

User Guide Azure Open AI Studio

Introduction

Data: Please refer to the data folder in the GitHub repository shared with you during the event. Please refer to the Readme file in the GitHub repository for the data description and the problem statement

Overview: We'll build a chatbot powered by Azure AI services such as Azure AI Search, Azure Open AI, and Azure Open AI Studio. This chatbot will be capable of ingesting any document, responding to questions based on only the information available in the documents, and handling multi-turn conversations. Assuming this is our first AI chatbot project, we will detail the steps. We'll use Azure AI Studio to develop this solution and deploy a sample interactive web application.

If time permits, we'll also investigate how to change prompts and parameters like top_k and their impact on responses. We will also try to give you a glimpse at other developer features, such as prompt flow and evaluation flow.

Note: In this step-by-step guide, we have used the resource configuration in table-1. You will receive your temporary values at the venue for this hackathon including a username.

Please follow the naming convention in Table 2 to create the names of the remaining Azure resources.

S No	Configuration	Name used in the documentation	Description
1	User name	Thailand01	Azure User name provided
2	Azure subscription	Tiger analytics-01	Name of Azure Subscription
3	Location	Australia East	Location to create Azure Resources
4	Azure OpenAI	openai-tigeranalytics-01	Name of Azure OpenAI service

Table - 1

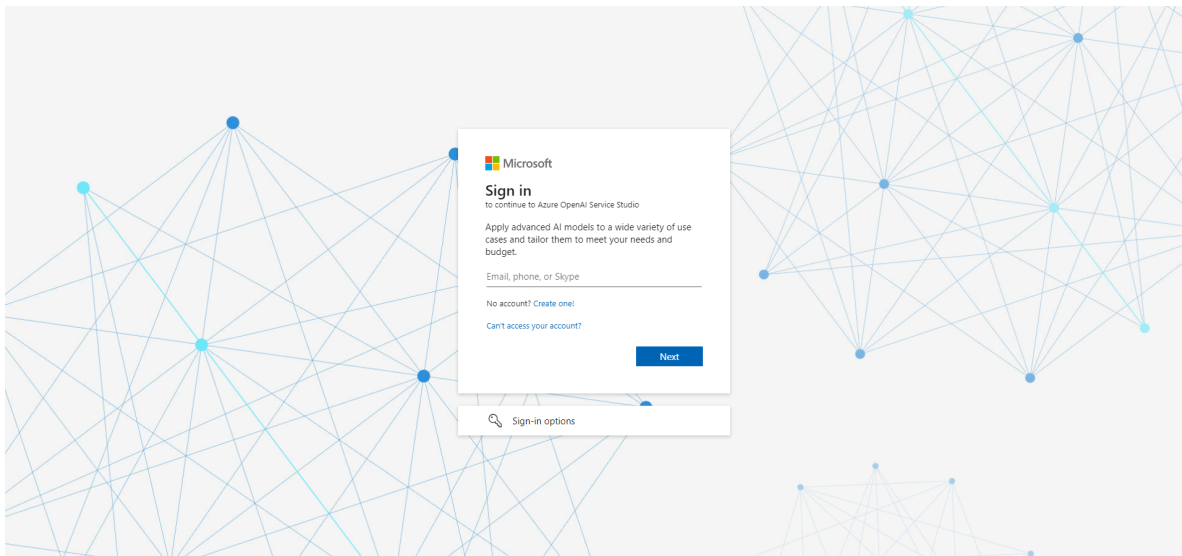
S No	Configuration	Naming Convention	Name used in documentation
1	Resource group name	rg-<User name>ai	rg-Thailand01ai
2	Azure AI search	<User name>aisearch	Thailand01aisearch
3	Connection storage name	<User name>docs	Thailand01docs

4	Index name	ai-build-<User name>-index-<version>	ai-build-Thailand01-index-v8
5	Web app name	<User name>-web app	Thailand01-web app

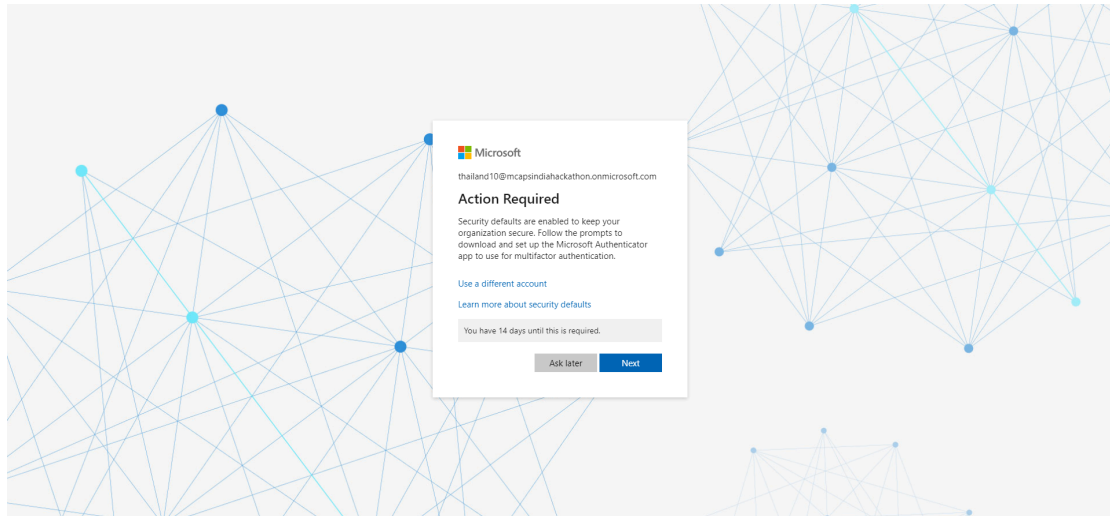
Table 2

1. Log in to the Azure Open AI Studio

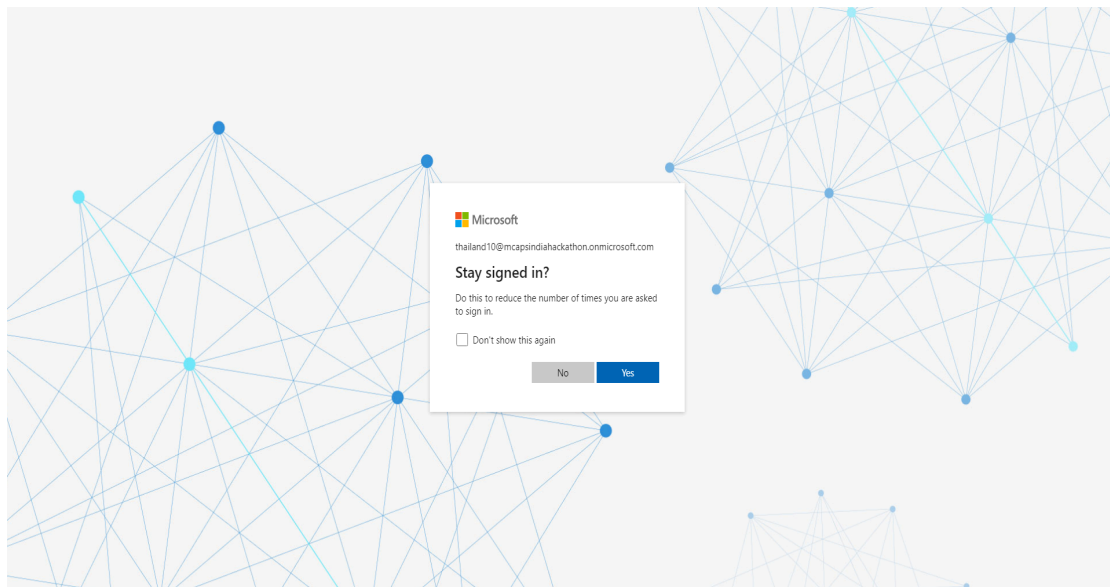
1. Please go to [AOAI Studio](#) from your browser (please use Incognito/ private mode to avoid clashes with your existing work profile)
2. Click on the Sign-in Button and enter the credentials shared with you at the registration desk.
(Please keep your slip with you till the end of the event)



3. After successful login, **if you are prompted to change the password**, please do so.
4. Setup MFA/authenticator: Skip this step by clicking **“Ask Later.”** (If **authenticator is a mandatory** requirement, please choose **“other method”** for login and provide your phone number, you can sign-in through OTP)

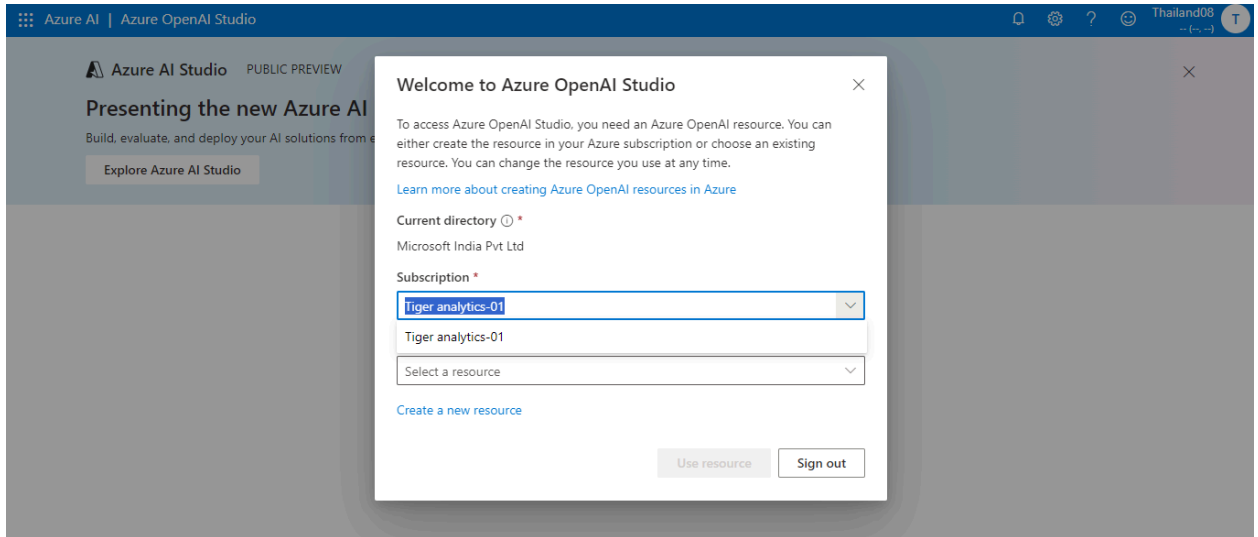


5. Click on **Yes** in the next step.

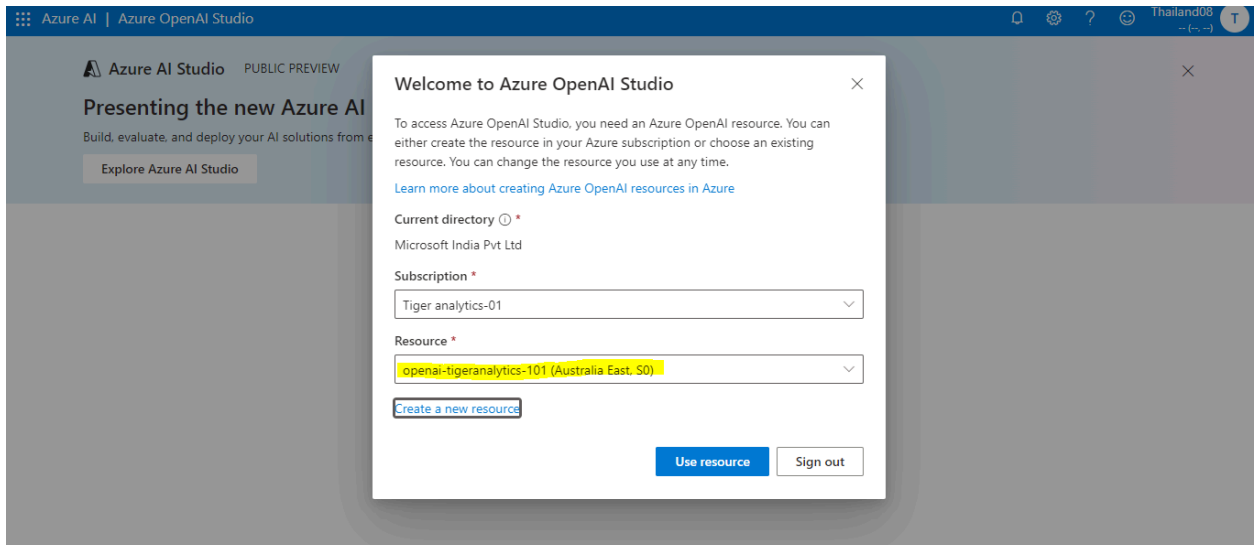


2. Welcome Page

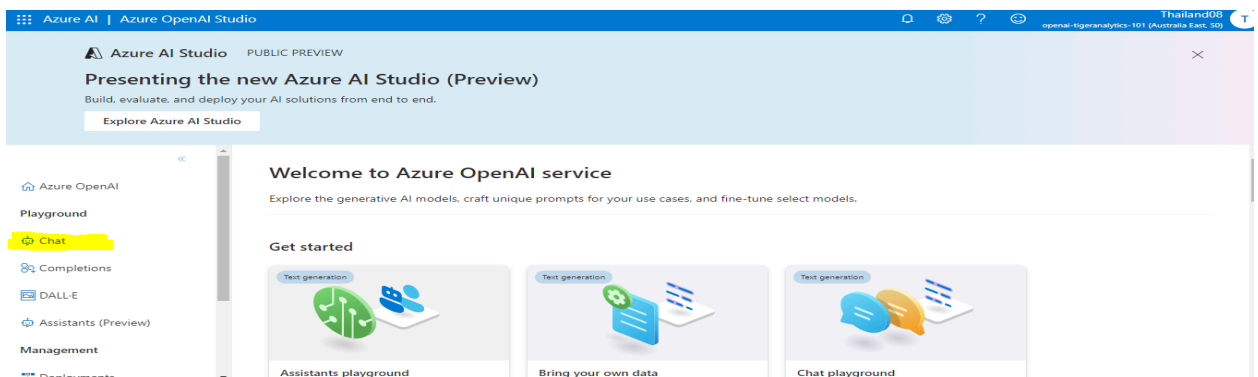
1. Under **Subscription**, select the default option. It will look like **Tiger analytics-<number>**, e.g. Tiger analytics-01



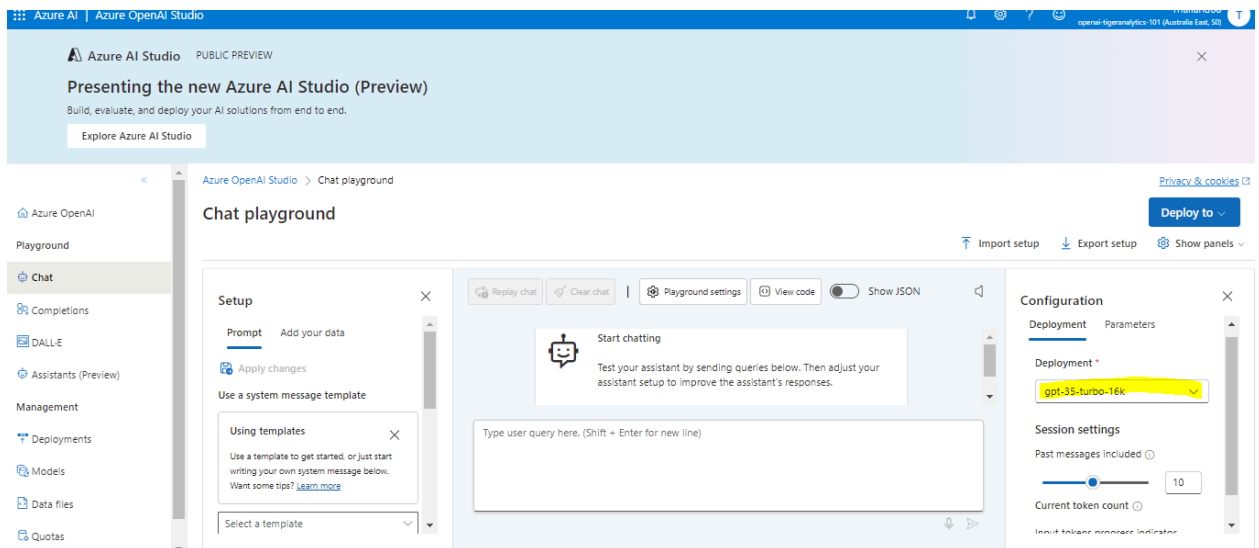
2. Under **Resource**, select the default option. It will look like **openai - Tiger analytics-<number>**, e.g. **open ai Tiger analytics-101**



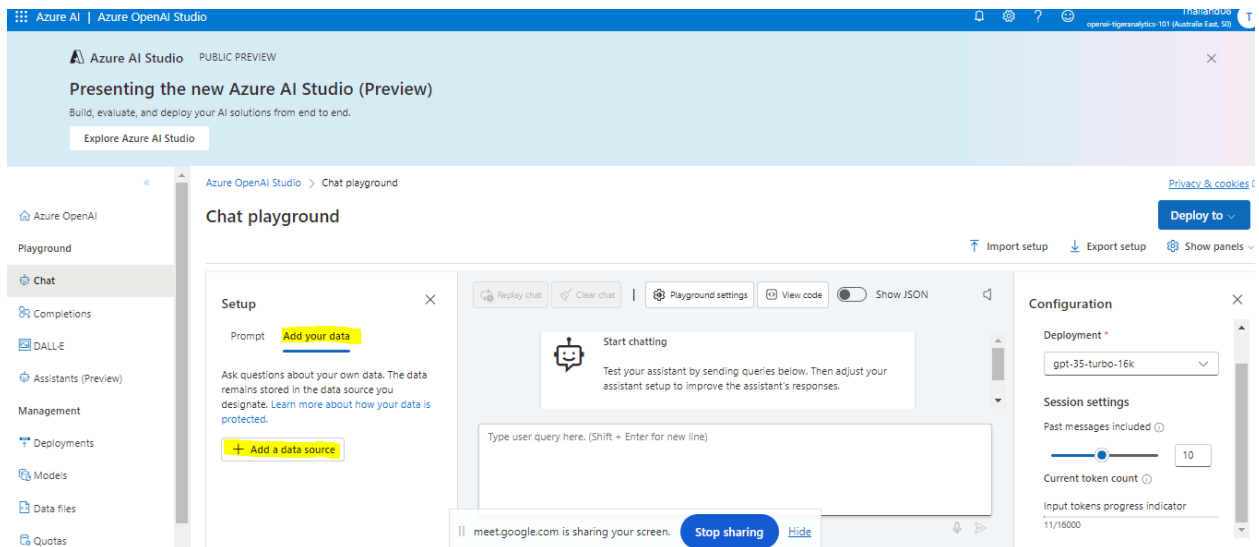
3. Select “chat” option in the AOAI studio home page



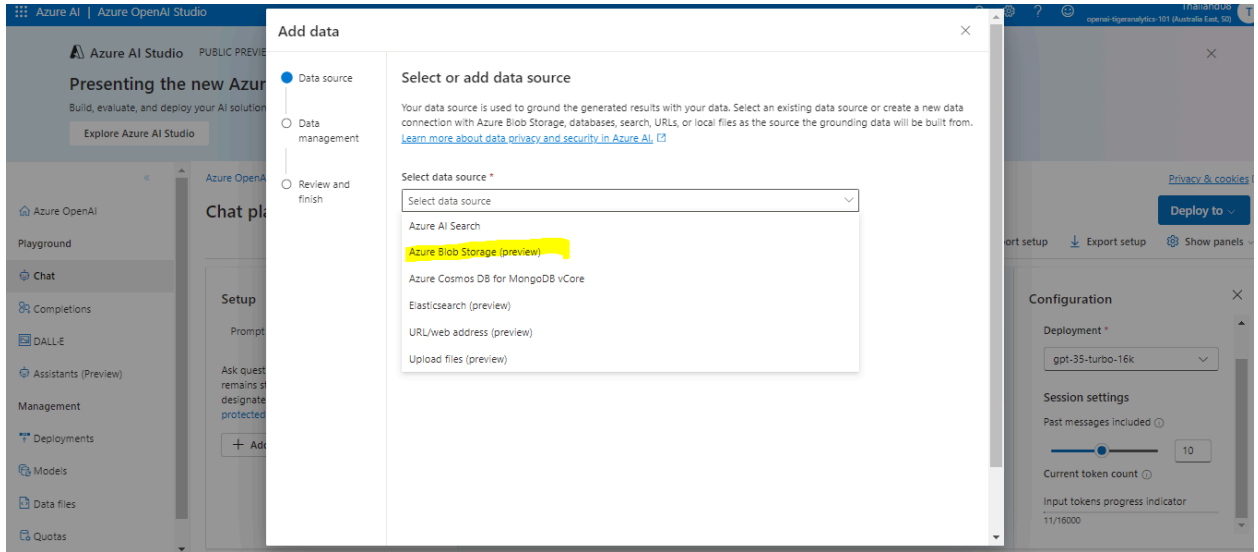
4. You will land up in the Playground page. Let's confirm if we are connected to the Azure OpenAI Service. On the right-hand side panel of the project playground, select the model from the dropdown **gpt-35-turbo16k**, and change the **past messages to include 10**.



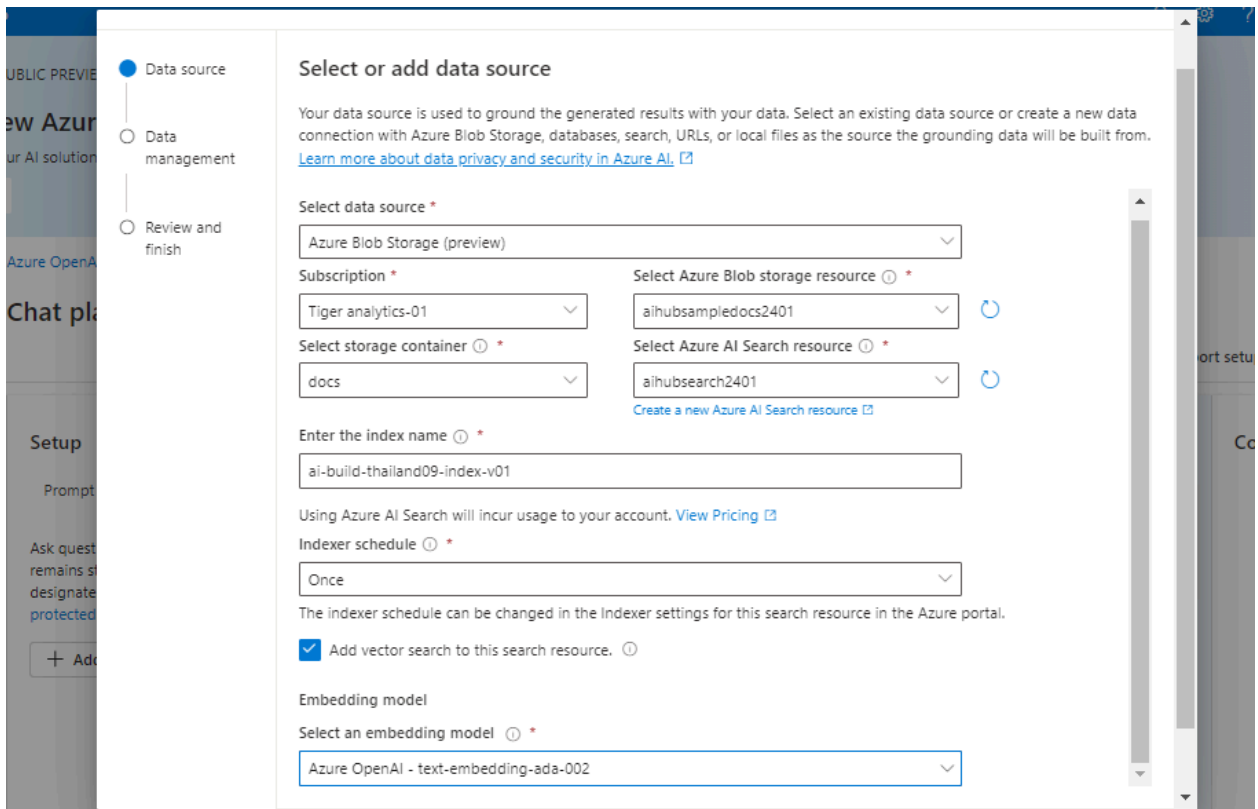
5. Now, let's add our data. On the left-hand side of the playground, go to **Add your Data** and click on **Add your Data**.



6. Select your "data source" and choose **"Azure Blob storage"**



7. Fill in the “data source details”. The **subscription ID** will be populated by default. Select the **storage account**, which will be in the format **aihubsampledocs<numericid>** and **blob container** as **docs**. For Azure AI search resource, choose the default option. Indexer schedule should be selected as “**once**”. Select “**add vector search tick mark**”. Choose the default **embedding model** from the drop down. In the **Index Settings**, use the index name as follows: “**ai-build-<userid>-index-<version>**”. Example **ai-build-Thailand01-index-v8**.



8. **Update the Data Management screen as below.** Ensure the search type is selected as “Hybrid + Semantic”. Keep the other settings as default.

The screenshot shows the 'Data management' configuration screen in the Azure AI Search portal. On the left, a sidebar contains three steps: 'Data source' (checked), 'Data management' (active), and 'Review and finish'. The main content area is titled 'Data management' and includes the following elements:

- A sub-header: 'Set up specific configurations for your data and how the model will respond to requests. [Learn more about data privacy and security in Azure AI.](#)
- A 'Search type' dropdown menu set to 'Hybrid + semantic'.
- Two informational messages: 'Using semantic search will incur usage to your Azure AI Search account. [View Pricing](#)' and 'Adding vector embeddings will incur usage to your account. [View Pricing](#)'.
- A 'Chunk Size' section with a description: 'Chunking is the process of breaking down your documents into smaller segments for search and retrieval. Chunk size is measured in tokens. If the selected chunk size results in low accuracy, re-ingest your data with a different size. [Learn more about selecting a chunk size](#)'.
- A 'Select a size' section with four radio button options: 256, 512, 1024 (default), and 1536.

9. **Review & Finish** - Validate your inputs and save

✓ Data source

✓ Data management

● Review and finish

Review and finish

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

Data source
 Azure Blob Storage

Storage resource
 aihubsampledocs2401

Storage container
 docs

Azure AI Search resource
 aihubsearch2401

Index name
 ai-build-thailand09-index-v01

Search type
 Hybrid + semantic

Embedding model
 text-embedding-ada-002

Semantic search configuration
 default

Back

Save and close

Cancel

10. You will be redirected to the playground page with the data index created.

Azure AI | Azure OpenAI Studio

Thailand10

openai-gpt-35-turbo-16k (Australia East, 50)

Azure OpenAI

Playground

Chat

Completions

DALL·E

Assistants (Preview)

Management

Deployments

Models

Data files

Quotas

Content filters (Preview)

Chat playground

Setup

Prompt

Add your data

Gain insights into your own data source. Your data is stored securely in your Azure subscription. [Learn more about how your data is protected.](#)

Data source: Azure AI Search

Search Resource: aihubsearch2401

Index: ai-build-thailand10-index-v99

Advanced settings >

Remove data source

Replay chat

Clear chat

Playground settings

View code

Show JSON

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 *

gpt-35-turbo-16k

Session settings

Past messages included

10

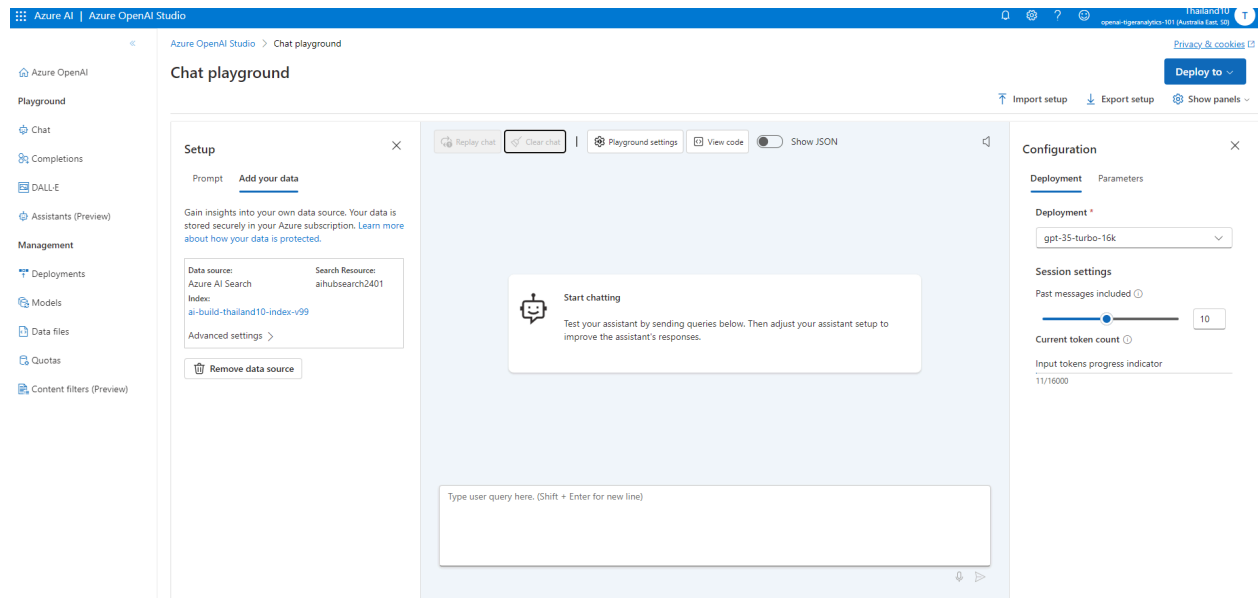
Current token count

11/16000

8

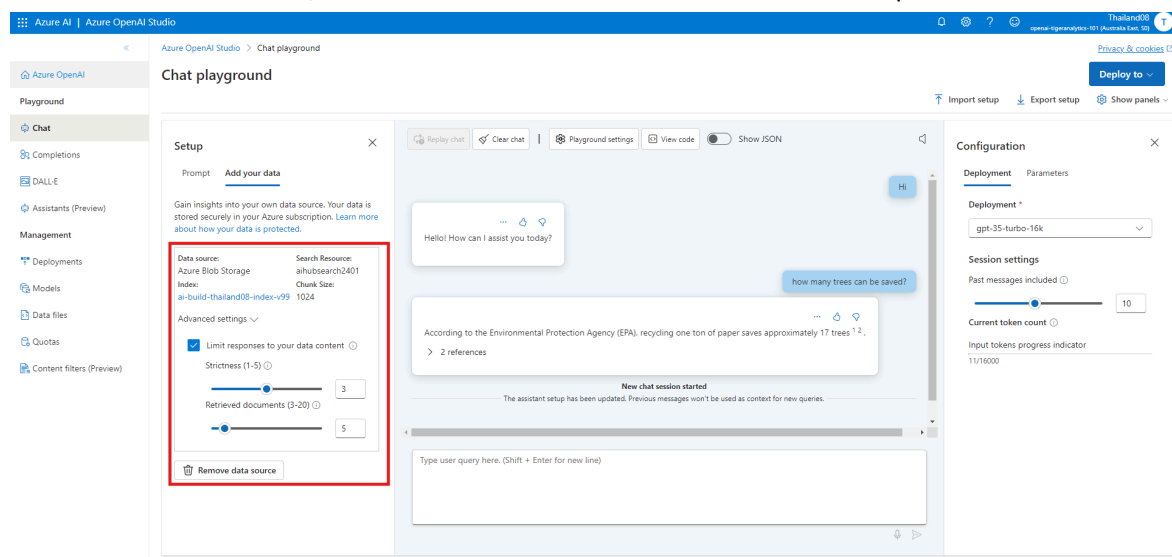
3. Start asking questions

1. Check on the responses from the data and ask questions. Some sample questions on the data can be found [here](#) (Please clone the repository or download the zip and navigate to the sample QnA spreadsheet in the **Data** folder to view sample questions).
2. Please click the **Clear Chat** button in the chat box before asking an unrelated follow-up question.

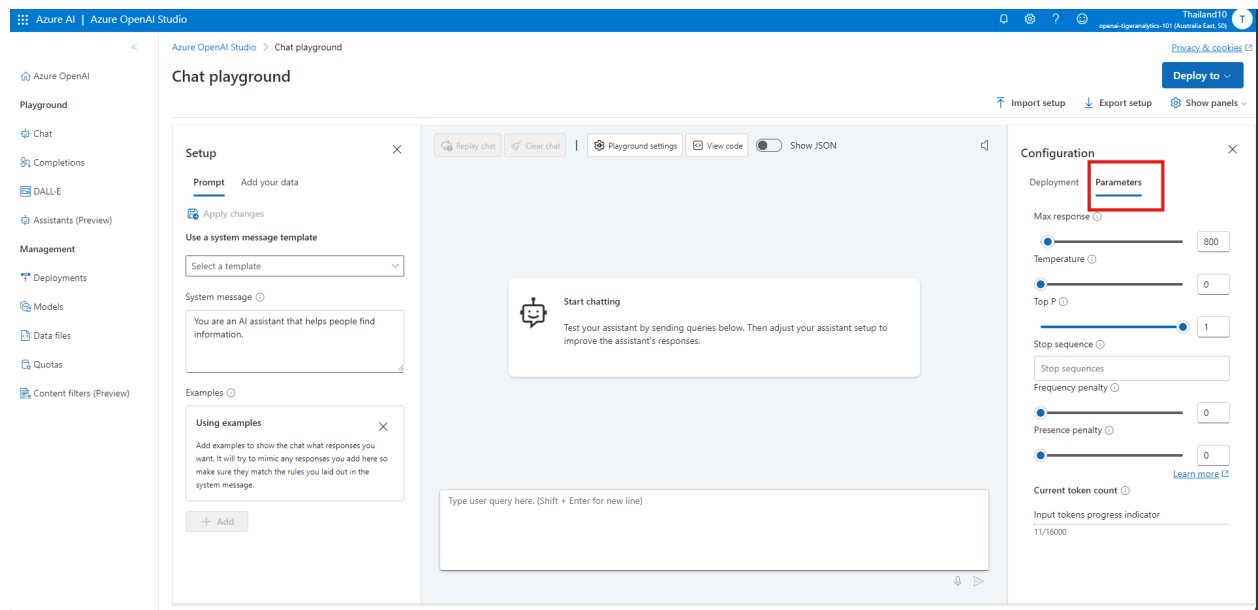


4. Tweak the parameters

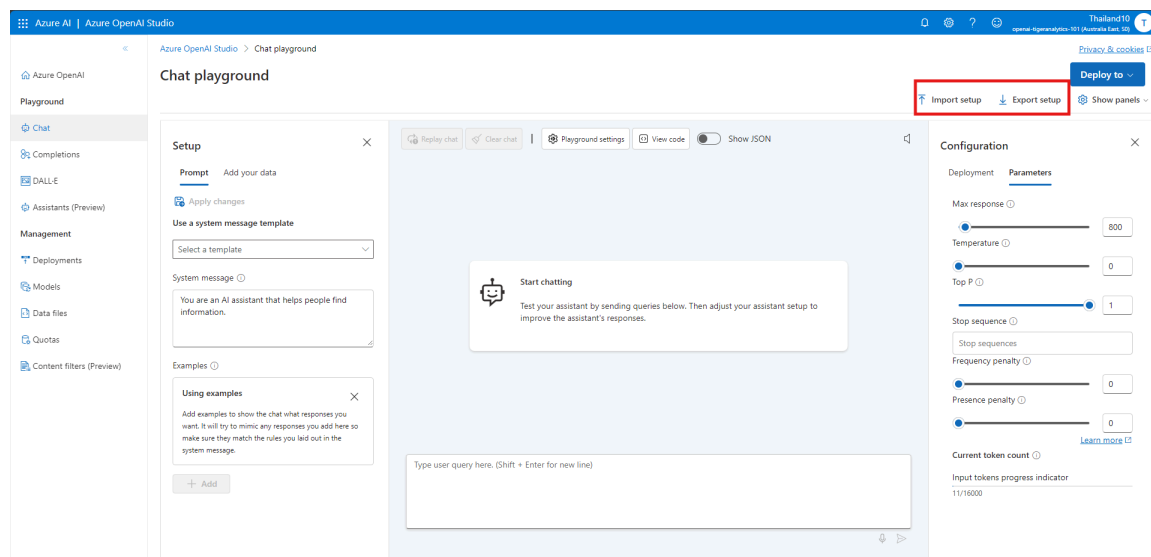
1. Suppose you want the responses to come not only from the source data alone but also from the external knowledge of the Azure OpenAI models. In that case, you can click on **Advanced Settings** under **Add your data** tab and uncheck **Limit responses to your data content**.
2. **Strictness** helps you set the threshold for relevant documents, and **Retrieved documents** help you set the number of chunks/documents that will be retrieved to answer the questions.



- Please check this guide on [how to generate text with Azure OpenAI Service - Azure OpenAI | Microsoft Learn](#)
- Tweak the model parameters from the **Parameters** tab. You can find the definition of the parameters [here](#).
- Make changes to the prompts to change the response to how you like (for example, we can add to give the response in bullet points or a happy tone). Remember to click on **Apply Changes** after changing the prompt. You can click the **Replay chat** button to get responses to the existing questions in the updated prompt.



- You can also **import** and **export** the chat settings (prompts, examples, parameters) to a JSON. You can try the same using this JSON. This feature helps share your workspace settings while working as a team.



7. You can also click on View Code to get executable code in the programming language of your choice with the Azure OpenAI, data, and index connections already set.

The screenshot displays the Azure OpenAI Studio interface, which is divided into several panels. On the left is a sidebar with navigation options: Azure OpenAI, 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 contains a 'Setup' section with tabs for 'Prompt' and 'Add your data'. Below this is a 'System message' field with the text 'You are an AI assistant that helps people find information.' and an 'Examples' section with a text area for adding chat examples. A 'Sample Code' panel is open, showing a Python script for integrating the current prompt and settings into an application. The code includes imports for 'openai' and 'os', sets the 'api_type' to 'azure', and configures the 'api_base', 'api_key', and 'deployment_id'. It also sets up the 'search_endpoint' and 'search_index_name'. A 'Configuration' panel is also visible, showing deployment parameters such as 'Max response' (set to 800), 'Temperature' (set to 0), 'Top P' (set to 1), 'Stop sequence', 'Frequency penalty' (set to 0), and 'Presence penalty' (set to 0). The 'Current token count' is shown as 11/16000.

Sample Code

```
1 import openai, os, requests
2
3 openai.api_type = "azure"
4 # Azure OpenAI on your own data is only supported by the
5 openai.api_version = "2023-08-01-preview"
6
7 # Azure OpenAI setup
8 openai.api_base = "https://openai-tigeranalytics-101.openai.azure.com/"
9 openai.api_key = os.getenv("OPENAI_API_KEY") # Add your
10 deployment_id = "gpt-35-turbo-16k" # Add your deployment ID
11 here
12 # Azure AI Search setup
13 search_endpoint = "https://alhubsearch2401.search.windows.net"; # Add your Azure AI
14 search_key = os.getenv("SEARCH_KEY"); # Add your Azure AI
15 search_index_name = "ai-build-thailand18-index-v99"; # Add
16 your Azure AI Search index name here
17
18 def setup_byod(deployment_id: str) -> None:
19     """Sets up the OpenAI Python SDK to use your own data for
20     the chat endpoint.
21     """
22     # Add your deployment ID and search index name here
```

Configuration

Deployment Parameters

- Max response: 800
- Temperature: 0
- Top P: 1
- Stop sequence:
- Frequency penalty: 0
- Presence penalty: 0

Current token count: 11/16000