atmosera®

Gold Microsoft Partner Microsoft | Azure Expert MSP

AZURE COGNITIVE SERVICES

Jeff Prosise

Azure Cognitive Services

- AI as a Service (AIaaS) - Deep-learning models offered thru REST APIs
- Object detection, sentiment analysis, neural machine translation (NMT), real-time speech translation, anomaly detection, and more

Vision Services
Computer Vision
Custom Vision
Face*

Language Services
Language Service
Translator Service

Speech Service
Text to Speech
Speech to Text
Speech Translation
Speaker Recognition

Decision Services
Anomaly Detector
Content Moderator
Personalizer

OpenAI
Managed version of OpenAI models and APIs for generative AI

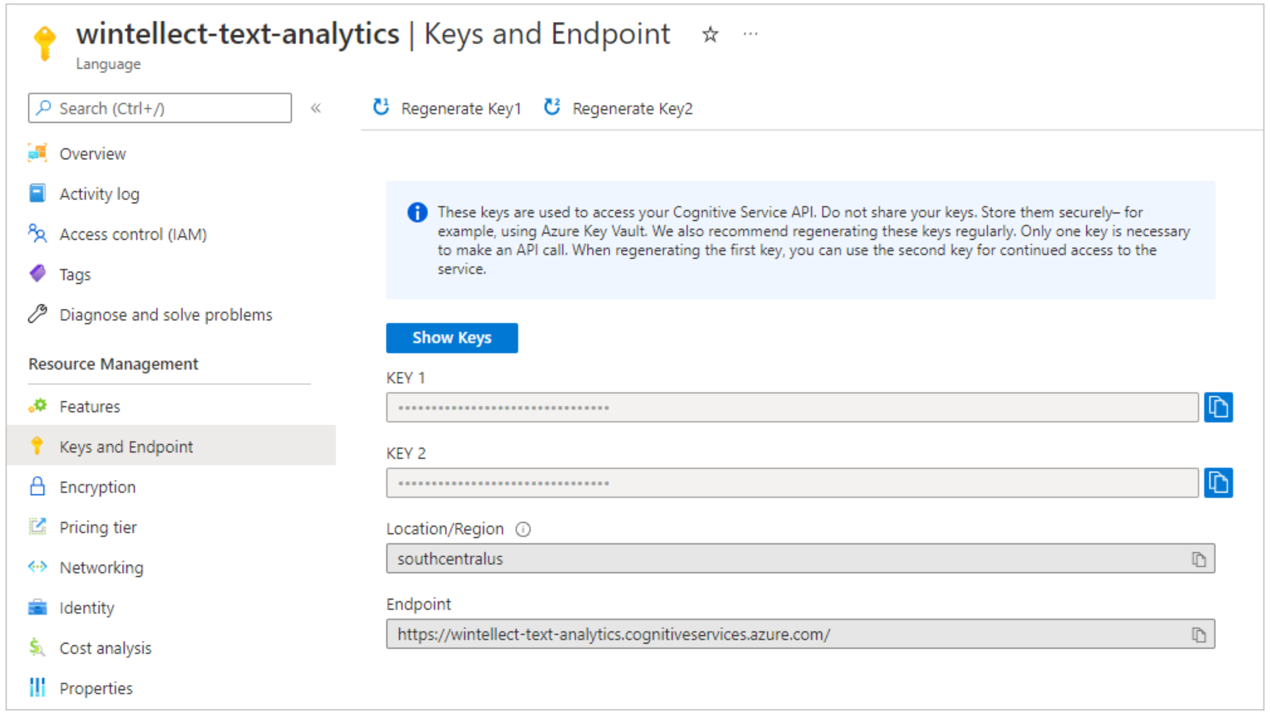
* No longer available to the general public. Available to managed customers and partners subject to approval by Microsoft.

Authentication

- Calls to Azure Cognitive Services must be authenticated so Microsoft can bill an Azure subscription for calls when appropriate
  - Most services have free tiers which are ideal for development
- All Cognitive Services support authentication with subscription keys obtained from the Azure Portal or the Azure CLI
  - Keys can be single-service or multi-service
- Most Cognitive Services support Azure Active Directory (AAD) authentication, too
  - Enhances security when caller is hosted in Azure

Key Management

- Guard subscription keys as if they were passwords
- Keys should be rotated periodically for security
- Azure provides two keys for each Cognitive Services resource you deploy so you can regenerate a key without affecting apps that use the service



These keys are used to access your Cognitive Service API. Do not share your keys. Store them securely– for example, using Azure Key Vault. We also recommend regenerating these keys regularly. Only one key is necessary to make an API call. When regenerating the first key, you can use the second key for continued access to the service.

KEY 1
••••••••••••••••••••••••••••••••

KEY 2
••••••••••••••••••••••••••••••••

Location/Region
southcentralus

Endpoint
https://wintellect-text-analytics.cognitiveservices.azure.com/

atmosera®

Gold Microsoft Partner Microsoft | Azure Expert MSP

Demo

Keys and Endpoints

Calling Cognitive Services APIs

```python
import requests

input = {'documents': [{'id': '1', 'text': 'Programming is fun, but the hours are long'}]}
headers = {
    'Ocp-Apim-Subscription-Key': KEY,
    'Content-type': 'application/json'
}

uri = ENDPOINT + 'text/analytics/v3.0/sentiment'
response = requests.post(uri, headers=headers, json=input)
results = response.json()

for result in results['documents']:
    print(result['confidenceScores'])
```

```
{'positive': 0.85, 'neutral': 0.03, 'negative': 0.12}
```

Cognitive Services SDKs

- Free SDKs for a variety of programming languages simplify your code and are available for most Cognitive Services

Language (Text Analytics)	C#	Java	JavaScript	Python				
Computer Vision	C#	Java	JavaScript	Python	Go			
Speech	C#	Java	JavaScript	Python	Go	C++	Swift	Objective-C
Anomaly Detector	C#	Java	JavaScript	Python	Go			

Using the Python Text-Analytics SDK

```python
from azure.core.credentials import AzureKeyCredential
from azure.ai.textanalytics import TextAnalyticsClient

client = TextAnalyticsClient(ENDPOINT, AzureKeyCredential(KEY))
input = [{ 'id': '1', 'text': 'Programming is fun, but the hours are long' }]
response = client.analyze_sentiment(input)

for result in response:
    print(result.confidence_scores)
```

```
{'positive': 0.85, 'neutral': 0.03, 'negative': 0.12}
```

Using the .NET Text-Analytics SDK

```
using Azure;
using Azure.AI.TextAnalytics;
using System;

var client = new TextAnalyticsClient(new Uri(ENDPOINT), new AzureKeyCredential(KEY));
var response = client.AnalyzeSentiment("Programming is fun, but the hours are long");
var result = response.Value.ConfidenceScores.Positive;
Console.WriteLine(result);
```

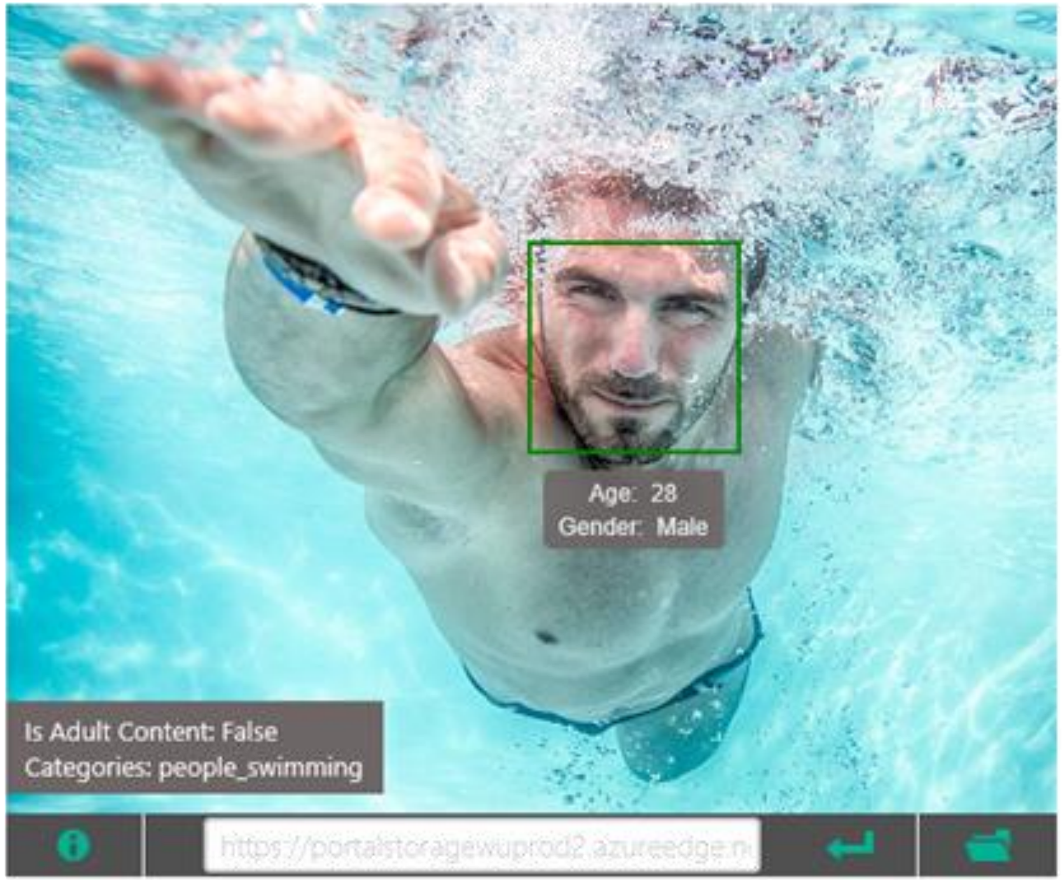
Handling Errors

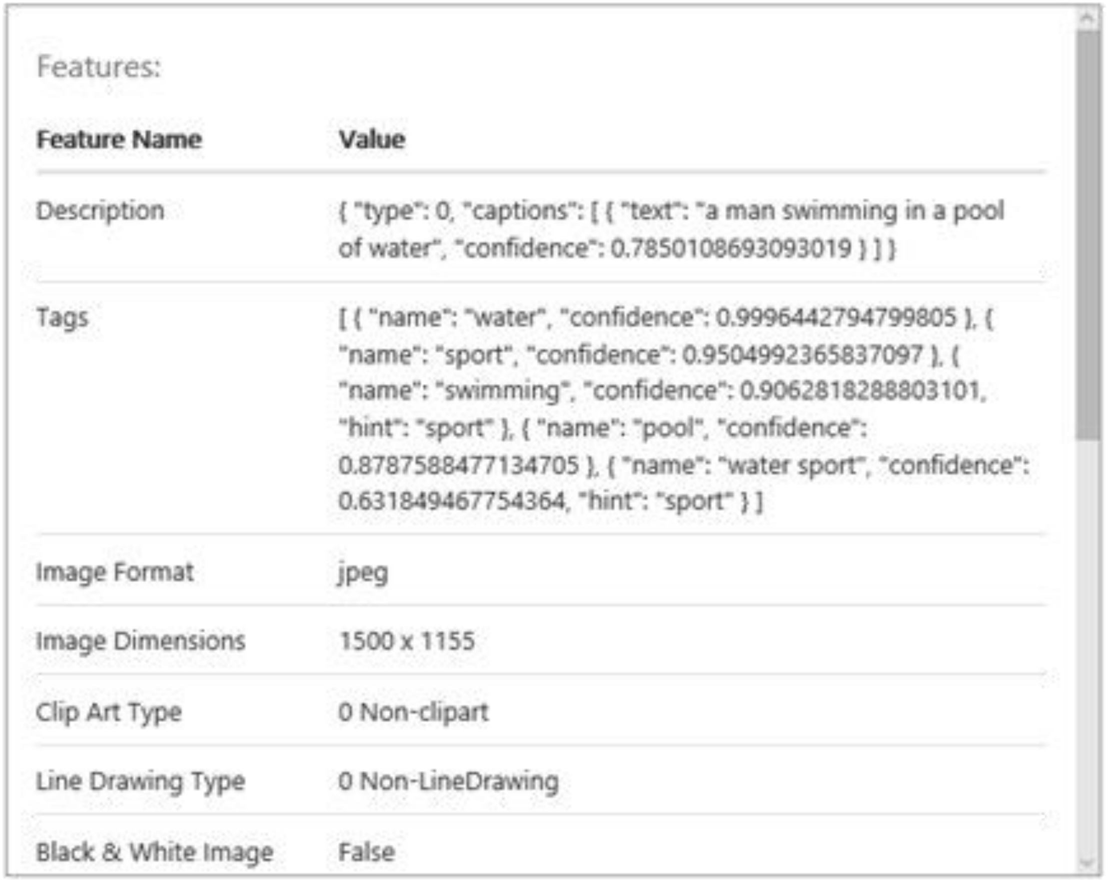
```python
from azure.core.credentials import AzureKeyCredential
from azure.ai.textanalytics import TextAnalyticsClient
from azure.core.exceptions import AzureError

try:
    client = TextAnalyticsClient(ENDPOINT, AzureKeyCredential(KEY))
    input = [{ 'id': '1', 'text': 'Programming is fun, but the hours are long' }]
    response = client.analyze_sentiment(input)

    for result in response:
        print(result.confidence_scores)

except AzureError as e:
    print(e.message)
```

Computer Vision Service

- Analyze images and extract information from them
  - Caption photos
  - Generate tags describing the contents of a photo
  - Detect and identify objects in photos (including bounding boxes)
  - Detect faces in photos and identify ages and genders
  - Identify images containing adult or inappropriate content
  - Generate "smart thumbnails," extract text from images, and more
- Python SDK in **azure-cognitiveservices-vision-computervision**
- **ComputerVisionClient** class provides wrapper around REST API

Computer Vision Service in Action


Age: 28
Gender: Male
Is Adult Content: False
Categories: people_swimming
https://portalstoragewuprod2.azureedge.n



Features:

Feature Name	Value
Description	{ "type": 0, "captions": [{ "text": "a man swimming in a pool of water", "confidence": 0.7850108693093019 }] }
Tags	[{ "name": "water", "confidence": 0.9996442794799805 }, { "name": "sport", "confidence": 0.9504992365837097 }, { "name": "swimming", "confidence": 0.9062818288803101, "hint": "sport" }, { "name": "pool", "confidence": 0.8787588477134705 }, { "name": "water sport", "confidence": 0.631849467754364, "hint": "sport" }]
Image Format	jpeg
Image Dimensions	1500 x 1155
Clip Art Type	0 Non-clipart
Line Drawing Type	0 Non-LineDrawing
Black & White Image	False

Captioning an Image

```python
client = ComputerVisionClient(ENDPOINT, CognitiveServicesCredentials(KEY))

# Submit image from local file system
with open('IMAGE_PATH', mode='rb') as image:
    result = client.describe_image_in_stream(image)
    for caption in result.captions:
        print(f'{caption.text} ({caption.confidence:.1%})')

# Submit image URL
result = client.describe_image('IMAGE_URL')
for caption in result.captions:
    print(f'{caption.text} ({caption.confidence:.1%})')
```

Generating Keywords from Images

```
client = ComputerVisionClient(ENDPOINT, CognitiveServicesCredentials(KEY))

# Submit image from local file system
with open('IMAGE_PATH', mode='rb') as image:
    result = client.tag_image_in_stream(image)
    for tag in result.tags:
        print(f'{tag.name} ({tag.confidence:.1%})')
```

Detecting Objects in Images

```
client = ComputerVisionClient(ENDPOINT, CognitiveServicesCredentials(KEY))

# Submit image from local file system
with open('IMAGE_PATH', mode='rb') as image:
    result = client.detect_objects_in_stream(image)
    for obj in result.objects:
        print(f'{obj.object_property}, {obj.confidence:.1%}, {obj.rectangle})')
```

Object name (person, dog, bicycle, etc.)

Bounding box

atmosera®

Gold Microsoft Partner Microsoft | Azure Expert MSP

Demo

The Computer Vision Service

Language Service

- Supports text analytics (e.g., sentiment analysis, named-entity recognition, and key-phrase extraction), question answering (semantic search), conversational language understanding, and more
- Helper classes available in Python SDKs
  - **TextAnalyticsClient** class in **azure-ai-textanalytics**
  - **QuestionAnsweringClient** class in **azure-ai-language-questionanswering**
  - **ConversationAnalysisClient** class in **azure-ai-language-conversations**
- Used by LaLiga to boost fan engagement with conversational AI

https://customers.microsoft.com/en-us/story/laliga-media-entertainment-azure

Analyzing Sentiment

```python
client = TextAnalyticsClient(ENDPOINT, AzureKeyCredential(KEY))
input = [{ 'id': '1', 'text': 'Programming is fun, but the hours are long' }]
response = client.analyze_sentiment(input)

for result in response:
    print(result.confidence_scores)
```

```
{'positive': 0.85, 'neutral': 0.03, 'negative': 0.12}
```

Analyzing Sentiment in Batches

```
input = [
    { 'id': '1', 'text': 'Programming is fun, but the hours are long' },
    { 'id': '2', 'text': 'Great food and excellent service' },
    { 'id': '3', 'text': 'The product worked as advertised but is overpriced' },
    { 'id': '4', 'text': 'Moving to the cloud was the best decision we ever made' }
]

client = TextAnalyticsClient(ENDPOINT, AzureKeyCredential(KEY))
response = client.analyze_sentiment(input)

for result in response:
    text = ''.join([x.text for x in result.sentences])
    print(f'{text} => {result.confidence_scores.positive}')
```

Question Answering

```python
from azure.ai.language.questionanswering import models as qna
from azure.ai.language.questionanswering import QuestionAnsweringClient

client = QuestionAnsweringClient(ENDPOINT, AzureKeyCredential(KEY))
question = 'What is the minimum age required to serve as a United State Senator?'

context = 'No Person shall be a Senator who shall not have attained to the Age of ' \
          'thirty Years, and been nine Years a Citizen of the United States, and ' \
          'who shall not, when elected, be an Inhabitant of that State for which ' \
          'he shall be chosen.'

input = qna.AnswersFromTextOptions(question=question, text_documents=[context])
results = client.get_answers_from_text(input)
answer = results.answers[0].short_answer
print(f'{answer.text} ({answer.confidence:.1%})')
```

```
thirty Years (88.0%)
```

Translator Service

- Uses neural machine translation (NMT) to translate text to other languages and understands >100 written languages and dialects
- Supports two APIs (one service, two endpoints)
  - Text translation API - No Python SDK available at this time
  - Document translation API - SDK available in **azure-ai-translation-document** to batch-translate documents (including PDFs) at scale in Azure Blob Storage
- Used by Volkswagen to translate onscreen instruction in cars and translate documentation into more than 40 languages

https://customers.microsoft.com/en-us/story/779468-volkswagen-azure-automotive-en

Translating French to English

```
input = [{ 'text': 'Quand votre nouveau livre sera-t-il disponible?' }]

headers = {
    'Ocp-Apim-Subscription-Key': KEY,
    'Ocp-Apim-Subscription-Region': REGION,
    'Content-type': 'application/json'
}

uri = ENDPOINT + 'translate?api-version=3.0&from=fr&to=en'
response = requests.post(uri, headers=headers, json=input)
results = response.json()

print(results[0]['translations'][0]['text'])
```

```
When will your new book be available?
```

atmosera®

Gold Microsoft Partner Microsoft | Azure Expert MSP

Demo

Language Services

Speech Service

- Converts text to speech and speech to text, identifies spoken languages, and translates speech to other languages in real time
- Python SDK available in **azure-cognitiveservices-speech**
  - **SpeechRecognizer** class converts speech to text
  - **SpeechSynthesizer** class converts text to speech in more than 300 voices
  - **TranslationRecognizer** class converts speech to text and translates it, too
- Used by Airbus to build voice-enabled apps for pilots and by KPMG to transcribe calls and reduce compliance costs for clients

https://customers.microsoft.com/en-us/story/858578-airbus-defense-and-intelligence-azure

Converting Text to Speech

```
from azure.cognitiveservices import speech

config = speech.SpeechConfig(KEY, REGION)
config.speech_synthesis_voice_name = 'en-US-JennyNeural'

synthesizer = speech.SpeechSynthesizer(config)
synthesizer.speak_text_async('When will your new book be published?').get()

# For a complete list of more than 300 neural voices, see https://docs.microsoft.com/en-
us/azure/cognitive-services/speech-service/language-support?tabs=speechtotext#prebuilt-
neural-voices
```

atmosera®

Gold Microsoft Partner Microsoft | Azure Expert MSP

Demo

The Speech Service

OpenAI

- Creator of ChatGPT, DALL·E 2, and other cutting-edge models
- Deploy models to Azure and access REST APIs for generating text, images, code, text embeddings, and more

GPT-3 **GPT-3.5** **GPT-4**	**Codex**	**DALL·E 2**	**Embeddings**	**Whisper**
Models that support **text and code generation**. Includes **ChatGPT**.	Models that support **code generation**, converting code to other languages, and more. **Discontinued on March 23, 2023**.	Produces **images from natural-language prompts**. Supports inpainting, outpainting, image variations, and more.	Models that generate **embedding vectors** from text. Used for semantic search, recommender systems, and more.	General-purpose **speech-to-text engine** trained on 680,000 hours of audio. Also available via open source.

Generating Text with ChatGPT

```python
import openai
openai.api_type = 'azure'
openai.api_version = '2023-03-15-preview'
openai.api_base = 'API_ENDPOINT' # Obtained from Azure portal
openai.api_key = 'API_KEY' # Obtained from Azure portal

messages = [{ 'role': 'user', 'content': 'Write a poem about deep learning' }]

response = openai.ChatCompletion.create(
    engine='my-chatgpt-instance', # Azure deployment name
    messages=messages
)
```

Streaming the Response

```python
messages = [{ 'role': 'user', 'content': 'Write a poem about deep learning' }]

chunks = openai.ChatCompletion.create(
    engine='my-chatgpt-instance',
    messages=messages,
    stream=True
)

for chunk in chunks:
    content = chunk['choices'][0].get('delta', {}).get('content')
    if content is not None:
        print(content, end='')
```

Translating Text

```
content = 'Translate the following text from English to French: Best food ever!'

messages = [{ 'role': 'user', 'content': content }]

response = openai.ChatCompletion.create(
    engine='my-chatgpt-instance',
    messages=messages,
    temperature=0
)
```

Generating Code

```python
content = 'Create a Python function that accepts an array of numbers as ' \
          'input, bubble sorts the numbers, and returns a sorted array'

messages = [{ 'role': 'user', 'content': content }]

response = openai.ChatCompletion.create(
    engine='my-chatgpt-instance',
    messages=messages,
    temperature=0
)
```

Handling Errors

```python
try:
    response = openai.ChatCompletion.create(
        engine='my-chatgpt-instance',
        messages=messages
    )

    print(response.choices[0].message.content)

except AuthenticationError as e:
    print('Invalid API key')
except ServiceUnavailableError as e:
    print('ChatGPT temporarily unavailable')
except Exception as e:
    print('Call failed')
```

atmosera®

Gold Microsoft Partner Microsoft

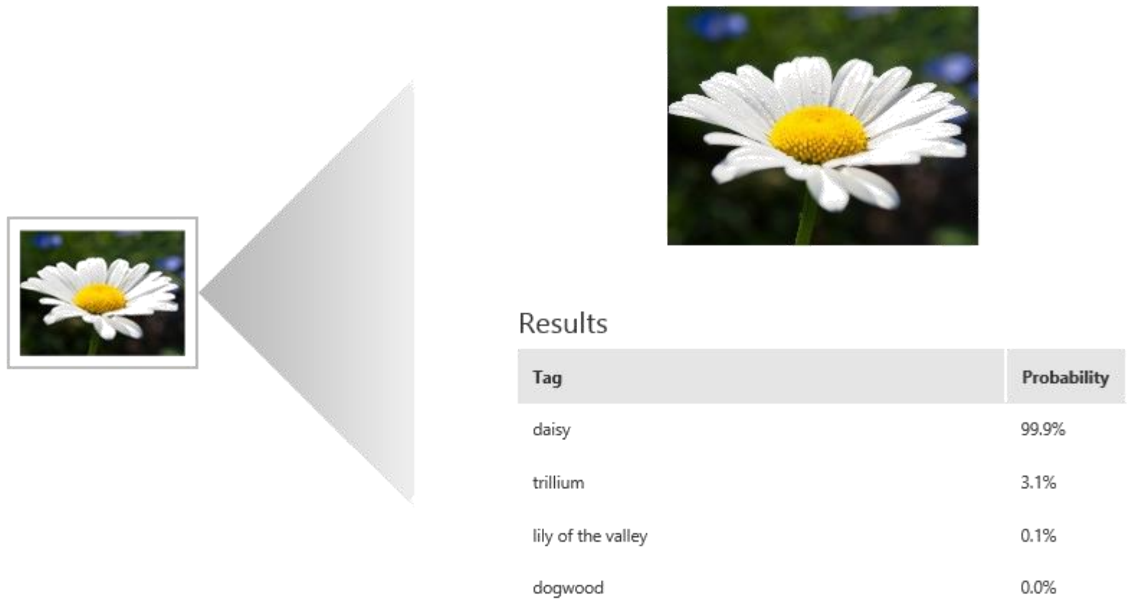
Azure Expert MSP

Demo

OpenAI

Custom Vision Service

- Build custom image-classification and object-detection models
  - Get acceptable accuracy with as few as 50 to 100 training images
- Build intelligent apps that invoke models using REST API
- Or export to CoreML or TensorFlow and run models locally

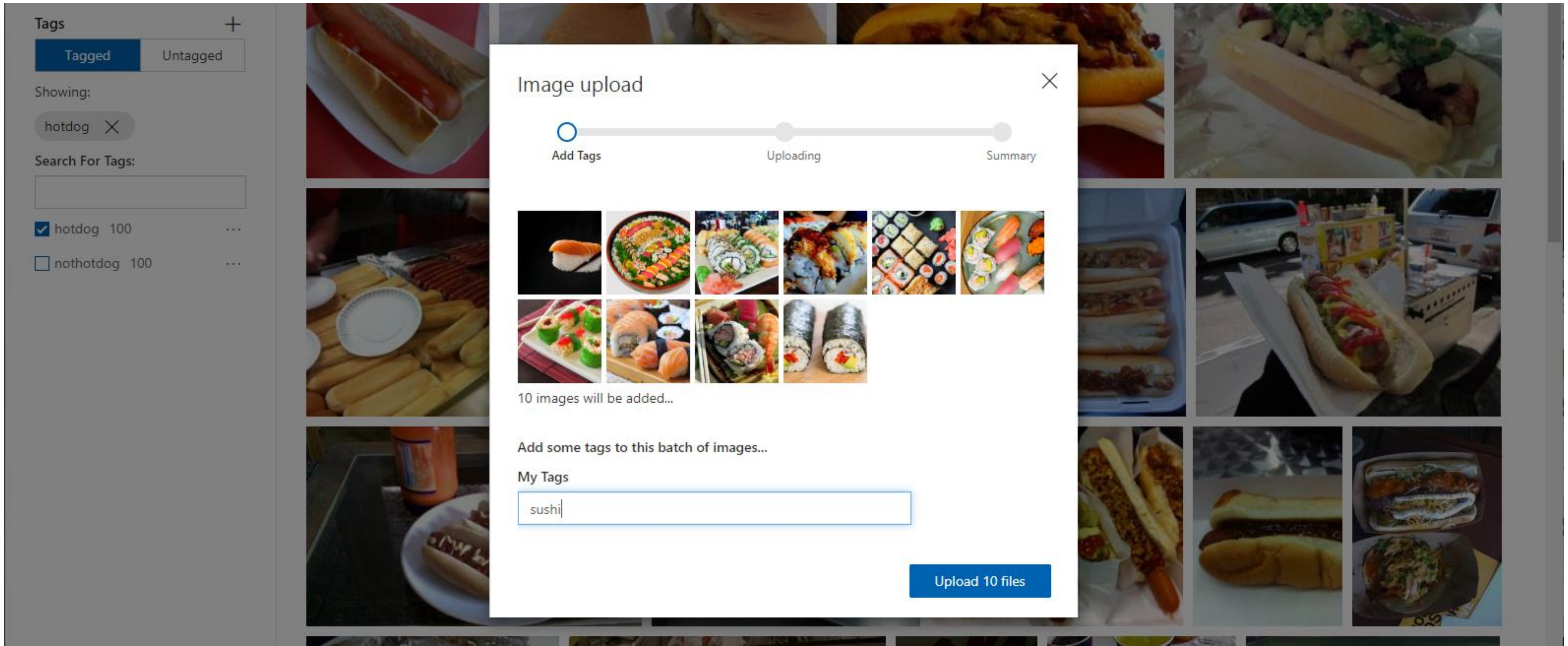


Results

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

Training a Custom Vision Service Model

- Use Custom Vision portal (www.customvision.ai) to create projects, upload and tag images, and train and publish models



10 images will be added...

Add some tags to this batch of images...

My Tags

sushi

atmosera®

Gold Microsoft Partner Microsoft | Azure Expert MSP

Demo

The Custom Vision Service

Cognitive Services Containers

- Most Cognitive Services can be hosted locally in Docker containers
  - Keep data on-premises (avoid sending it to Azure)
  - Insulate your code from API changes and updates
  - Run Cognitive Services physically close to apps that use them
  - Use Kubernetes or other container orchestration platforms to control scaling
- Containerized services do NOT prevent you from being billed
  - Containers send metering information to Microsoft using port 443

https://docs.microsoft.com/en-us/azure/cognitive-services/cognitive-services-container-support

Disconnected Containers

- Containers that run without an Internet connection
  - Do not send metering information to Microsoft
  - For environments where connectivity is non-existent or spotty or policies and regulations forbid apps that use Cognitive Services from accessing the Internet
- Only available for subsets of certain services (e.g., text extraction via Computer Vision service)
- Only available to "strategic customers and partners" with Enterprise Agreements (EAs) in place and must be approved by Microsoft

https://docs.microsoft.com/en-us/azure/cognitive-services/containers/disconnected-containers