Microsoft

AI-100T01A

Module 01: Introducing Azure Cognitive Services

Agenda

- Overview of Azure Cognitive Services
- Creating Cognitive Services
- Testing Cognitive Services
- Lab: Implement Computer Vision

Lesson 01

Overview of Azure Cognitive Services

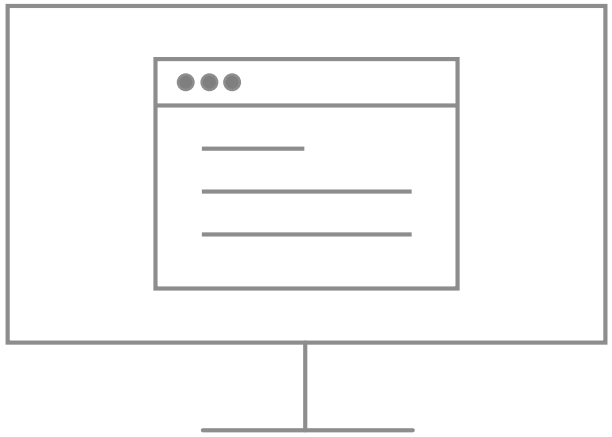
Lesson Objectives

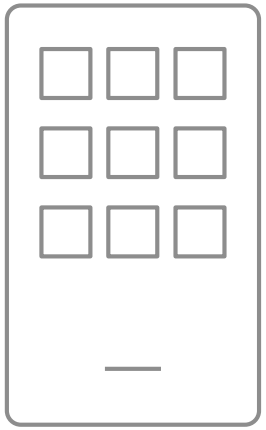
- L01 – Introduce Computer Vision
- L02 – Introduce Speech Services
- L03 – Introduce Language Services
- L04 – Introduce Search

Azure Cognitive Services and AI

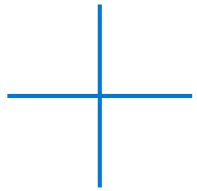
- Building an AI application from scratch can be daunting
- Microsoft Azure Cognitive Services consist of
  - APIs
  - SDKs
  - Services
- Core features includes:
  - Speech
  - Language
  - Vision
  - Search
  - Decision

Access Pre-Built Services





Your Apps



Vision

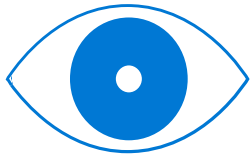
Speech

Language

Search

Decision

Introducing Vision

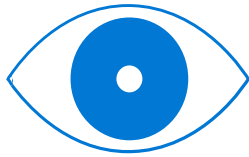


Vision

The current list of vision APIs are:

- Computer Vision - The Computer Vision service provides you with access to advanced algorithms for processing images and returning information
- Custom Vision Service - The Custom Vision Service allows you to build custom image classifiers.
- Face API - Face API provides access to advanced face algorithms, enabling face attribute detection and recognition.
- Video Indexer - Video Indexer enables you to extract insights from your video.

Introducing Vision, Cont'd



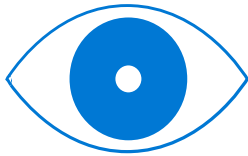
Vision

Services currently in preview:

- Ink Recognizer - Used to recognize digital ink and handwriting
- Form Recognizer - Extract text, key-value pairs, and tables from documents.

Computer Vision APIs – Tagging Images

Identifying and tagging visual features in an image

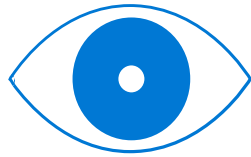


Vision


```
{
    "tags": [
        {
            "name": "grass",
            "confidence": 0.9999995231628418
        },
        {
            "name": "outdoor",
            "confidence": 0.99992108345031738
        },
        {
            "name": "house",
            "confidence": 0.99685388803482056
        }
    ],
    "requestId": "06f39352-e445-42dc-96fb-0a1288ad9cf1",
    "metadata": {
        "height": 200,
        "width": 300,
        "format": "Jpeg"
    }
}
```


Computer Vision APIs – Detecting Objects

Returns the bounding box coordinates for each tag applied

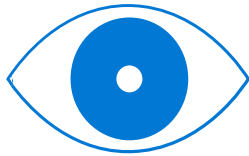


Vision


```
{
  "objects":[
    {
      "rectangle":{
        "x":730,
        "y":66,
        "w":135,
        "h":85
      },
      "object":"kitchen appliance",
      "confidence":0.501
    },
    {
      "rectangle":{
        "x":523,
        "y":377,
        "w":185,
        "h":46
      },
      "object":"computer keyboard",
      "confidence":0.51
    },
    {
      "rectangle":{
        "x":471,
        "y":218,
        "w":289,
        "h":226
      },
      "object":"Laptop",
      "confidence":0.85,
      "parent":{
        "object":"computer",
        "confidence":0.851
      }
    }
```


Computer Vision APIs – Detecting Brands

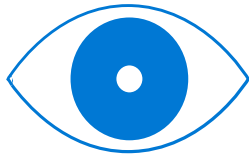
Identify commercial brands in images or videos



Vision


```
{
  "brands":[
    {
      "name":"Microsoft",
      "confidence":0.657,
      "rectangle":{
        "x":436,
        "y":473,
        "w":568,
        "h":267
      }
    },
    {
      "name":"Microsoft",
      "confidence":0.85,
      "rectangle":{
        "x":101,
        "y":561,
        "w":273,
        "h":263
      }
    }
  ],
  "requestId":"10dcd2d6-0cf6-4a5e-9733-dc2e4b08ac8d",
  "metadata":{
    "width":1286,
    "height":1715,
    "format":"Jpeg"
  }
}
```


Computer Vision APIs – Detecting Brands



Vision

- Identify and categorize an entire image, using a category taxonomy with parent/child hereditary hierarchies.
- Categories can be used alone, or with our new tagging models.
- Currently, English is the only supported language for tagging and categorizing images.

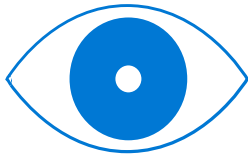
Computer Vision APIs – Describe an Image

Generate a description of an entire image in human-readable language, using complete sentences

Vision


```
{
   "description": {
      "tags": ["outdoor", "building", "photo", "city", "white", "black", "large", "sitting", "old",
"water", "skyscraper", "many", "boat", "river", "group", "street", "people", "field", "tall", "bird",
"standing"],
      "captions": [
         {
            "text": "a black and white photo of a city",
            "confidence": 0.95301952483304808
         },
         {
            "text": "a black and white photo of a large city",
            "confidence": 0.94085190563213816
         },
         {
            "text": "a large white building in a city",
            "confidence": 0.93108362931954824
         }
      ]
   },
   "requestId": "b20bfc83-fb25-4b8d-a3f8-b2a1f084b159",
   "metadata": {
      "height": 300,
      "width": 239,
      "format": "Jpeg"
   }
}
```

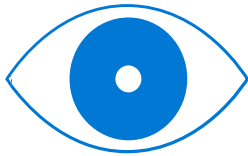

Computer Vision APIs – Detecting Other Aspects



- Computer Vision can determine whether an image is black & white or color and, for color images, identify the dominant and accent colors.
- Computer Vision can also detect adult material in images. Content flags are applied with a score between zero and one with values closer to 1 representing a greater confidence that the content is adult oriented*.

*also offered by the Azure Content Moderator service

Computer Vision APIs – Recognize Text



- Read API - detects textual content in the supplied images and converts identified text into a machine-readable character stream.

- Optical Character Recognition (OCR) - this feature is similar to the Read API but it operates synchronously and is not really optimized to perform OCR on large documents.

Introducing Speech

The speech APIs that are a part of the Azure Cognitive Services, provide specific functionality that can help your applications to work with speech

- **Text-to-speech**
- **Speech-to-text**
- **Speech translation**

Text-to-Speech

- Provides neural text-to-speech voices nearly indistinguishable from human speech (English).
- Convert text to natural-sounding speech.
- Offer multiple genders and/or dialects for many supported languages.
- Support plain text input or Speech Synthesis Markup Language (SSML).

Speech-to-Text

- Transcribe continuous real-time speech into text.
- Batch-transcribe speech from audio recordings.
- Support intermediate results, end-of-speech detection, automatic text formatting, and profanity masking.
- Call on Language Understanding (LUIS) to derive user intent from transcribed speech.

Speech Translation

Speech Translation is designed to perform **real-time speech translation** for scenarios such as:

- Live presentation translation
- In-person or remote translated communications
- Customer support
- Business intelligence
- Media subtitling
- Multilingual AI interactions

Speech Translation

Walkthrough – Create a Speech Translation Subscription

Introducing Language

The Language services available in the Azure Cognitive Services cover four main areas of focus and one preview feature:

- Text Analytics
- Translator Text
- Language Understanding (LUIS)
- QnA Maker
- Immersive Reader (In Preview)

Introducing Language – Text Analytics

The Text Analytics API is a Cognitive Service designed to help you extract information from text.

Through the service you can:

- Identify language
- Discover sentiment
- Extract key phrases
- Detect well-known entities from text

Introducing Language – Translator Text

There are two aspects to this service, translator text and custom translator.

- Translator Text is a cloud-based machine translation service you can use to translate text in near real-time through a simple REST API call
- The customized translation system can be used to translate text with the Translator Text API or Microsoft Speech Services

** The API has support for more than 60 languages*

Introducing Language – LUIS

Language Understanding (LUIS) is a cloud-based API service that applies custom machine-learning intelligence to a user's conversational, natural language text to predict overall meaning, and pull out relevant, detailed information

LUIS makes use of three key aspects for understanding language:

- Intent - An intent represents a task or action the user wants to perform. It is a purpose or goal expressed in a user's utterance.
- Utterance - Utterances are input from the user that your app needs to interpret.
- Entities - The entity represents a word or phrase inside the utterance that you want extracted

Introducing Language – QnA Maker

- Question and answer extraction from unstructured text
- Knowledge base creation from collections of Q&As
- Semantic matching for knowledge bases

Introducing Language

Walkthrough - Call the Text Analytics API from the Online Testing Console

Introducing Search

Search

The Search aspect of Cognitive Services, at a high level, utilizes Bing as the search engine. Key aspects of the Search service are:

- Bing New Search
- Bing Video Search
- Bing Web Search
- Bing Autosuggest
- Bing Custom Search
- Bing Entity Search
- Bing Image Search
- Bing Visual Search
- Bing Local Business Search (Preview)

Demo: Bing Search Services

- Bing Autosuggest – See how the autosuggest feature works when entering text
- Bing Image Search – See an example of Bing results based on image searches and understand the various attributes available

Introducing Decision

- **Content Moderator** - Use this service to detect potentially offensive content such as unwanted images, filter text for profanity or other unwanted text, and moderate content that might be considered adult or racy content.
- **Anomaly Detector (Preview)** - This feature set can help you to monitor business health in real-time, leverage interactive data analytics for your business, and even help to conduct Internet of Things (IoT) remote monitoring.
- **Personalizer (Preview)** - Personalizer helps to deliver richer personalized experiences in the applications that you develop. It can understand and manage the reinforcement learning loop and can be deployed anywhere from the cloud to the edge.

Content Moderator

- **Text Moderation -** The service can evaluate written text to determine if it contains the following:
  - Profanity
  - Inappropriate text
  - Personally Identifiable Information (PII)
  - Information in custom term lists

- **Image Moderation**
  - Use machine-assisted image moderation and human-in-the-loop Review tool to moderate images for adult and racy content.
  - Scan images for text content and extract the text.
  - Detect faces to aid in detecting personal data that may be contained in the images due to the recognition of faces in the images. You can match images against custom lists, and take further action.
  - The Optical Character Recognition (OCR) operation predicts the presence of text content in an image and extracts it for text moderation.
  - The Match operation allows fuzzy matching of incoming images against any of your custom lists, created and managed using the List operations. If a match is found, the operation returns the identifier and the moderation tags of the matched image.

Content Moderator, Cont'd

- **Video Moderation**
  - Video-trained classifier (preview) - Microsoft's adult and racy video classifier is trained with videos
  - Shot detection - rather than just outputting frames, the service also provides shot-level data so you can review the video using shot-level or frame-level data
  - Key frame detection - the service identifies and outputs only potentially complete (good) frames. The feature makes frame-level adult and racy analysis easier and more efficient.
  - Visualization for human review - using the review tool, users can evaluate the insights of the video, change tags related to the video and perform many features associated with reviewing and evaluating video.
  - Transcript moderation - video files that will have transcripts or closed captioning require moderation to search for offensive speech. This option requires the use of the Azure Media Indexer to convert the speech to text, if the transcript files are not available, and then use the Content Moderator review API to scan the text in the review tool

- **Review Tool**
  - A common set of tools to moderate text, image, and video.
  - Automate the creation of human reviews when moderation API results come in.

Demo: Content Moderator

Is this a garbage or crap email abcdef@abcd.com, phone: 6657789887, IP: 255.255.255.255, 1 Microsoft Way, Redmond, WA 98052. These are all UK phone numbers, the last two being Microsoft UK support numbers: 44 870 608 4000 or 0344 800 2400 or 0800 820 3300.

text.Language
eng

text.detectedIPA
[{"SubType":"IPV4","Text...

text.detectedAddress
[{"Text":"1 Microsoft Wa...

text.HasProfanity
True

text.Category1
6.08019627179601E-06

text.Category3
0.987999975681305

text.detectedEmail
[{"Detected":"abcdef@abc...

text.detectedPhoneNumber
[{"CountryCode":"US","Te...

text.hasPII
True

matchTerms
[{"Index":12,"OriginalIn...

text.Category2
0.142411544919014

a r

Clear

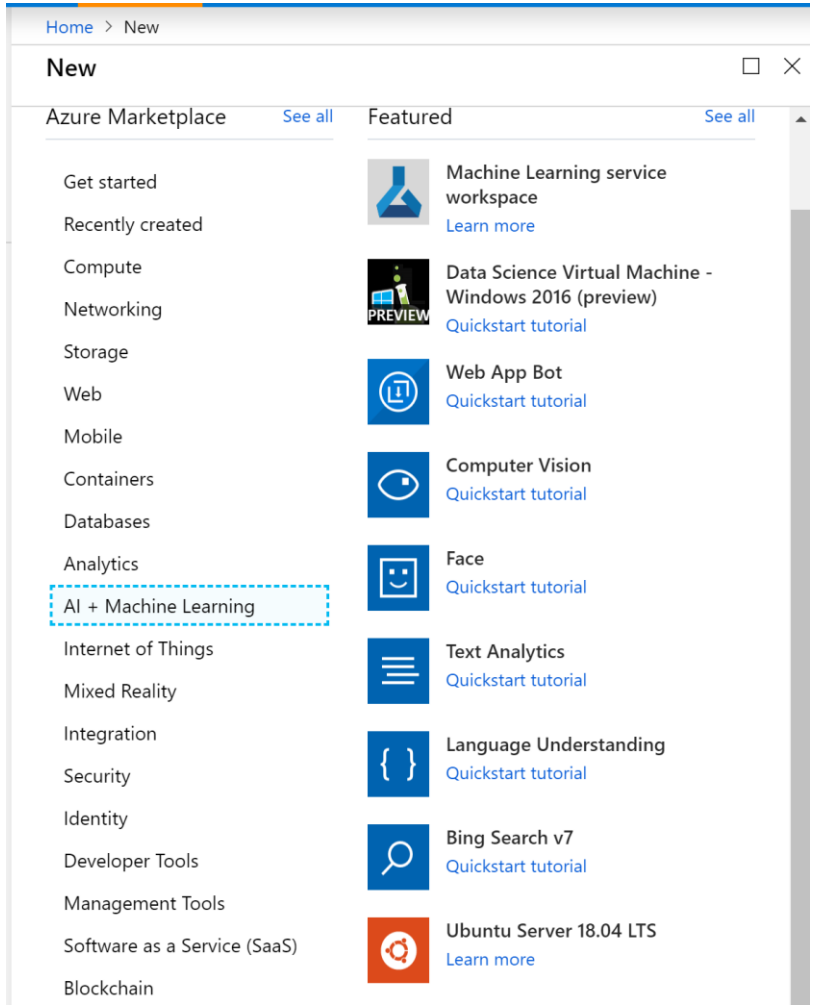
Lesson 02

Creating Cognitive Services using the Azure Portal

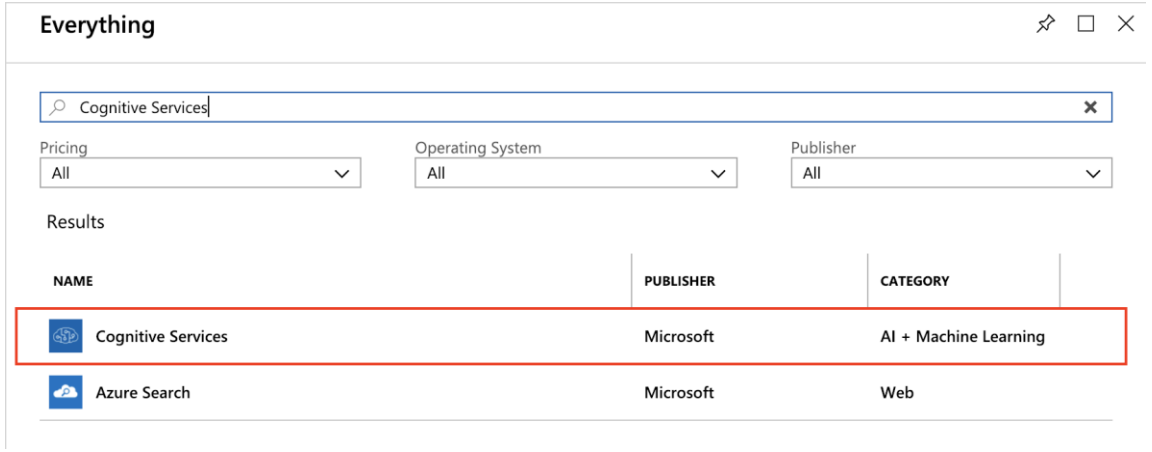
Lesson Objectives

- L01 – Creating and Managing Cognitive Services Accounts

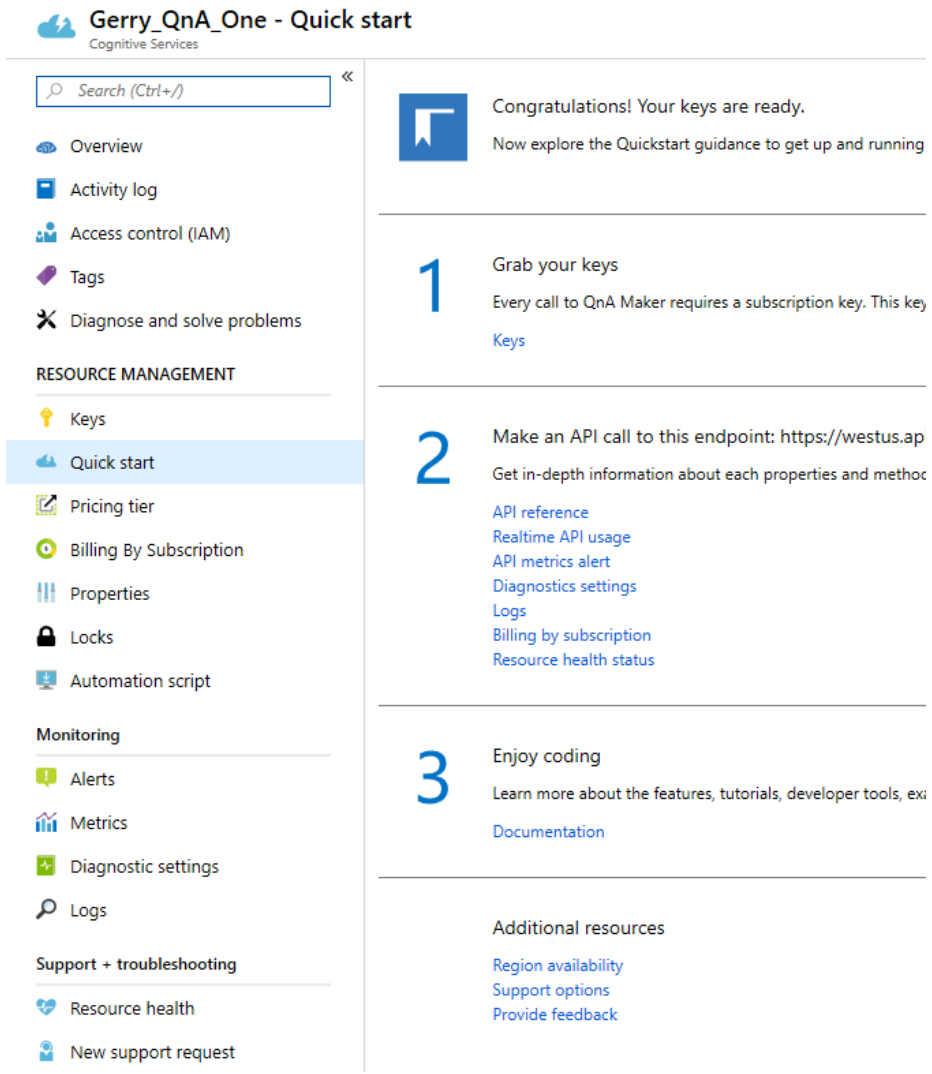
Cognitive Services Accounts



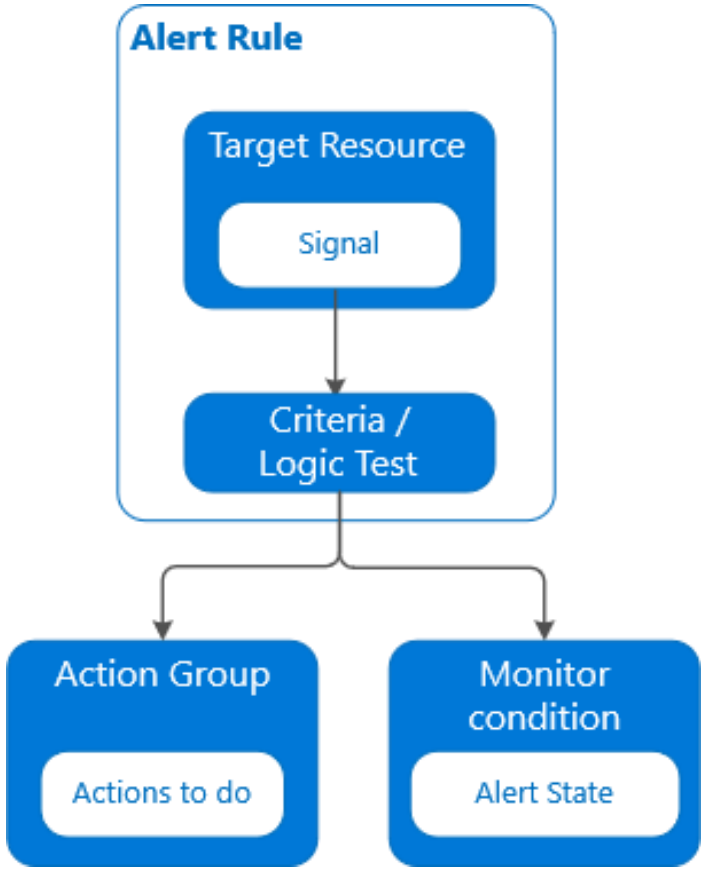
Create for single-service or multi-service usage



Manage Cognitive Service Accounts



Monitor Cognitive Service Accounts



Cognitive Services

Walkthrough – Create and Manage a Cognitive Services Account

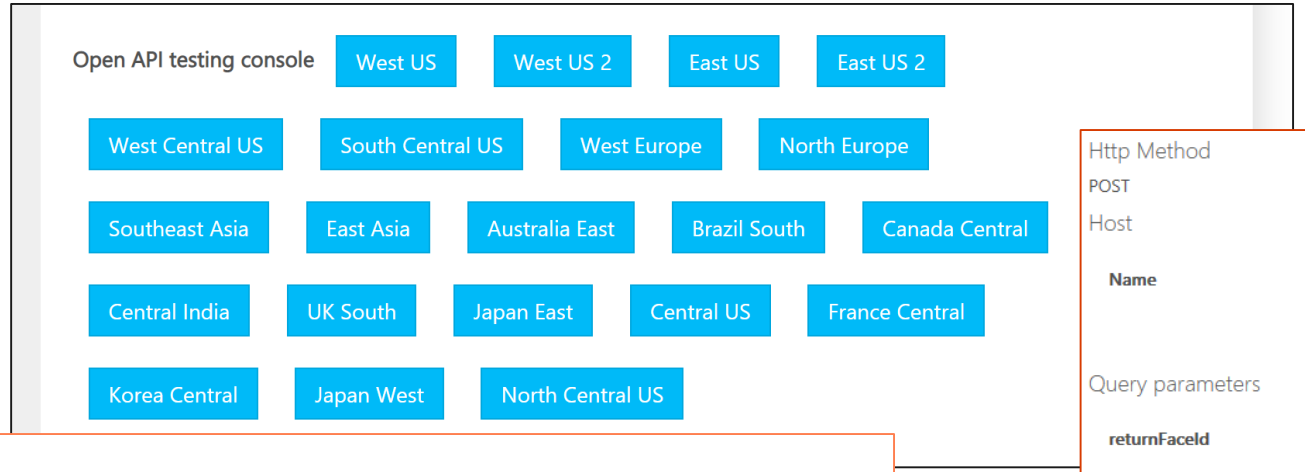
Lesson 03

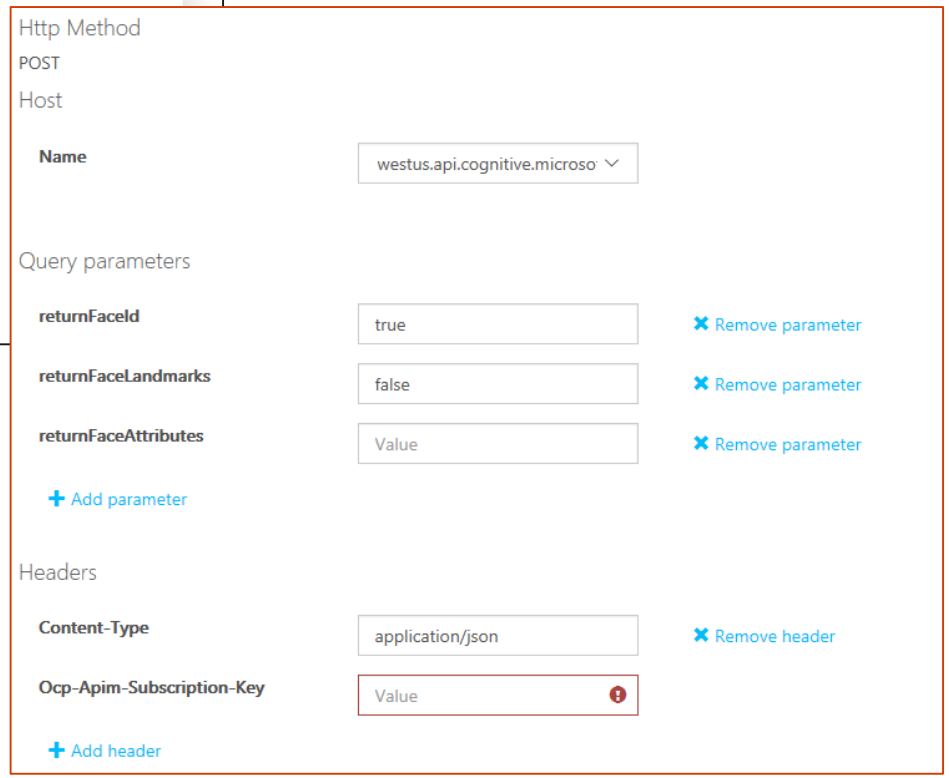
Testing Cognitive Services using the API Testing Console

Lesson Objectives

- L01 – Using the3 API Testing Console

The API Testing Console





Lab 1: Meeting the Technical Requirements

Lab Objectives

- L01 – Configure the lab environment
- L02 – Download necessary support files and assets

Lab 2: Implement Computer Vision

Lab Objectives

- L01 – Work with the Computer Vision API

Lab Scenario

- In this lab you learn how to implement the Computer Vision API. We use the ImageProcessing portable class library (PCL), discussing its contents and how to use it in your own applications.

Module Summary

In this module, you have learned about:

- Microsoft Azure Cognitive Services
- What Computer Vision is
- What the Speech Services are
- What the Language Services are
- Bing Search Services
- What Decision Services are available

Next steps

After the course, consider researching the other aspects of the Azure Cognitive Services and trying the samples presented on the site.