

# Deploy your Model to a WebApp on Azure

This file is meant for personal use by [cwdownloads68@gmail.com](mailto:cwdownloads68@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.**

This file is meant for personal use by cwwdowns68@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.

## 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

This file is meant for personal use by cwwdowns68@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.

# Provision an Azure OpenAI Resource

This file is meant for personal use by cwdownloads68@gmail.com only.

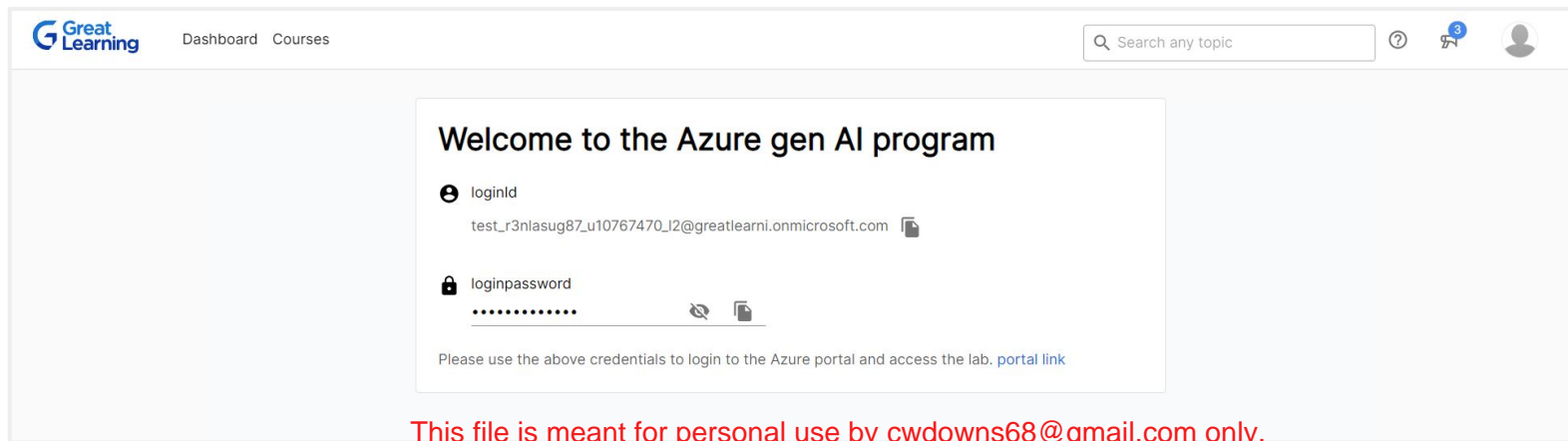
Sharing or publishing the contents in part or full is liable for legal action.

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.



Great Learning Dashboard Courses

Search any topic

### Welcome to the Azure gen AI program

loginId  
test\_r3nlasug87\_u10767470\_l2@greatlearni.onmicrosoft.com

loginpassword  
.....

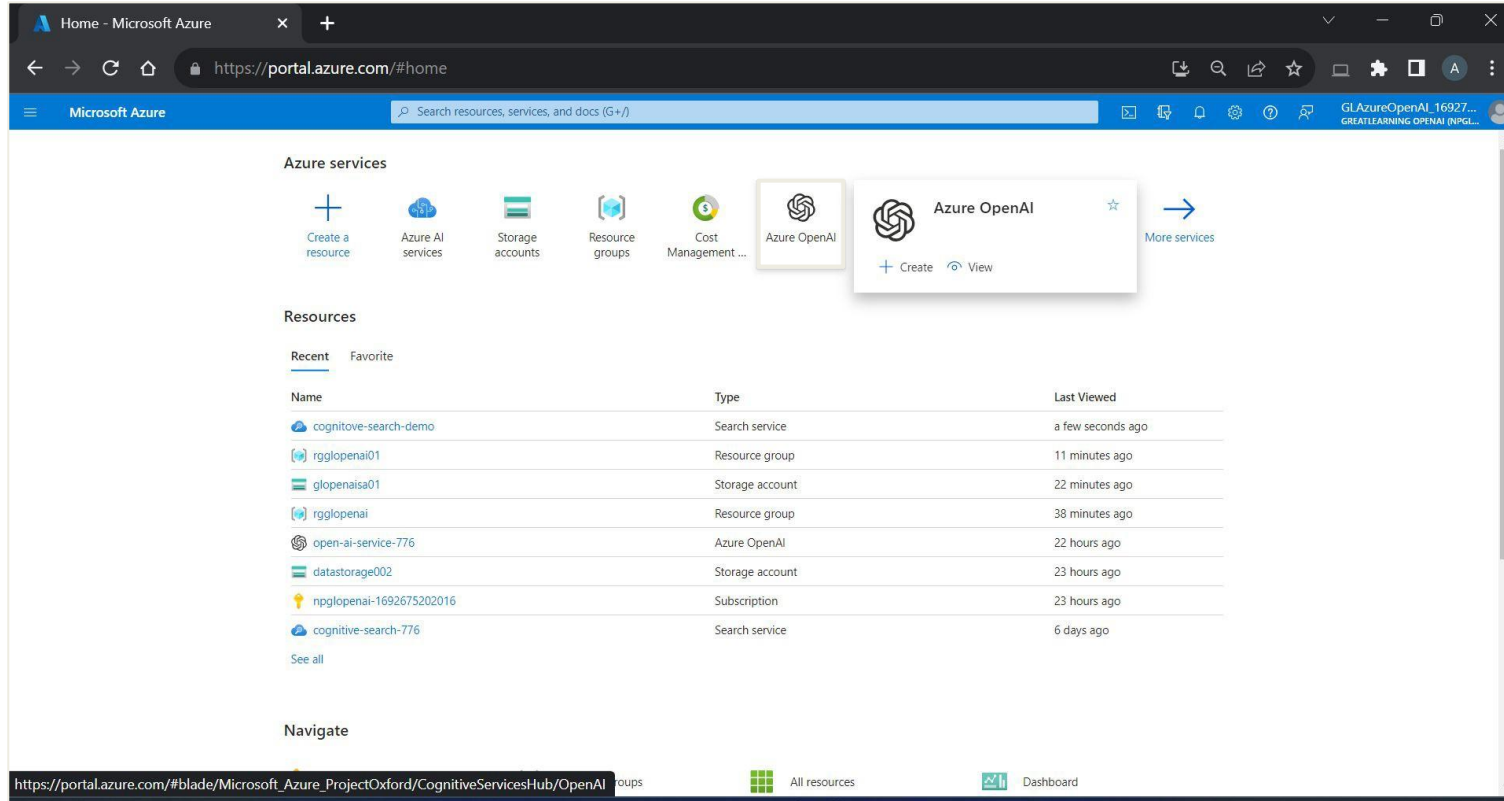
Please use the above credentials to login to the Azure portal and access the lab. [portal link](#)

This file is meant for personal use by cwwdowns68@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.

# Click on Azure OpenAI service



The screenshot shows the Microsoft Azure portal interface. At the top, the browser address bar displays 'https://portal.azure.com/#home'. The Microsoft Azure logo is in the top left, and a search bar is in the top center. The main content area is divided into sections: 'Azure services' at the top, followed by 'Resources' with 'Recent' and 'Favorite' tabs, and 'Navigate' at the bottom. In the 'Azure services' section, several tiles are visible: 'Create a resource', 'Azure AI services', 'Storage accounts', 'Resource groups', 'Cost Management...', 'Azure OpenAI', and 'More services'. The 'Azure OpenAI' tile is highlighted with a yellow border, and a tooltip is displayed over it. The tooltip contains the Azure OpenAI logo, the text 'Azure OpenAI', a star icon, and two buttons: '+ Create' and 'View'. The 'Resources' section shows a table of recent resources.




Name	Type	Last Viewed
cognitove-search-demo	Search service	a few seconds ago
rgglopenai01	Resource group	11 minutes ago
glopenaisa01	Storage account	22 minutes ago
rgglopenai	Resource group	38 minutes ago
open-ai-service-776	Azure OpenAI	22 hours ago
datastorage002	Storage account	23 hours ago
npglopenai-1692675202016	Subscription	23 hours ago
cognitive-search-776	Search service	6 days ago

This file is meant for personal use by cwwdowns68@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 \* ⓘ  
Resource group \* ⓘ

GlazureaiuserFeb28

default\_resource\_group

[Create new](#)

**Instance Details**

Region ⓘ  
Name \* ⓘ  
Pricing tier \* ⓘ

East US


TestOpenAI101

Standard S0






[View full pricing details](#)

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

 **2 Network**

 3 Tags

 4 Review + submit

 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



[Home](#) > [Azure AI services](#) | [Azure OpenAI](#) >

## Create Azure OpenAI



✓ Basics   ✓ Network   **3** Tags   4 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

:

Azure AI services

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

# Click on Create

 Microsoft Azure

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

Home >


## Create Azure OpenAI ...

✓ Basics

✓ Network

✓ Tags

4 Review + submit

 [View automation template](#)

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**

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

**Network**

Previous

Next

Create

# Deployment will take few mins

The screenshot shows the Microsoft Azure portal interface. At the top, there's a blue header with the Microsoft Azure logo and a search bar. Below the header, the main content area displays the deployment details for 'Microsoft.CognitiveServicesOpenAI-20240322163603'. A red box highlights the 'Your deployment is complete' message, which includes the deployment name, subscription, resource group, start time, and correlation ID. Another red box highlights the 'Go to resource' button under the 'Next steps' section. The left sidebar shows navigation options like Overview, Inputs, Outputs, and Template. The right sidebar contains links for Cost management, Microsoft Defender for Cloud, Free Microsoft tutorials, and Work with an expert.

Microsoft Azure

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

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 Start time : 22/03/2024, 16:38:47

Subscription : [GlazureaiuserFeb28](#) Correlation ID : 028e6f70-f34b-438e-9b2c-6ec87917cba3

Resource group : [default\\_resource\\_group](#)

> 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

# Click on Explore once the deployment is complete

The screenshot shows the Microsoft Azure portal interface. At the top, the header includes the Microsoft Azure logo, a search bar, and user information. The main content area is titled 'TestOpenAI101' and is highlighted with a red box. Below the title, there are tabs for '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', and 'Metrics'. The 'Overview' tab is selected, showing details about the resource group, status, location, subscription, and tags. A 'Get Started' section is visible, containing three cards: 'Monitor your Azure OpenAI usage', 'Develop', and 'Explore and deploy'. The 'Explore and deploy' card is highlighted with a red box and contains the text 'Explore and deploy the generative AI models, craft unique prompts for your use cases, and fine-tune select models.' and a button 'Go to Azure OpenAI Studio'.

# Deploy a model to use Chat Playground

This file is meant for personal use by cwwdowns68@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 to deploy any model.

This file is meant for personal use by cwwdowns68@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. On the left, the 'Deployments' tab is selected in the sidebar. The main panel displays the 'Deployments' page with a 'Create new deployment' button. A modal window is open for creating a new deployment. The modal contains the following fields and options:

- Deployment name:** A text input field containing 'checktoday123'.
- Select a model:** A dropdown menu with 'gpt-35-turbo-16k' selected.
- Model version:** A dropdown menu with '0613 (Default)' selected.
- Deployment type:** A dropdown menu with 'Standard' selected.
- Current resource:** A dropdown menu with 'eastus' selected.
- Tokens per Minute Rate Limit (thousands):** A slider control set to '5K'.
- Content Filter:** A dropdown menu with 'Default' selected.
- Enable Dynamic Quota:** A toggle switch that is currently 'Enabled'.

At the bottom of the modal, there are 'Create' and 'Cancel' buttons. The background shows the 'Deployments' page with a list of existing deployments and a search bar.

This file is meant for personal use by cwwdowns68@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

Privacy & cookies

Explore Azure AI Studio →

Azure OpenAI

Playground

Early access playground

Chat

Completions

DALL·E

Assistants (Preview)

Management

Deployments

Models

Data files

Quotas

Content filters (Preview)

Deployments

Deployments 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

Search

Deployment name	Model name	Model v...	Deploye...	Capacity	Status	Model reti...	Content Fil...	Rate limit (...)	Updated at
✓ checktoday123	gpt-35-turbo-16k	0613	Standard	5K TPM	✓ Succeeded	10/1/2024	Microsoft.Default	5000	8/1/2024 7:29 PM

This file is meant for personal use by cwwdowns68@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

This file is meant for personal use by [cwdownloads68@gmail.com](mailto:cwdownloads68@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:**

# 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”**.

The screenshot shows the Azure OpenAI Studio interface. The top navigation bar includes 'Azure AI | Azure OpenAI Studio' and a 'Use the new version' toggle. The left sidebar contains a 'Chat playground' breadcrumb and a list of options: 'Azure OpenAI', 'Playground', 'Early access playground', 'Chat' (highlighted with a red box), 'Completions', 'DALL·E', 'Assistants (Preview)', 'Management', 'Deployments', 'Models', 'Data files', 'Quotas', and 'Content filters (Preview)'. The main area is titled 'Chat playground' and features a 'Setup' panel on the left, a central chat area, and a 'Configuration' panel on the right. The 'Setup' panel includes a 'Prompt' section with an 'Apply changes' button (highlighted with a red box), a 'System message' section with a text area (highlighted with a red box), and an 'Examples' section. The central chat area displays a 'Start chatting' message and a text input field. The 'Configuration' panel shows a 'Deployment' dropdown menu (highlighted with a red box) and 'Session settings'.

This file is meant for personal use by cwwdowns68@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

This file is meant for personal use by cwdownloads68@gmail.com only.

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

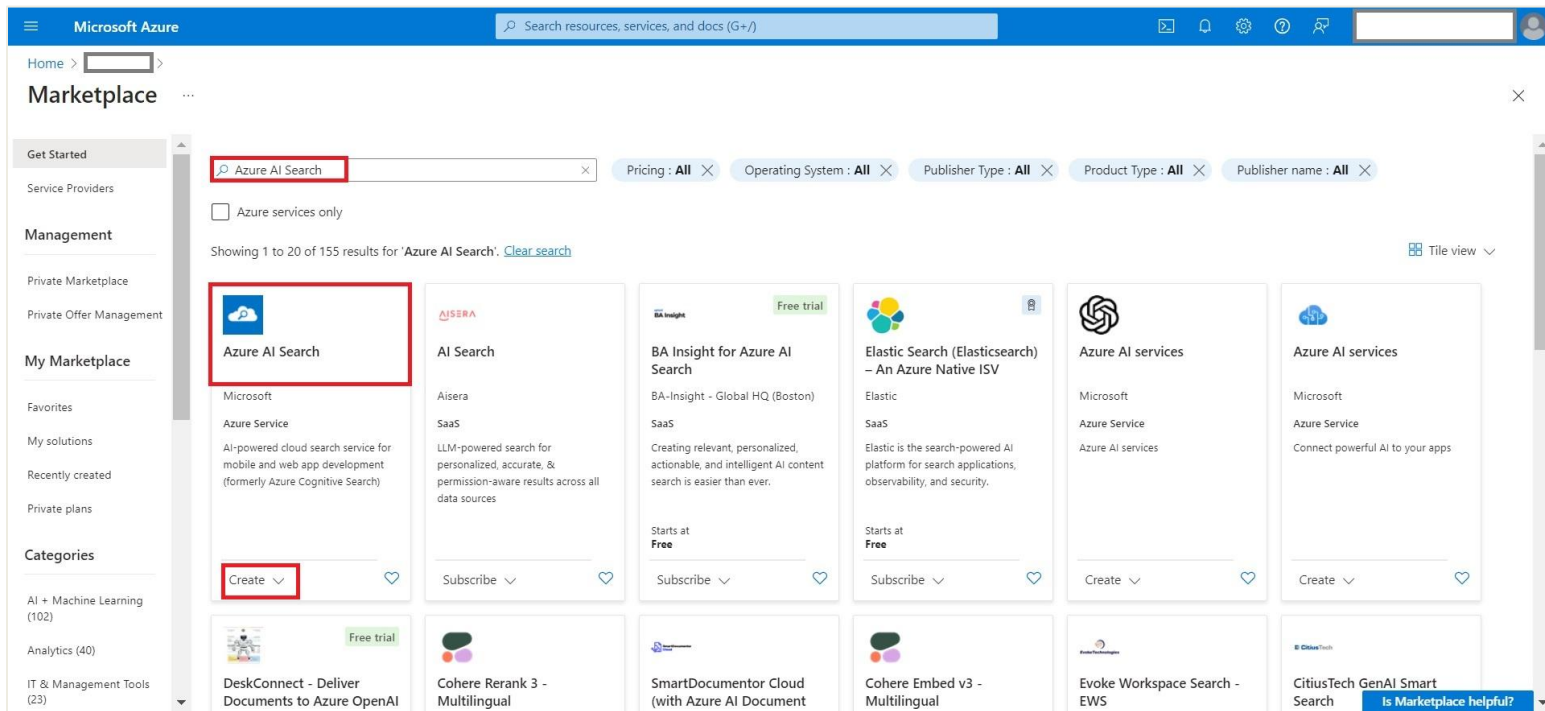
Proprietary content © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# 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.



This file is meant for personal use by cwwdowns68@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

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

Microsoft Azure

Search resources, services, and docs (G+)

Home > > Marketplace >

### Create a search service

Basics Scale Networking Tags Review + create

**Project details**

Subscription \*

Resource Group \*  [Create new](#)

**Instance Details**

Service name \*

Location \*

Pricing tier \*   
15 GB/Partition, max 3 replicas, max 3 partitions, max 9 search units  
[Change Pricing Tier](#)

[Review + create](#) [Previous](#) [Next: Scale](#)

This file is meant for personal use by cwdownloads68@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

Click on “**Create**” button.

Microsoft Azure

Search resources, services, and docs (G+)

Home > Marketplace >

## Create a search service

Validation Success

Basics Scale Networking Tags Review + create

**Basics**

Subscription	
Resource Group	
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
------------------------------	--------

**Create** Previous Next Download a template for automation

This file is meant for personal use by cwdownloads68@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 displays the Microsoft Azure portal interface. At the top, the navigation bar shows 'Microsoft Azure' and a search bar. Below this, the breadcrumb trail indicates 'Home > search-service-checkai123 | Overview'. The main content area is titled 'Deployment' and features a search bar and action buttons: 'Delete', 'Cancel', 'Redeploy', 'Download', and 'Refresh'. A red box highlights a success message: 'Your deployment is complete'. Below this message, deployment details are listed: 'Deployment name : search-service-checkai123', 'Subscription : [redacted]', 'Resource group : [redacted]', 'Start time : 01/08/2024, 20:20:31', and 'Correlation ID : 1364b034-e244-457f-8abc-83dc15edb59d'. A 'Go to resource' button is visible. The right sidebar contains sections for 'Cost management', 'Microsoft Defender for Cloud', 'Free Microsoft tutorials', and 'Work with an expert'.

Microsoft Azure

search-service-checkai123 | Overview

Deployment

Search

Overview

Inputs

Outputs

Template

✓ 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 >

# Create Azure AI search service

You will be able to view the search service created.

The screenshot displays the Microsoft Azure portal interface for a search service named 'checkai123'. The service is highlighted with a red box in the top left. The left sidebar contains a navigation menu with options like Activity log, Access control (IAM), Tags, and Search management. The main content area shows the 'Overview' tab for the search service, which includes a search bar and a list of actions: Add index, Import data, Import and vectorize data, Search explorer, Refresh, Delete, and Move. Below this, the 'Essentials' section provides key details about the service, including its resource group, location, subscription, and status. The 'Pricing tier' is set to 'Basic', which is also highlighted with a red box. The 'Search units' are set to 1. The 'Url' is https://checkai123.search.windows.net. The 'Replicas' are set to 1 (No SLA). The 'Partitions' are set to 1. The 'Search units' are set to 1. The 'Tags' section shows 'Add tags'. The 'Get started' tab is selected, and the 'Properties' sub-tab is active.

Essentials	
Resource group ( <a href="#">move</a> )	: [REDACTED]
Location ( <a href="#">move</a> )	: East US
Subscription ( <a href="#">move</a> )	: [REDACTED]
Subscription ID	: 6cc13ccc-43a9-461d-abcc-6e482c3e8031
Status	: Running
Tags ( <a href="#">edit</a> )	: <a href="#">Add tags</a>

Properties	
Url	: https://checkai123.search.windows.net
Pricing tier	: <b>Basic</b>
Replicas	: 1 (No SLA)
Partitions	: 1
Search units	: 1

# Case Study

This file is meant for personal use by cwwdowns68@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.

**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 cwwdowns68@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

Use the new version

Privacy & cookies

Explore Azure AI Studio

Deploy to

Import setup

Export setup

Show panels

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.

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

Configuration

Deployment

Parameters

Deployment \*

checktoday123

Session settings

Past messages included

10

Current token count

Input tokens progress indicator

115/16000

This file is meant for personal use by cwwdowns68@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

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



# Add your Data

**Add data**

**Data source**

Select or add data source

Your data source is used to ground the generated results with your data. Select an existing data source or create a new data connection with Azure Blob Storage, databases, search, URLs, or local files as the source the grounding data will be built from. [Learn more about data privacy and security in Azure AI.](#)

Select data source \*

URL/web address (preview)

Resource

Select Azure AI Search resource \*

checkai123

Create a new Azure AI Search resource

Select Azure Blob storage resource \*

Index name

Enter the index name \*

document-index

Data source location

Enter a url/web address \*

https://huggingface.co/datasets/P...

Using Azure AI Search will incur usage to your account. [View Pricing](#)

☐ Add vector search to this search resource.

**Next** Cancel

# Add your Data

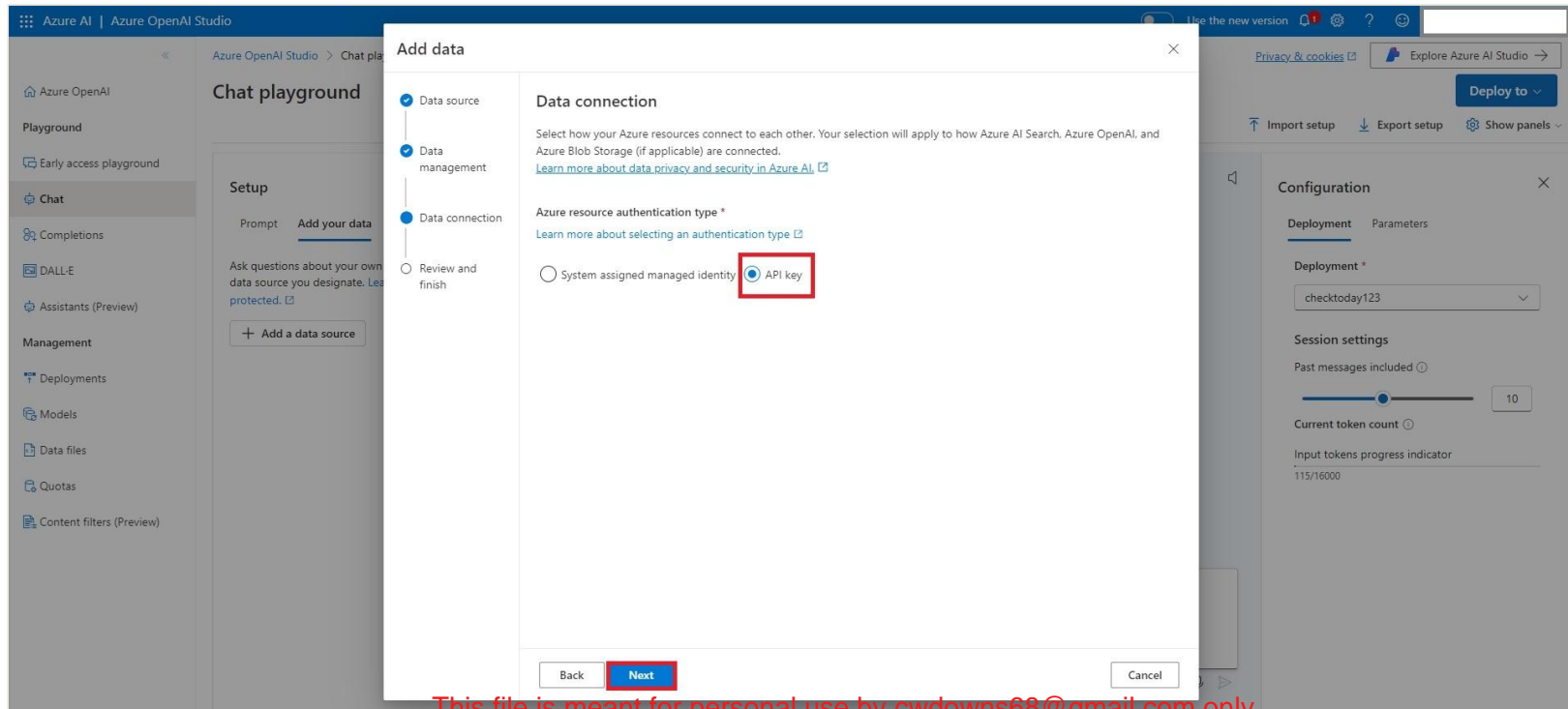
Select “**Keyword**” in Search type and **1024** as Chunk size.  
Click on **Next**

The screenshot shows the 'Add data' dialog in the Azure OpenAI Studio interface. The dialog has a sidebar with four steps: 'Data source' (selected), 'Data management', 'Data connection', and 'Review and finish'. The 'Data management' section is active, showing instructions to set up configurations for data and how the model will respond to requests. A link to 'Learn more about data privacy and security in Azure AI' is provided. The 'Search type' dropdown menu is set to 'Keyword'. Below this, the 'Chunk Size' section explains that chunking is the process of breaking down documents into smaller segments for search and retrieval, and that the chunk size is measured in tokens. It also provides a link to 'Learn more about selecting a chunk size'. The 'Select a size' section has three radio buttons: '256', '512', and '1024 (default)' (which is selected). The '1536' option is also visible. At the bottom of the dialog, there are three buttons: 'Back', 'Next' (highlighted in red), and 'Cancel'. The background shows the 'Chat playground' setup page with a 'Deploy to' button and a 'Configuration' panel on the right.

This file is meant for personal use by cwdownloads68@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 **"API Key"** in resource authentication type  
Click on **Next**



This file is meant for personal use by cwwdowns68@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

Review the settings.

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

**Add data**

**Review and finish**

Review the configurations you set for your data  
[Learn more about data privacy and security in Azure AI](#)

**Azure AI Search resource**  
checkai123

**Data source**  
Web page

**Storage resource**  
checktoday123

**Index name**  
document-index

**Web page**  
huggingface.co/datasets/PolyAI/banking77

**Search type**  
Keyword

**Azure resource authentication type**  
API key

**Semantic search configuration**  
default

**Buttons:** Back, Save and close, Cancel

This file is meant for personal use by cwwdowns68@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 displays the Azure OpenAI Studio interface. On the left, a sidebar contains navigation links: Azure OpenAI, Playground, Early access playground, Chat (highlighted), Completions, DALL-E, Assistants (Preview), Management, Deployments, Models, Data files, Quotas, and Content filters (Preview). The main area is titled 'Chat playground' and features a 'Setup' panel on the left with a red border. This panel has a 'Prompt' tab and an 'Add your data' sub-tab. It contains the text: 'Gain insights into your own data source. Your data is stored securely in your Azure subscription. Learn more about how your data is protected.' Below this, a table lists the data source details: 'Data source: Search Resource: checkai123', 'Web page: document-index', and 'Index: Chunk Size: 1024'. An 'Advanced settings >' link is also present. At the bottom of the panel is a 'Remove data source' button. To the right of the setup panel are controls for 'Clear chat', 'Chat capabilities', 'View code', and a 'Show JSON' toggle. The main chat area contains a 'Start chatting' prompt with a robot icon and the instruction: 'Test your assistant by sending queries below. Then adjust your assistant setup to improve the assistant's responses.' At the bottom is a text input field with the placeholder 'Type user query here. (Shift + Enter for new line)'. The top of the interface shows the 'Azure AI | Azure OpenAI Studio' header, a 'Use the new version' toggle, and various utility icons. The bottom right corner indicates '115/16000 tokens to be sent'.

Data source:	Search Resource:
Web page	checkai123
Index:	Chunk Size:
document-index	1024

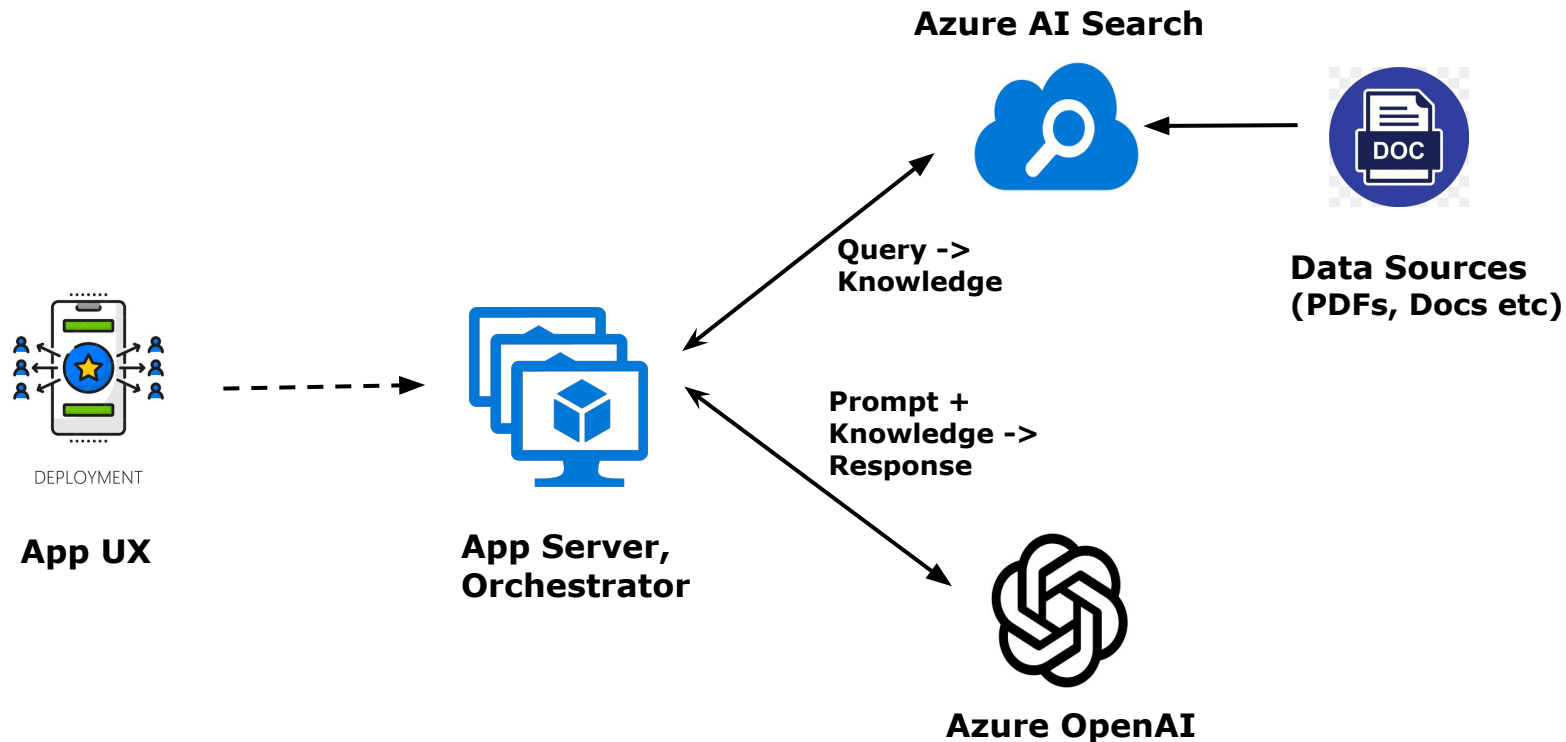
# Architecture of Web App Deployment

This file is meant for personal use by [cwdownloads68@gmail.com](mailto:cwdownloads68@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 cwwdowns68@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. On the left is a sidebar with navigation options: Azure OpenAI, Playground, Early access playground, Chat, Completions, DALL-E, Assistants (Preview), Management, Deployments, Models, Data files, Quotas, and Content filters (Preview). The main area is titled 'Chat playground' and has a 'Setup' section with a 'Prompt' tab and an 'Add your data' tab. Below this, there's a table for 'Data source' and 'Search Resource' with fields for 'Web page', 'Index', and 'Chunk Size'. A 'Remove data source' button is at the bottom of this section. On the right side of the main area, there's a 'Deploy to' button highlighted with a red box. A modal dialog box titled 'Deploy to a web app' is open in the center. It contains a yellow warning box at the top stating: '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.' Below the warning, there's a section 'Pick your configurations to deploy a web app. [Learn more about web apps](#)'. There are two radio buttons: 'Create a new web app' (selected) and 'Update an existing web app'. Below these are several dropdown menus: 'Name' (checkwebapp123), 'Subscription' (a blurred selection), 'Resource group' (a blurred selection), 'Location' (East US, highlighted with a red box), and 'Pricing plan' (Free (F1), highlighted with a red box). There's also a checkbox for 'Enable chat history in the web app' which is unchecked. At the bottom of the dialog, there's a 'Deploy' button highlighted with a red box and a 'Cancel' button. A footer note says 'Web apps will incur usage to your account [View Pricing](#)'. The background interface shows a 'Chat' section with a 'Prompt' input and a 'Completions' section with a 'Generate' button. The top right of the interface has a 'Use the new version' toggle and a 'Privacy & cookies' link.

This file is meant for personal use by cwwdowns68@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 displays the Azure OpenAI Studio interface. On the left is a sidebar with navigation options: Azure OpenAI, Playground, Early access playground, Chat (selected), Completions, DALL-E, Assistants (Preview), Management, Deployments, Models, Data files, Quotas, and Content filters (Preview). The main area is titled 'Chat playground' and contains a 'Setup' panel on the left and a chat area on the right. The 'Setup' panel has tabs for 'Prompt' and 'Add your data'. Under 'Add your data', it shows 'Data source: Search Resource' and 'Web page: checkai123'. Below this, it lists 'Index: document-index' and 'Chunk Size: 1024'. There is an 'Advanced settings >' link and a 'Remove data source' button. The chat area has buttons for 'Clear chat', 'Chat capabilities', 'View code', and 'Show JSON'. It also features a 'Start chatting' button and a text input field with the placeholder 'Type user query here. (Shift + Enter for new line)'. On the right side of the interface, a 'Notifications' panel is open, showing a message: 'Web app deployed' with a green checkmark icon, followed by 'checkwebapp123' and a red-bordered button labeled 'Launch web app'.

This file is meant for personal use by cwwdowns68@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

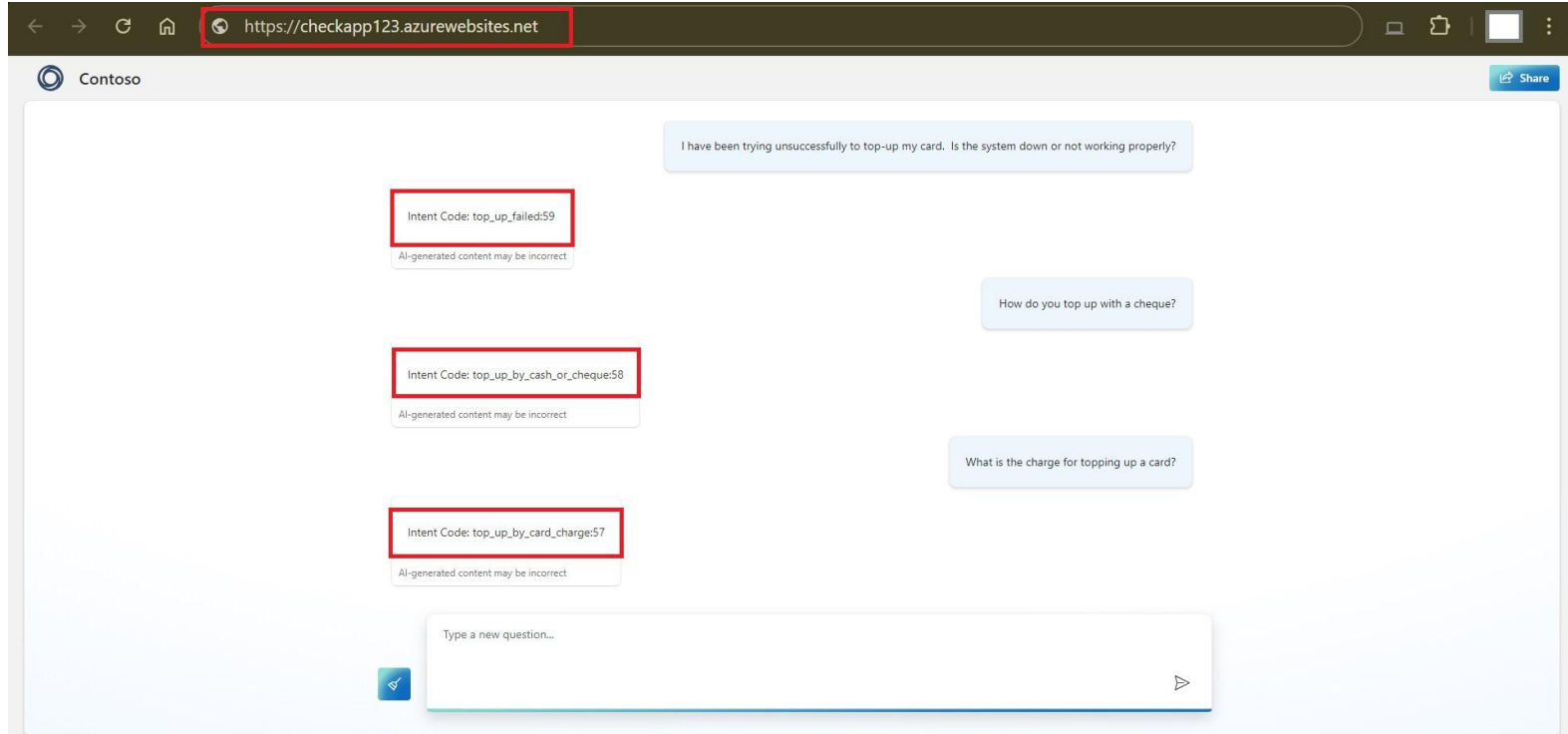
This file is meant for personal use by [cwdownloads68@gmail.com](mailto:cwdownloads68@gmail.com) only.

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.



# 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.**

# Summary

This file is meant for personal use by cwwdowns68@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 cwwdowns68@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 !

