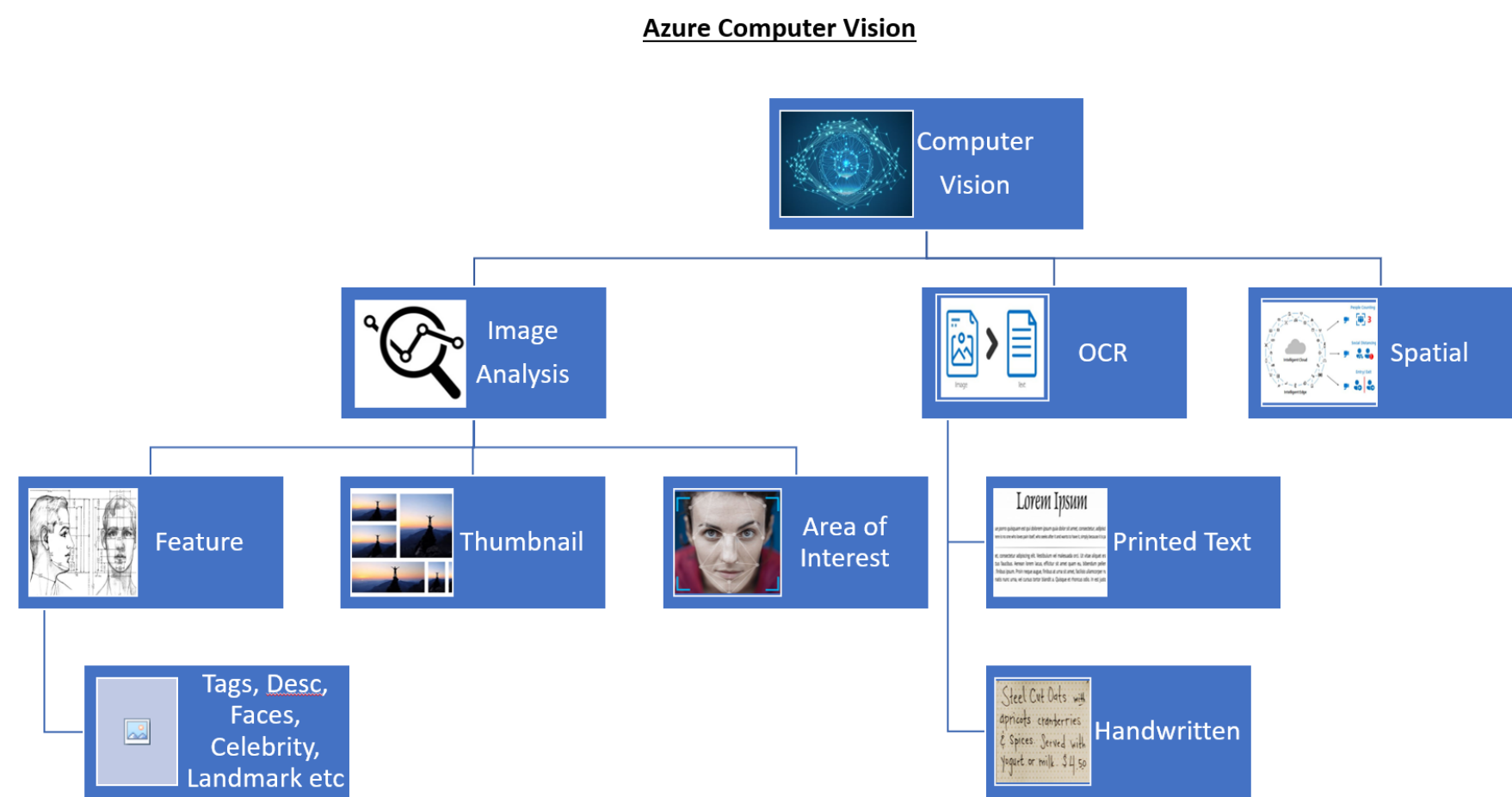


Document Name	HOL – Azure Computer Vision Service v3.2
Author	Shiva S Tomar & Anupreet Kaur
Reviewer	
Executive Summary	Azure Cognitive APIs enable the developers of all skill levels to add human intelligence in their applications. The services are designed for developers interested in pursuing DS/AI/ML skills and people who want to acquire the deep technical knowledge on the Cognitive APIs of Azure, despite not having Machine Learning expertise.
Purpose	This document is created to help you gain level 350 working knowledge on Azure Computer Vision Service. You will be able to explore each functionality offered by the service through the API and observe the outcomes. We have also shared a sample dataset to replicate what we have used to create the content of this workshop. Once you complete these labs, you’ll go from <i>Zero to Hero</i> on the respective Azure Cognitive service and should be able to <i>Demo, Develop and Deploy</i> your own custom use cases. The important thing to note here is that you don’t need to refer any other documents to complete this workshop.
Intent of Guide	This workshop is designed to help you explore all the features of a service offered through their APIs. The diagram shown in the beginning of the document is its functional Architecture; talking about the functionalities offered by the service in a flow. It also covers the Concepts, How-to and best practices about the service. This document is not intended to enable you with scenarios of deployment in production.

Service brief: Azure Computer Vision Service

Azure's Computer Vision service gives you access to advanced algorithms that process images and return information based on the visual features you're interested in. For Ex: a user can perform OCR – to extract text information in an image; image processing – to identify the content of the image & spatial Analysis – to analyse the presence and movement of people on a video feed and produces events that other systems can respond to.

Diagram: Functional Architecture



The service offers various functionalities under 3 major categories, which are as follows :

1. **Image Analysis** : Here, the model allows us to extract features from images, such as Tags, Description, Objects, Faces, Celebrities, Landmarks etc. The model also gives us the Area of Interest and Thumbnails for the image, by intelligently understanding the important areas in the image. Your model can then focus on the important aspects of the image while processing it further, thereby, giving you better outcomes.
2. **Optical Character Recognition** : This allows you to read text from printed or handwritten images & documents, thereby, allowing you to digitise and extract information from them. You can use this to process invoices, bills, financial reports, articles etc.
3. **Spatial Analyses** : This allows you to analyse your physical spaces by understanding people's movements and presence within a given area. This can be leveraged for scenarios like tracking people, security surveillance etc.

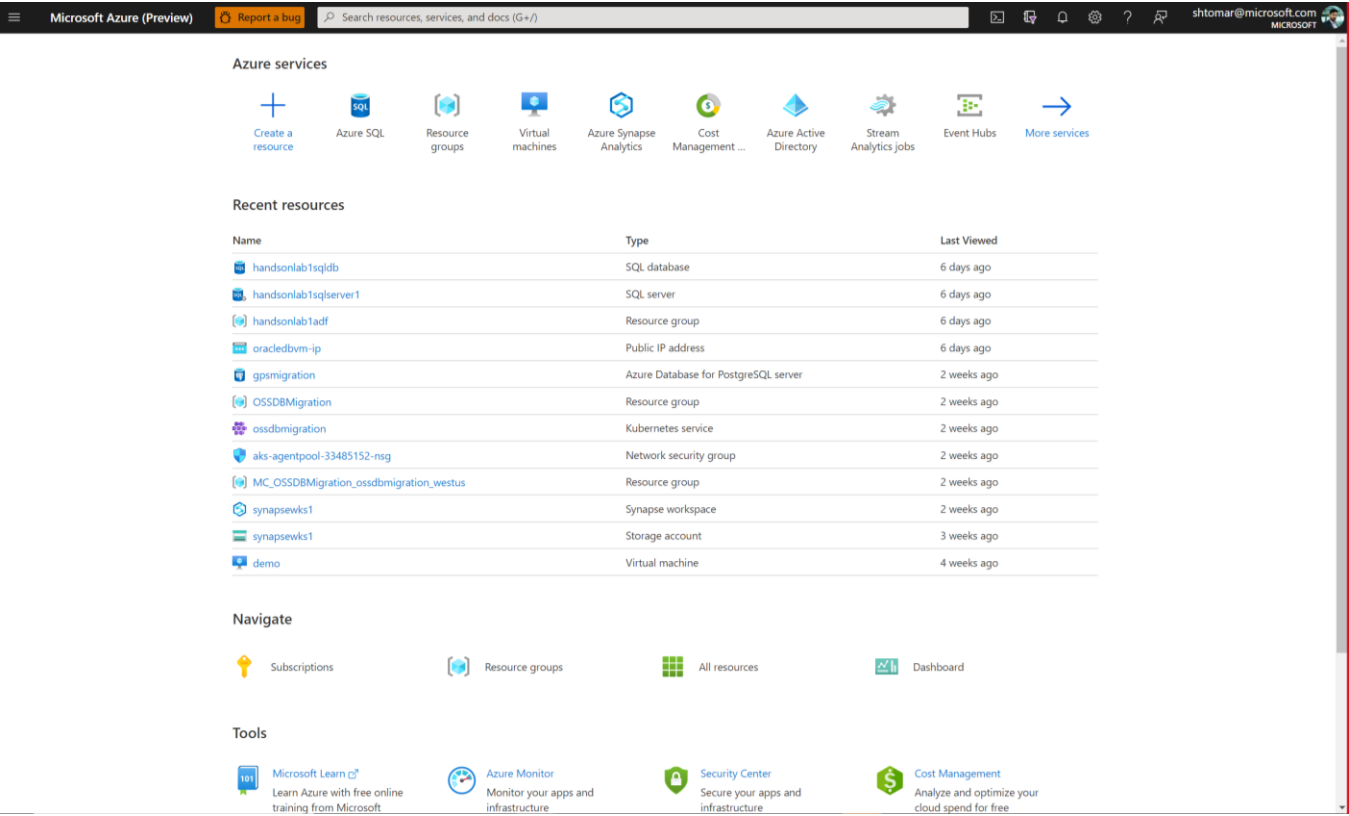
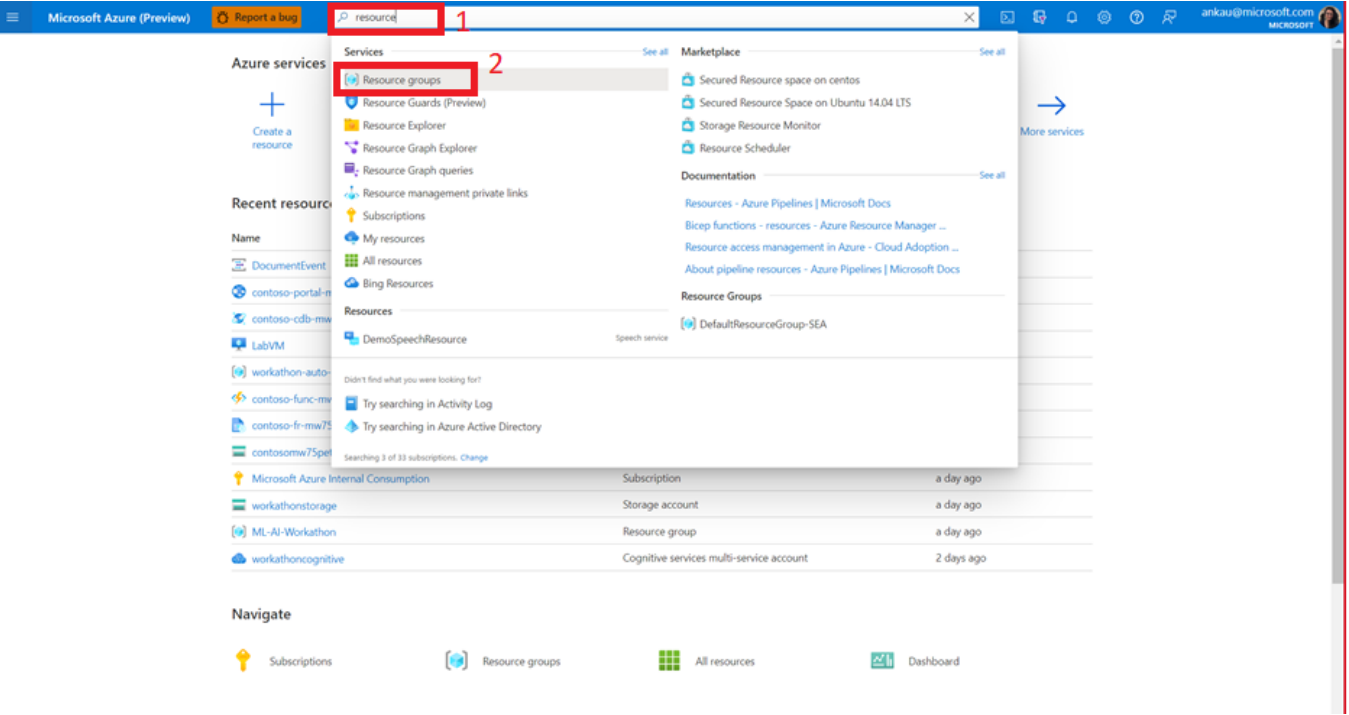
These APIs are available both as REST APIs and language specific SDKs.
You can deploy the models on the cloud or on the edge.

Step by step hands on guide to go from *Zero to Hero*

Pre-requisites

- Download & Install Postman
 - Postman is a free tool which allows you to make API calls
 - You can download the desktop application or get started using the web version ([Download Postman | Try Postman for Free](#))
- An active Azure Account
 - You can use your current Azure Subscription or get started by creating a free trial account (<https://azure.microsoft.com/en-in/free>)

Let’s get started!

Screenshots	Steps & Significance
	Sign into your Azure Portal.
	<p><u>Create a Resource Group</u></p> <p>Follow steps 1 & 2 to create a resource group.</p>

Microsoft Azure (Preview)

Report a bug

Search resources, services, and docs (G+I)

Home >

Resource groups

+ Create

Manage view

Refresh

Export to CSV

Open query

Assign tags

Feedback

Filter for any field...

Subscription == 3 of 33 selected

Location == all

Add filter

Showing 1 to 24 of 24 records.

No groupingList view

Name	Subscription	Location
AzureMachineLearningDemo	anu_internalSubscription	Southeast Asia
BotsDemo	anu_internalSubscription	Central India
cloud-shell-storage-centralindia	anu_internalSubscription	Central India
CognitiveServicesDemo	anu_internalSubscription	East US
databricks-rg-DatabricksDemo1-xm1ang22kqe	anu_internalSubscription	Southeast Asia
DatabricksMain	anu_internalSubscription	Southeast Asia
DefaultResourceGroup-SEA	anu_internalSubscription	Southeast Asia
formrecogdemo1	anu_internalSubscription	Central India
FormRecognizerDemo	anu_internalSubscription	East US
handsontab1adff	anu_internalSubscription	Central India
handsontab1adff-synapse	anu_internalSubscription	Central India
ML-AI-Workathon	Microsoft Azure Internal Consumption	Central India
NetworkWatcherRG	anu_internalSubscription	East US
ossdb	anu_internalSubscription	Central India
QnAMakerDemo	anu_internalSubscription	West US
rg0906	anu_internalSubscription	Southeast Asia
synapse-random	anu_internalSubscription	Central India
synapseworkshop	anu_internalSubscription	Central India

Click create to create a new resource group.

Microsoft Azure (Preview)

Report a bug

Search resources, services, and docs (G+I)

Home > Resource groups >

Create a resource group

Basics

Tags

Review + create

Resource group - A container that holds related resources for an Azure solution. The resource group can include all the resources for the solution, or only those resources that you want to manage as a group. You decide how you want to allocate resources to resource groups based on what makes the most sense for your organization. [Learn more](#)

Project details

Subscription *

anu_internalSubscription

Resource group *

ML-AI-Workathon

Resource details

Region *

(Asia Pacific) Central India

Review + create

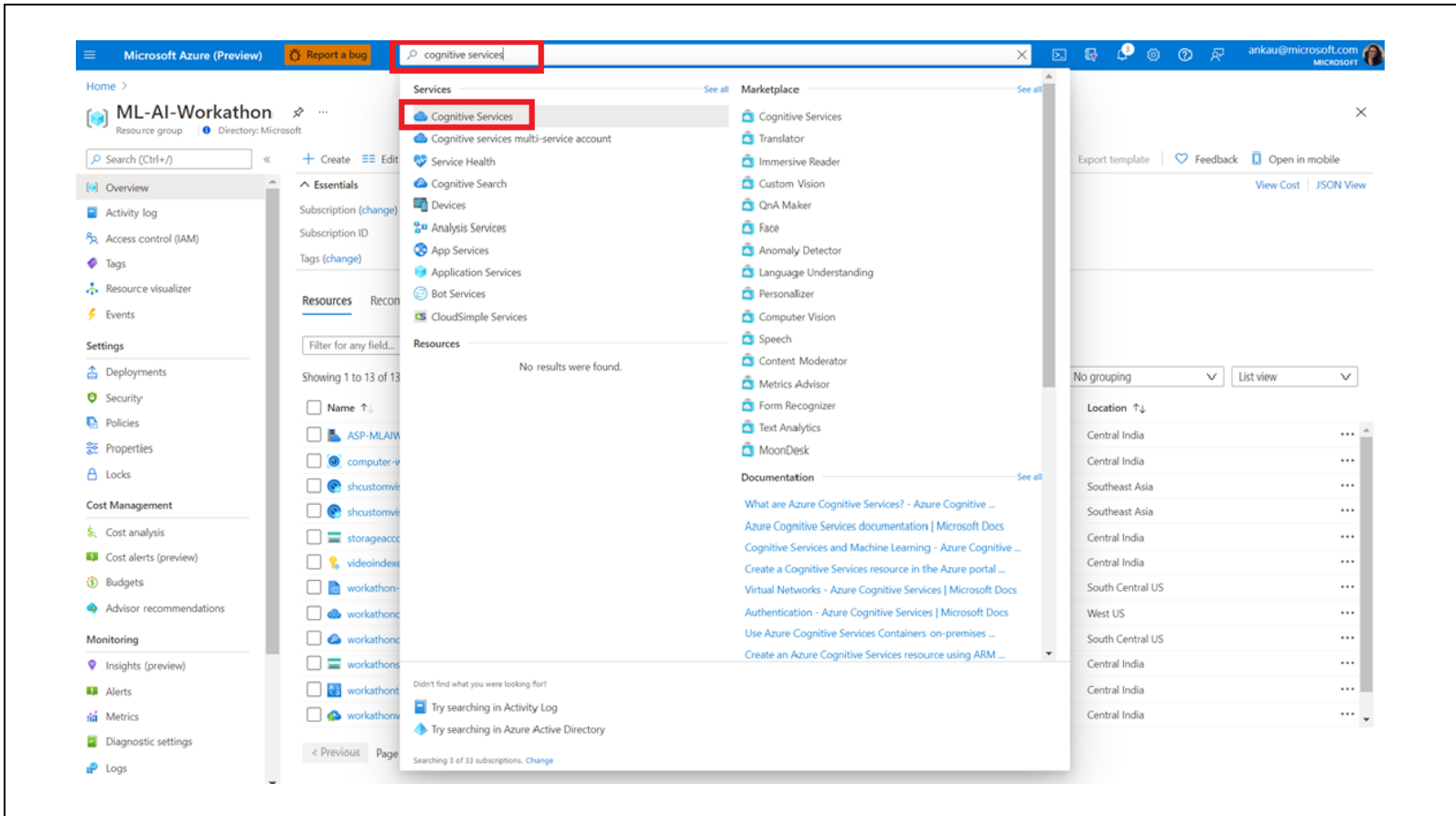
< Previous

Next > Tags

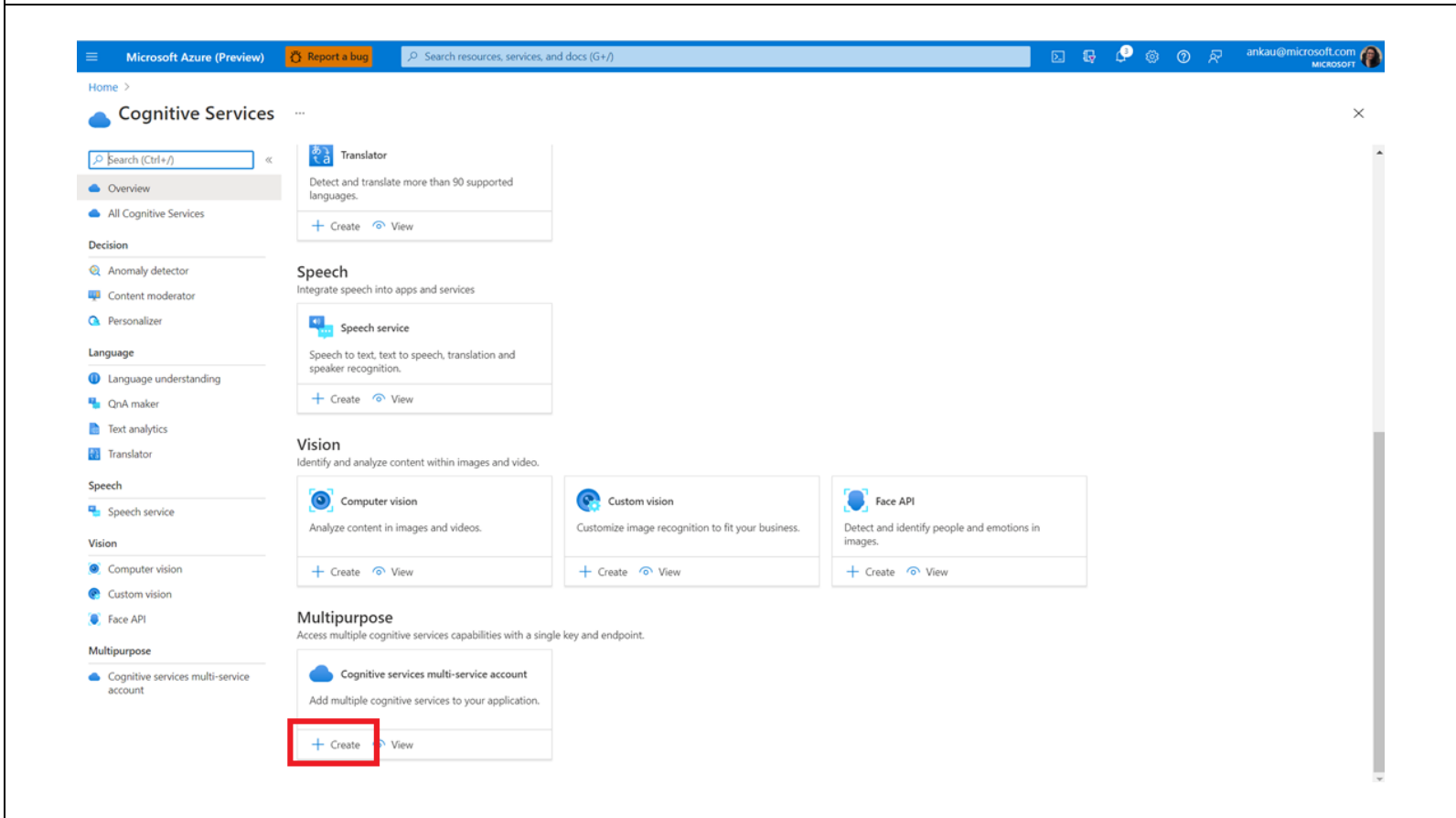
Enter the details –

- Subscription : Azure subscription in which you want to deploy the resource group
- Resource Group : Name of your choice for the resource group
- Region : Region where you want to deploy the resource group

Click Review + Create.



Once the resource group is created, search for Cognitive Services in the search bar above and select Cognitive Services.



Create a multipurpose cognitive service.

Significance : A multipurpose Cognitive Service account allows you to leverage the same resource for many cognitive services, which include :

- [Computer Vision](#) - Analyze images
- [Content Moderator](#) - Check text, image or videos for offensive or undesirable content
- [Face](#) - Recognize people and their attributes in an image
- [Form Recognizer](#) - Identify and extract text, key/value pairs and table data from form documents
- [Language Understanding](#) - Extract meaning from natural language
- [Speech](#) - Transform speech-to-text, text-to-speech and recognize speakers
- [Text Analytics](#) - Detect sentiment, key phrases, entities and human language type in text

In this lab, we used a multipurpose Cognitive Service account since we would be learning about all the above-mentioned services. However, you can also spin up individual services to execute these labs or for your development / production scenarios. The only difference is spinning up individual services allows logical separation from workspace standpoint and easy monitoring of billability.

Microsoft Azure (Preview)

Report a bug

Search resources, services, and docs (G+)

ankau@microsoft.com

Home > Cognitive Services >

Create Cognitive Services

BasicsIdentityTagsReview + create

Get access to Vision, Language, Search, and Speech Cognitive Services with a single API key. Quickly connect services together to achieve more insights into your content and easily integrate with other services like Azure Search. [Learn more](#)

Project details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription *

anu_internalSubscription

Resource group *

(New) ML-AI-Workathon

Create new

Instance details

Region *

West US

Location specifies the region only for included regional services. This does not specify a region for included non-regional services. Custom Vision service is not available in this region. Click here for more details.

Name *

workathoncognitive

Pricing tier *

Standard S0

[View full pricing details](#)

By checking this box I acknowledge that I have read and understood all the terms below.☒

Review + create< PreviousNext: Identity >

Enter the details to create a new cognitive service as follows -

Project details	Description
Subscription	Select one of your available Azure subscriptions.
Resource group	The Azure resource group that will contain your Cognitive Services resource. You can create a new group or add it to a pre-existing group.
Region	The location of your cognitive service instance. Different locations may introduce latency but have no impact on the runtime availability of your resource.
Name	A descriptive name for your cognitive services resource.
Pricing tier	The cost of your Cognitive Services account depends on the options you choose and your usage.

Click Review + Create.

Microsoft Azure (Preview)

Report a bug

Search resources, services, and docs (G+)

ankau@microsoft.com

Home > Cognitive Services >

Create Cognitive Services

BasicsIdentityTagsReview + create

Validation Passed

TERMS

By clicking "Create", I (a) agree to the legal terms and privacy statement(s) associated with the Marketplace offering(s) listed above; (b) authorize Microsoft to bill my current payment method for the fees associated with the offering(s), with the same billing frequency as my Azure subscription; and (c) agree that Microsoft may share my contact, usage and transactional information with the provider(s) of the offering(s) for support, billing and other transactional activities. Microsoft does not provide rights for third-party offerings. See the [Azure Marketplace Terms](#) for additional details.

Basics

Subscriptionanu_internalSubscription

Resource groupML-AI-Workathon

RegionWest US

Nameworkathoncognitive

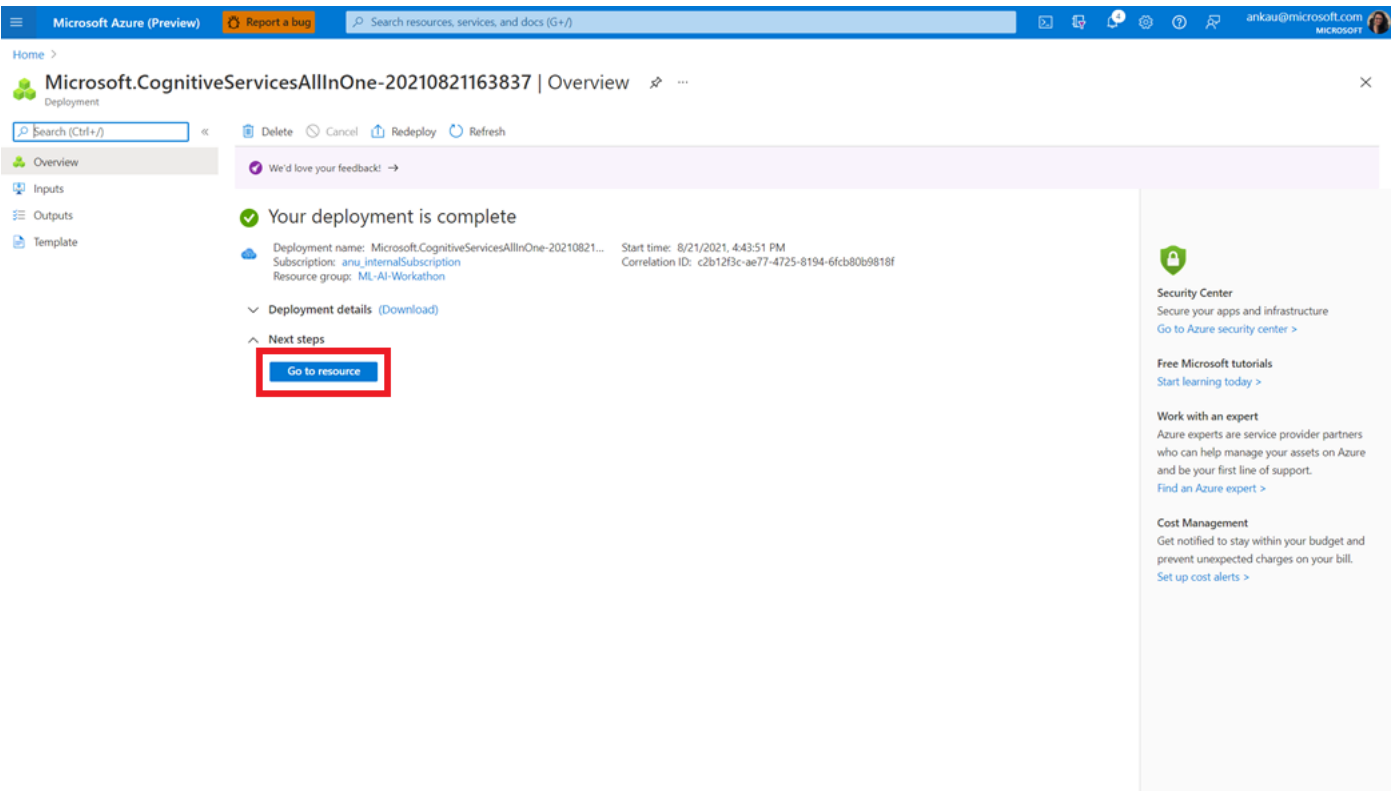
Pricing tierStandard S0

Identity

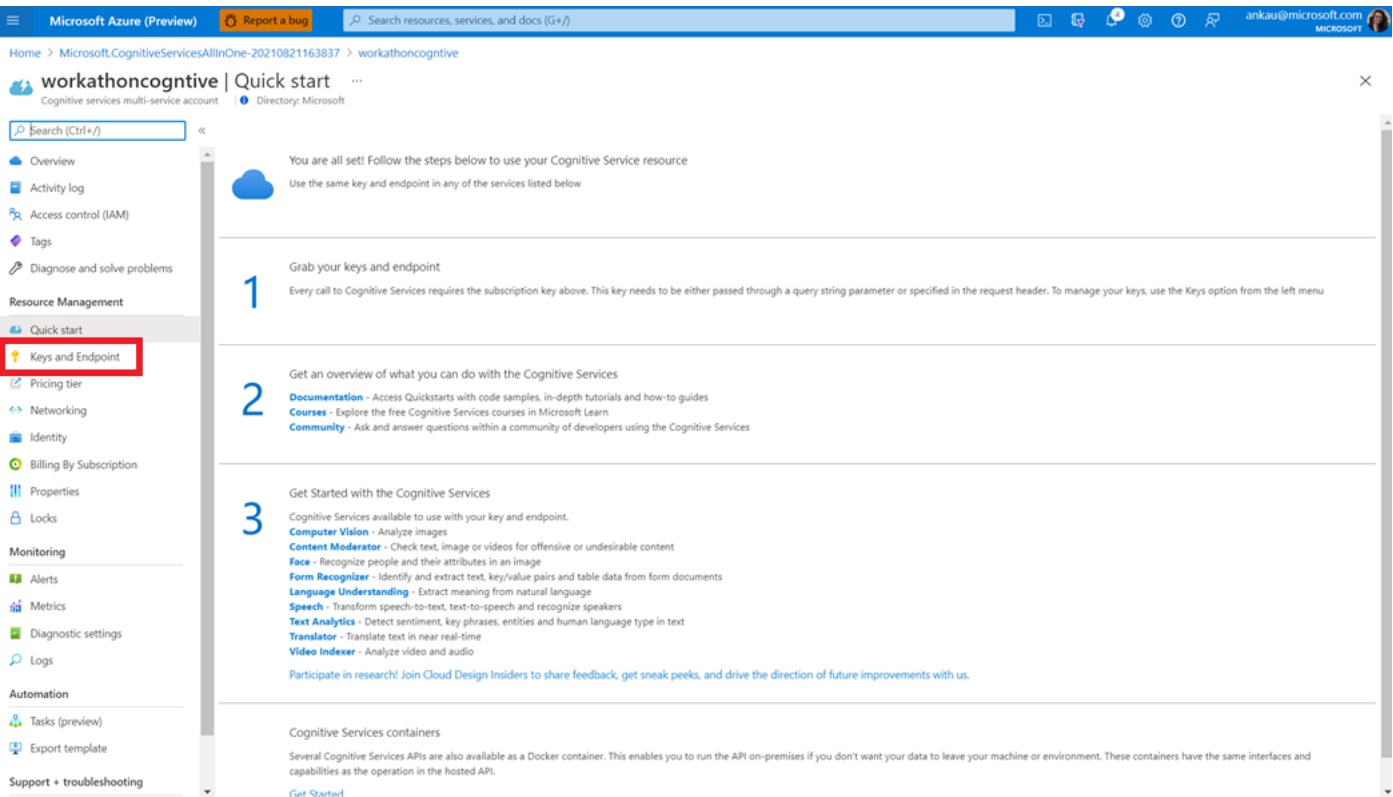
Identity typeNone

Create< PreviousNextDownload a template for automation

Verify the details and click Create.



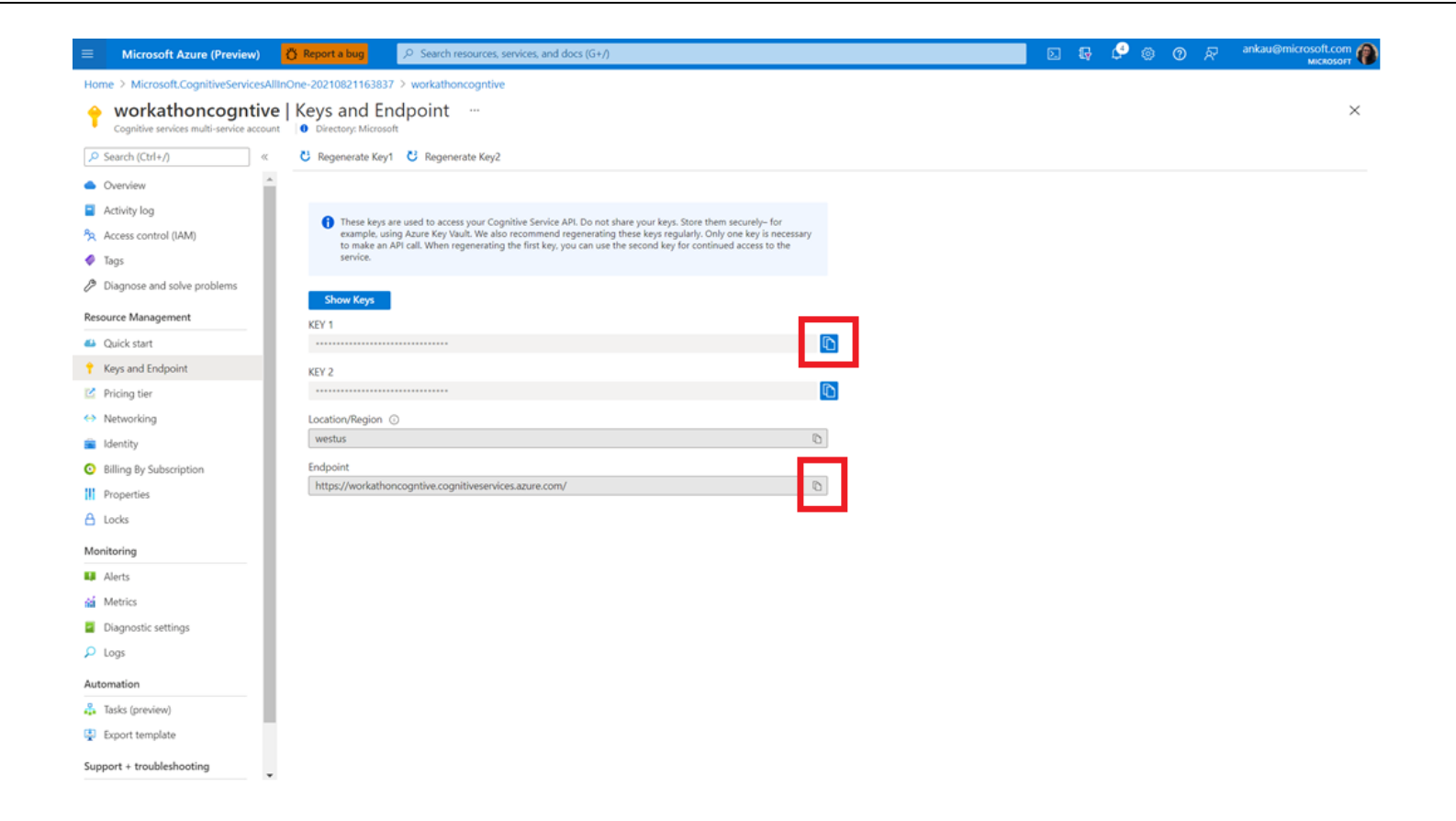
After the resource has been deployed, click Go to Resource.



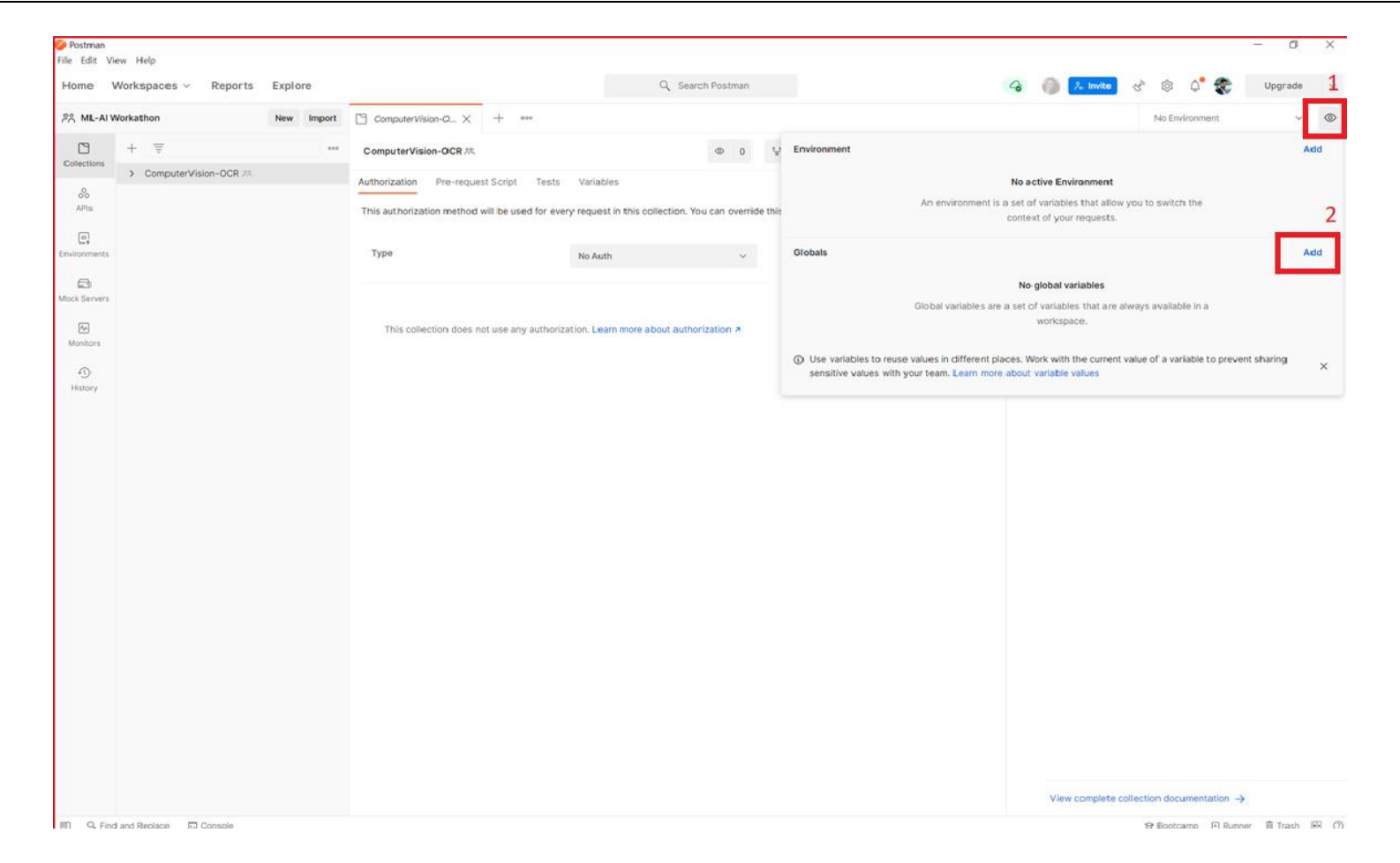
Copy Keys & Endpoints

On the Quick start page, you can find details about different cognitive services and can click the hyperlinks to learn more.

Click Key and Endpoints.



Copy the Key and Endpoint. Paste these in a notepad. You will leverage these at a later step, while setting up the global variables in Postman.

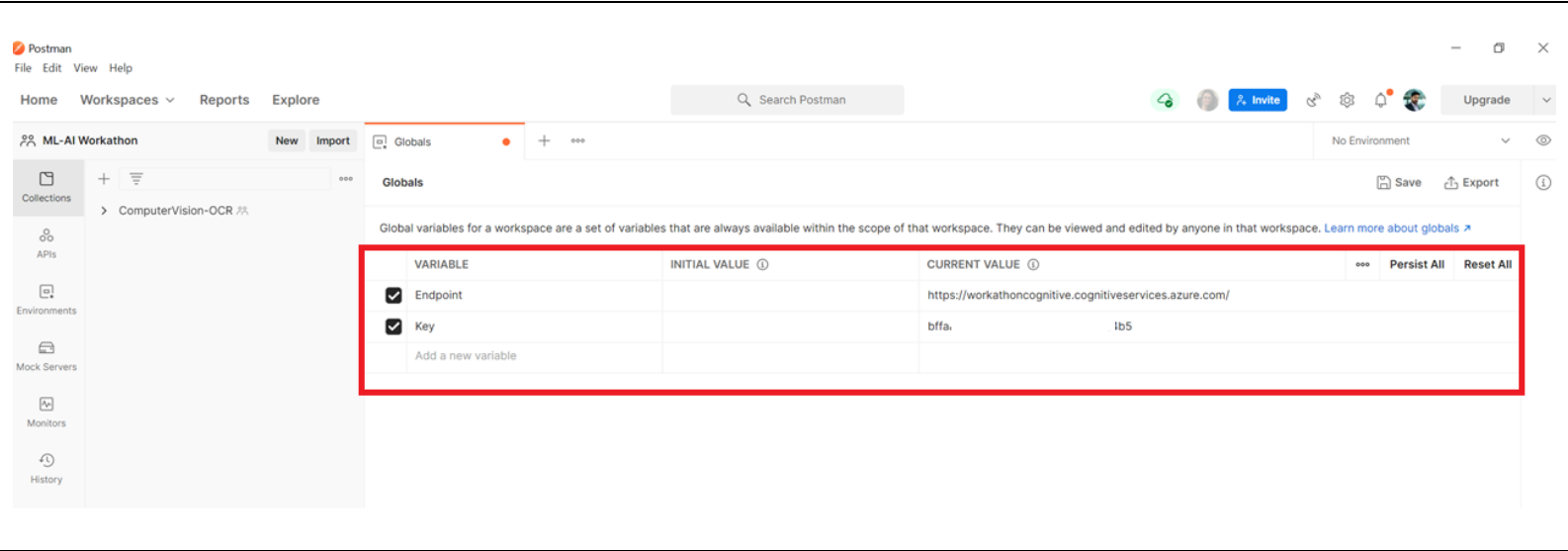


We have now switched the interface to Postman. If you haven't downloaded the Postman client, you can use web version.

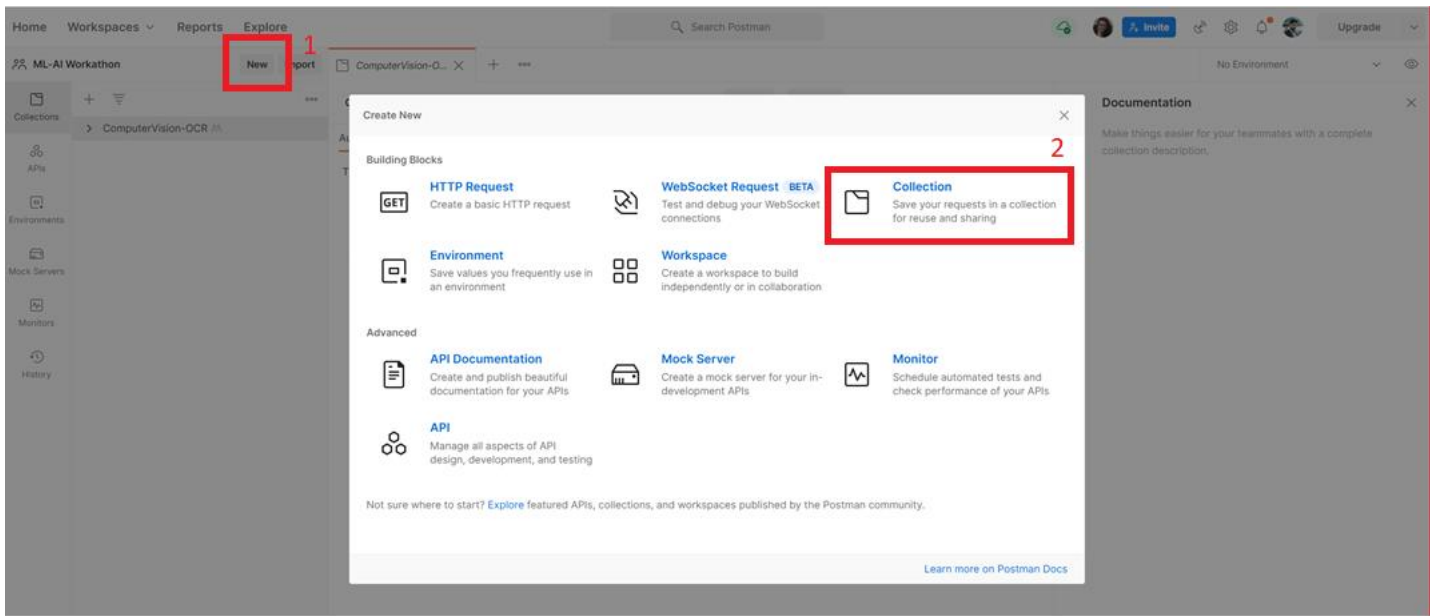
Configure global variables in Postman

Significance :
The global variables will be created once and leveraged time & again, in each API request that we make through Postman.
This way, you will not have to hard code the Endpoint & Key for every request you make, thereby, making it more secure. This will also save time and effort.

- Follow step 1 & 2 to add global variables for :
1. Endpoint – Paste the endpoint of the Multipurpose service account copied earlier
 2. Key - Paste the Key of the Multipurpose service account copied earlier



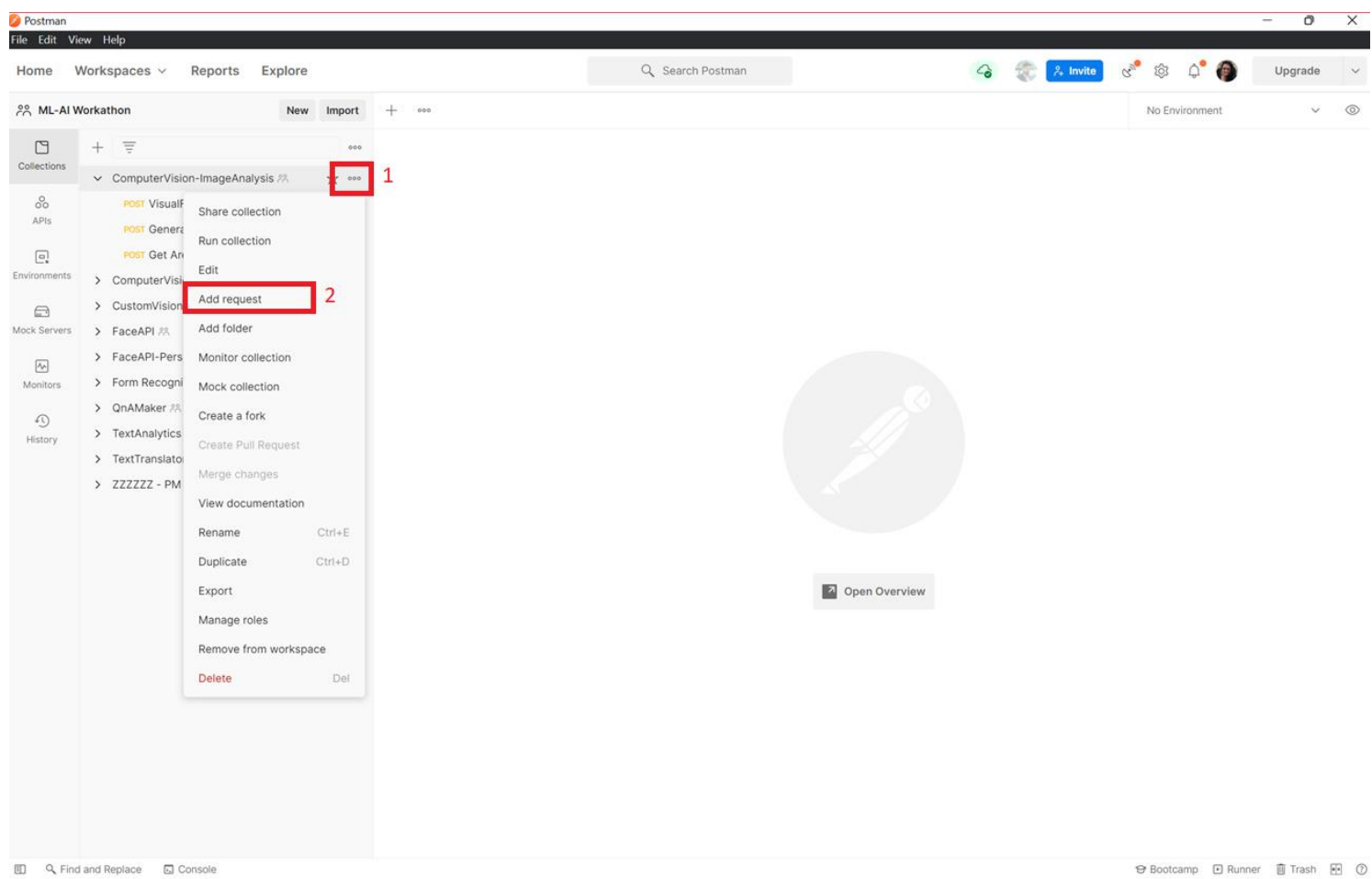
Once you have added the global variables, they will appear this way in your global variables section.



Create new collection in Postman

Open Postman > select New.
On the pop up select Collection.
Name the collection ComputerVision-ImageAnalysis.

Collection is like a folder for managing the API call requests.



Once you have created the collection, follow steps 1 & 2, to create a new request.

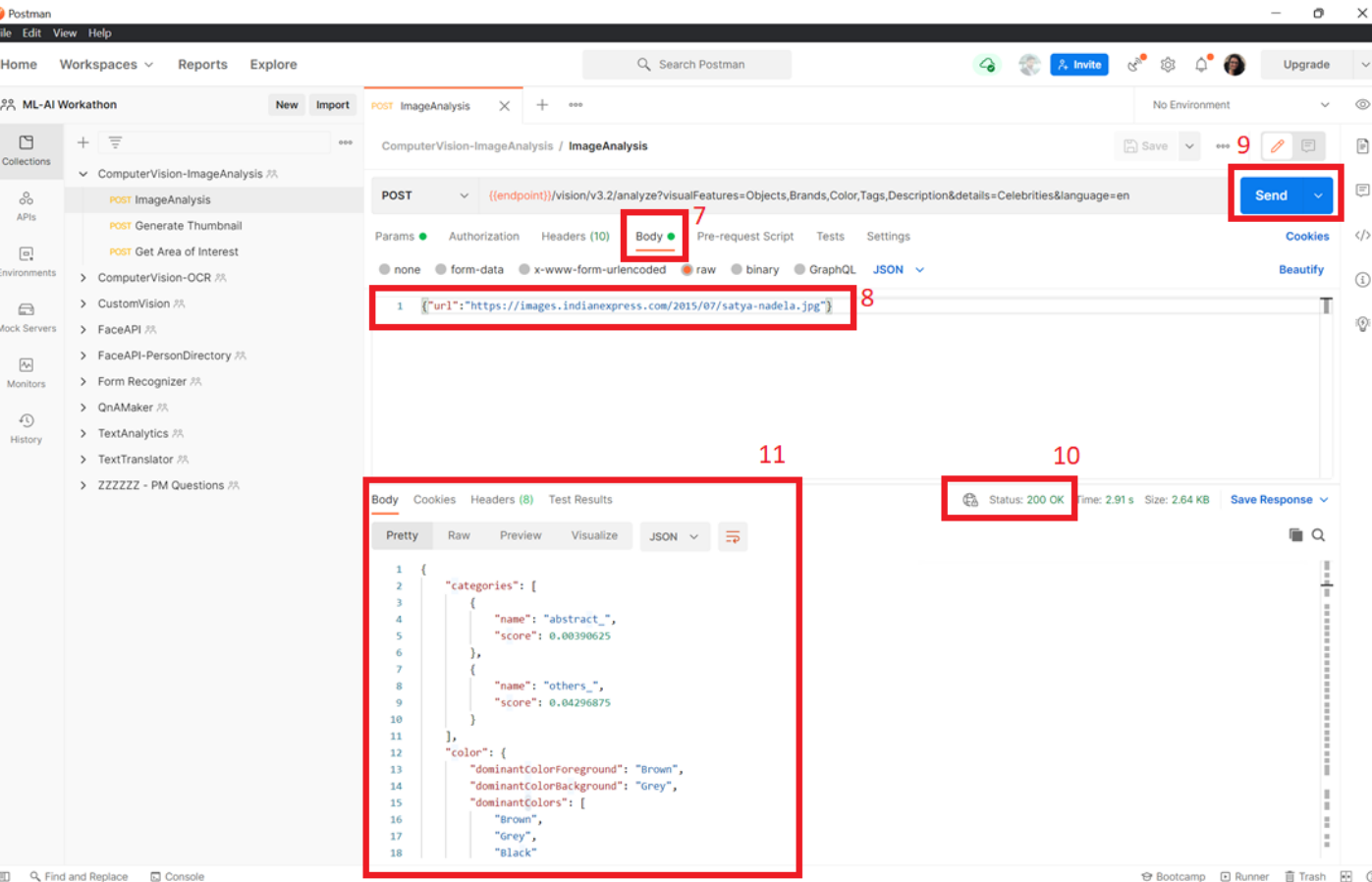
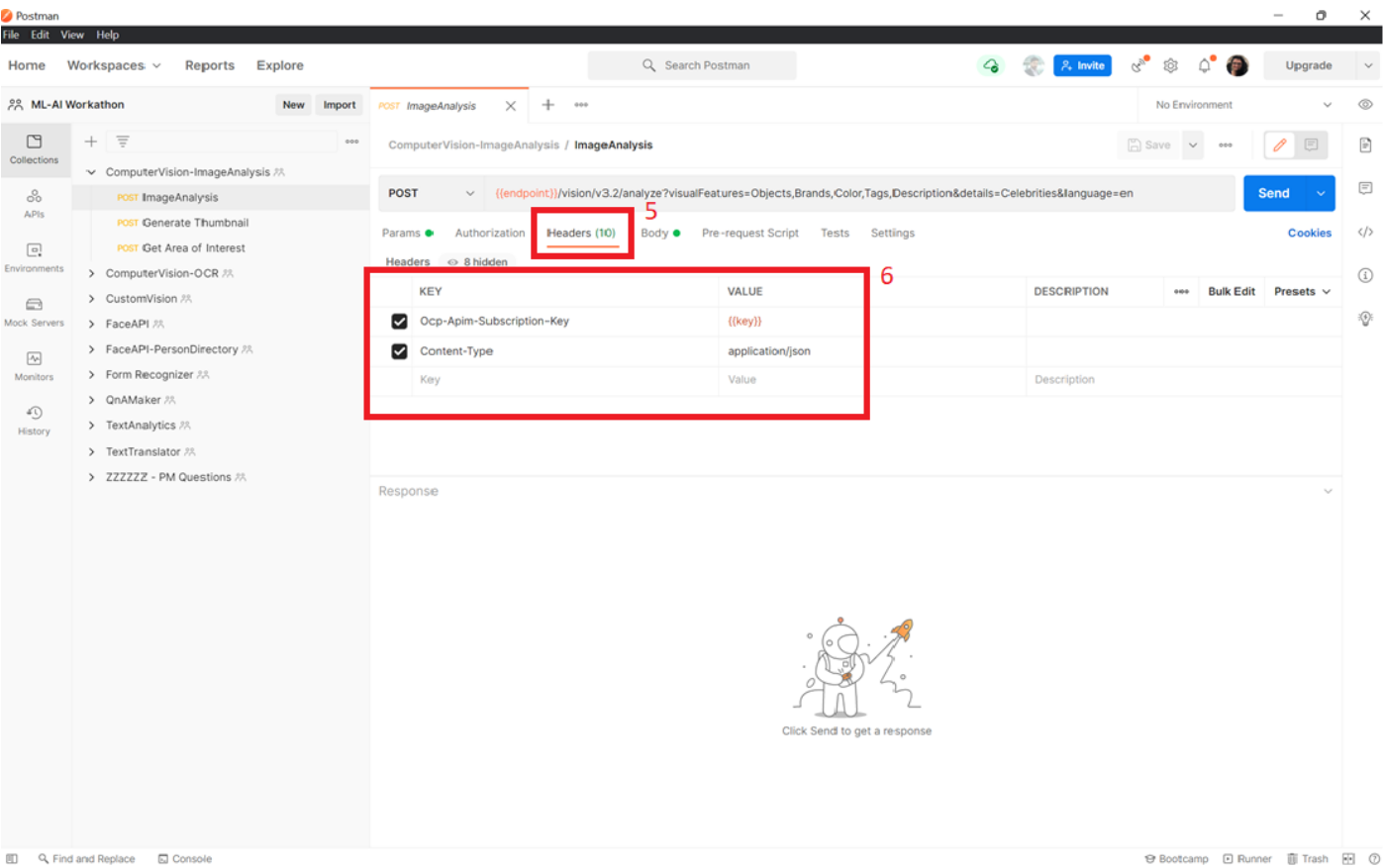
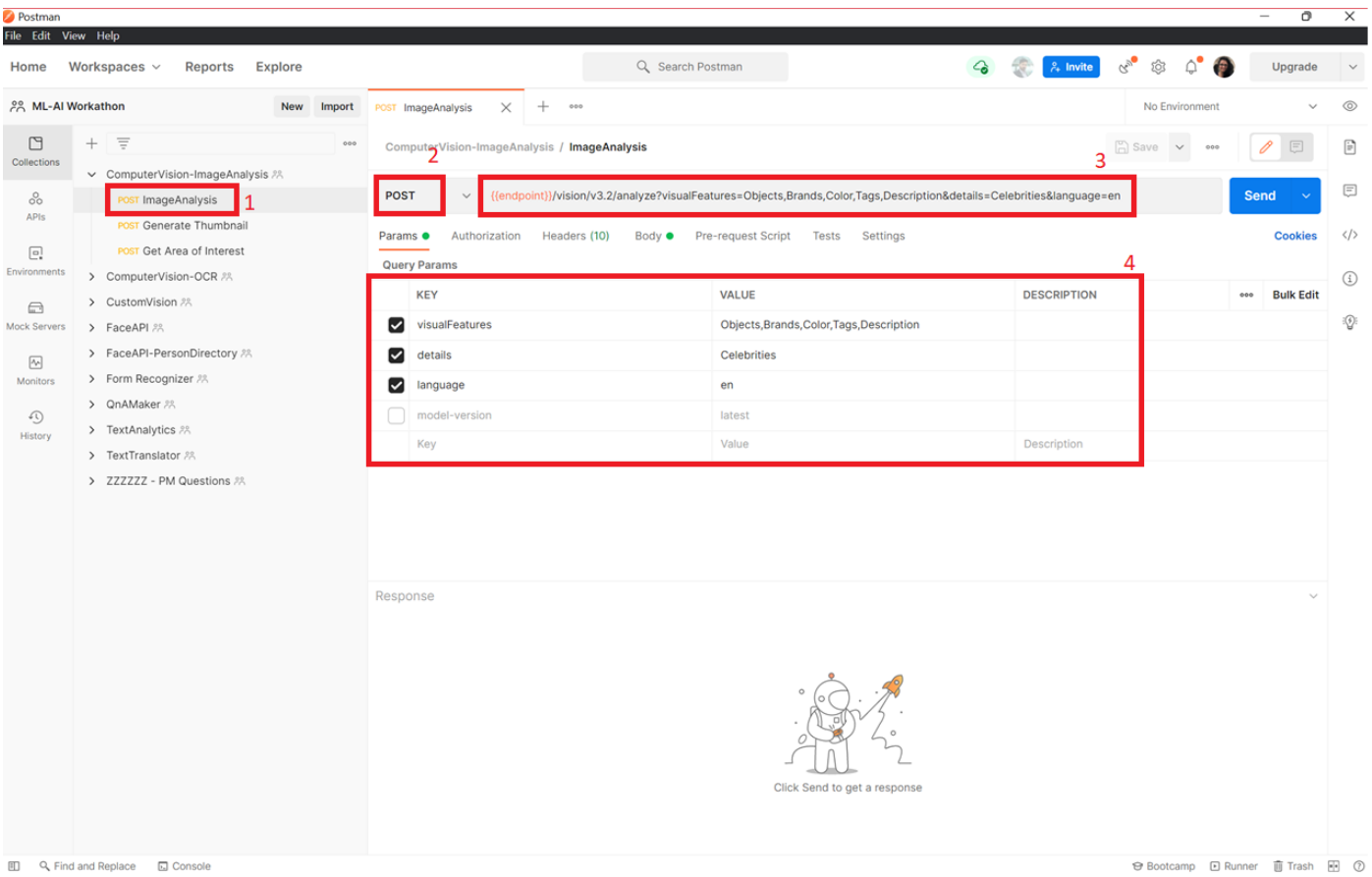


Image Analysis

This request upon successful execution will return the details that you requested through visualFeatures & details parameters, in the language mentioned.

URL : {{endpoint}}/vision/v3.2/analyze

Headers :
Ocp-Apim-Subscription-Key : {{key}}
Content-Type : application/json

Body :
{"url":"https://images.indianexpress.com/2015/07/satya-nadela.jpg"}

You should also try exploring with different URLs & local images.

Params :
visualFeatures : Objects,Brands,Color,Tags,Description
details : Celebrities
language : en

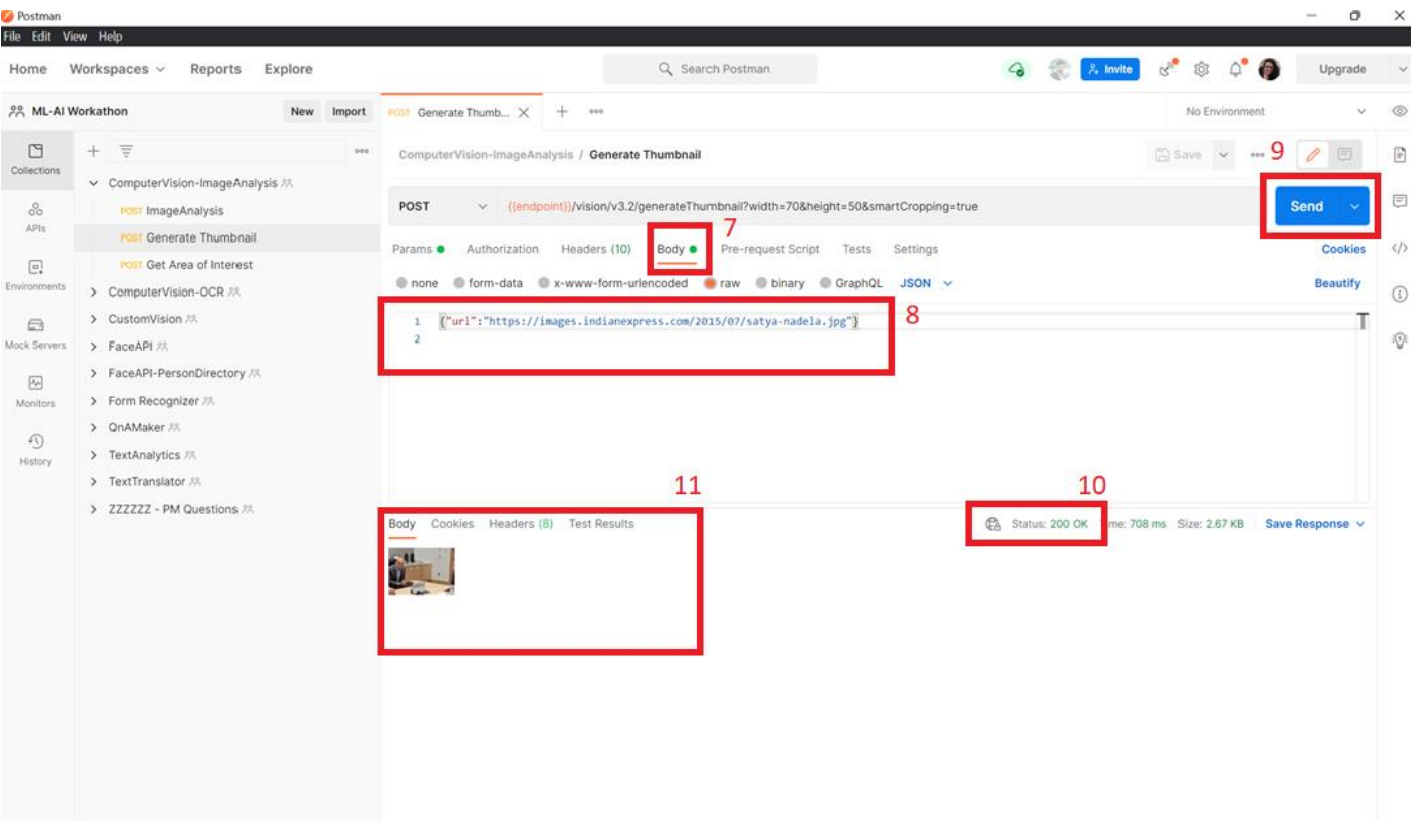
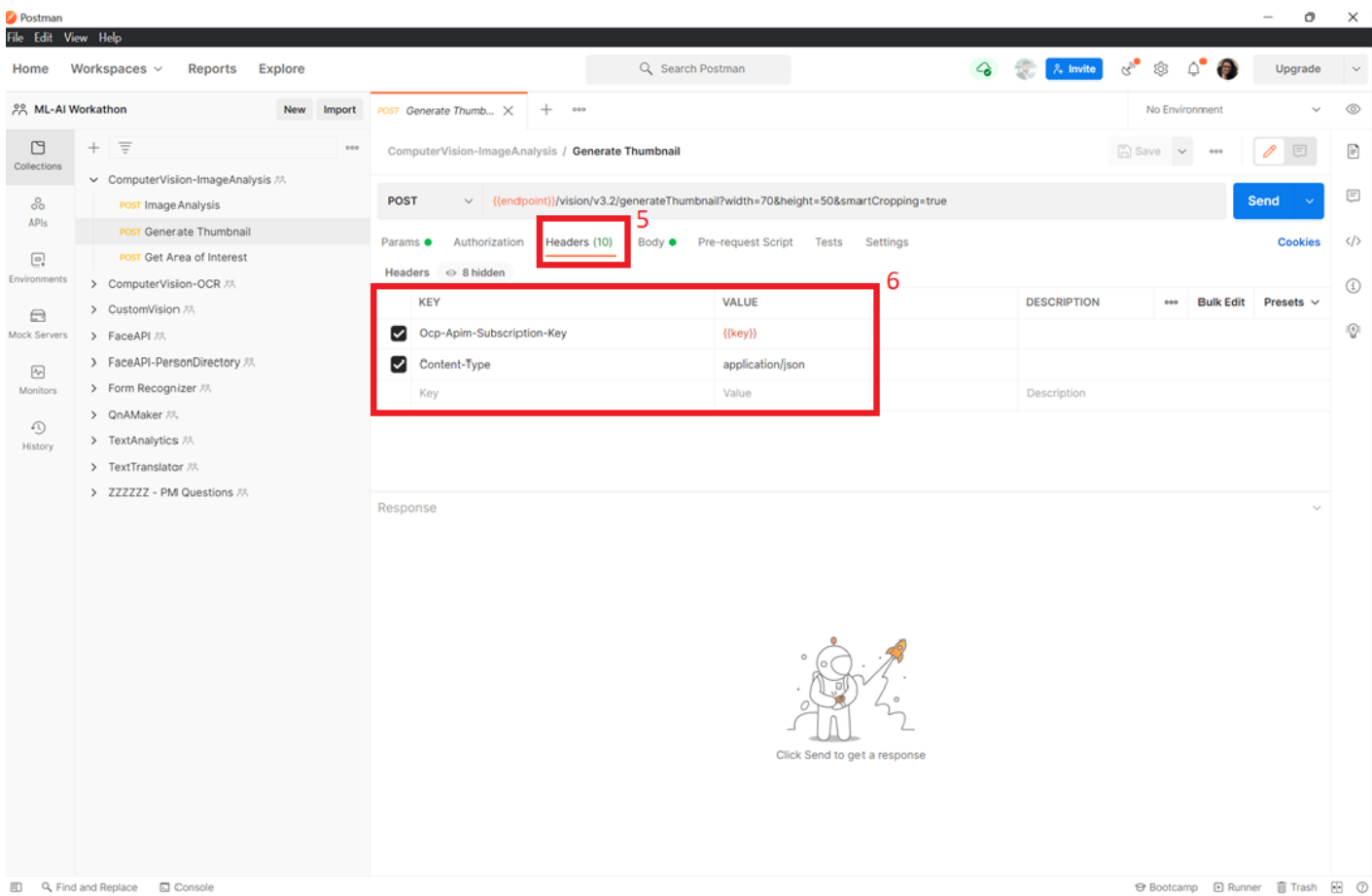
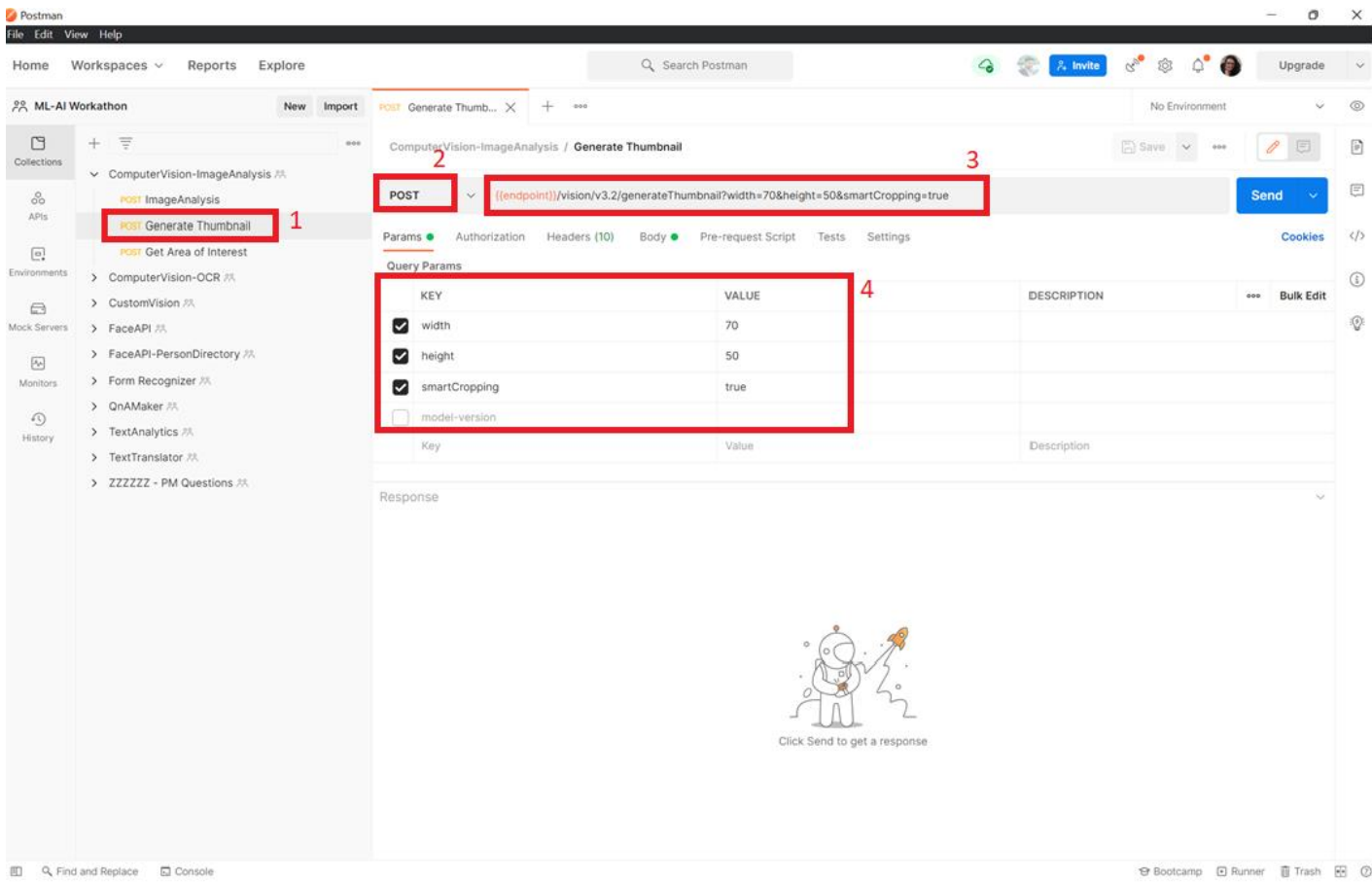
Once you have added the Params, they will be appended to the URL and it will look similar to the one in the screenshot.

Try exploring different values for the parameters :

- faces, Adult, ImageType in visualFeatures
- Landmarks in details

Significance of input & output

- {{endpoint}} , {{key}} : Values being picked from global variables
- Ocp-Apim-Subscription-Key : This is the Azure Cognitive service key, that will authenticate the request.
Content-Type : This refers to the input type that you provide in the body, for eg application/json allows you to enter body text in JSON format. Change the content-type on the basis of input you provide.
- Once you have added the Params, they will be appended to the URL and it will look similar to the one in the screenshot. Try exploring different values for the parameters:
 - Faces, Adult, ImageType in visualFeatures
 - Landmarks in details
- After you execute the call, observe the status returned, as shown in step 10. This should reflect 200 OK.
- Observe the JSON output and observe the keys and their corresponding values that you get. Observe how close these are to how a human would have interpreted the image.



Generate Thumbnail

This request upon successful execution will return an image binary, with user specified height & width. The model smartly identifies the important features in the image and crops it accordingly. You can also choose a different aspect ratio than the input image.

URL : `{{endpoint}}/vision/v3.2/generateThumbnail`

Headers :
Ocp-Apim-Subscription-Key : `{{key}}`
Content-Type : application/json

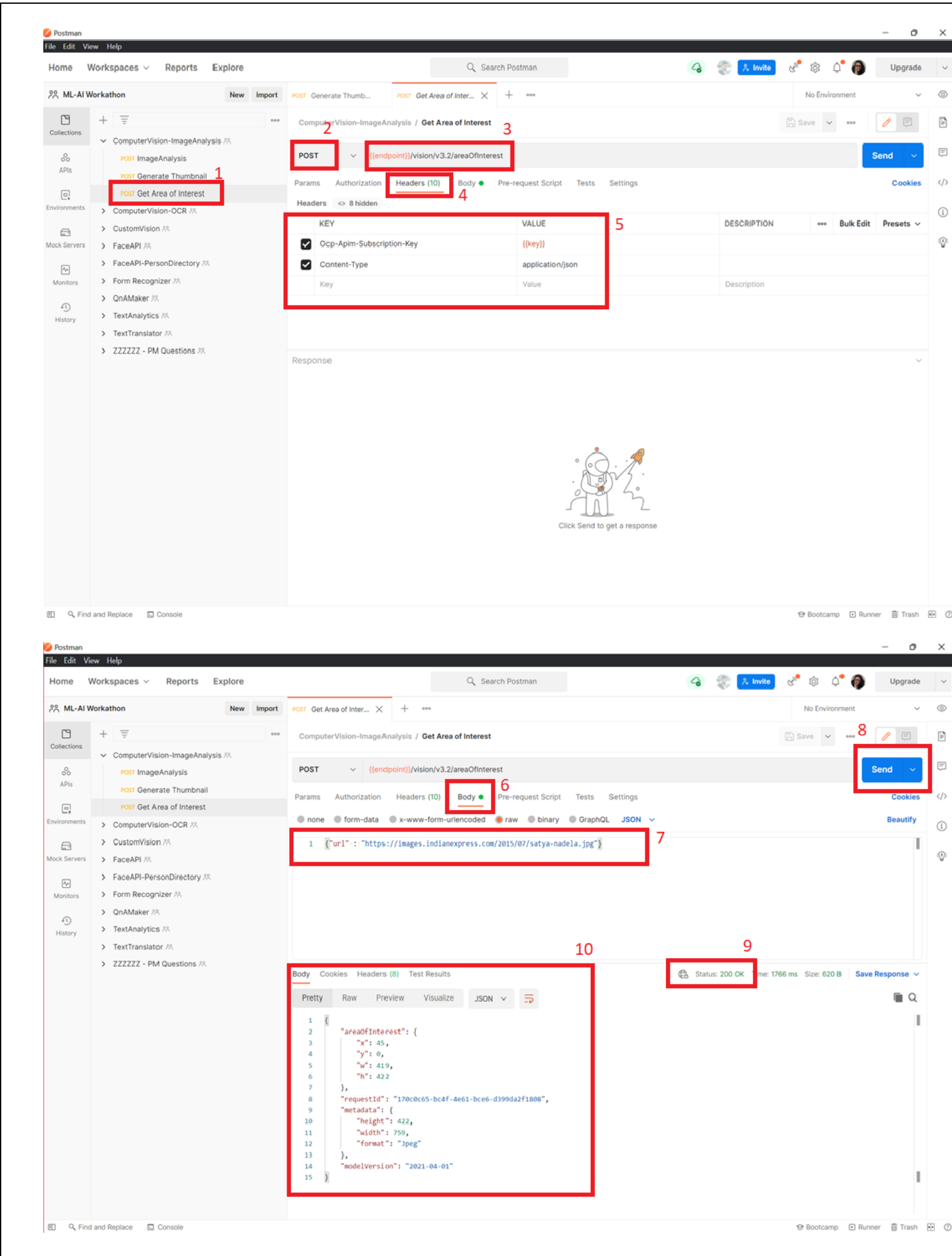
Body :
`{"url":"https://images.indianexpress.com/2015/07/satya-nadela.jpg"}`

You should also try exploring with different URLs & local images.

Params :
Width : 70
Height : 50
smartCropping : true

Significance of input & output

1. `{{endpoint}}` , `{{key}}` : Values being picked from global variables
2. Ocp-Apim-Subscription-Key : This is the Azure Cognitive service key, that will authenticate the request.
Content-Type : This refers to the input type that you provide in the body, for eg application/json allows you to enter body text in JSON format. Change the content-type on the basis of input you provide.
3. Once you have added the Params, they will be appended to the URL and it will look similar to the one in the screenshot. Try exploring different values for the parameters.
4. Setting smartCropping to true allows you to specify a different aspect ratio than the input image.
5. After you execute the call, observe the status returned, as shown in step 10. This should reflect 200 OK.
6. Observe the JSON output and observe the keys and their corresponding values that you get. Observe how close these are to how a human would have interpreted the image.



Get Area of Interest

This request upon successful execution will return the coordinates for the most important area of the image. You can leverage that for further image processing, cropping etc.

URL : `{{endpoint}}/vision/v3.2/areaOfInterest`

Headers :

Ocp-Apim-Subscription-Key : `{{key}}`

Content-Type : `application/json`

Body :

`{"url":"https://images.indianexpress.com/2015/07/satya-nadela.jpg"}`

You should also try exploring with different URLs & local images.

Significance of input & output

- `{{endpoint}}` , `{{key}}` : Values being picked from global variables
- Ocp-Apim-Subscription-Key : This is the Azure Cognitive service key, that will authenticate the request.
Content-Type : This refers to the input type that you provide in the body, for eg application/json allows you to enter body text in JSON format. Change the content-type on the basis of input you provide.
- After you execute the call, observe the status returned, as shown in step 9. This should reflect 200 OK.
- The `areaOfInterest` value return the x & y coordinate of the top left point of the bounding box. W & h are the width & height respectively. Hence, you can calculate all the 4 corners of the bounding box.

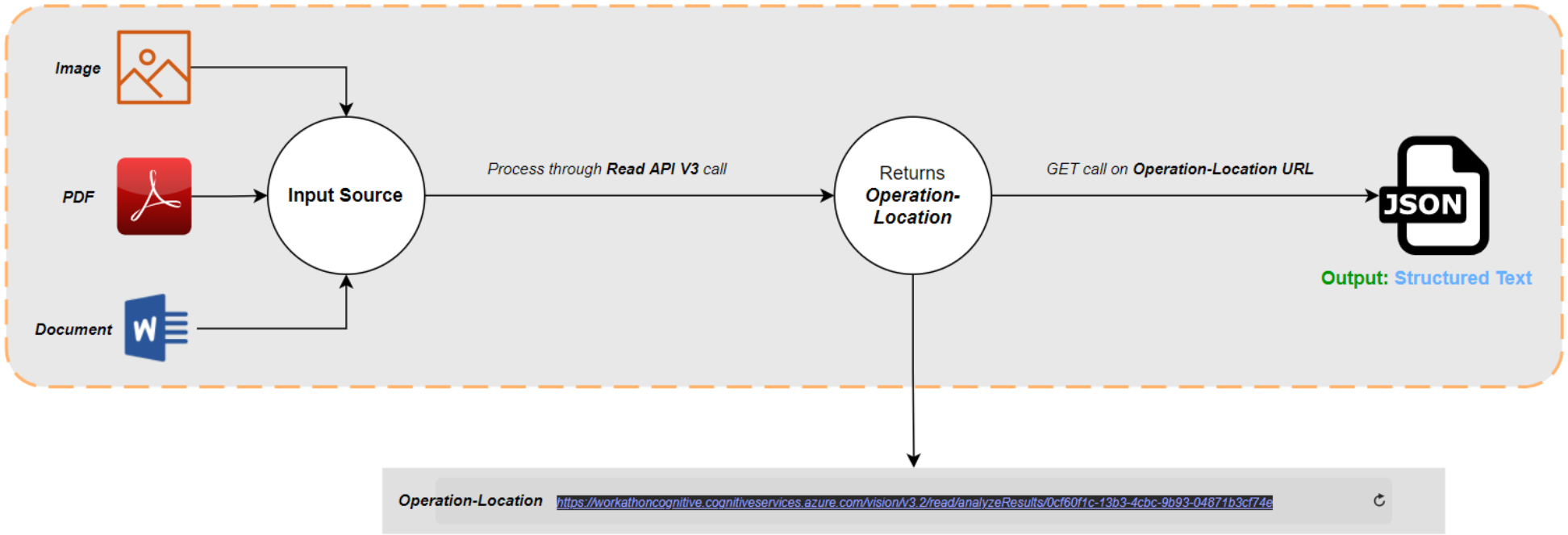
Optical Character Recognition

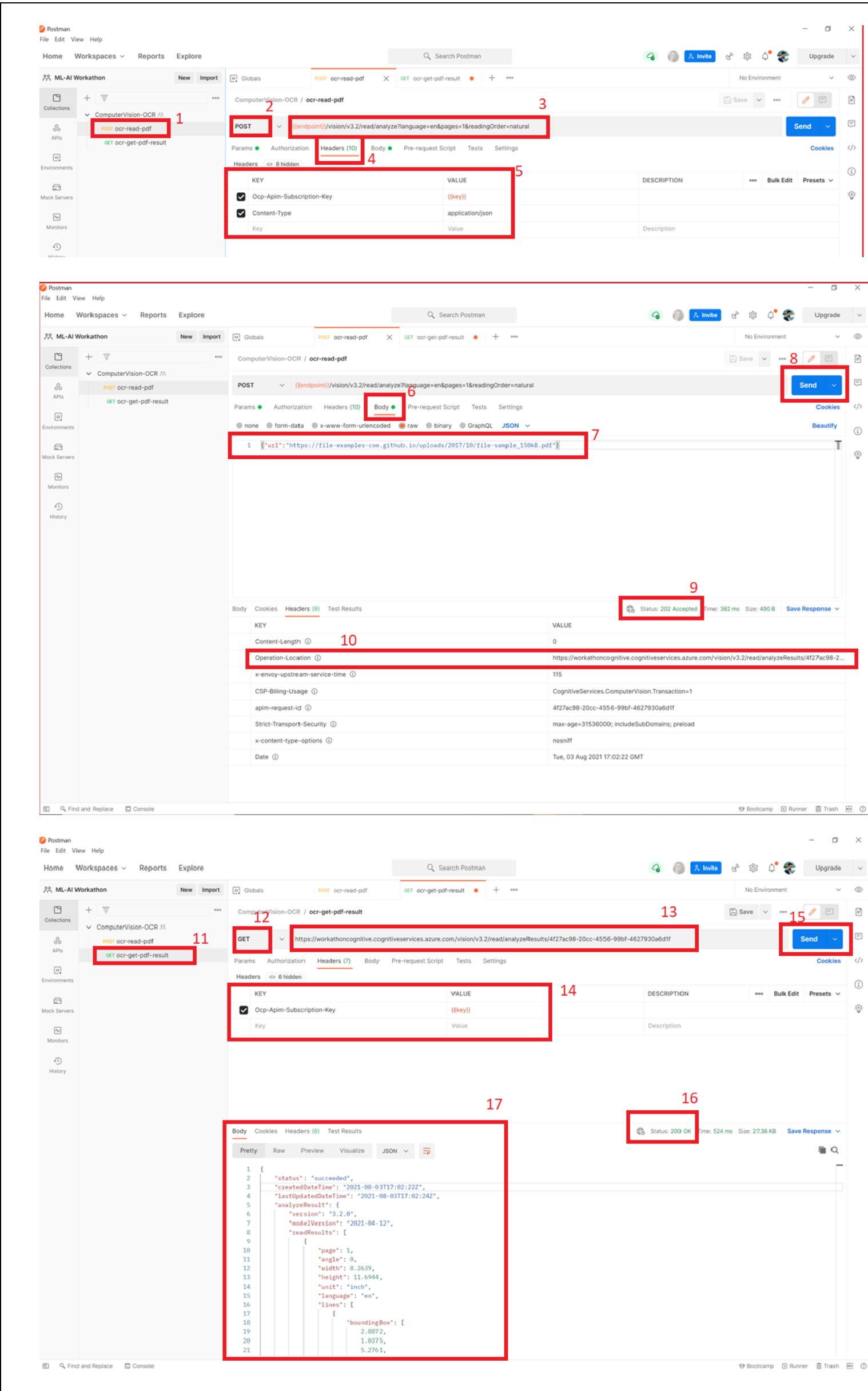
Optical Character Recognition refers to extracting text from documents and images, both printed and handwritten. Azure offers 2 major APIs for Optical Character Recognition :

- 1. OCR API : This is used for images that contain relatively lesser text
- 2. Read API : This is used for images & documents that are text heavy or you have multi-page documents

In the scenario where you are picking up a mix of images or documents with varied quantities of text, we suggest you use Read API, for consistency, simplicity & improved accuracy. In this workshop, we will be exploring OCR features using the Read API, for different input types such as documents, images, handwritten text.

OPTICAL CHARACTER RECOGNITION WORKFLOW





OCR From Document

This request upon successful execution will return the contents of the file in a structured & hierarchical format such as pages, lines, words. This also returns the exact coordinates of where the line or word is found & the confidence score.

This is an async process, hence requires 2 API calls. The first call, as shown in step 3, generates the Operation-Location URL, shown in step 10. The second call is a GET operation to this URL.

URL 1 : `{{endpoint}}/vision/v3.2/read/analyze`

Headers :
Ocp-Apim-Subscription-Key : `{{key}}`
Content-Type : application/json

Body :
`{"url": "https://file-examples-com.github.io/uploads/2017/10/file-sample_150kB.pdf"}`

This is a multi-page document. You should also try exploring with different document URLs & local documents.

Params :
Language : en
Pages : 1
readingOrder : natural

URL2 : Operation-Location fetched from step 10. Make sure to change the call to GET.

Headers :
Ocp-Apim-Subscription-Key : `{{key}}`
Content-Type : application/json

Significance of input & output

- `{{endpoint}}` , `{{key}}` : Values being picked from global variables
- Ocp-Apim-Subscription-Key : This is the Azure Cognitive service key, that will authenticate the request.
Content-Type : This refers to the input type that you provide in the body, for eg application/json allows you to enter body text in JSON format. Change the content-type on the basis of input you provide.
- Once you have added the Params, they will be appended to the URL and it will look similar to the one in the screenshot. Try exploring different values for the parameters.
- Params :
 - Pages : If you have multiple pages but want to process selective pages, you can specify them here. Such as 1,2,5 or 1-4, 9 in a document longer than 9 pages, in this case.
 - readingOrder : This has 2 options – Basic & Natural. In case of selecting Natural & there are multiple columns or a bit unstructured text in the document, the model will itself follow a human like reading order. This is where you can observe the cognitive capabilities visually.
Try using the basic parameter to observe the difference in outcome.
- After the 2nd API call, you should get status 200 OK as in step 16.
- Observe the output received in step 17 and notice the hierarchy and objects returned. You can also download the JSON file for better readability, by using Save Response option on the right corner.

Lorem Ipsum

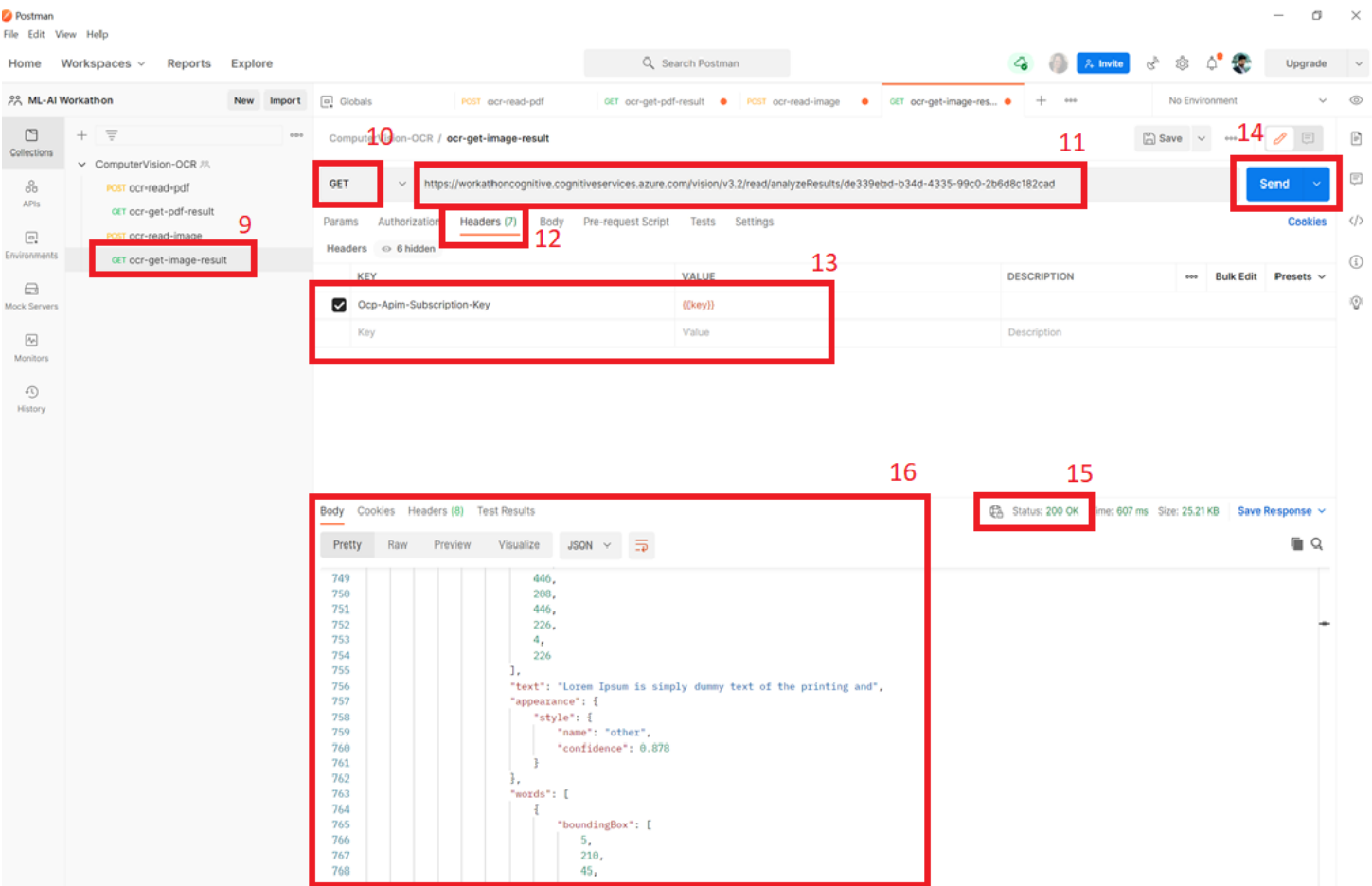
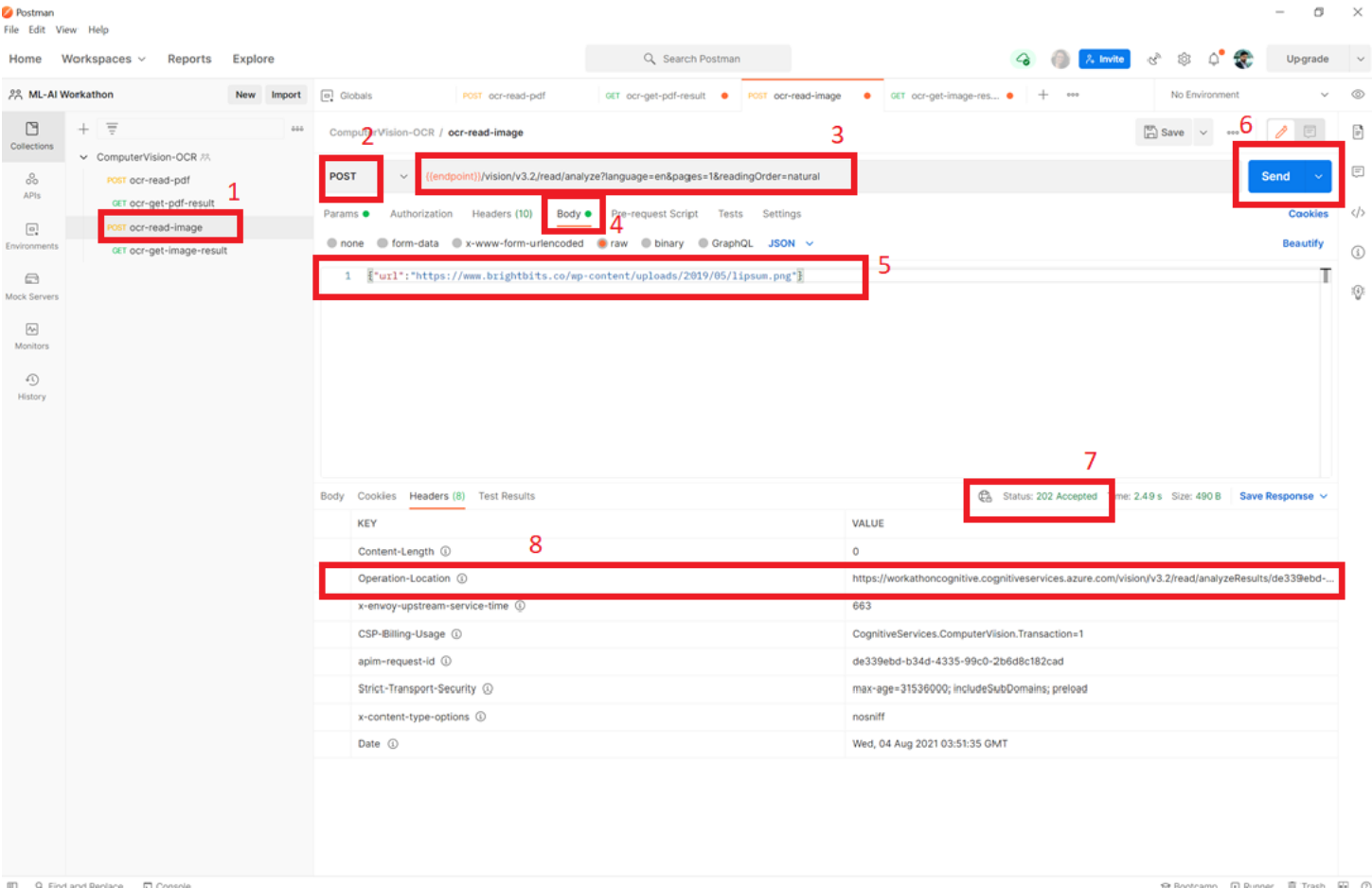
"Neque porro quisquam est qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit..."
 "There is no one who loves pain itself, who seeks after it and wants to have it, simply because it is pain..."

What is Lorem Ipsum?

Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged. It was popularised in the 1960s with the release of Letraset sheets containing Lorem Ipsum passages, and more recently with desktop publishing software like Aldus PageMaker including versions of Lorem Ipsum.

Why do we use it?

It is a long established fact that a reader will be distracted by the readable content of a page when looking at its layout. The point of using Lorem Ipsum is that it has a more-or-less normal distribution of letters, as opposed to using 'Content here, content here', making it look like readable English. Many desktop publishing packages and web page editors now use Lorem Ipsum as their default model text, and a search for 'lorem ipsum' will uncover many web sites still in their infancy. Various versions have evolved over the years, sometimes by accident, sometimes on purpose (injected humour and the like).



OCR From Image

This request upon successful execution will return the contents of the image in a structured & hierarchical format such as lines, words. This also returns the exact coordinates of where the line or word is found & the confidence score.

This is an async process, hence requires 2 API calls. The first call, as shown in step 3, generates the Operation-Location URL, shown in step 8. The second call is a GET operation to this URL.

URL 1 : {{endpoint}}/vision/v3.2/read/analyze

Headers :
 Ocp-Apim-Subscription-Key : {{key}}
 Content-Type : application/json

Body :
 {"url":"https://www.brightbits.co/wp-content/uploads/2019/05/lipsum.png"}

You should also try exploring with different image URLs & local images.

Params :
 Language : en
 readingOrder : natural

URL2 : Operation-Location fetched from step 8. Make sure to change the call to GET.

Headers :
 Ocp-Apim-Subscription-Key : {{key}}
 Content-Type : application/json

Significance of input & output

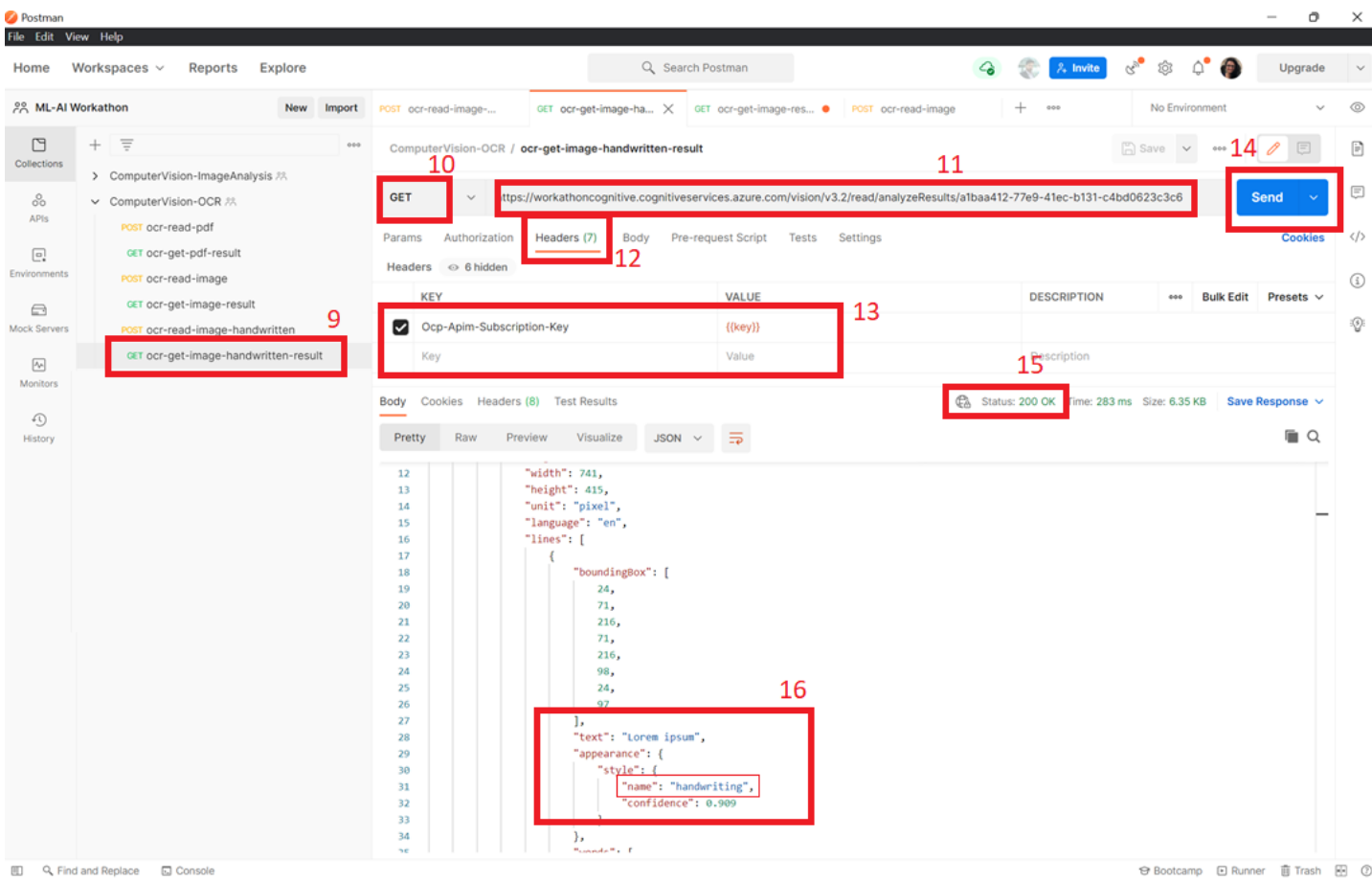
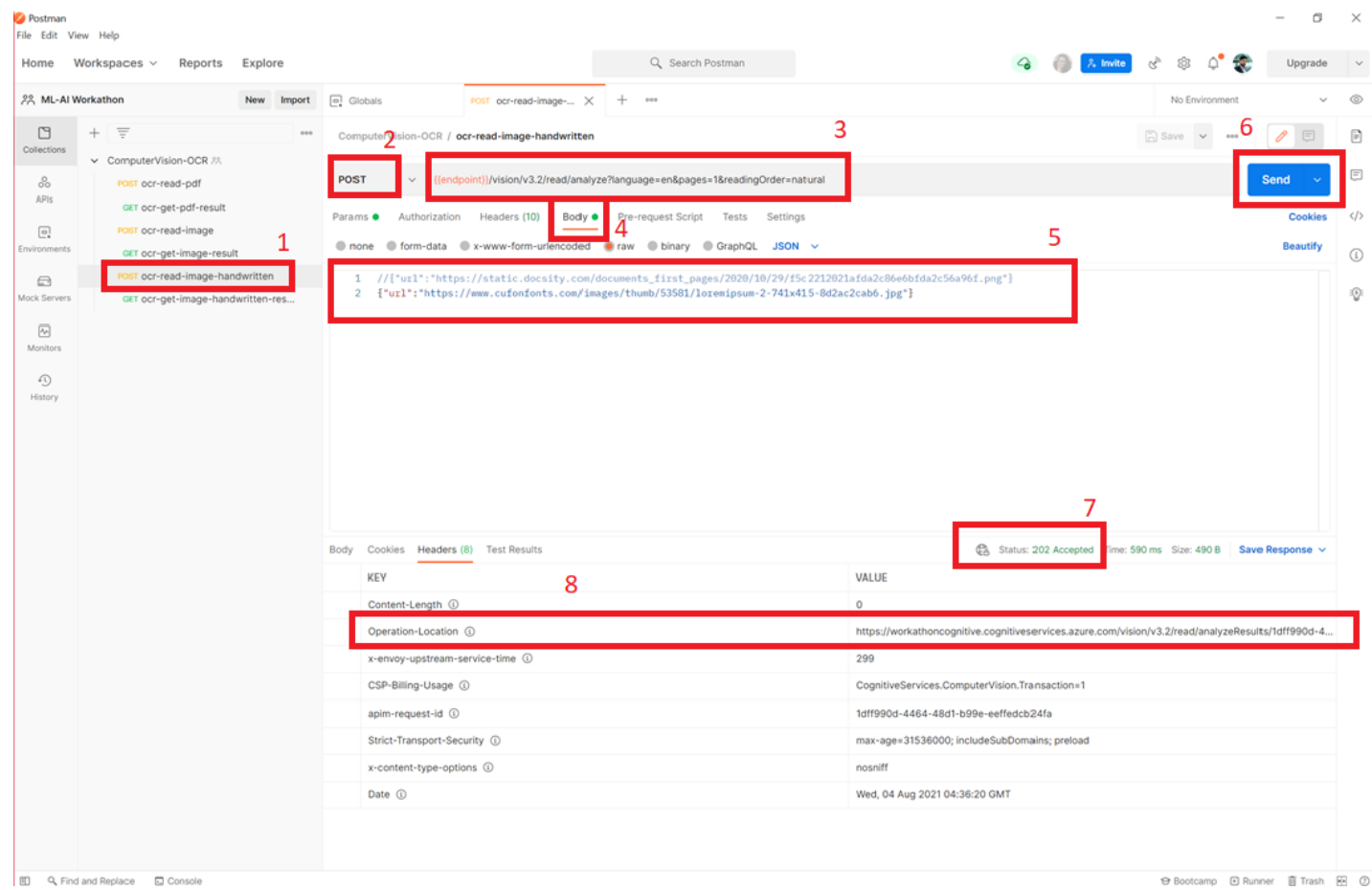
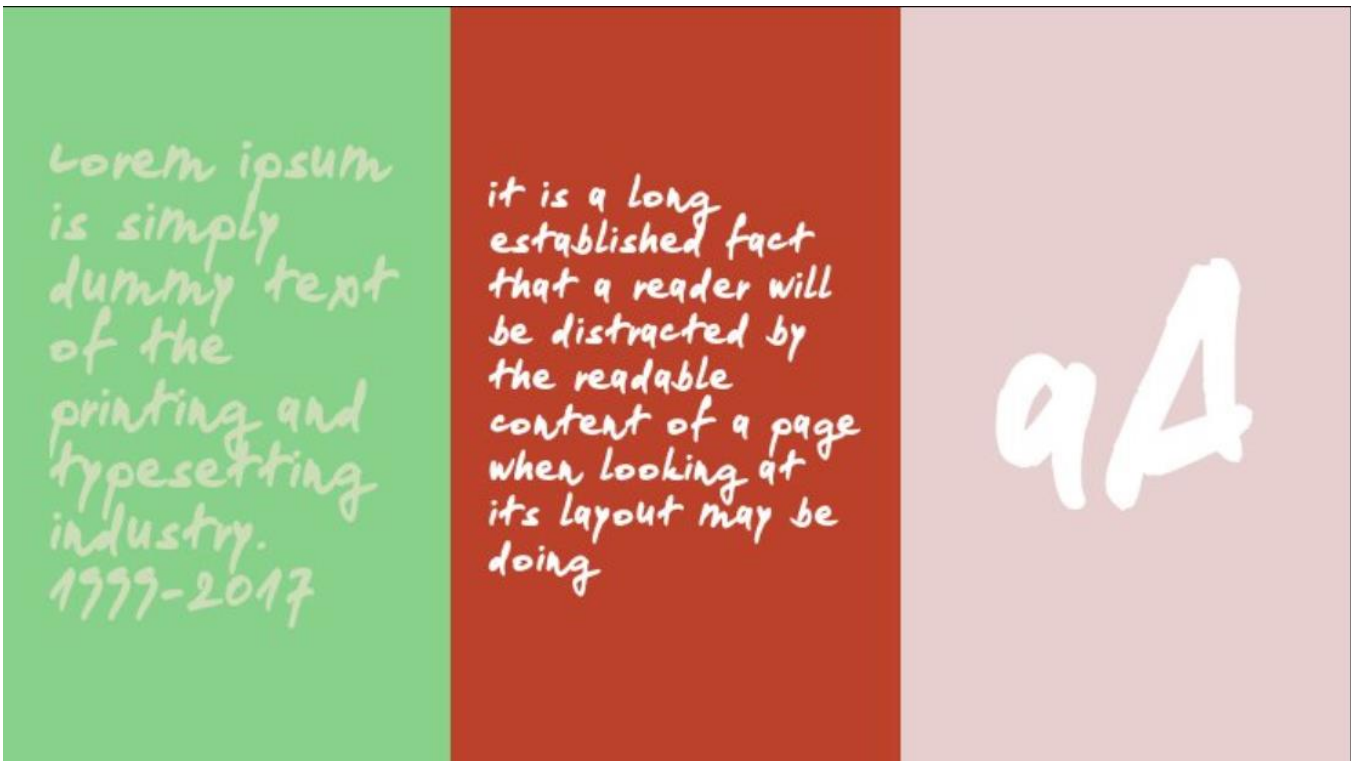
1. {{endpoint}}, {{key}} : Values being picked from global variables
2. Ocp-Apim-Subscription-Key : This is the Azure Cognitive service key, that will authenticate the request.

Content-Type : This refers to the input type that you provide in the body, for eg application/json allows you to enter body text in JSON format. Change the content-type on the basis of input you provide.

3. Once you have added the Params, they will be appended to the URL and it will look similar to the one in the screenshot. Try exploring different values for the parameters.
4. Params :
 - a. readingOrder : This has 2 options – Basic & Natural. In case of selecting Natural & there are multiple columns or a bit unstructured text in the document, the model will itself follow a human like reading order. This is where you can observe the cognitive capabilities visually. Try using the basic parameter to observe the difference in outcome.

5. After the 2nd API call, you should get status 200 OK as in step 15.

6. Observe the output received in step 16 and notice the hierarchy and objects returned. You can also download the JSON file for better readability, by using Save Response option on the right corner.



OCR From Handwritten Content in Image

This request upon successful execution will return the contents of the image or document in a structured & hierarchical format such as lines, words. This also returns a parameter saying that the text is handwritten, the exact coordinates of where the line or word is found & the confidence score.

This is an async process, hence requires 2 API calls. The first call, as shown in step 3, generates the Operation-Location URL, shown in step 8. The second call is a GET operation to this URL.

URL 1 : `{{endpoint}}/vision/v3.2/read/analyze`

Headers :
Ocp-Apim-Subscription-Key : `{{key}}`
Content-Type : application/json

Body :
`{"url":"https://www.cufonfonts.com/images/thumb/53581/loremipsum-2-741x415-8d2ac2cab6.jpg"}`

You should also try exploring with different documents & images containing handwritten text.

Params :
Language : en
readingOrder : natural

URL2 : Operation-Location fetched from step 8. Make sure to change the call to GET.

Headers :
Ocp-Apim-Subscription-Key : `{{key}}`
Content-Type : application/json

Significance of input & output

- `{{endpoint}}` , `{{key}}` : Values being picked from global variables
- Ocp-Apim-Subscription-Key : This is the Azure Cognitive service key, that will authenticate the request.

Content-Type : This refers to the input type that you provide in the body, for eg application/json allows you to enter body text in JSON format. Change the content-type on the basis of input you provide.

- Once you have added the Params, they will be appended to the URL and it will look similar to the one in the screenshot. Try exploring different values for the parameters.
- Params :
 - readingOrder : This has 2 options – Basic & Natural. In case of selecting Natural & there are multiple columns or a bit unstructured text in the document, the model will itself follow a human like reading order. This is where you can observe the cognitive capabilities visually. Try using the basic parameter to observe the difference in outcome.

- After the 2nd API call, you should get status 200 OK as in step 15.

- Observe the output received in step 16 and notice the hierarchy and objects returned. You can also download the JSON file for better readability, by using Save Response option on the right corner.

- Notice the text > style > name hierarchy than mentions the text being handwritten.

Homework

1. For Image Analysis, try exploring different values for the parameters:

Faces, Adult, ImageType in visualFeatures

Landmarks in details

2. For OCR

Try out different images and documents with both handwritten & printed text.

Try out different parameter settings such as Pages, readingOrder etc, to observe the changes in the outcomes.

Additional recommended resources

Service Name	Category	Links
Computer Vision	Programming Language	C#, Python, Java, Javascript, Go, REST API
	Tiers	Free (not for production), Standard (Differences between tiers)
	Pricing	https://azure.microsoft.com/en-in/pricing/details/cognitive-services/computer-vision/#pricing
	Limits	Image Analysis and OCR
	Language Support	https://docs.microsoft.com/en-in/azure/cognitive-services/computer-vision/language-support
	Sample Apps	Image Processing App , Web App Javascript
	Regional Availability & Support	https://azure.microsoft.com/en-us/global-infrastructure/services/?products=cognitive-services&regions=all
	SLAs for Cognitive Services	https://azure.microsoft.com/en-in/support/legal/sla/cognitive-services/v1_1/
	Compliance & Certificates	https://azure.microsoft.com/en-us/support/legal/cognitive-services-compliance-and-privacy/
	Cognitive Services Updates	https://azure.microsoft.com/en-us/updates/?product=cognitive-services

Security best practices

1. [Azure Cognitive Services security](#)

2. [Networking](#)

3. [Authentication](#)

4. [Key Management](#)

5. [Data loss prevention](#)

6. [Azure security baseline](#)

7. [Regulatory Compliance controls](#)

Responsible AI being a part of best practices, we encourage you to read [this](#).

[Computer Vision Documentation](#)

[API & Error references](#)