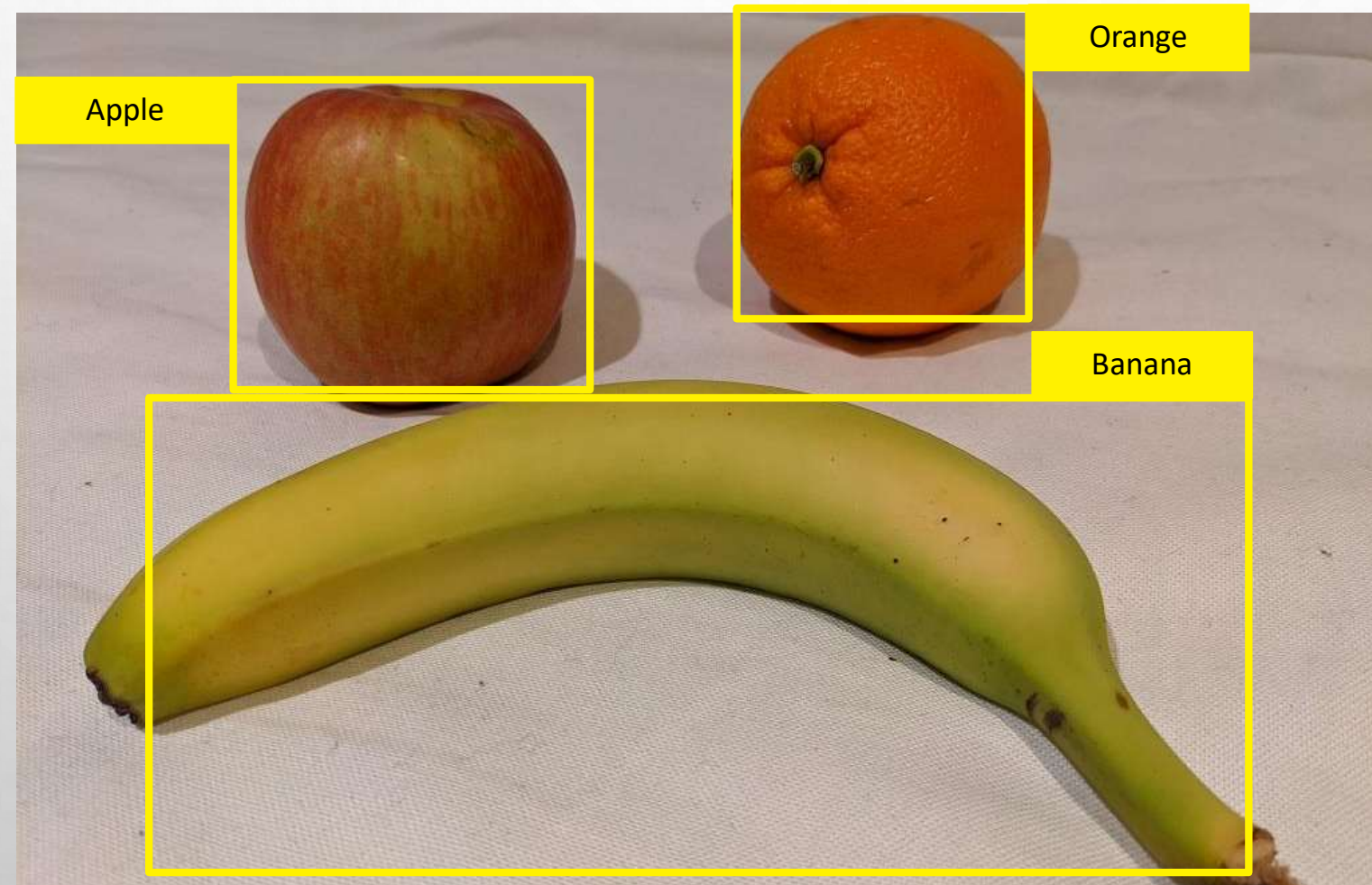
Train a model to predict the class label for the image

In other words, what is this a picture of?



What is Object Detection?

Train a model to predict the class label for the image, but it returns the coordinates in the image where the applied label(s) are found.



Question

You have a Custom Vision service project that performs object detection. The project uses the General domain for classification and contains a trained model.

You need to export the model for use on a network that is disconnected from the internet.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Actions

Change the classification type.

Export the model.

Retrain the model.

Change Domains to **General (compact)**.

Create a new classification model.

Answer Area



Answer

Change Domains to **General (compact)**.

Retrain the model.

Export the model.

Azure AI Custom Vision – Export Model (Important Note)

<https://learn.microsoft.com/en-us/azure/ai-services/custom-vision-service/export-your-model>

Face Service

Face Service

- Azure AI Vision Service has limited face Capability. It can only detect faces and Provides the location in the images.
- Azure AI Face Service provides more comprehensive facial detection, analysis, and recognition capabilities.

Face Service

Face recognition : It is the process of verifying or identifying individuals by their faces

Verification (one-to-one" matching)

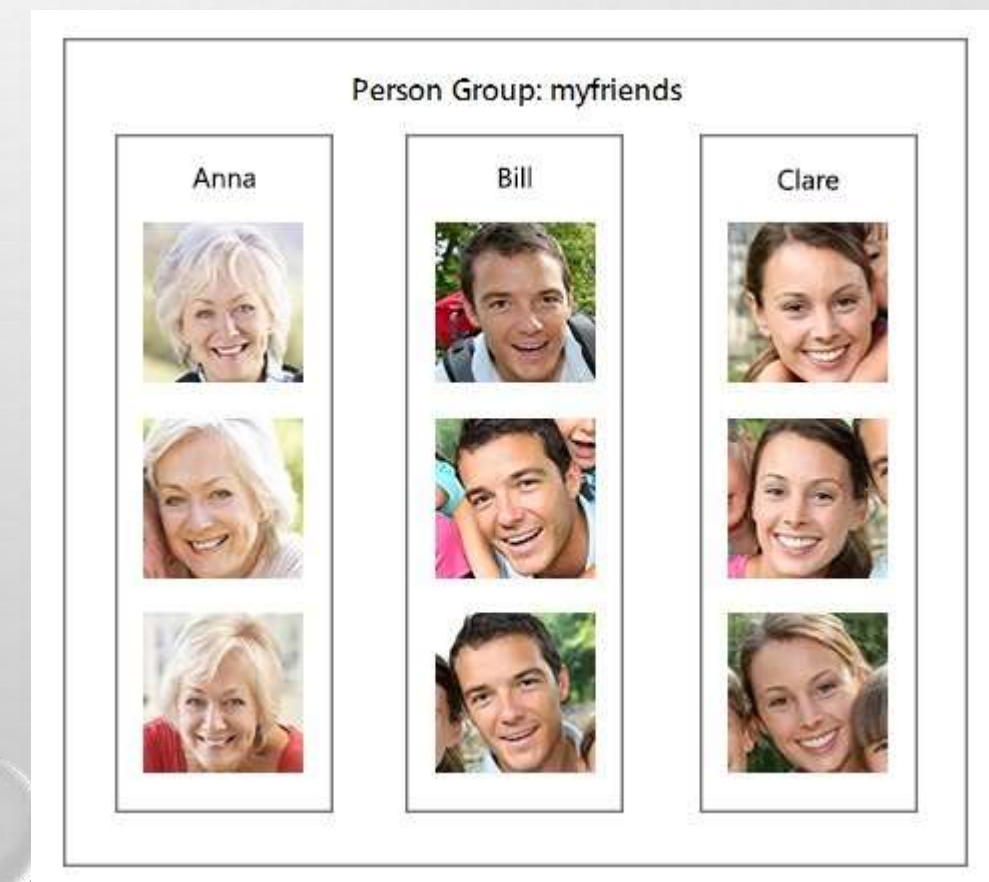
Do these two faces belong to the same person?"

Verification is a "one-to-one" matching of a face in an image to a single face from a secure repository or photo to verify that they're the same individual.



Identification (one-to-many" matching)

Identification is a "one-to-many" matching of a face in an image to multiple faces from a secure repository or photo gallery to identify which individual, if any, matches.



The background of the slide is a light gray gradient. It is decorated with several realistic water droplets of various sizes, primarily located in the top-left, top-right, and bottom-right corners. The droplets have highlights and shadows, giving them a three-dimensional appearance.

Analyzing Videos

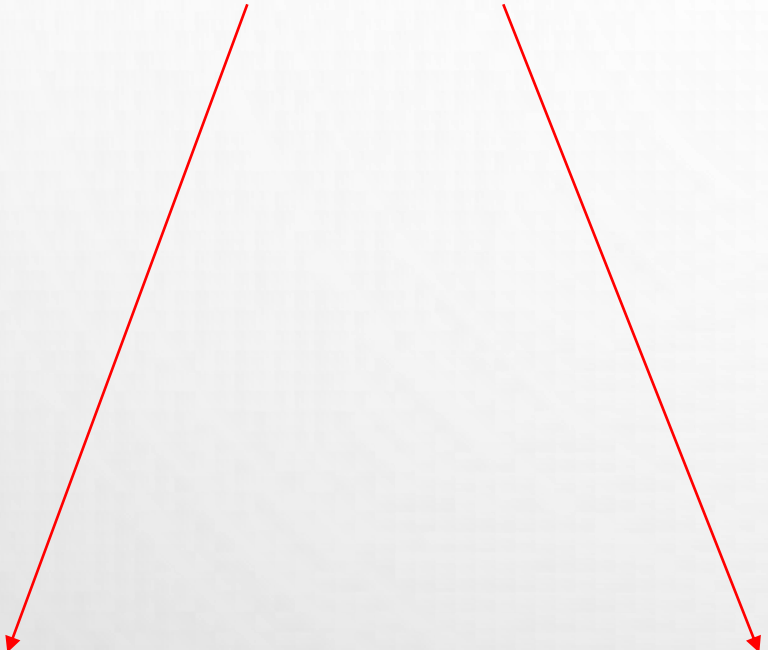
Azure AI Video Indexer

The **Azure AI Video Indexer** service is designed to help you extract information from videos.

- **Facial recognition** - detecting the presence of individual people in the image. This requires [Limited Access](#) approval.
- **Optical character recognition** - reading text in the video.
- **Speech transcription** - creating a text transcript of spoken dialog in the video.
- **Topics** - identification of key topics discussed in the video.
- **Sentiment** - analysis of how positive or negative segments within the video are.
- **Labels** - label tags that identify key objects or themes throughout the video.
- **Content moderation** - detection of adult or violent themes in the video.
- **Scene segmentation** - a breakdown of the video into its constituent scenes.

Azure AI Video Indexer API

- To use the Azure Video Indexer API, you need some information to authenticate requests:



Account ID
(Video Indexer Portal)

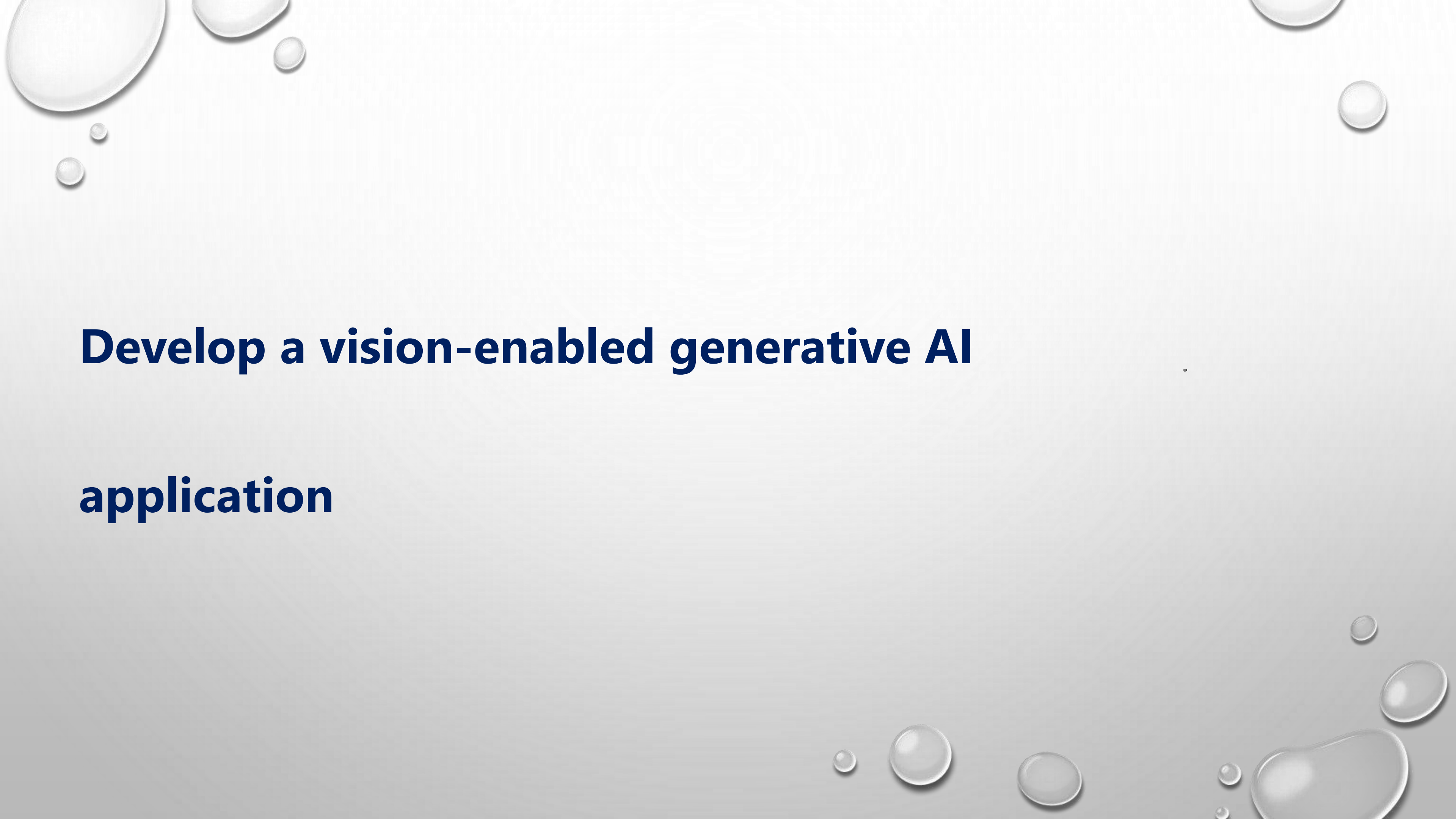
API Key
Azure Video Indexer Developer Portal
<https://api-portal.videoindexer.ai>

Question

You need to analyze video content to identify any mentions of specific company names.
Which three actions should you perform in sequence?

Sign in to the Azure Video Indexer web portal
Sign in to the Azure AI Custom Vision website
From Content model customization, select Language
Add the specific company names to the include list
Add the specific company names to the exclude list
From Content model customization, select Brands

Sign in to the Azure Video Indexer web portal
From Content model customization, select Brands
Add the specific company names to the include list



Develop a vision-enabled generative AI application

Multimodal Model

- Generative AI models enable you to develop chat-based applications that reason over and respond to input. Often this input takes the form of a text-based prompt, but increasingly **multimodal models** that can respond to visual input are becoming available.
- **multimodal models** - a model that supports not only text-based input, but image-based (and in some cases, audio-based) input as well.
- Multimodal models available in **Azure AI Foundry** include:
 - Microsoft **Phi-4-multimodal-instruct**
 - OpenAI **gpt-4o**
 - OpenAI **gpt-4o-mini**