

# Deploy your Model to a WebApp on Azure

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

## Agenda

- **Analyze the deployment process of an LLM model using a banking dataset to understand customer intents.**
- **Step-by-step guide on deploying the LLM model as a web app on Azure Chat Playground.**
- **Review and test the responses generated by the deployed web app.**

## Steps Performed

- **Provision the Azure OpenAI resource**
- **Deploy a model**
- **Use the chat playground**
- **Create Azure AI search service**
- **Add your data**
- **Case Study**
- **Architecture of Web App Deployment**
- **Deploy your model to a Web App**
- **Test the Web App**
- **Summary & Clean Up**

# Provision an Azure OpenAI Resource

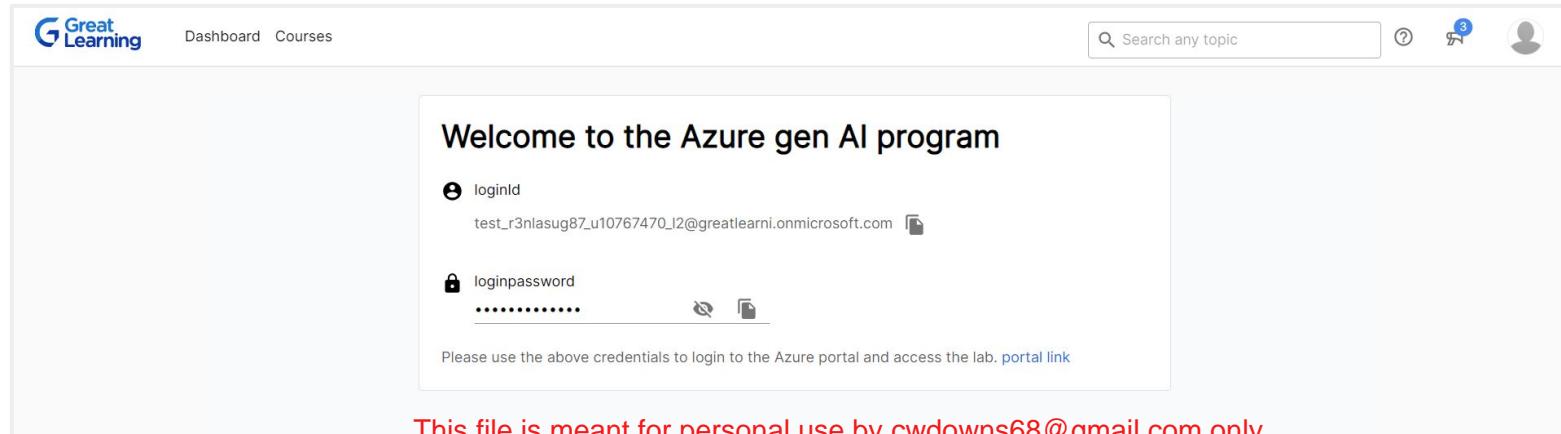
This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

Proprietary Content © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Ground Rules : Important

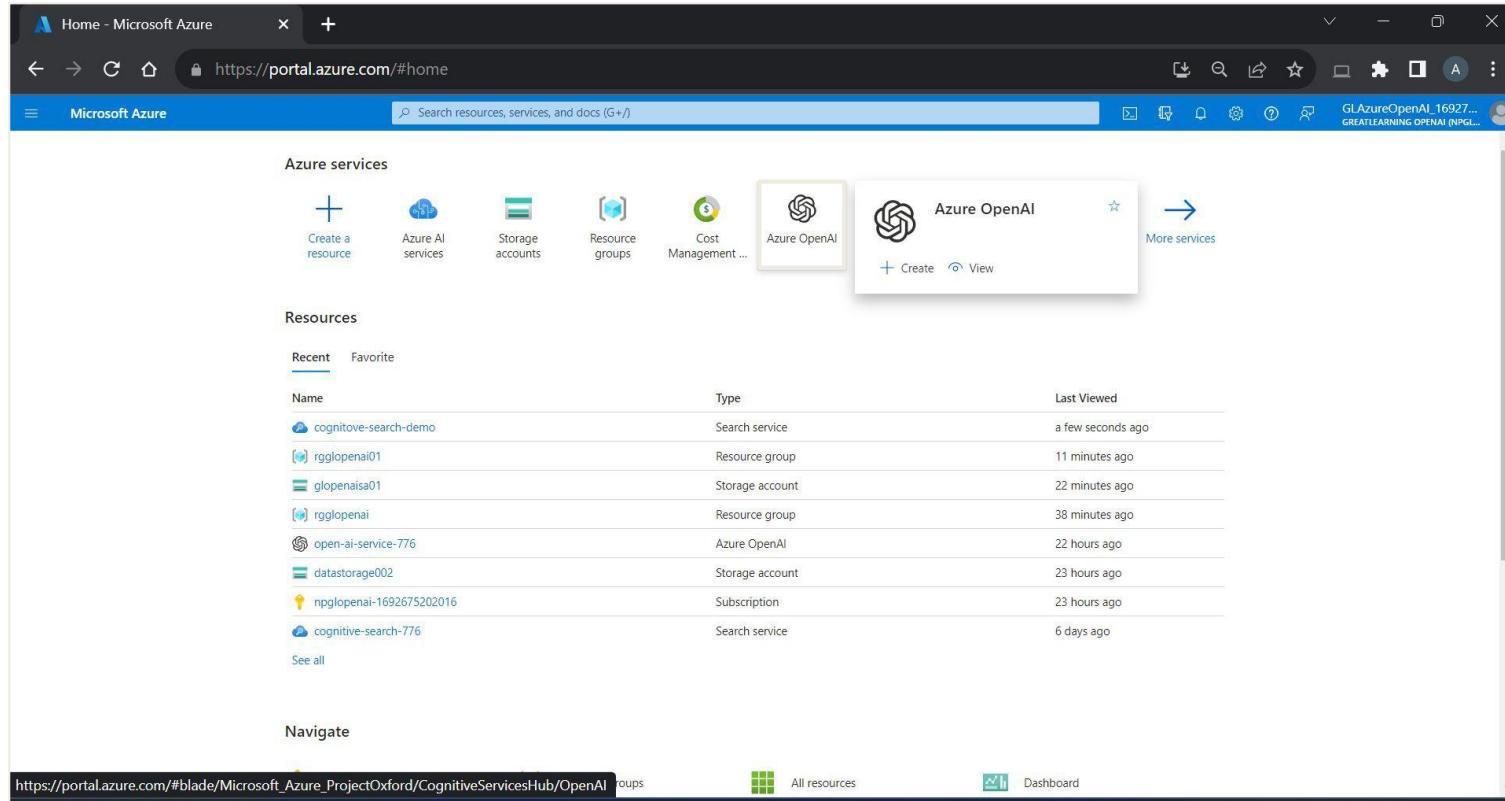
Ground rules for setting up an OpenAI Resource:

- Utilize the **default resource group** for all your tasks.
- Select **Region = East US** for the Open AI resource
- **Pricing tier: Standard S0 pricing tier**
- **Allow 10 to 15 minutes** for the resource to become fully operational following its creation.
- If the microsoft credentials are required at any point, employ the credentials found within the **"Welcome to the Azure Gen AI" section on Olympus** to complete the Microsoft Login process.
- It is advised to create and use a **single resource** for all future tasks within the program.



The screenshot shows a web interface for the Azure gen AI program. At the top, there is a navigation bar with the Great Learning logo, Dashboard, Courses, a search bar, and user icons. The main content area has a title "Welcome to the Azure gen AI program". Below the title, there are two input fields: "loginId" containing "test\_r3nlasug87\_u10767470\_l2@greatlearni.onmicrosoft.com" and "loginpassword" containing a masked password. Below these fields is a note: "Please use the above credentials to login to the Azure portal and access the lab. [portal link](#)".

# Click on Azure OpenAI service



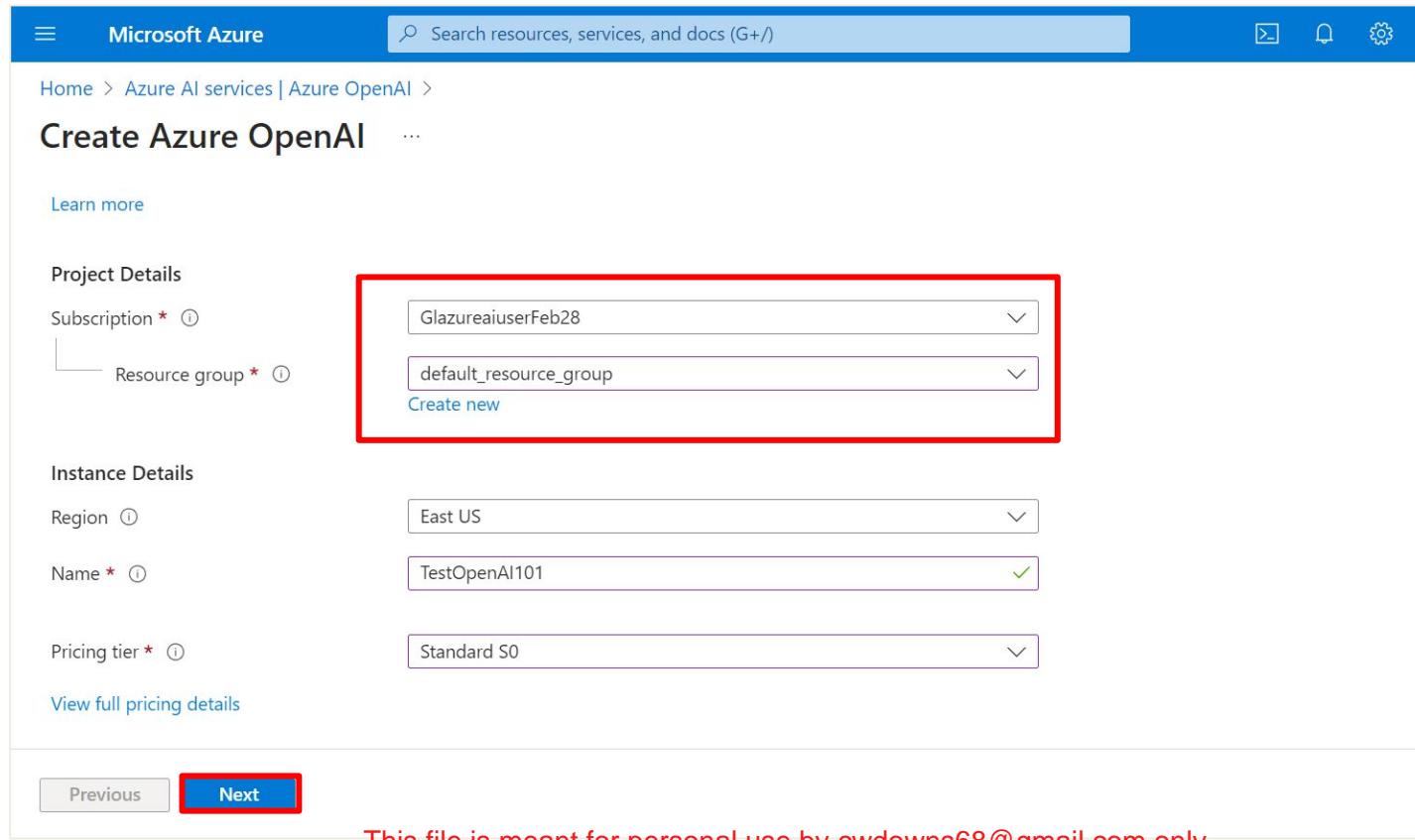
The screenshot shows the Microsoft Azure portal home page. At the top, there is a search bar and a navigation bar with icons for download, search, and other portal functions. The main area features a "Azure services" section with icons for Create a resource, Azure AI services, Storage accounts, Resource groups, Cost Management ..., Azure OpenAI (which is highlighted with a yellow box), and More services. Below this is a "Resources" section with tabs for Recent and Favorite. The "Recent" tab is selected, displaying a list of resources with their names, types, and last viewed times:

Name	Type	Last Viewed
cognitove-search-demo	Search service	a few seconds ago
rgopenai01	Resource group	11 minutes ago
glopenaisa01	Storage account	22 minutes ago
rgopenai	Resource group	38 minutes ago
open-ai-service-776	Azure OpenAI	22 hours ago
datastorage002	Storage account	23 hours ago
npgopenai-1692675202016	Subscription	23 hours ago
cognitive-search-776	Search service	6 days ago

At the bottom, there is a "Navigate" section with links for All resources and Dashboard.

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.  
Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Select the details as below



Microsoft Azure Search resources, services, and docs (G+) ☰ 🔍 🔔 🔋

Home > Azure AI services | Azure OpenAI >

## Create Azure OpenAI

[Learn more](#)

**Project Details**

Subscription \* ⓘ GlazureaiuserFeb28

Resource group \* ⓘ default\_resource\_group

[Create new](#)

**Instance Details**

Region ⓘ East US

Name \* ⓘ TestOpenAI101

Pricing tier \* ⓘ Standard S0

[View full pricing details](#)

[Previous](#) **Next**

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.  
Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

Microsoft Azure Search resources, services, and docs (G+) ≡ test\_r3nlasug87\_u10767... GREAT LEARNING (GREATLEARNI... 👤

Home > Azure AI services | Azure OpenAI >

## Create Azure OpenAI ...

Basics 2 Network Tags Review + submit

i Configure network security for your Azure AI services resource.

Type \*

All networks, including the internet, can access this resource.  
 Selected networks, configure network security for your Azure AI services resource.  
 Disabled, no networks can access this resource. You could configure private endpoint connections that will be the exclusive way to access this resource.

Previous Next

Microsoft Azure  Search resources, services, and docs (G+/       test\_r3nlasug87\_u10767...  
GREAT LEARNING (GREATLEARNI...)

Home > Azure AI services | Azure OpenAI >

## Create Azure OpenAI

 Basics  Network  Tags  Review + submit

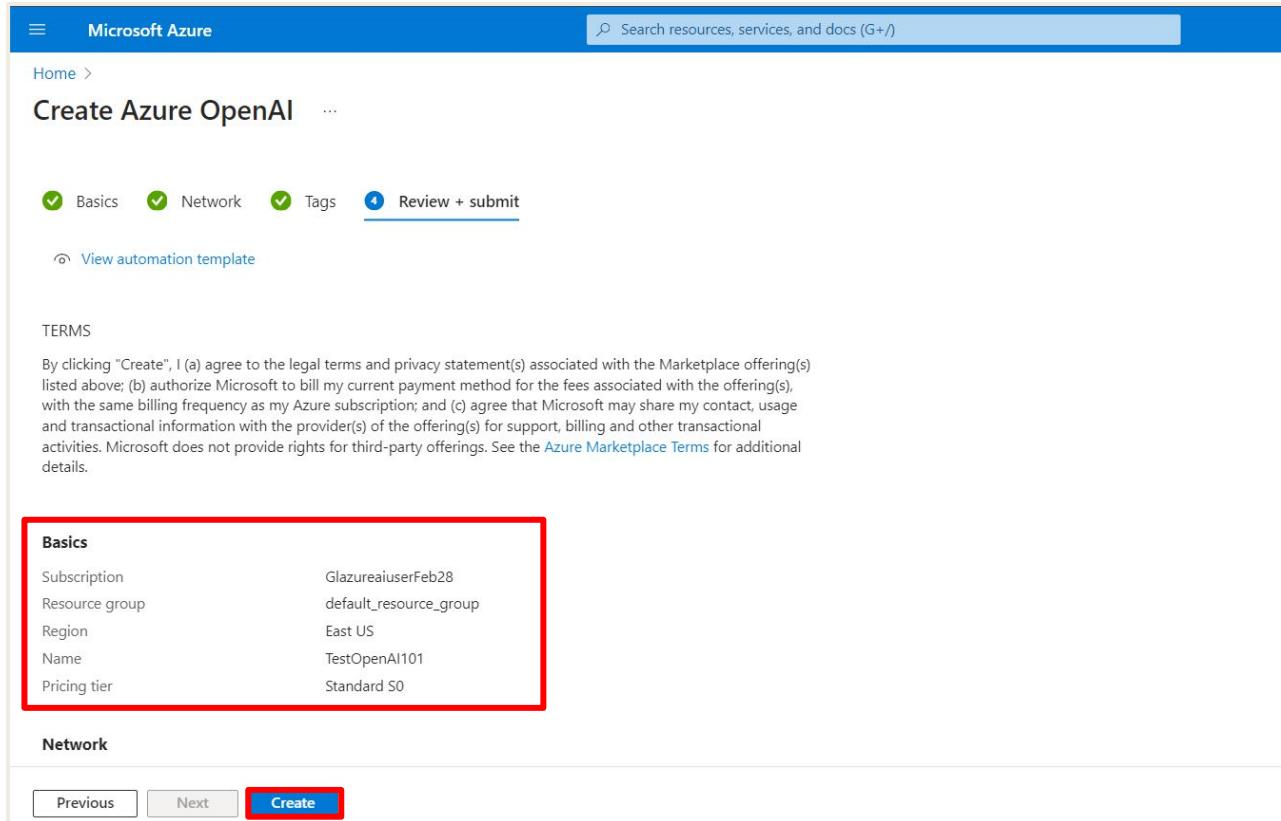
Tags are name/value pairs that enable you to categorize resources and view consolidated billing by applying the same tag to multiple resources and resource groups. [Learn more about tags](#)

Note that if you create tags and then change resource settings on other tabs, your tags will be automatically updated.

Name 	Value 	Resource
<input type="text"/>	:	Azure AI services

[Previous](#) [Next](#)

# Click on Create



The screenshot shows the Microsoft Azure 'Create Azure OpenAI' wizard. The top navigation bar includes the Microsoft Azure logo, a search bar, and a 'Review + submit' button which is underlined, indicating it's the current step. Below the navigation, the breadcrumb path 'Home > Create Azure OpenAI' is shown. The main content area has tabs for 'Basics', 'Network', 'Tags', and 'Review + submit'. The 'Basics' tab is selected and highlighted with a red box. It contains the following configuration details:

Subscription	GlazureaiuserFeb28
Resource group	default_resource_group
Region	East US
Name	TestOpenAI101
Pricing tier	Standard S0

Below the 'Basics' section, there is a 'Network' section. At the bottom of the page are buttons for 'Previous', 'Next', and 'Create'.

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.  
Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Deployment will take few mins

Microsoft Azure

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

test r3nlasug87 u10767...  
GREAT LEARNING (GREATLEARNI...)

Home > Microsoft.CognitiveServicesOpenAI-20240322163603 | Overview

Deployment

Search Delete Cancel Redeploy Download Refresh

Overview

Inputs

Outputs

Template

Your deployment is complete

Deployment name : Microsoft.CognitiveServicesOpenAI-20240322163603  
Subscription : GlazureaiuserFeb28  
Resource group : default\_resource\_group

Start time : 22/03/2024, 16:38:47  
Correlation ID : 028e6f70-f34b-438e-9b2c-6ec87917cba3

Deployment details

Next steps

Go to resource

Cost management

Get notified to stay within your budget and prevent unexpected charges on your bill.  
[Set up cost alerts >](#)

Microsoft Defender for Cloud

Secure your apps and infrastructure  
[Go to Microsoft Defender for Cloud >](#)

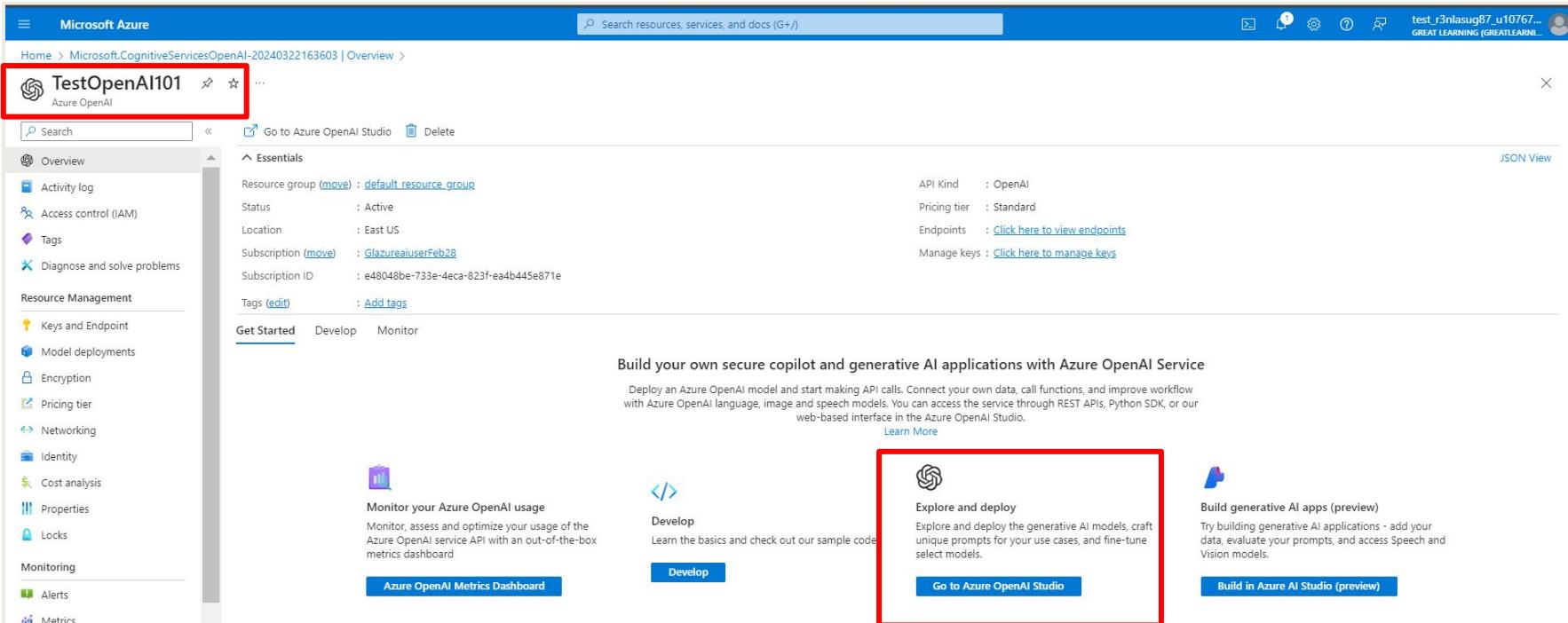
Free Microsoft tutorials

[Start learning today >](#)

Work with an expert

Azure experts are service provider

# Click on Explore once the deployment is complete



Microsoft Azure

Home > Microsoft.CognitiveServicesOpenAI-20240322163603 | Overview >

## TestOpenAI101

Azure OpenAI

Search

Go to Azure OpenAI Studio Delete

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Resource Management

Keys and Endpoint

Model deployments

Encryption

Pricing tier

Networking

Identity

Cost analysis

Properties

Locks

Monitoring

Alerts

Metrics

Get Started Develop Monitor

Essentials

Resource group ( <a href="#">move</a> ) : <a href="#">default_resource_group</a>	API Kind : OpenAI
Status : Active	Pricing tier : Standard
Location : East US	Endpoints : <a href="#">Click here to view endpoints</a>
Subscription ( <a href="#">move</a> ) : <a href="#">glazureaiuserFeb28</a>	Manage keys : <a href="#">Click here to manage keys</a>
Subscription ID : e48048be-733e-4eca-823f-ea4b445e871e	
Tags ( <a href="#">edit</a> ) : <a href="#">Add tags</a>	

Build your own secure copilot and generative AI applications with Azure OpenAI Service

Deploy an Azure OpenAI model and start making API calls. Connect your own data, call functions, and improve workflow with Azure OpenAI language, image and speech models. You can access the service through REST APIs, Python SDK, or our web-based interface in the Azure OpenAI Studio.

[Learn More](#)

Monitor your Azure OpenAI usage

Monitor, assess and optimize your usage of the Azure OpenAI service API with an out-of-the-box metrics dashboard.

[Azure OpenAI Metrics Dashboard](#)

Develop

Explore and deploy

Explore and deploy the generative AI models, craft unique prompts for your use cases, and fine-tune select models.

[Go to Azure OpenAI Studio](#)

Build generative AI apps (preview)

Try building generative AI applications - add your data, evaluate your prompts, and access Speech and Vision models.

[Build in Azure AI Studio \(preview\)](#)

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.  
Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Deploy a model to use Chat Playground

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

Proprietary Content © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Ground Rules : Important

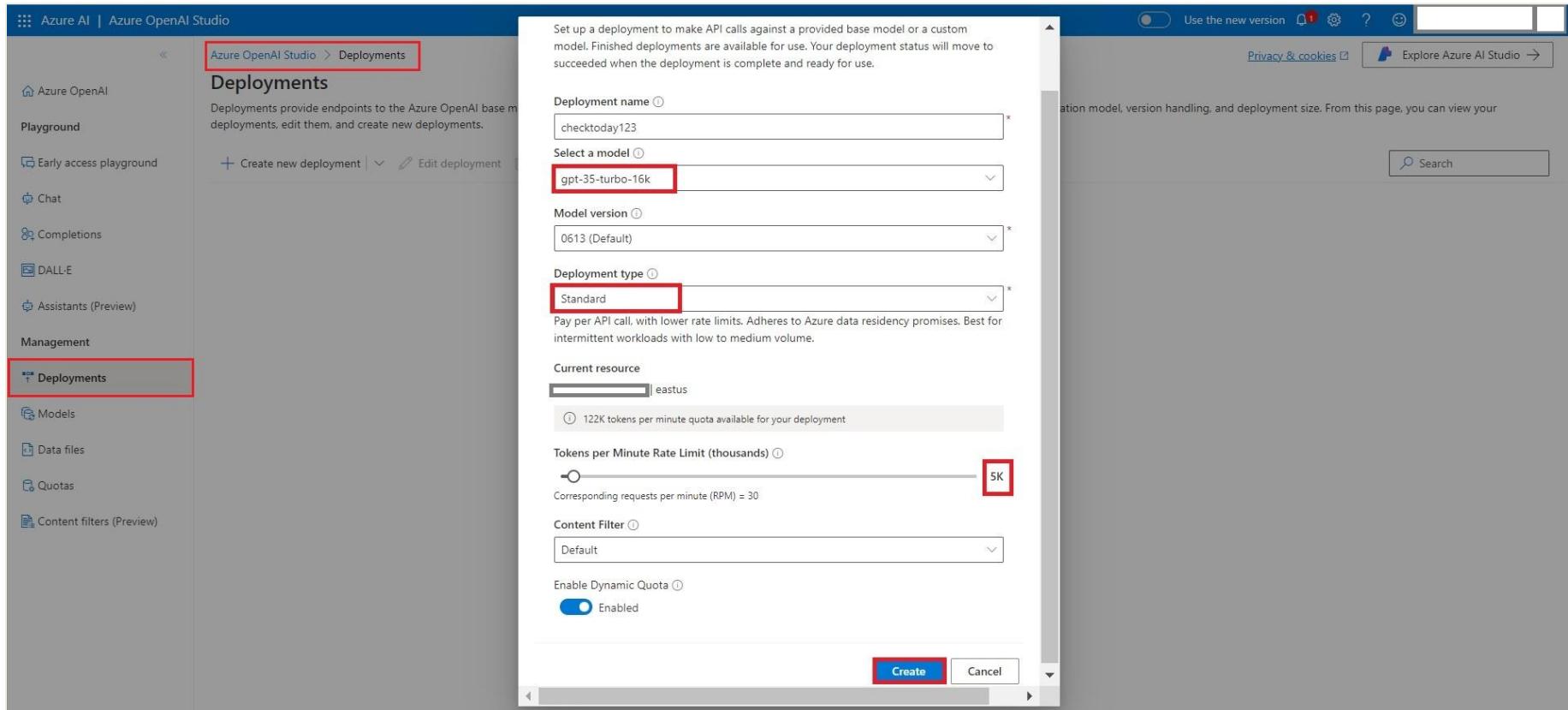
Ground rules for deploying a model in Azure OpenAI Resource:

- **Deployment name:** A unique name of your choice
- Select the **Model:** gpt-35-turbo-16k only
- Model version: **Auto-update to default**
- Deployment type: **Standard**
- **Tokens per minute rate limit: 5K**
- Content filter: Default
- Enable dynamic quota: Enabled
- **Allow 10 to 15 minutes** for the deployment to become fully operational following its creation.
- Please ensure that you have **only 1 ACTIVE DEPLOYMENT**. This is because the total token limit summed up across all the deployments is only **240k** and if you exceed this, you will not be able deploy any model.

This file is meant for personal use by cwdowns68@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.  
Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Deploy a model



The screenshot shows the Azure OpenAI Studio interface with the 'Deployments' page open. The 'Deployments' section is highlighted with a red box. Within this section, the 'Create new deployment' button is also highlighted with a red box. The main content area displays a form for creating a new deployment:

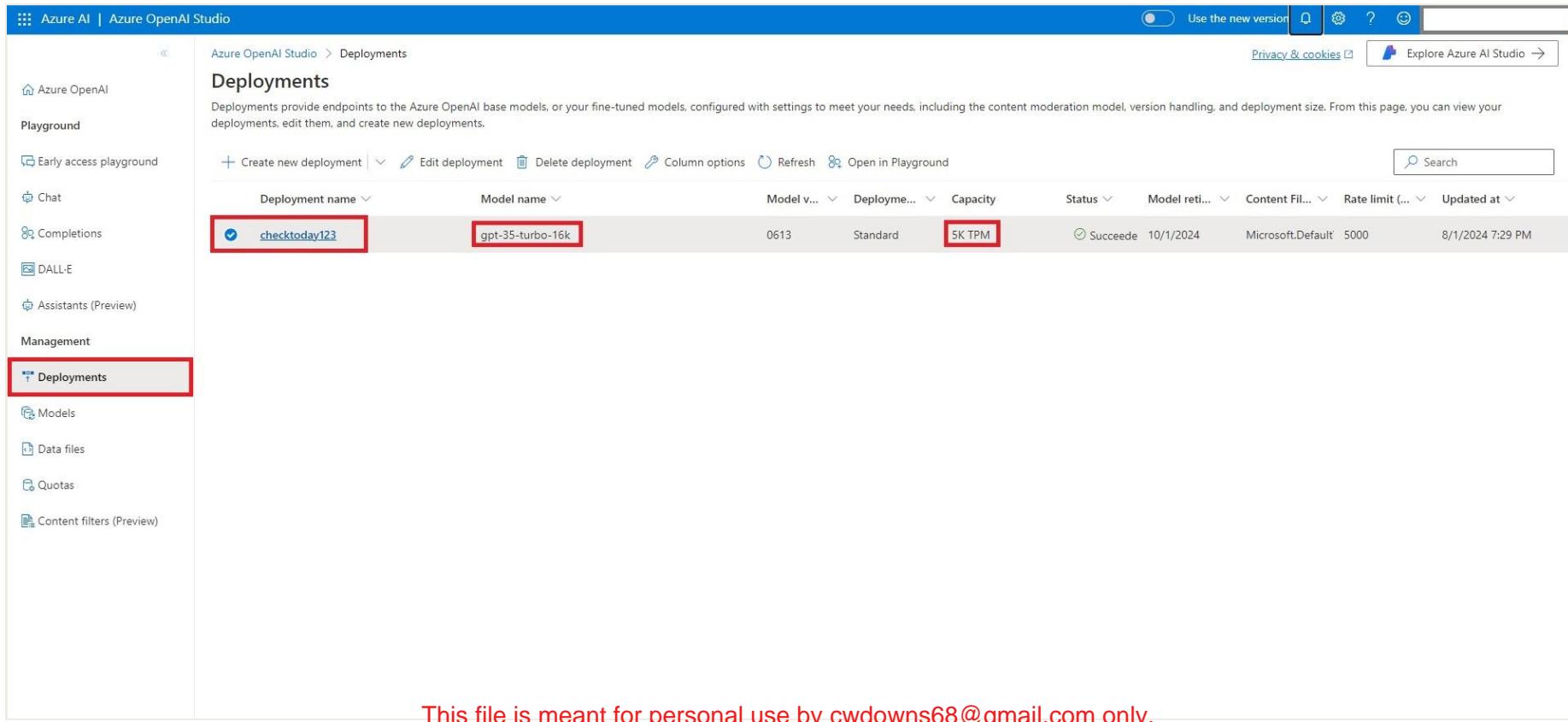
- Deployment name:** checktoday123
- Select a model:** gpt-35-turbo-16k
- Model version:** 0613 (Default)
- Deployment type:** Standard
- Current resource:** eastus
- Tokens per Minute Rate Limit (thousands):** 5K (highlighted with a red box)
- Content Filter:** Default
- Enable Dynamic Quota:** Enabled

At the bottom right of the dialog are 'Create' and 'Cancel' buttons.

This file is meant for personal use by cwdowns68@gmail.com only.  
 Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# You will be able to view this



Azure AI | Azure OpenAI Studio

Azure OpenAI Studio > Deployments

Deploys provide endpoints to the Azure OpenAI base models, or your fine-tuned models, configured with settings to meet your needs, including the content moderation model, version handling, and deployment size. From this page, you can view your deployments, edit them, and create new deployments.

+ Create new deployment | Edit deployment | Delete deployment | Column options | Refresh | Open in Playground

Deployment name Model name Model v... Deployme... Capacity Status Model reti... Content Fil... Rate limit (...) Updated at

Deployment name	Model name	Model v...	Deployme...	Capacity	Status	Model reti...	Content Fil...	Rate limit (...)	Updated at
<input checked="" type="checkbox"/> checktoday123	gpt-35-turbo-16k	0613	Standard	5K TPM	 Succeeded	10/1/2024	Microsoft.Default	5000	8/1/2024 7:29 PM

Navigation menu:

- Azure OpenAI
- Playground
- Early access playground
- Chat
- Completions
- DALL-E
- Assistants (Preview)
- Management
  - Deployments (selected)
  - Models
  - Data files
  - Quotas
  - Content filters (Preview)

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

# Use the Chat Playground

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

Proprietary Content © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Use the Chat Playground

1. In the Playground section, select the Chat page. The Chat playground page consists of three main panels (which may be arranged right-to-left horizontally, or top-to-bottom vertically depending on your screen resolution):
  - **Setup** - used to set the context for the model's responses.
  - **Chat session** - used to submit chat messages and view responses.
  - **Configuration** - used to configure settings for the model deployment.
2. In the Configuration panel, ensure that your **gpt-35-turbo-16k model deployment** is selected.
3. In the Setup panel, review the default System message, which should be "**You are an AI assistant that helps people find information**". The system message is included in prompts submitted to the model, and provides context for the model's responses; setting expectations about how an AI agent based on the model should interact with the user.

# Use the system message

In the previous weeks, you've engaged in a chat conversation with your model based on the default system message. You can customize the system setup to have more control over the kinds of responses generated by your model with respect to the case study covered that will be covering here.

You are an assistant trained to classify customer intents based on provided text.

Below is a list of customer intents along with their intent codes separated by a ":":

- top\_up\_by\_bank\_transfer\_charge:56
- top\_up\_by\_card\_charge:57
- top\_up\_by\_cash\_or\_cheque:58
- top\_up\_failed:59
- top\_up\_limits:60
- top\_up\_reverted:61
- topping\_up\_by\_card:62

Given the following text, classify it and output only one intent code:

Customer: ```{text}```

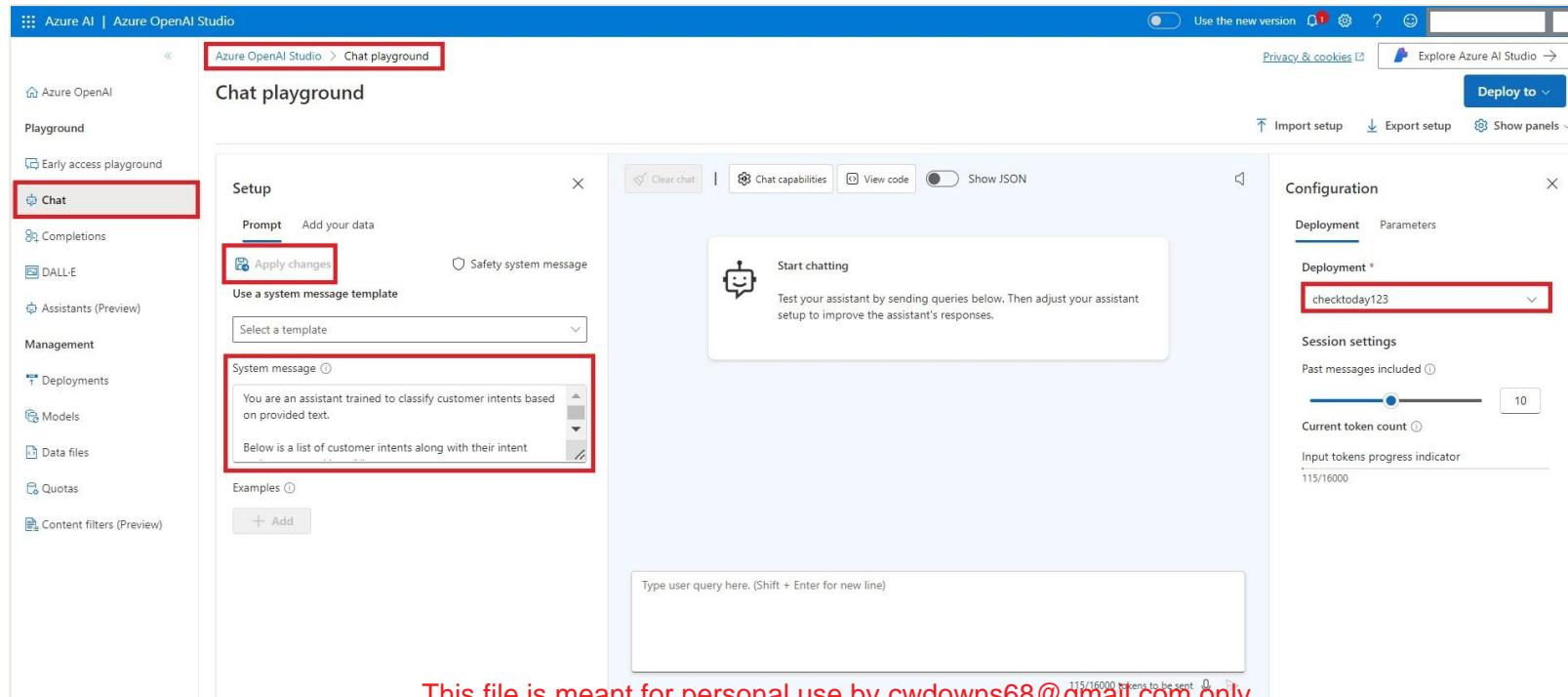
Intent Code:

This file is meant for personal use by cwdowns68@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

# Use the Chat Playground

When the model deployment is completed, navigate to the “Chat Playground”. Enter the system message provided and then click on “Apply Changes”.



Azure AI | Azure OpenAI Studio

Azure OpenAI Studio > Chat playground

Chat playground

Setup

Prompt Add your data

Apply changes Safety system message

Use a system message template

Select a template

System message

You are an assistant trained to classify customer intents based on provided text.  
Below is a list of customer intents along with their intent

Start chatting

Test your assistant by sending queries below. Then adjust your assistant setup to improve the assistant's responses.

Configuration

Deployment Parameters

Deployment \* checktoday123

Session settings

Past messages included 10

Current token count

Input tokens progress indicator 115/16000

Type user query here. (Shift + Enter for new line)

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

# Create Azure AI search service

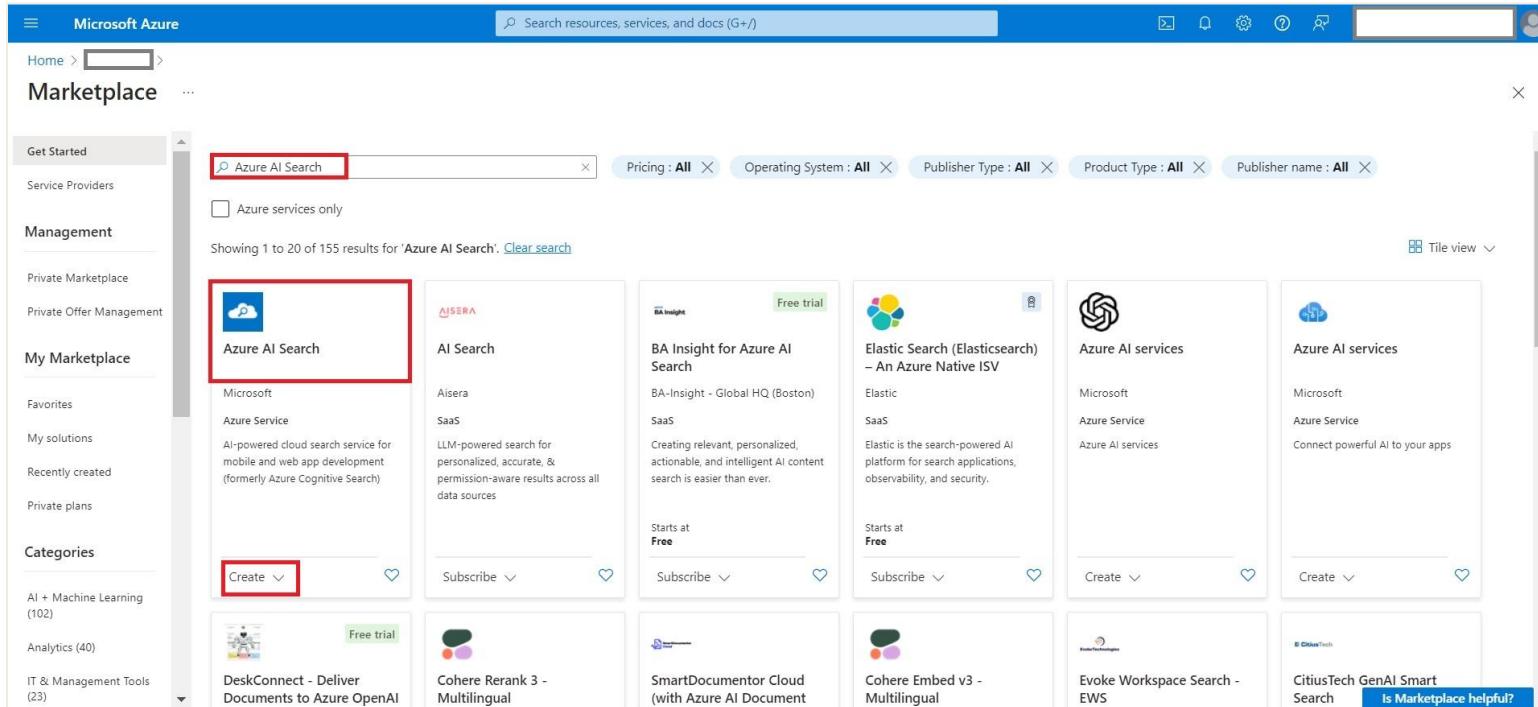
This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.  
Proprietary Content © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Need of Azure AI search service

- Integrating Azure AI Search Service with your deployed webapp significantly enhances its capability to retrieve and utilize data efficiently.
- This service enables quick access to relevant information from large datasets, supporting powerful search functionalities like full-text search, filtering, and navigation. By providing precise and contextually relevant answers, it improves the overall user experience, ensuring users receive accurate and helpful responses.
- Seamlessly integrating with extensive datasets from Hugging Face, Azure AI Search Service ensures optimal data utilization and real-time information retrieval. Its scalability and flexibility allows the deployed webapp to handle large volumes of data and adapt to complex queries, making it an indispensable tool for enhancing the webapp's performance and user satisfaction.

# Create Azure AI search service

1. Navigate to **Azure AI search service** under the resource group provisioned.
2. Click on “**Create**” button.

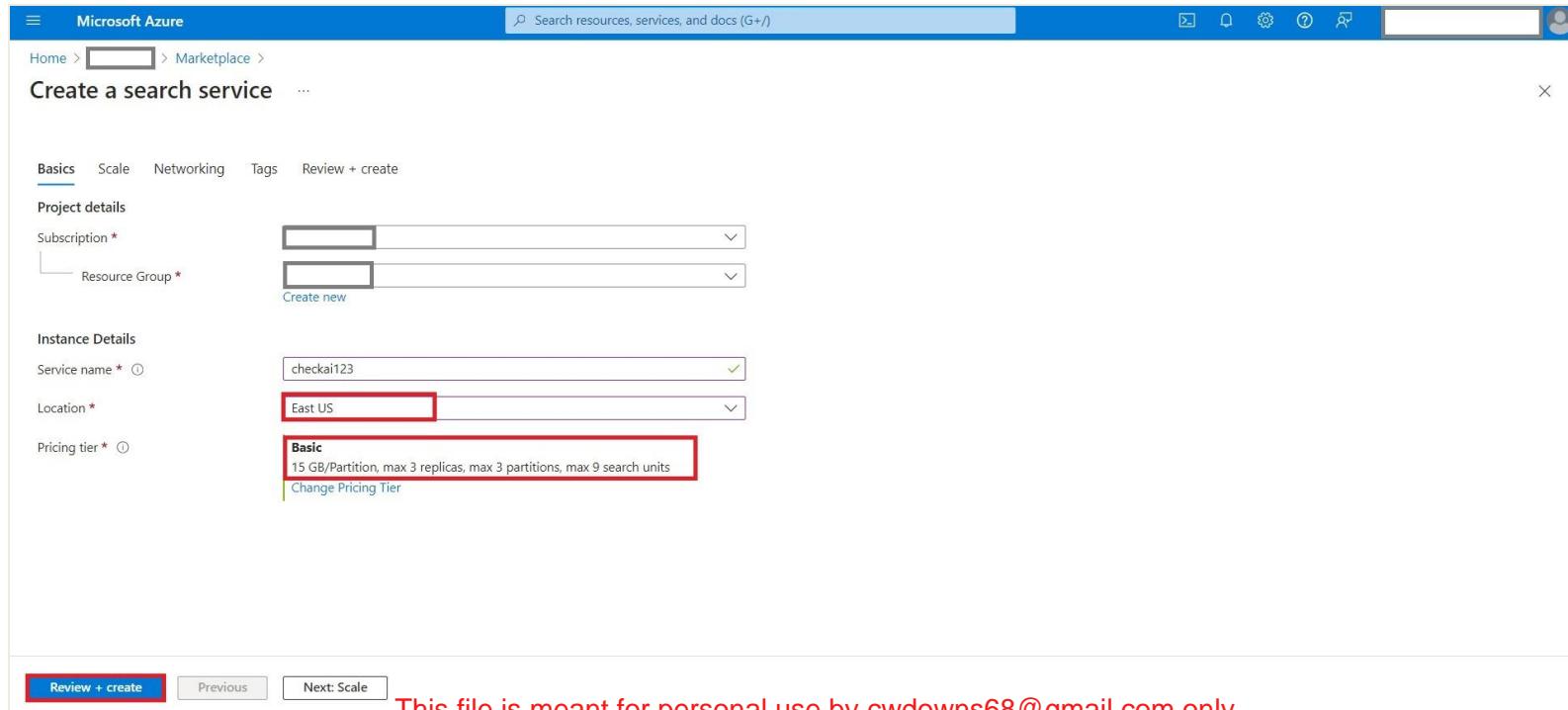


The screenshot shows the Microsoft Azure Marketplace interface. The search bar at the top contains the query "Azure AI Search". Below the search bar, there are several filter options: Pricing : All, Operating System : All, Publisher Type : All, Product Type : All, and Publisher name : All. A checkbox for "Azure services only" is also present. The results section displays 1 to 20 of 155 results for 'Azure AI Search'. The first result, "Azure AI Search" by Microsoft, is highlighted with a red box. This item has a "Create" button with a dropdown arrow below it. Other results include "AI Search" by Aisera, "BA Insight for Azure AI Search" by BA Insight, "Elastic Search (Elasticsearch) – An Azure Native ISV" by Elastic, "Azure AI services" by Microsoft, and "Azure AI services" by Microsoft again. Each result card includes a thumbnail, the provider name, the service name, a brief description, and a "Subscribe" or "Create" button.

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

# Create Azure AI search service

1. Use **East US** location ad **Basic Pricing Tier**.
2. Click on “**Review + create**” button.

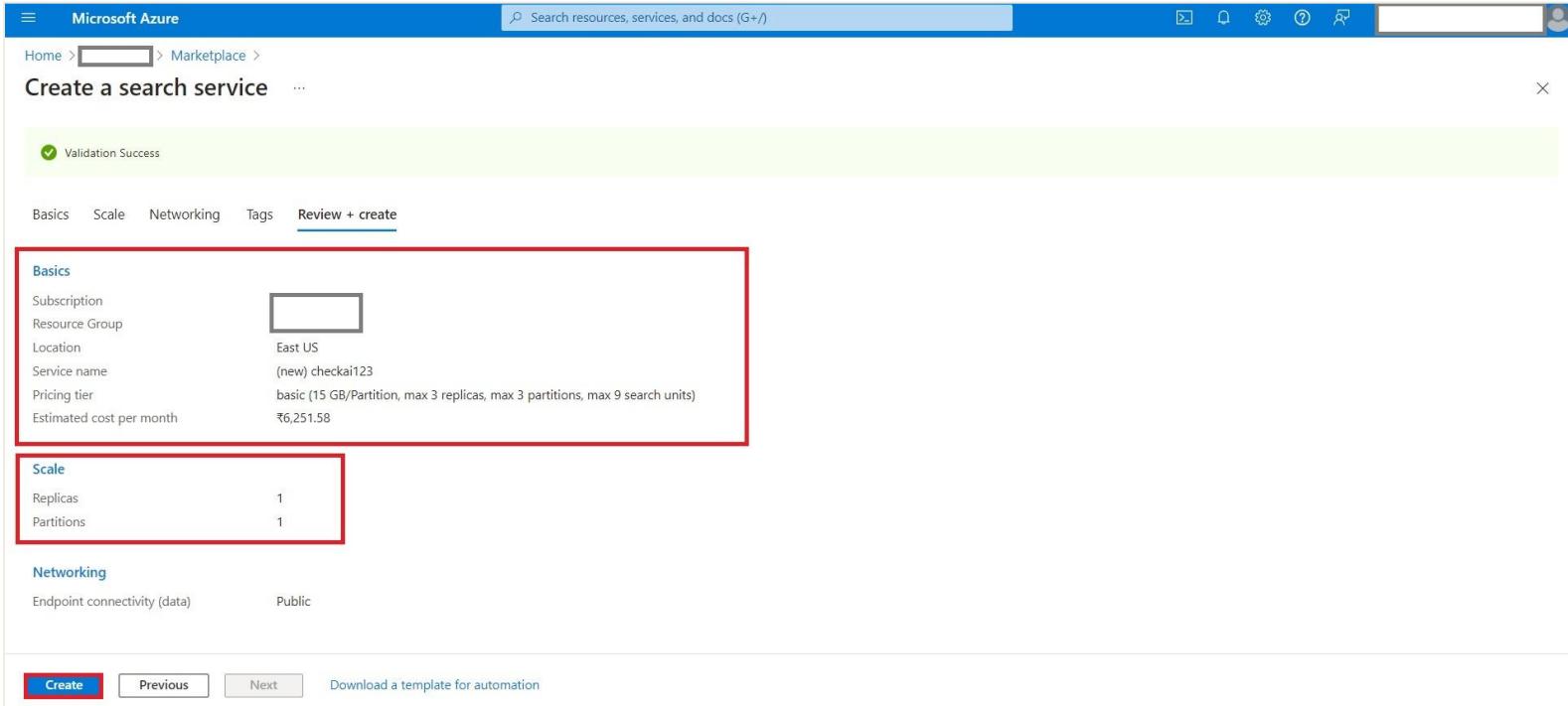


The screenshot shows the Microsoft Azure 'Create a search service' wizard in progress. The 'Basics' tab is selected. In the 'Project details' section, the 'Subscription' dropdown is populated with a value. Below it, the 'Resource Group' dropdown also has a value, with a link to 'Create new' below it. In the 'Instance Details' section, the 'Service name' field contains 'checkai123'. The 'Location' field is highlighted with a red box and contains 'East US'. The 'Pricing tier' field is also highlighted with a red box and contains 'Basic', with a note below stating '15 GB/Partition, max 3 replicas, max 3 partitions, max 9 search units' and a 'Change Pricing Tier' link. At the bottom of the wizard, there are buttons for 'Review + create' (which is highlighted with a red box), 'Previous', and 'Next: Scale'.

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

# Create Azure AI search service

Click on “Create” button.



The screenshot shows the Microsoft Azure portal interface for creating a search service. The top navigation bar includes 'Microsoft Azure', a search bar, and various icons. The main page shows a breadcrumb path: Home > [redacted] > Marketplace > Create a search service. A green success message 'Validation Success' is displayed. The 'Review + create' tab is selected. The configuration is as follows:

Category	Setting	Value
Basics	Subscription	[redacted]
	Resource Group	[redacted]
	Location	East US
	Service name	(new) checkai123
	Pricing tier	basic (15 GB/Partition, max 3 replicas, max 3 partitions, max 9 search units)
Estimated cost per month	₹6,251.58	
Scale	Replicas	1
	Partitions	1
Networking	Endpoint connectivity (data)	Public

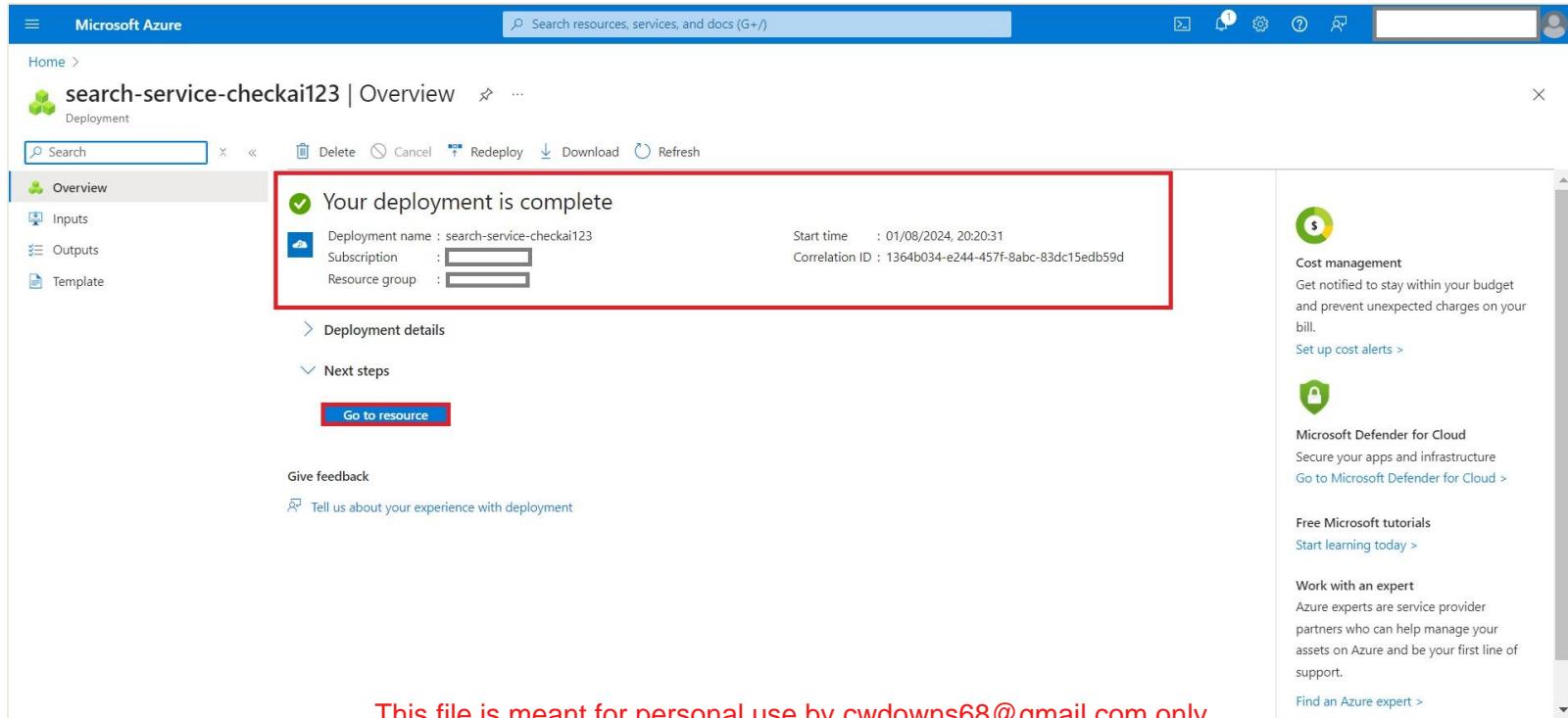
At the bottom, there are 'Create', 'Previous', and 'Next' buttons, along with a link to 'Download a template for automation'.

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Create Azure AI search service

Deployment will take few minutes. Once it is done, click on “**Go to resource**”.



The screenshot shows the Microsoft Azure Deployment Overview page for a deployment named "search-service-checkai123". A red box highlights the main message: "Your deployment is complete". Below this, deployment details are listed: Deployment name: search-service-checkai123, Subscription: [redacted], Resource group: [redacted]. To the right, there are links for Cost management, Microsoft Defender for Cloud, Free Microsoft tutorials, and Work with an expert. At the bottom, a red banner displays the text: "This file is meant for personal use by cwdowns68@gmail.com only. Sharing or publishing the contents in part or full is liable for legal action."

Microsoft Azure

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

Home > search-service-checkai123 | Overview

Deployment

Search Delete Cancel Redeploy Download Refresh

Overview

Your deployment is complete

Deployment name : search-service-checkai123  
Subscription : [redacted]  
Resource group : [redacted]

Start time : 01/08/2024, 20:20:31  
Correlation ID : 1364b034-e244-457f-8abc-83dc15edb59d

Deployment details

Next steps

Go to resource

Give feedback

Tell us about your experience with deployment

Cost management

Get notified to stay within your budget and prevent unexpected charges on your bill.  
Set up cost alerts >

Microsoft Defender for Cloud

Secure your apps and infrastructure  
Go to Microsoft Defender for Cloud >

Free Microsoft tutorials

Start learning today >

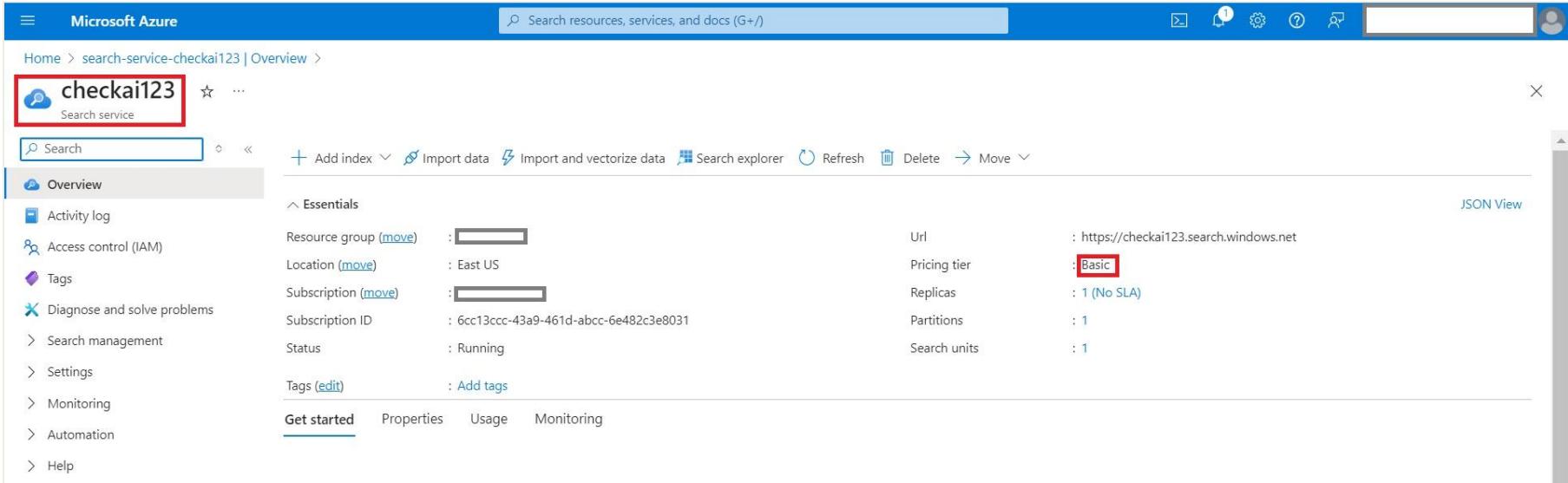
Work with an expert

Azure experts are service provider partners who can help manage your assets on Azure and be your first line of support.  
Find an Azure expert >

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

# Create Azure AI search service

You will be able to view the search service created.



The screenshot shows the Microsoft Azure search service overview page for a service named 'checkai123'. The service is highlighted with a red box. The URL is https://checkai123.search.windows.net. The pricing tier is set to 'Basic' (highlighted with a red box). The replicas are 1 (No SLA), partitions are 1, and search units are 1. The status is 'Running'. The location is 'East US'. The resource group and subscription details are listed. The left sidebar shows navigation options like Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Search management, Settings, Monitoring, Automation, and Help. The top navigation bar includes a search bar, a refresh button, and other account-related icons.

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Case Study

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.  
Proprietary Content © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Case Study

**Problem Statement:** In the banking sector, efficiently managing customer queries is vital for enhancing service quality and customer satisfaction. Given the diversity of customer intents, ranging from general inquiries to specific service requests, accurately identifying these intents is crucial. This case study focuses on a subset of customer intents related to credit card top-ups. The challenge is to deploy a large language model (LLM) based multi-class classification model and utilize few-shot learning prompts to accurately classify customer queries related to credit card top-ups.

**Objective:** To improve customer service efficiency and satisfaction by:

1. **Deploying a Multi-Class Classification Model:** Implement a large language model (LLM) capable of classifying a range of customer intents, specifically targeting queries related to credit card top-ups.
2. **Utilizing Few-Shot Learning Prompts:** Apply few-shot learning techniques to train the model on a banking dataset, enabling it to effectively understand and classify customer intents with minimal examples.
3. **Real-Time Deployment:** Deploy the model using the Azure Chat Playground for real-time analysis of customer queries and categorizing the intents.

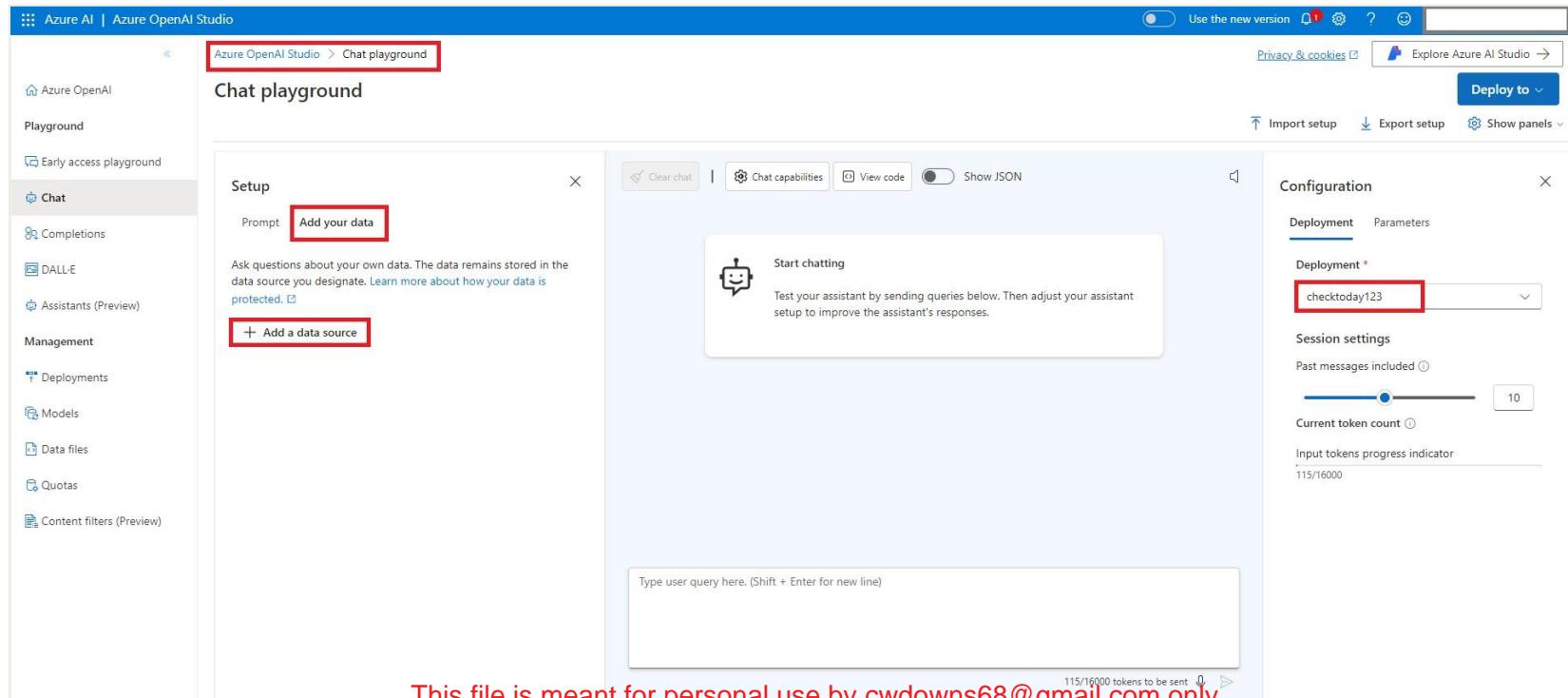
By achieving these objectives, the model aims to enhance the accuracy of intent identification and streamline the handling of credit card top-up queries, leading to improved customer service and satisfaction.

# Add your Data

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.  
Proprietary Content © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Add your Data

Next, navigate back to the Azure OpenAI Chat Playground.  
Click on “+Add a data source”



Azure AI | Azure OpenAI Studio

Azure OpenAI Studio > Chat playground

Chat playground

Setup

Prompt **Add your data**

Ask questions about your own data. The data remains stored in the data source you designate. Learn more about how your data is protected.

+ Add a data source

Start chatting

Test your assistant by sending queries below. Then adjust your assistant setup to improve the assistant's responses.

Configuration

Deployment Parameters

Deployment \* **checktoday123**

Session settings

Past messages included

Current token count 10

Input tokens progress indicator 115/16000

Type user query here. (Shift + Enter for new line)

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

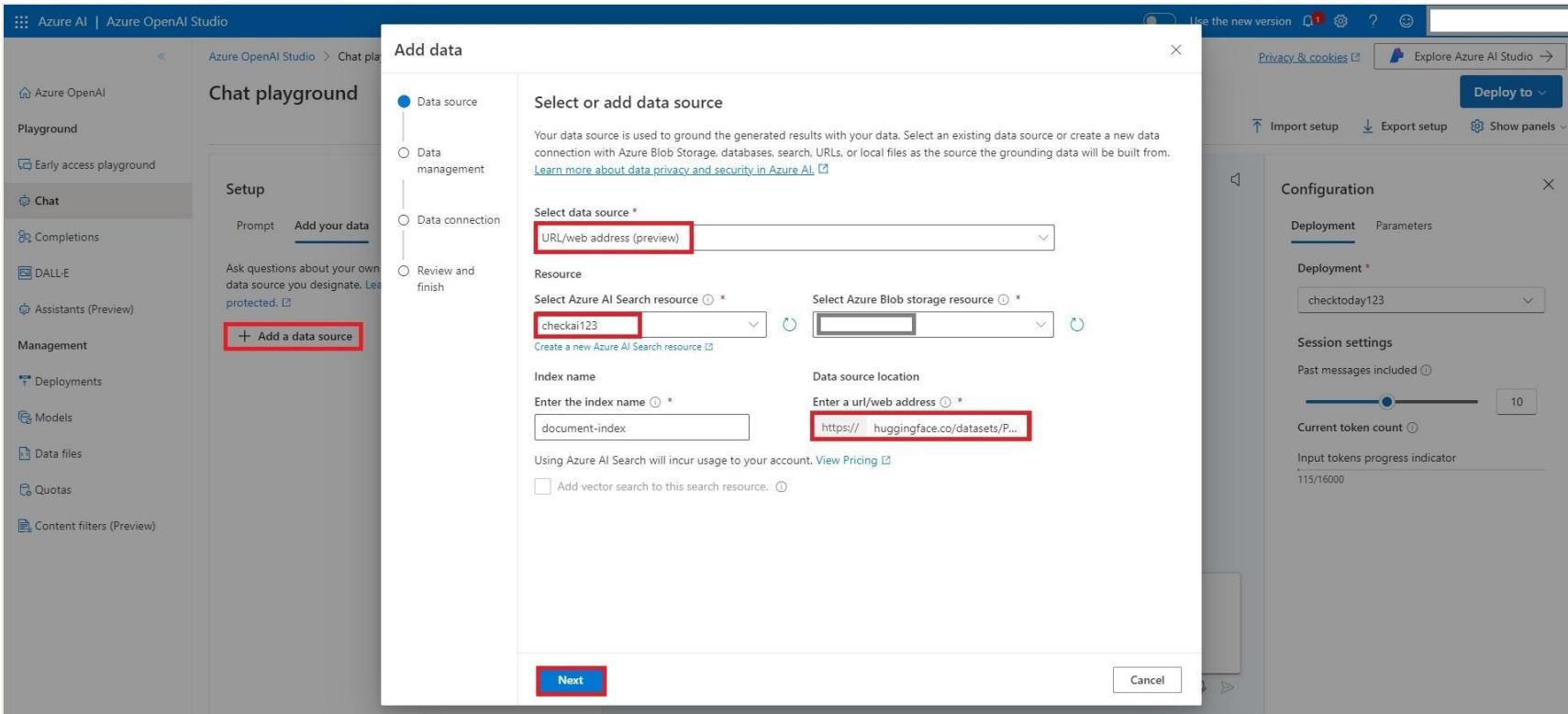
# Add your Data

1. We will be using the a **hugging face dataset called banking77** from their dataset repository.
2. Explore the dataset at <https://huggingface.co/datasets/PolyAI/banking77>
3. There are a total of 77 classes or intents which have been defined. For the purpose of learning, we will pick a few classes (7) which are related to credit card top-ups.
4. **Select data source :** URL/web address (preview)
5. Select the Azure AI search resource created previously
6. Select the Blob Storage resource
7. Enter an index-name of your choice
8. Copy and paste the dataset mentioned in **Step2** in the section “**Enter a url/web address**”
9. Click on **Next**

56	top_up_by_bank_transfer_charge
57	top_up_by_card_charge
58	top_up_by_cash_or_cheque
59	top_up_failed
60	top_up_limits
61	top_up_reverted
62	topping_up_by_card

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

# Add your Data



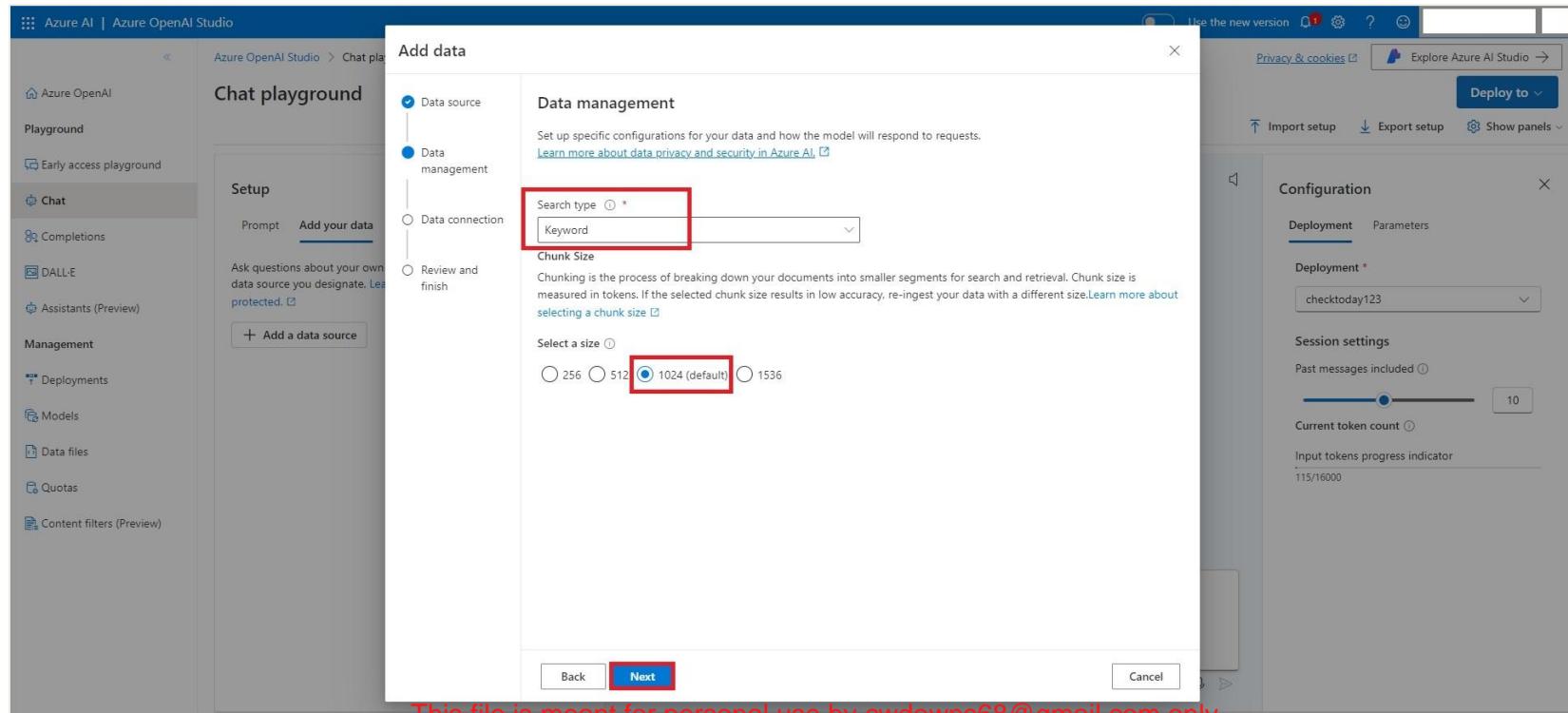
The screenshot shows the 'Add data' dialog box in the Azure OpenAI Studio. The 'Data source' tab is active. The 'Select or add data source' section lists four options: 'Data source' (selected), 'Data management', 'Data connection', and 'Review and finish'. The 'Data source' section includes a 'URL/web address (preview)' input field containing 'https://huggingface.co/datasets/P...', which is highlighted with a red box. Below this are dropdowns for 'Select Azure AI Search resource' (set to 'checkai123') and 'Select Azure Blob storage resource'. The 'Index name' field is set to 'document-index'. The 'Data source location' field has 'Enter a url/web address' and 'https://huggingface.co/datasets/P...' highlighted with a red box. At the bottom are 'Next' and 'Cancel' buttons.

This file is meant for personal use by cwdowns68@gmail.com only.  
 Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Add your Data

Select “**Keyword**” in Search type and **1024** as Chunk size.  
Click on **Next**

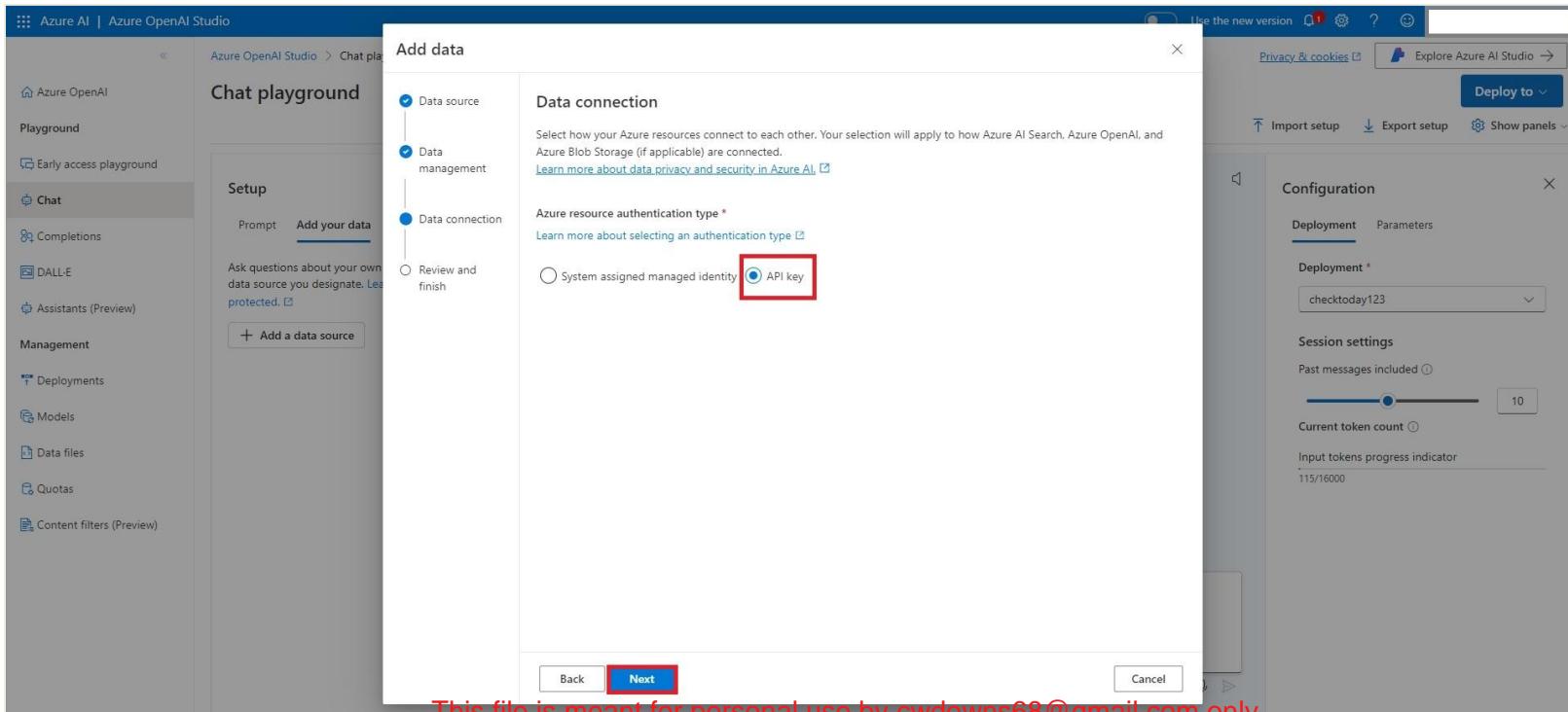


The screenshot shows the Azure OpenAI Studio interface with the 'Chat playground' section selected. A modal window titled 'Add data' is open, specifically the 'Data management' setup step. In the 'Search type' dropdown, the option 'Keyword' is selected. In the 'Select a size' dropdown, the option '1024 (default)' is selected and highlighted with a red box. The 'Next' button at the bottom of the modal is also highlighted with a red box.

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

# Add your Data

Select “API Key” in resource authentication type  
Click on **Next**

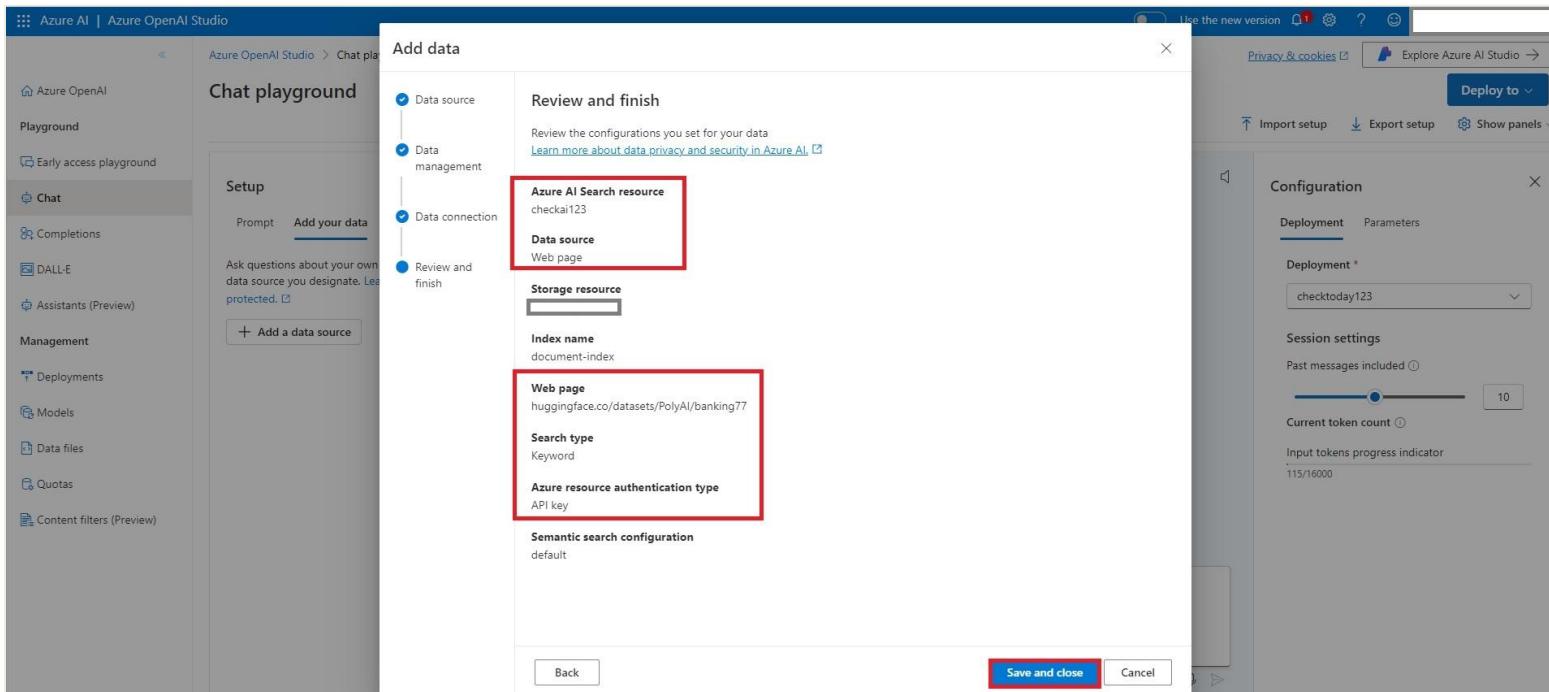


This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

# Add your Data

Review the settings.

Click on **Save and Close**. Allow 10-15 minutes for the ingestion process.



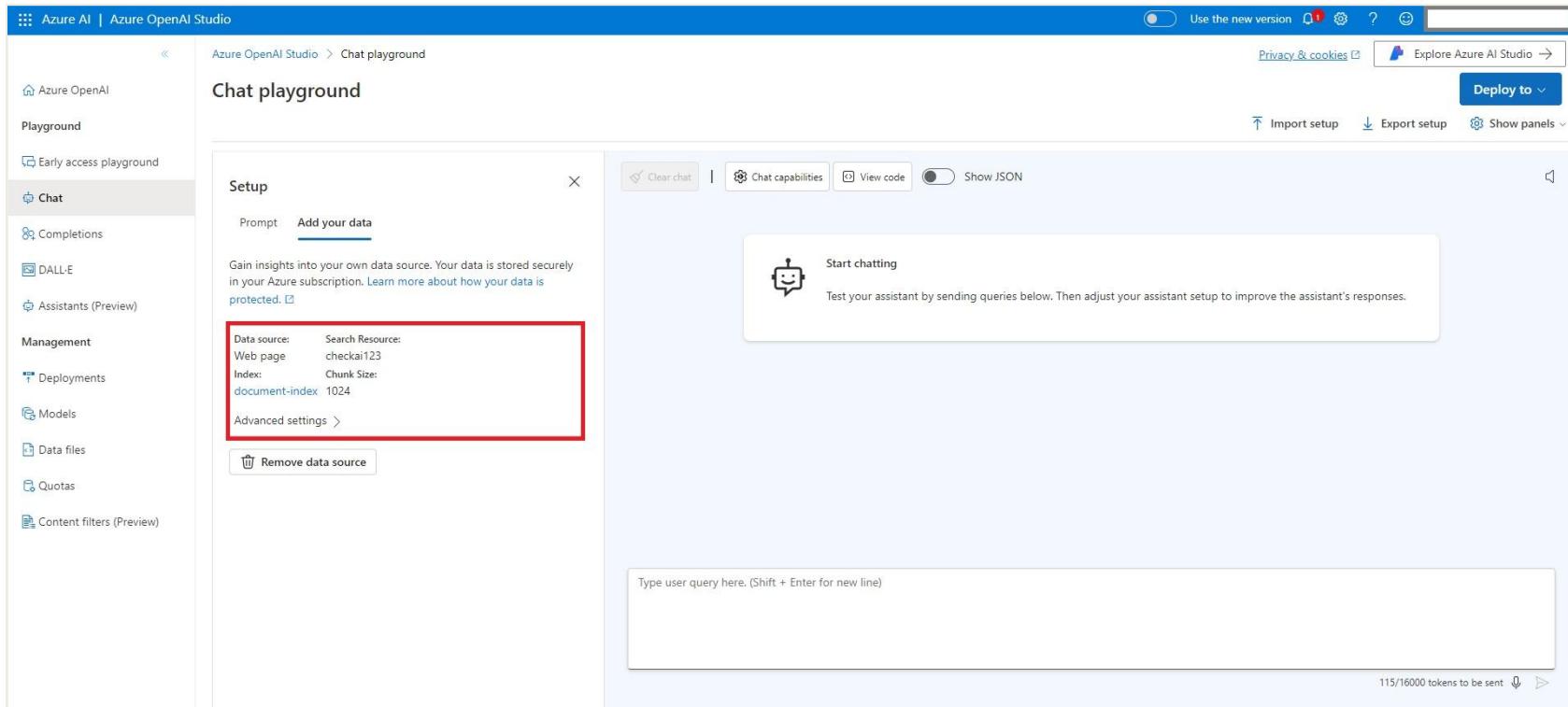
The screenshot shows the Azure OpenAI Studio interface with the 'Chat playground' selected. A modal window titled 'Add data' is open, specifically the 'Review and finish' step. The left sidebar shows various AI models and configurations like Chat, Completions, DALL-E, Assistants (Preview), Management, Deployments, Models, Data files, Quotas, and Content filters (Preview). The main area displays the configuration for a 'Data source' (Web page) under 'Azure AI Search resource'. The 'Data source' section is highlighted with a red border. It includes fields for 'Index name' (document-index), 'Web page' (huggingface.co/datasets/PolyAI/banking77), 'Search type' (Keyword), and 'Azure resource authentication type' (API key). Below this, there's a 'Semantic search configuration' set to 'default'. At the bottom of the modal are 'Back', 'Save and close', and 'Cancel' buttons. To the right of the modal, a 'Configuration' panel is visible with tabs for 'Deployment' and 'Parameters', showing deployment details and session settings.

This file is meant for personal use by cwdowns68@gmail.com only.  
 Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Add your Data

You will now be able to see the data source added under the “Add your data” section.



The screenshot shows the Azure OpenAI Studio interface with the "Chat playground" selected in the sidebar. The main area displays the "Setup" configuration for a chat assistant. A red box highlights the "Data source" configuration section, which includes fields for "Search Resource" (checkbox123), "Web page" (checkai123), "Index" (Chunk Size: document-index), and "Chunk Size" (1024). Below this is an "Advanced settings >" link and a "Remove data source" button. To the right, there's a "Start chatting" section with a robot icon and a text input field for user queries. The top navigation bar includes links for "Privacy & cookies", "Explore Azure AI Studio", "Deploy to", and various setup export/import options.

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

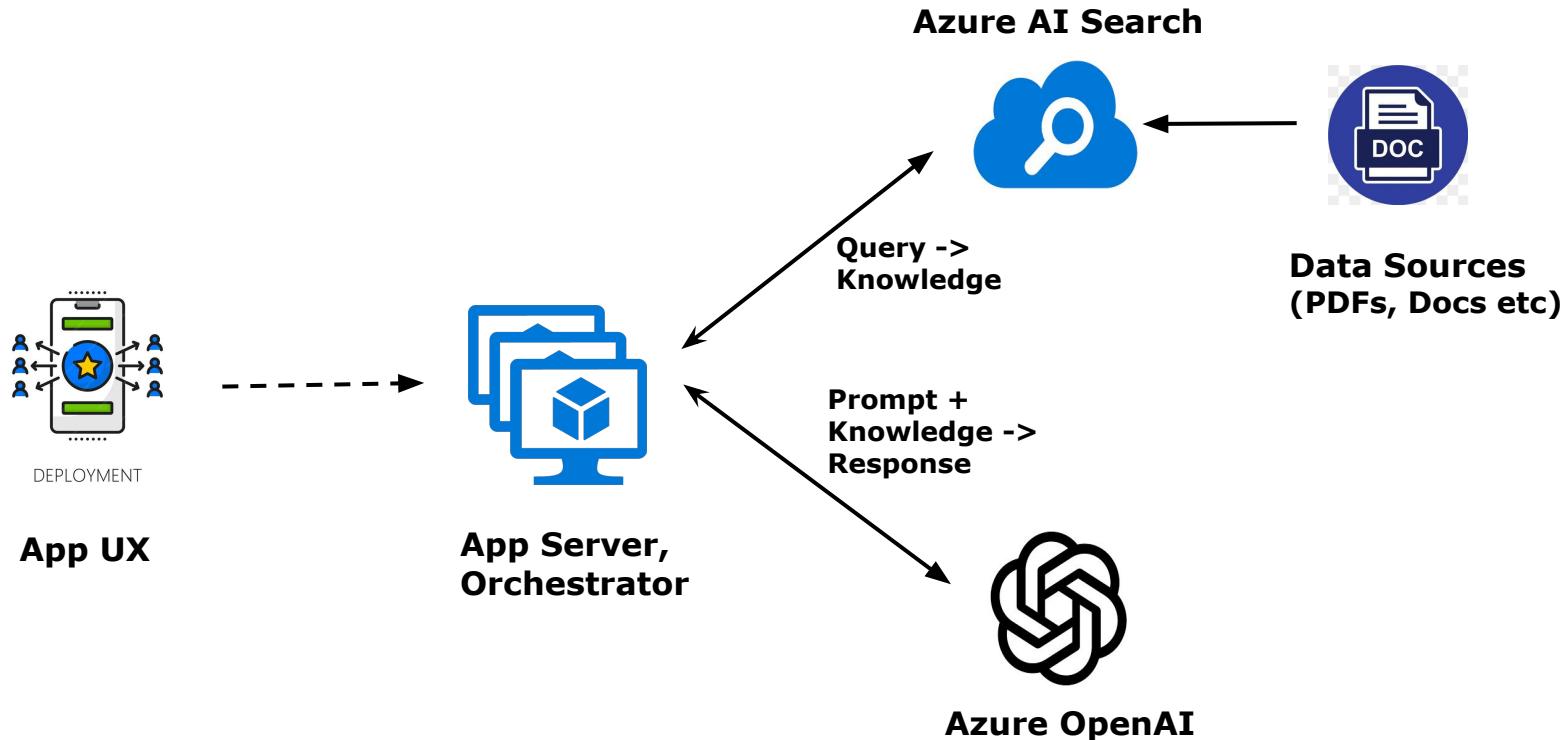
Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Architecture of Web App Deployment

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

Proprietary Content © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Architecture of Web App Deployment



# Architecture of Web App Deployment

- 1. User Query Submission**
  - Users enter queries through the web app interface (App UX).
- 2. Query Routing**
  - The App Server & Orchestrator receive and direct the query to Azure AI Search.
- 3. Knowledge Retrieval**
  - Azure AI Search processes the query, searches through data sources (e.g., PDFs, documents), and retrieves relevant information.
- 4. Prompt Creation & Response Generation**
  - The retrieved knowledge is combined with the original query to create a prompt. Azure OpenAI processes this prompt to generate a detailed response.
- 5. Response Delivery**
  - The response is sent back through the App Server & Orchestrator, then delivered to the user via the App UX.

# Deploy your model to a web app

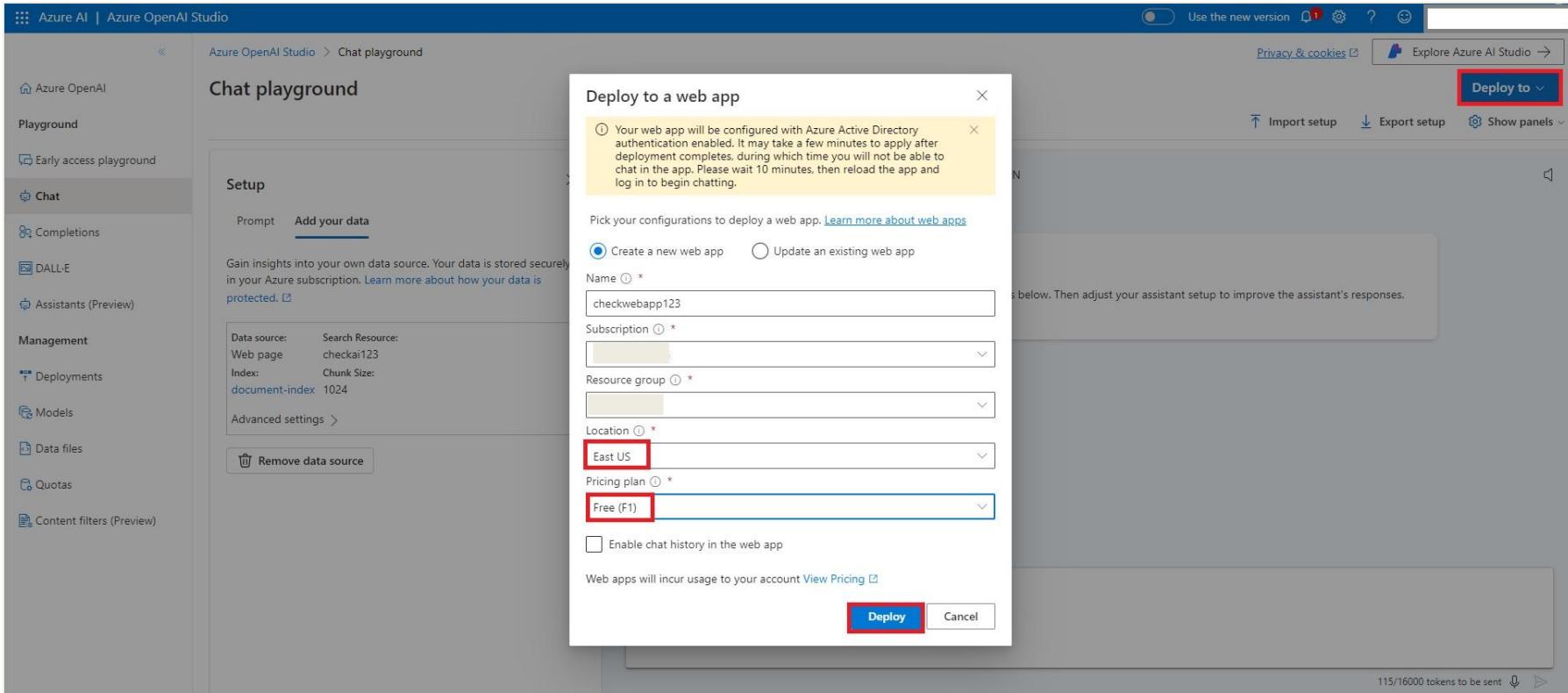
This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

Proprietary Content © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Deploy your model to a WebApp

1. At the top right of the Chat playground page, in the **Deploy to menu**, select A new web app.
2. In the Deploy to a web app dialog box, create a new web app with the following settings:
  - Name: *A unique name*
  - Subscription: *Your Azure subscription*
  - Resource group: *The resource group in which you provisioned your Azure OpenAI resource*
  - Locations: *The region where you provisioned your Azure OpenAI resource*
  - Pricing plan: **Free (F1)**
  - Enable chat history in the web app: Unselected
  - I acknowledge that web apps will incur usage to my account: Selected
3. Deploy the new web app and wait for deployment to complete (which may take 15 minutes or so)
4. After your web app has deployed successfully, use the button at the top right of the Chat playground page to launch the web app. The app may take a few minutes to launch. If prompted, accept the permissions request.

# Deploy your model to a WebApp



The screenshot shows the Azure OpenAI Studio interface with the 'Chat playground' selected in the sidebar. A modal window titled 'Deploy to a web app' is open, prompting the user to configure a new web app. The configuration fields include:

- Name:** checkwebapp123
- Subscription:** (dropdown menu)
- Resource group:** (dropdown menu)
- Location:** East US
- Pricing plan:** Free (F1)
- Enable chat history in the web app:** (checkbox)

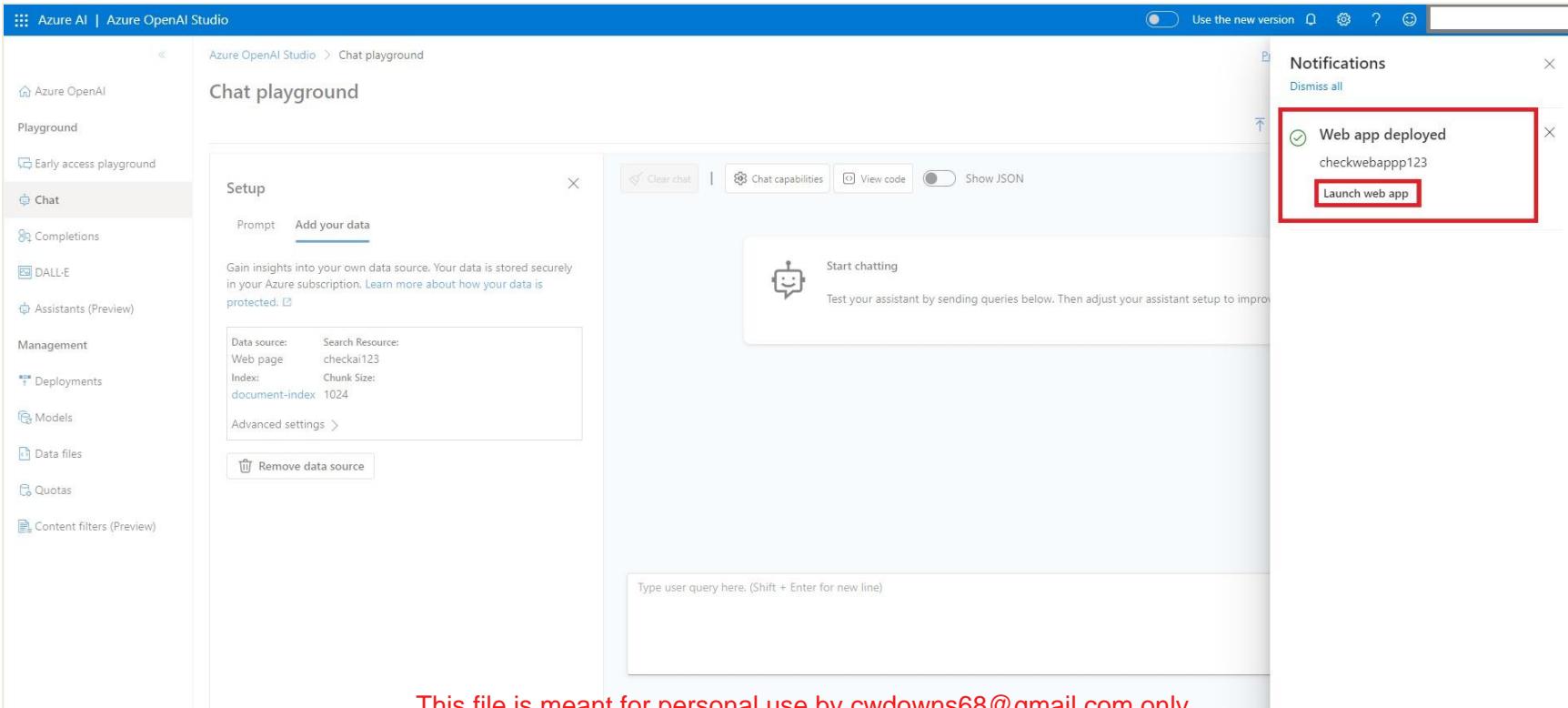
A yellow informational box at the top of the modal states: "Your web app will be configured with Azure Active Directory authentication enabled. It may take a few minutes to apply after deployment completes, during which time you will not be able to chat in the app. Please wait 10 minutes, then reload the app and log in to begin chatting."

This file is meant for personal use by cwdowns68@gmail.com only.  
 Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Deploy your model to a Web App

Click on “Launch web app”. Wait for 15-20 minutes to launch it.



The screenshot shows the Azure OpenAI Studio interface with the "Chat playground" selected in the sidebar. A notification window is open, indicating that the "Web app deployed" and providing a link to "Launch web app".

**Notifications**

- Web app deployed  
checkwebapp123  
[Launch web app](#)

This file is meant for personal use by cwdowns68@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Test the Web App

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.  
Proprietary Content © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Test the Web App

Now, it's time to test the web app's response by interacting with it.

The screenshot shows a web browser window with the URL <https://checkapp123.azurewebsites.net> in the address bar. The page content is a conversational AI interface. The user has asked three questions, and the AI has responded with three messages. Each message includes an intent code highlighted with a red box:

- User message: "I have been trying unsuccessfully to top-up my card. Is the system down or not working properly?"
- AI response: "Intent Code: top\_up\_failed:59  
AI-generated content may be incorrect"
- User message: "How do you top up with a cheque?"
- AI response: "Intent Code: top\_up\_by\_cash\_or\_cheque:58  
AI-generated content may be incorrect"
- User message: "What is the charge for topping up a card?"
- AI response: "Intent Code: top\_up\_by\_card\_charge:57  
AI-generated content may be incorrect"

At the bottom, there is a text input field with placeholder text "Type a new question..." and a blue send button icon.

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Understanding the Responses generated

The outputs you're getting indicate how well your model is classifying the customer intents based on the provided examples. Here's a brief overview of the results:

## 1. Top-Up Failure

- **Question:** "I have been trying unsuccessfully to top-up my card. Is the system down or not working properly?"
- **Expected Output:** **59 (top\_up\_failed)**
- **Explanation:** This question addresses issues related to failed top-up attempts. The model should recognize this intent and return the code **59**.

## 2. Top-Up by Cheque

- **Question:** "How do you top up with a cheque?"
- **Expected Output:** **58 (top\_up\_by\_cash\_or\_cheque)**
- **Explanation:** This question is about the process of topping up with a cheque. The model should classify this under the intent for cash or cheque top-ups, returning **58**.

## 3. Top-Up Card Charge

- **Question:** "What is the charge for topping up a card?"
- **Expected Output:** **57 (top\_up\_by\_card\_charge)**
- **Explanation:** This question inquires about fees associated with card top-ups. The model should classify this intent and return **57**.

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.

# Summary

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.  
Proprietary Content © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Summary

1. The model's outputs reveal its proficiency in categorizing customer queries related to card top-ups.
2. It effectively identifies issues with unsuccessful top-ups, questions about using cheques, and inquiries about card charges.
3. This accurate classification enables the business to promptly address customer concerns, streamline support processes, and provide precise information on transaction issues and methods.
4. By ensuring that customer queries are classified correctly, the model enhances user satisfaction and operational efficiency, allowing the business to respond more effectively to diverse customer needs.

# Clean Up

This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.  
Proprietary Content © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Clean Up

When you're done with your Azure OpenAI resource, remember to delete the deployment or the entire resource in the Azure portal at <https://portal.azure.com>.



# Happy Learning !



This file is meant for personal use by cwdowns68@gmail.com only.  
Sharing or publishing the contents in part or full is liable for legal action.