

# Azure OpenAI Services

December 2025



# What is OpenAI?

OpenAI is a leading research laboratory that develops safe and beneficial AI for personal and professional use. They've created some of the most advanced AI models, including GPT-5, Embeddings, DALL-E, and Whisper.

# What is Azure OpenAI Service?

Azure OpenAI Service is a Microsoft Azure cloud service that provides enterprise access to OpenAI's advanced language and generative models, such as GPT models, through Azure infrastructure.

[Introducing-OpenAI](#)



[Microsoft and OpenAI partnership unveils new AI opportunities](#)

[What is Azure OpenAI Service?](#)



With Azure OpenAI, customers get the security capabilities of Microsoft Azure while running the same models as OpenAI. Azure OpenAI offers private networking, regional availability, and responsible AI content filtering.

## The Microsoft Responsible AI Standard



# Comparing Azure OpenAI and OpenAI

Azure OpenAI offers superior security capabilities in comparison to OpenAI. Additionally, Azure OpenAI provides support for Azure Virtual Network (VNet) and private endpoints, which are not available in OpenAI. Another advantage of using Azure OpenAI is that it comes with Azure Cognitive Services, Service Level Agreement for Azure (this “SLA”), while OpenAI currently only offers a status page:

[SLA for Cognitive Services](#)

[Azure OpenAI vs OpenAI](#)

[Explore the benefits of Azure OpenAI Service with Microsoft](#)  
[Learn](#)



# Data, privacy, and security for Azure OpenAI Service

Azure OpenAI stores and processes data to provide the service and to monitor for uses that violate the applicable product terms. Please also see the [Microsoft Products and Services Data Protection Addendum](#), which governs data processing by the Azure OpenAI Service except as otherwise provided in the applicable [Product Terms](#).





## Important

The prompts (inputs) and completions (outputs):

- are **NOT** available to other customers.
- are **NOT** available to OpenAI.
- are **NOT** used to improve OpenAI models.
- are **NOT** used to improve any Microsoft or 3rd party products or services.
- are **NOT** used for automatically improving Azure OpenAI models for your use in your resource (The models are stateless, unless you explicitly fine-tune models with your training data).

The Azure OpenAI Service is fully controlled by Microsoft; Microsoft hosts the OpenAI models in Microsoft's Azure environment and the Service does NOT interact with any services operated by OpenAI (e.g. ChatGPT, or the OpenAI API).

[More details about Data, privacy, and security...](#)



# Create and deploy an Azure OpenAI Service resource

Please note that for the sake of simplicity, this document provides step-by-step instructions to create an Azure OpenAI Service resource and deploy a model via Azure Foundry. You can choose either resource creation via Azure Command-Line-Interface (CLI) or PowerShell. Please see [here](#) for more details.

## Create customized, AI-powered search with an Azure free account

Get started with 12 months of free services including Azure AI Services

[Start free](#)[Pay as you go](#)

### 1

## An Azure subscription

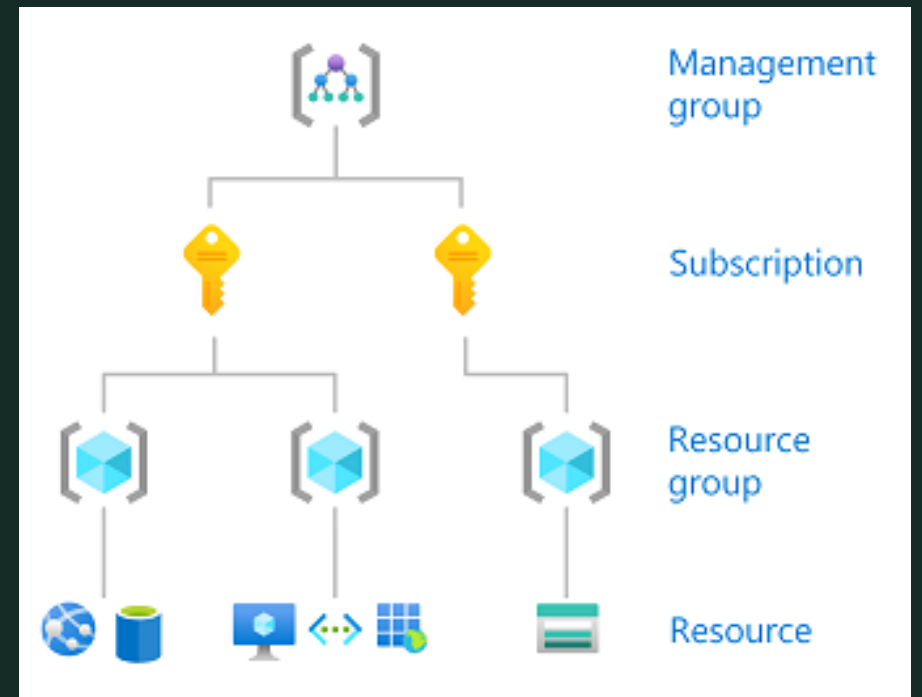
- [Create one for free](#)
- Start free. Get \$200 credit to use in 30 days. While you have your credit, get free amounts of popular services and 55+ other services.
  - Consider upgrading to the paid tier from the outset if you intend to utilize this service for production purposes.
- After your credit, move to pay as you go to keep getting popular services and 55+ other services. Only pay if you use more than the free monthly amounts.
- After 12 months, you'll continue getting 55+ services free always—and still only pay for what you use beyond the free monthly amounts.

## zapliance | Create and deploy an Azure OpenAI Service resource

2

Access permissions to create Azure OpenAI resources and to deploy models.

Please see [here](#) for more details.





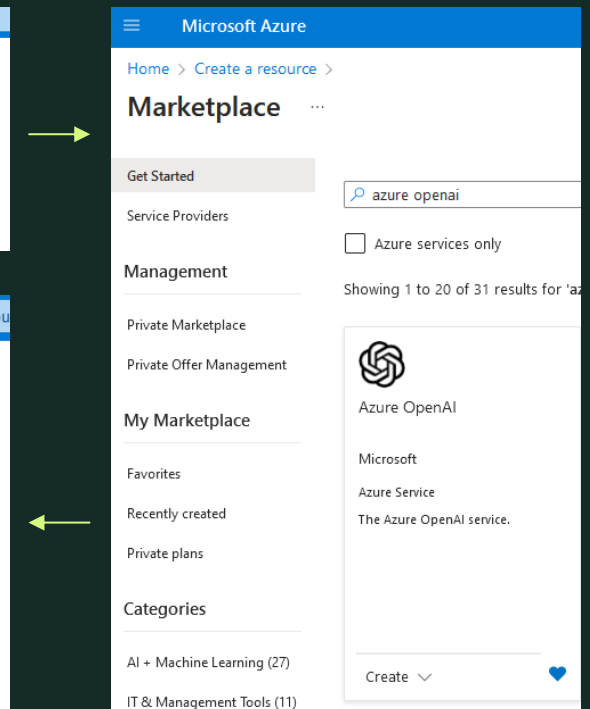
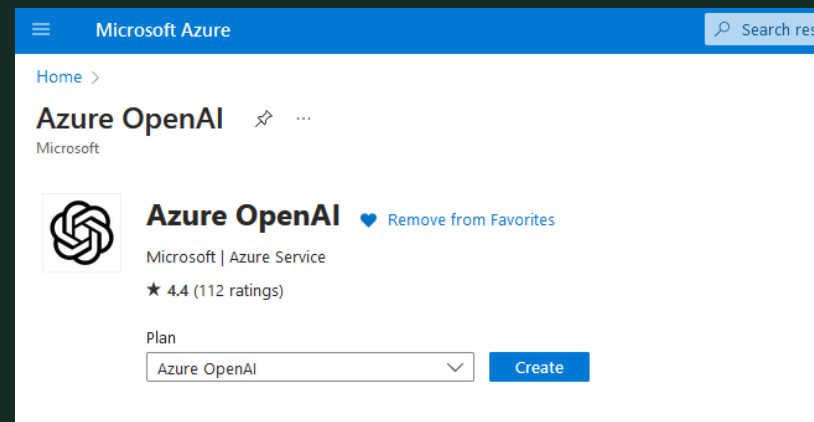
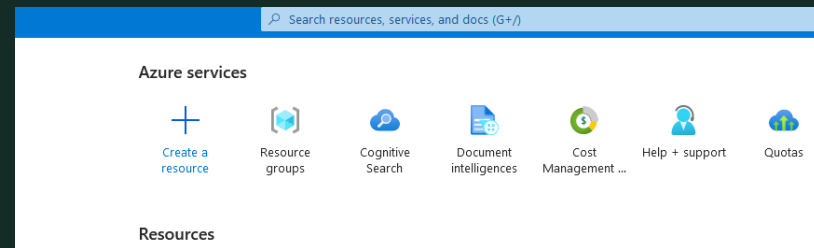
# zapliance | Create and deploy an Azure OpenAI Service resource

3

## Create a resource

Resource is a manageable item that is available through Azure. Virtual machines, storage accounts, web apps, databases, and virtual networks are examples of resources. Resource groups, subscriptions, management groups, and tags are also examples of resources.

- Sign in with your Azure subscription in the Azure portal.
- Select **Create a resource** and search for the **Azure OpenAI Service**. When you locate the service, select **Create**.



# zapliance | Create and deploy an Azure OpenAI Service resource

4

On the Create Azure OpenAI page, provide the following information for the fields on the Basics tab:

Field	Description
Subscription	The Azure subscription used in your Azure OpenAI Service onboarding application.
Resource group	The Azure resource group to contain your Azure OpenAI resource. You can create a new group or use a pre-existing group.
Region	The location of your instance. Different locations can introduce latency, but they don't affect the runtime availability of your resource.
Name	A descriptive name for your Azure OpenAI Service resource, such as <b>MyOpenAIResource</b> .
Pricing Tier	The pricing tier for the resource. Currently, only the Standard tier is available for the Azure OpenAI Service.

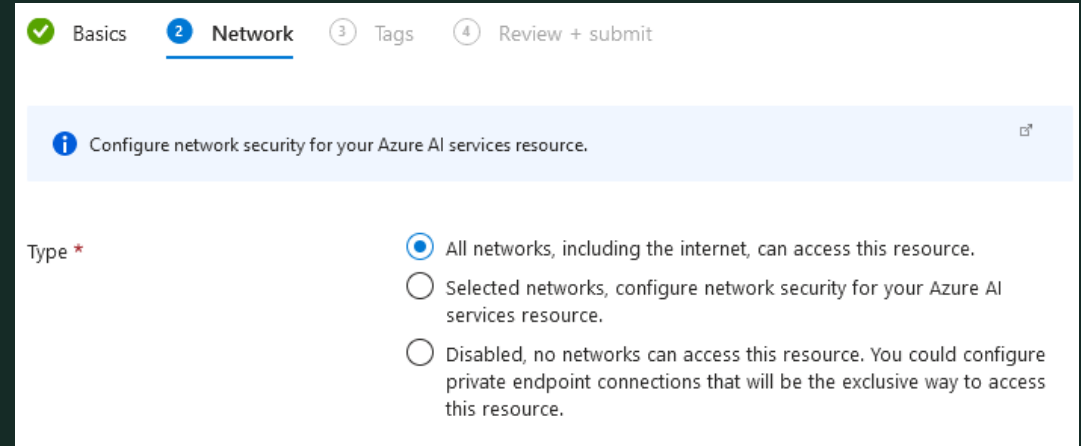
The screenshot shows the 'Create Azure OpenAI' page in the Microsoft Azure portal. The page has a blue header with the Microsoft Azure logo and a search bar. Below the header, there's a breadcrumb trail: Home > openai03 > Marketplace >. The main title is 'Create Azure OpenAI'. Below the title, there are four tabs: Basics (selected), Network, Tags, and Review + submit. The Basics tab contains a description of Azure OpenAI Service, a 'Learn more' link, and a 'Project Details' section. The 'Project Details' section has two rows: 'Subscription' with a dropdown menu showing 'Azure-Abonnement 1' and 'Resource group' with a dropdown menu showing 'openai03' and a 'Create new' link. Below this is the 'Instance Details' section with three rows: 'Region' with a dropdown menu showing 'Sweden Central', 'Name' with a text input field containing 'openai-swedencentral-instance' and a green checkmark, and 'Pricing tier' with a dropdown menu showing 'Standard S0'. There is a 'View full pricing details' link below the pricing tier. At the bottom, there is a 'Content review policy' section and two buttons: 'Previous' and 'Next'.

# zapliance | Create and deploy an Azure OpenAI Service resource

5

## Configure network security

Regarding your concerns and security, you can configure the network security in three types. More details about each type can be found [here](#).



The screenshot shows the 'Network' configuration step in the Azure OpenAI Service deployment wizard. At the top, there are four tabs: 'Basics' (checked), 'Network' (active), 'Tags', and 'Review + submit'. Below the tabs, a light blue banner contains an information icon and the text 'Configure network security for your Azure AI services resource.' Below this, the 'Type' section has a red asterisk and three radio button options: 'All networks, including the internet, can access this resource.' (selected), 'Selected networks, configure network security for your Azure AI services resource.', and 'Disabled, no networks can access this resource. You could configure private endpoint connections that will be the exclusive way to access this resource.'

6

Select **Next** and configure any **Tags** for your resource, as desired.

7

Select **Next** to move to the final stage in the process: Review + submit

8

Confirm your configuration settings, and select **Create**.

# zapliance | Create and deploy an Azure OpenAI Service resource

9

## Deploy a model

Before you can generate text or inference, you need to deploy a model. You can select from one of several available models in Azure Foundry Portal. To deploy a model, follow these steps:

### 9.1

Sign in Explore Foundry Portal

### 9.2

From Chat playground, create a new deployment with model GPT-4.1

#### 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 Foundry portal.

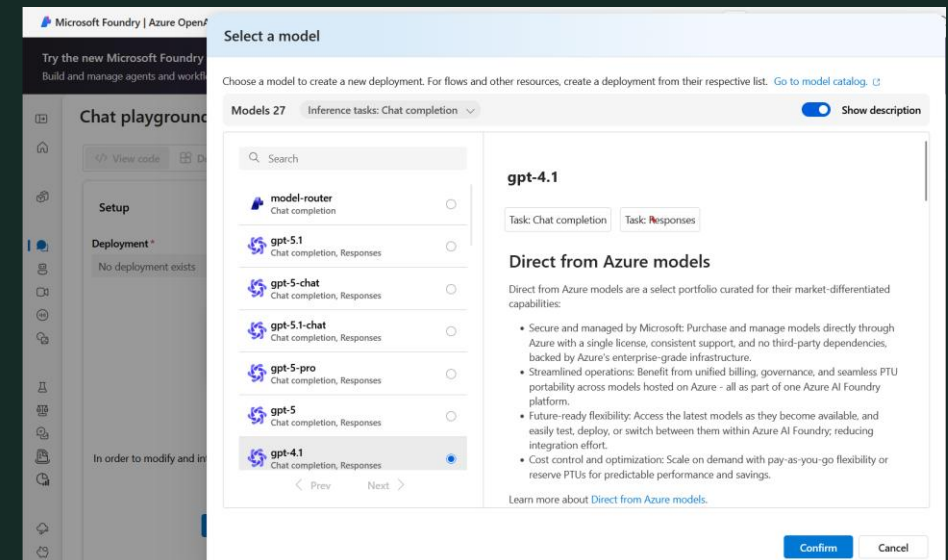
[Learn More](#)



#### Explore and deploy

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

[Explore Foundry portal](#)



# zapliance | Create and deploy an Azure OpenAI Service resource

## 9.3

Set the required field as shown in the screenshot:

- Deployment name: gpt-4.1
- Deployment type: Data Zone Standard
- Resource location: <your resource location in Europe, e.g. sweden central, west Europe, etc>

Field	Description
Select a model	Model availability varies by region. For a list of available models per region, see <a href="#">Model summary table and region availability</a> .
Deployment name	Choose a name carefully. The deployment name is used in your code to call the model by using the client libraries and the REST APIs.
Advanced options (Optional)	You can set optional advanced settings, as needed for your resource. <ul style="list-style-type: none"><li>- For the Content Filter, assign a content filter to your deployment.</li><li>- For the Tokens per Minute Rate Limit, adjust the Tokens per Minute (TPM) to set the effective rate limit for your deployment. You can modify this value at any time by using the <a href="#">Quotas</a> menu.</li></ul>

### Deploy gpt-4.1

Deployment name \*

gpt-4.1

Deployment type

Data Zone Standard

Data Zone Standard: Pay per API call with higher rate limits. Learn more about [Data Zone deployment types](#).

Data might be processed anywhere within the Microsoft specified data zone, outside of the AI resource's Azure geography, but data storage remains in the AI resource's Azure geography. Learn more about [data residency](#).

Deployment details

Collapse

Model version upgrade policy

Upgrade once new default version becomes available

Model version

2025-04-14 (Default)

Resource location

France Central

300K tokens per minute quota available for your deployment

Tokens per Minute Rate Limit ⓘ

300K

Corresponding requests per minute (RPM) = 300

# zapliance | Create and deploy an Azure OpenAI Service resource

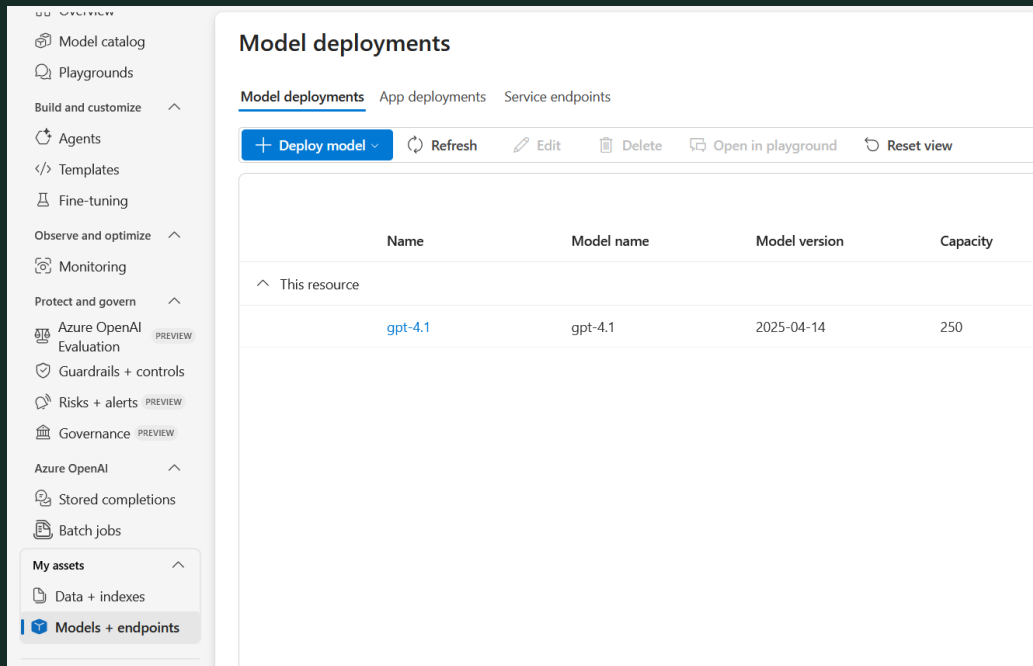
10

The deployments table shows a new entry that corresponds to your newly created model.

When the deployment completes, your model deployment status changes to **succeeded**.

## Retrieve key and endpoint

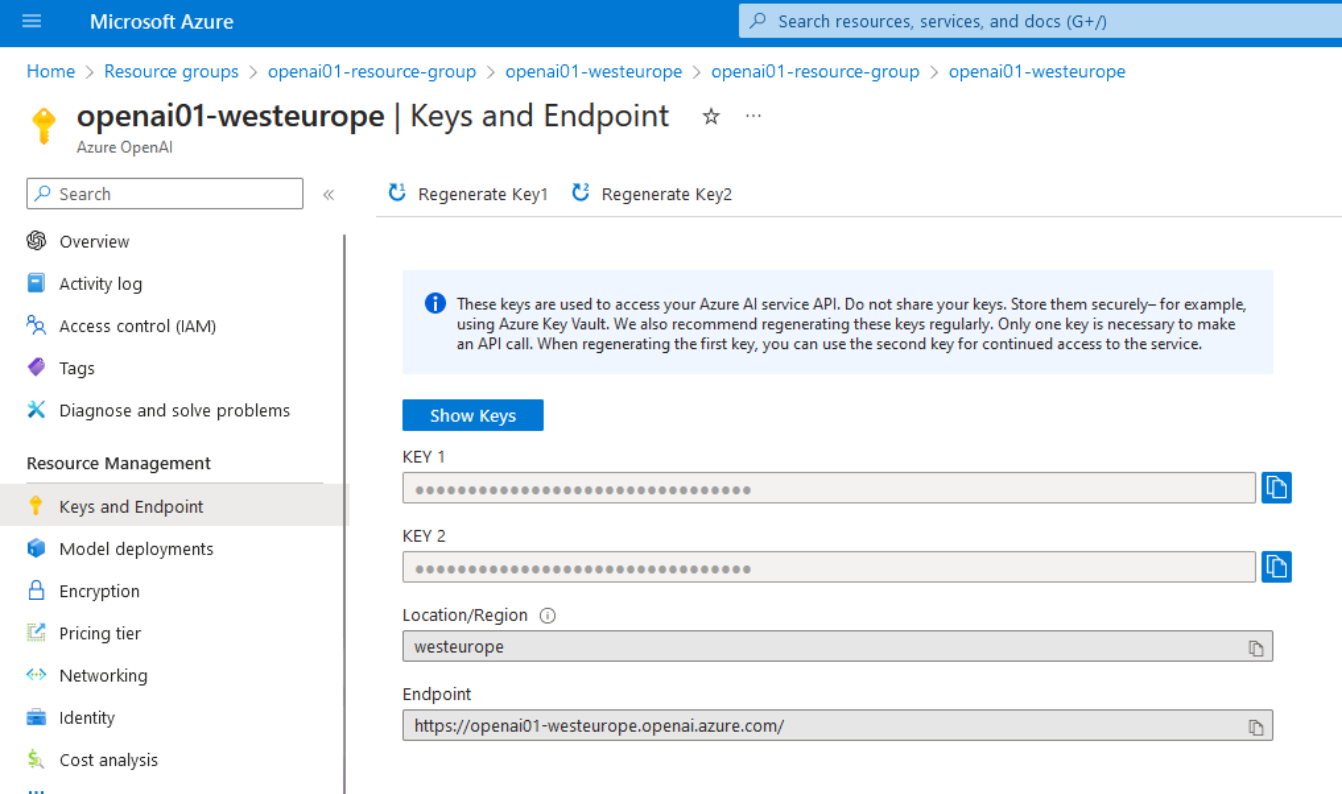
To successfully make a call against the Azure OpenAI service, you'll need the following:



Field	Description
Endpoint	This value can be found in the Keys & Endpoint section when examining your resource from the Azure portal. Alternatively, you can find the value in Azure OpenAI Studio > Playground > Code View. An example endpoint is: <a href="https://docs-test-001.openai.azure.com/">https://docs-test-001.openai.azure.com/</a> .
API-KEY	This value can be found in the Keys & Endpoint section when examining your resource from the Azure portal. You can use either KEY1 or KEY2.
Deployment - name	This value will correspond to the custom name you chose for your deployment when you deployed a model. This value can be found under Resource Management > Deployments in the Azure portal or alternatively under Management > Deployments in Azure OpenAI Studio.

11

Go to your resource in the Azure portal. The Endpoint and Keys can be found in the Resource Management section. Copy your endpoint and access key as you'll need both for authenticating your API calls. You can use either **KEY1** or **KEY2**. Always having two keys allows you to securely rotate and regenerate keys without causing a service disruption.



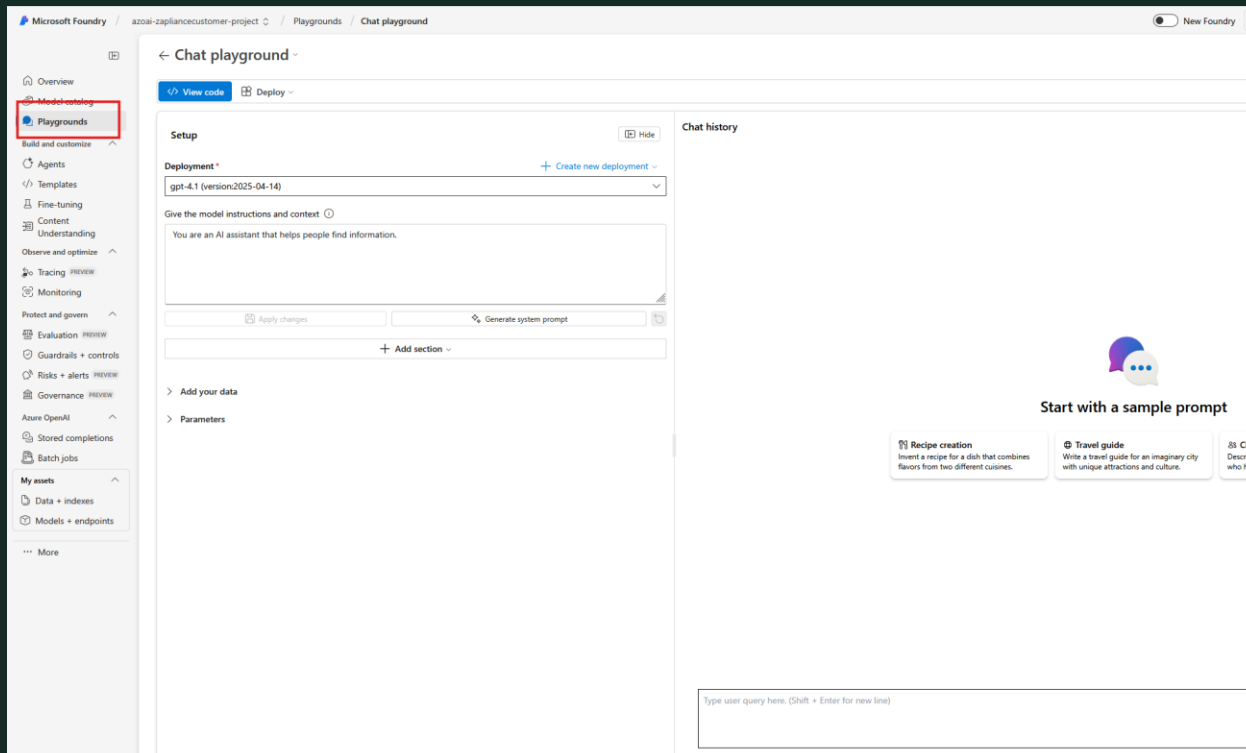
The screenshot displays the Microsoft Azure portal interface. The top navigation bar includes the 'Microsoft Azure' logo and a search bar. The breadcrumb trail indicates the path: Home > Resource groups > openai01-resource-group > openai01-west europe > openai01-resource-group > openai01-west europe. The main heading is 'openai01-west europe | Keys and Endpoint', with 'Azure OpenAI' noted below it. A left-hand sidebar lists various management options: Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, and a 'Resource Management' section containing Keys and Endpoint (highlighted), Model deployments, Encryption, Pricing tier, Networking, Identity, and Cost analysis. The main content area features a search bar, 'Regenerate Key1' and 'Regenerate Key2' buttons, and an informational message about key security. Below this, a 'Show Keys' button is present. The page then displays 'KEY 1' and 'KEY 2' as masked strings with copy icons, followed by the 'Location/Region' (west europe) and the 'Endpoint' (https://openai01-west europe.openai.azure.com/), both also with copy icons.



# Azure OpenAI Playground



In the Playground, you can start testing your deployments with different prompts.



You will also notice, on the right-hand side of the **Playground**, a set of parameters to be tuned. Those are necessary in order to make the model more tailored to different use cases

# Required API Configuration

We'll need to get the following information from your Azure OpenAI service:

```
# COMPANY_NAME
```

```
AZURE_OPENAI_ENDPOINT = <your azure openai endpoint>
```

```
AZURE_OPENAI_API_KEY = <your azure openai api key>
```

```
AZURE_OPENAI_API_TYPE = azure
```

```
AZURE_OPENAI_MODEL = <your openai model name>
```

```
AZURE_OPENAI_DEPLOYMENT_NAME = <your openai deployment name>
```



Thank you

