

# Personalised AI

Azure OpenAI on data allows developers to use supported AI chat models that can reference specific sources of data to ground the response. Adding this information allows the model to reference both the specific data provided and its pretrained knowledge to provide more effective responses. Azure OpenAI on personalized data utilizes the search ability of Azure Cognitive Search to add the relevant data chunks to the prompt.

Once the data is in a Cognitive Search Index, Azure OpenAI on personalized data goes through the following steps:

1. Receive user prompt.
2. Determine relevant content and intent of the prompt.
3. Query the search index with that content and intent.
4. Insert search result chunk into the Azure OpenAI prompt, along with system message and user prompt.
5. Send the entire prompt to Azure OpenAI.
6. Return response and data reference (if any) to the user.

By default, Azure OpenAI on personalised data, encourages, but doesn't require the model to respond only using personalised data. This setting can be unselected when connecting data, which may result in the model choosing to use its pretrained knowledge over personalised data.

## Fine tuning VS personalised data

Fine-tuning is a technique used to create a custom model by training an existing foundational model such as gpt-35-turbo with a dataset of additional training data. Fine-tuning can result in higher quality requests that prompt engineering alone, customize the model on examples larger than can fit in a prompt, and allow user to provide fewer examples to get the same high-quality response. However, the process for fine-tuning is both costly and time intensive and should only be used for *use cases* where it's necessary.

Azure OpenAI on personalized data still uses the stateless API to connect to the model, which removes the requirement of training a custom model with personalized data and simplifies the interaction with the AI model. Cognitive Search first finds the useful information to answer the prompt, then Azure OpenAI forms the response based on the information.

Adding data is done through the Azure AI Studio, in the Chat Playground. The data source added is then used to augment the prompt sent to the model. When adding data one can choose to upload data files, use data in blob storage account, or connect to an existing Cognitive Search index. If using large text files or forms, use the available data preparation script to improve the AI model's accuracy.

**Note:** If uploading or using files already in the storage account, Azure OpenAI supports .md, .txt, .html, .pdf, .docx and .pptx. If any of these files contain graphics or images, the response quality depends on how well text can be extracted from the visual content.

# STEPS

1. Provision an Azure OpenAI resource with the following credentials:

- i. Subscription: Path InfotechLtd
- ii. Resource group: Demo
- iii. Region: EastUS
- iv. Name: Demo4567
- v. Pricing: Standard S0

Microsoft Azure

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

All services > Azure AI services | Azure OpenAI >

### Create Azure OpenAI

[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	Path Infotech Ltd
Resource group	Demo
Region	East US
Name	Demo4567
Pricing tier	Standard S0

**Network**

Type	All networks, including the internet, can access this resource.
------	---

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

[Give feedback](#)

Wait for the deployment to complete then go to the deployed Azure OpenAI resource in Azure portal.

Microsoft Azure

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

All services >

### Microsoft.CognitiveServicesOpenAI-20230831125859 | Overview

Deployment

Search

Delete Cancel Redeploy Download Refresh

**Overview**

- Inputs
- Outputs
- Template

**Your deployment is complete**

Deployment name : Microsoft.CognitiveServicesOpenAI-20230831125859 Start time : 8/31/2023, 1:01:25 PM  
Subscription : Path Infotech Ltd Correlation ID : e6c2168b-a338-4795-865a-031a2f9baa90  
Resource group : Demo

**Deployment details**

Resource	Type	Status	Operation details
Demo4567	Azure AI service	OK	<a href="#">Operation details</a>

**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**

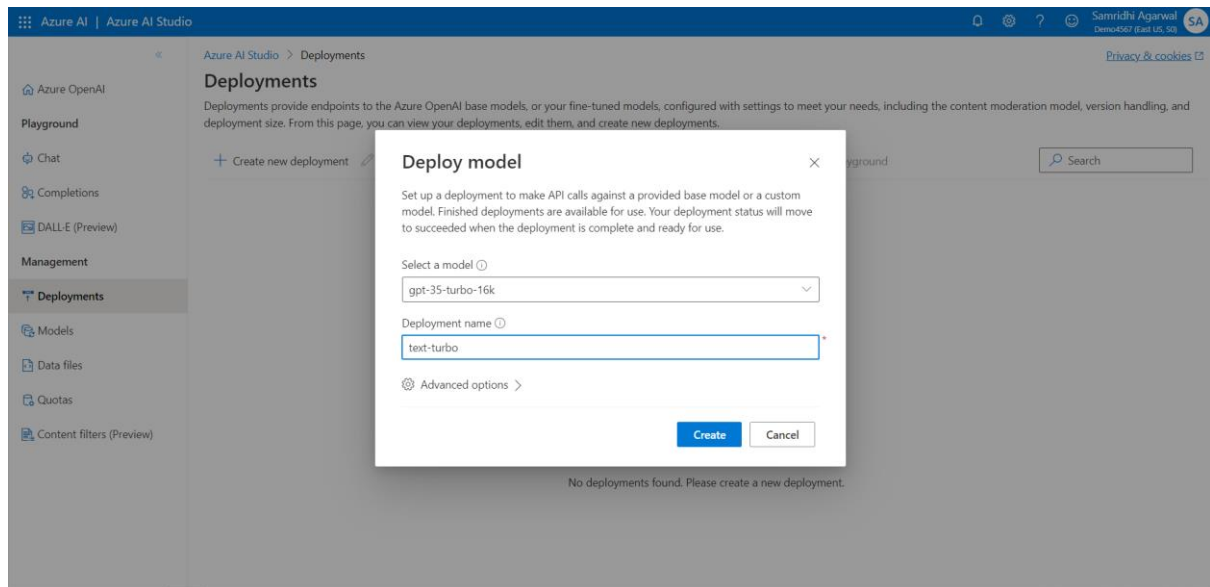
Azure experts are service provider partners who can help manage your assets on Azure and be your first line of support.  
[Find an Azure expert >](#)

## 2. Deploying a model

To use Azure OpenAI API for code generation, a model must be deployed first to use through the Azure OpenAI Studio. Once deployed, this model is used with the playground and reference that model in the app.

Deploy a model with the following credentials:

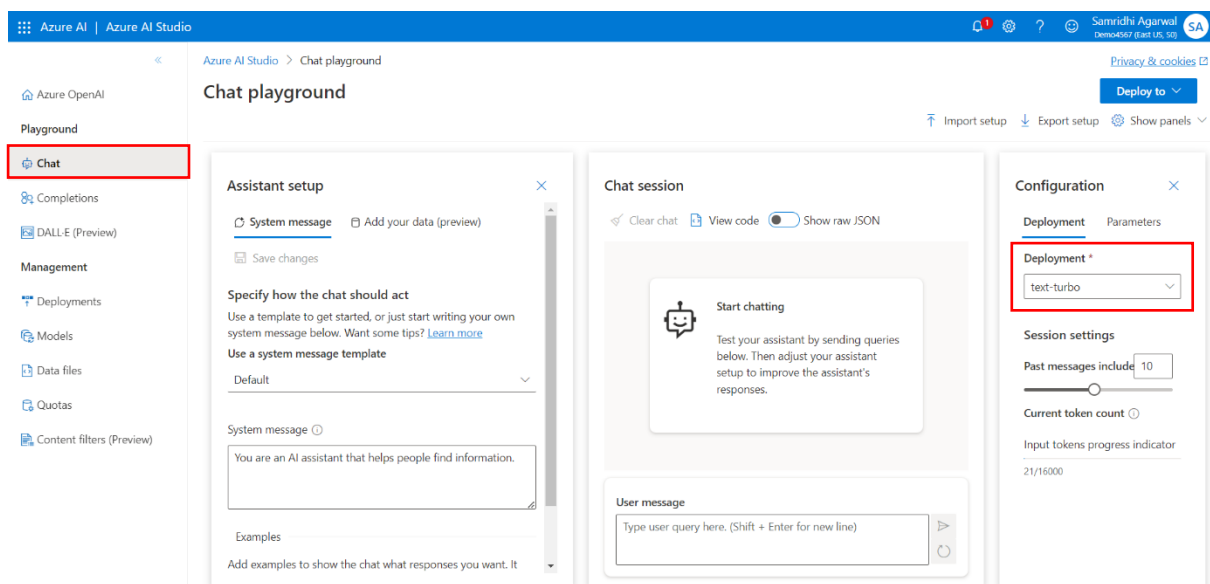
- i. Model: gpt-35-turbo-16k
- ii. Name: turbo-text



## Observing normal chat behavior without adding your own data

Before connecting personalized data to Azure OpenAI, let's observe the base model response to queries without any grounding data.

1. Navigate to the Chat playground, and select the deployed model in the configuration pane.



2. Enter random text prompt in the chat session, to judge the outputs without loading data source.

Chat session

[Clear chat](#) [View code](#) ☐ Show raw JSON

I'd like to take a trip to new york, where should I stay?

New York City offers a wide range of accommodations to suit different preferences and budgets. Some popular areas to consider for your stay include:

1. Manhattan: This borough is known for its iconic attractions like Times Square, Central Park, and the Theater District. It offers a variety of hotels, from luxury options in Midtown to more affordable choices in neighborhoods like the Lower East Side or Chelsea.
2. Brooklyn: This trendy borough offers a vibrant arts scene, diverse culinary options, and beautiful parks. Williamsburg and DUMBO are popular neighborhoods with boutique hotels and unique accommodations.
3. Queens: This borough is a more affordable option with easy access to Manhattan. It is home to attractions like Flushing Meadows-Corona Park and the Queens Museum. Long Island City is a popular neighborhood for hotels, offering great views of the Manhattan skyline.
4. Staten Island: If you prefer a quieter atmosphere, Staten Island might be a good choice. It is known for its beautiful parks, such as the Staten Island Greenbelt and the Staten Island Zoo.

Consider factors such as your budget, preferred location, and the attractions you plan to visit when choosing your accommodation.

User message

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

The responses currently are based on the pre-trained broad knowledge that the model has.

3. Create a storage account and a cognitive search resource as well.

Microsoft Azure

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

Home >

### Create a storage account

Basics Advanced Networking Data protection Encryption Tags Review

**Basics**

Subscription	Path Infotech Ltd
Resource Group	Demo
Location	eastus
Storage account name	pathlite
Deployment model	Resource manager
Performance	Standard
Replication	Read-access geo-redundant storage (RA-GRS)

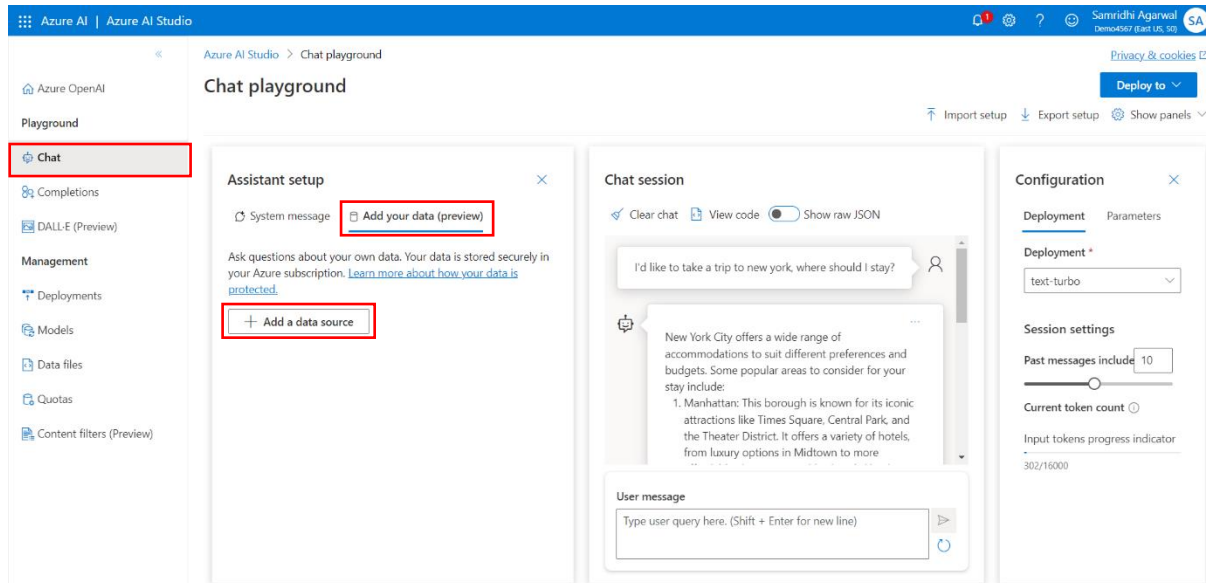
**Advanced**

Enable hierarchical namespace	Enabled
Enable network file system v3	Disabled
Allow cross-tenant replication	Disabled
Access tier	Hot
Enable SFTP	Enabled

[Create](#) [< Previous](#) [Next >](#) [Download a template for automation](#) [Give feedback](#)

Wait for the deployment to complete then go to the deployed Azure OpenAI resource in Azure portal.

4. Connecting data into chat playground
  - a. Navigate to the Chat playground in Azure AI Studio
  - b. Select “Add your data” tab in the Assistant Setup pane.
  - c. Select “Add a data source” to get data connected.



5. In the Add data, enter the following values for data source:

**Data source:** Upload files

**Azure Blob Storage Resource:** (resource created)

**CORS:** “ON”

**Azure Cognitive Search Resource:** (resource created)

**Index name:** margiestravel

## 6. Download the files which will be used as the personalized data.

Before you start

Provision an Azure OpenAI resource

Deploy a model

Observe normal chat behavior without adding your own data

Connect your data in the chat playground

Chat with a model grounded in your data

Clean up

### Connect your data in the chat playground

Next, add your data in the chat playground to see how it responds with your data as grounding

1. Download the **data** that you will use from GitHub. Extract the PDFs in the **.zip** provided.
2. Navigate to the **Chat** playground, and select **Add your data** in the Assistant setup pane.
3. Select **Add a data source** and choose **Upload files** from the dropdown.
4. You'll need to create a storage account and Azure Cognitive Search resource. Under the dropdown for the storage resource select **Create a new Azure Blob storage resource**, and create a storage account with the following settings. Anything not specified leave as the default.
  - **Subscription:** Same subscription as your Azure OpenAI resource
  - **Resource group:** Same resource group as your Azure OpenAI resource
  - **Storage account name:** Enter globally unique name
  - **Region:** Same region as your Azure OpenAI resource
  - **Redundancy:** Locally-redundant storage (LRS)
5. Once the resource is being created, come back to Azure OpenAI Studio and select **Create a new Azure Cognitive Search resource** with the following settings. Anything not specified leave as the default.

MicrosoftLearning/mslearn-openai

brochures.zip

The downloaded zip file contains all the necessary pdf files to be used as personalized data.

Search brochures							
Name	Type	Compressed size	Password pr...	Size	Ratio	Date modified	
Dubai Brochure	WPS PDF Document	365 KB	No	380 KB	5%	4/12/2022 11:24 AM	
Las Vegas Brochure	WPS PDF Document	541 KB	No	557 KB	3%	4/12/2022 11:24 AM	
London Brochure	WPS PDF Document	426 KB	No	441 KB	4%	4/12/2022 11:24 AM	
Margies Travel Company Info	WPS PDF Document	330 KB	No	345 KB	5%	4/12/2022 11:24 AM	
New York Brochure	WPS PDF Document	360 KB	No	374 KB	4%	4/12/2022 11:24 AM	
San Francisco Brochure	WPS PDF Document	362 KB	No	377 KB	4%	4/12/2022 11:24 AM	

- On the upload files page, upload the pdf files downloaded from the Microsoft learning page which had a hyperlink of github which directly got us downloaded sample pdf files which will be treated as personalised data.

### Assistant setup

System message

Add your data (preview)

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

☒ Limit responses to your data content ⓘ

Data source: Search Resource:

Upload Files travelqna

Index: margiestravel

Remove data source

- The chatbot is ready to be tested with personalized data.

Upload files

Review and finish

created or selected in the previous step.

Drag and drop.  
or  
Browse for a file

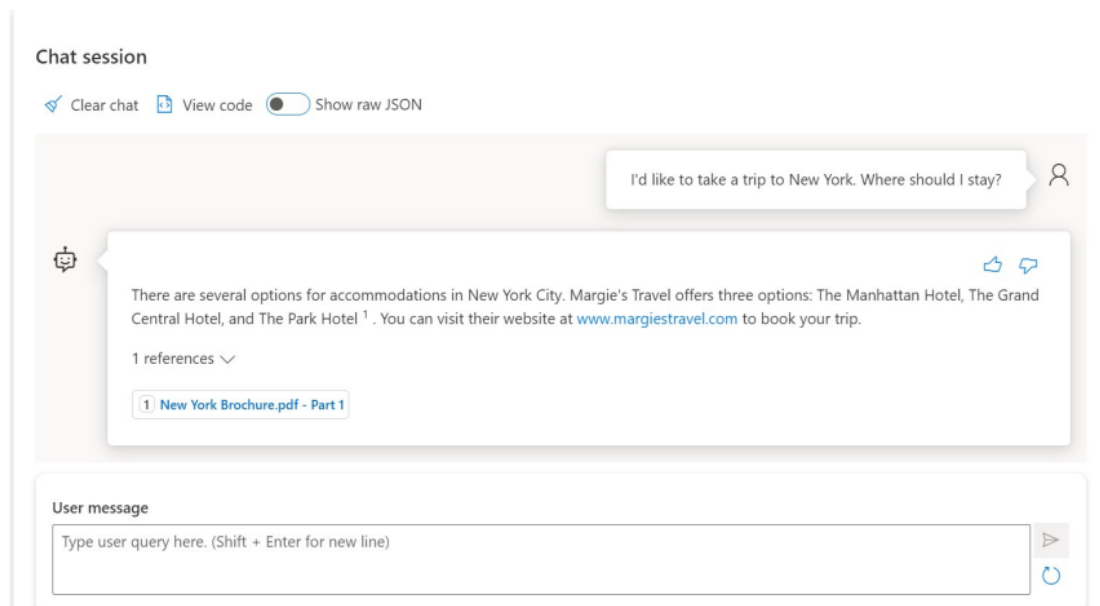
(.txt, .md, .html, .pdf, .docx, .pptx)  
16 MB size limit

Upload files

File name	Type	Size	State
London Brochure.pdf	PDF	440.07 ...	Uploaded
Margies Travel Company Inf...	PDF	344.01 ...	Uploaded
New York Brochure.pdf	PDF	373.45 ...	Uploaded

## Testing

Checking the chatbot for the personalized data.



The same question has been asked after providing personalized data and gives the answers based on the provided pdfs of data.

## Summary

In this module, we saw how Azure OpenAI on personalized data allows developers to use supported AI chat models to reference specific data sources. Connecting personalized data allows the model to reference the specific information provided and its pretrained knowledge to provide more effective responses.

## Reference

<https://microsoftlearning.github.io/mslearn-openai/instructions/Labs/06-use-own-data.html>