

Alfred: Givaudan's Internal ChatGPT Journey – From Zero to Hero

Cihan Cinar



Sponsors



We are a team of senior infrastructure specialists, software developers and data engineers who are experts in the Microsoft Azure Cloud. We partner with you to deliver innovative business solutions using Agile, DevOps and advanced Software Automation. Twenty years in business. Headquartered in Toronto. Work for customers across Canada and the US. Work across multiple industries and sectors. www.objectsharp.com

Microsoft is a proud sponsor of Global Azure in Toronto on April 20th. We're passionate about supporting the developer community and invite you to access valuable documentation and training resources by visiting docs.microsoft.com and the Microsoft Reactor.



Alfred Givaudan's Internal ChatGPT Journey

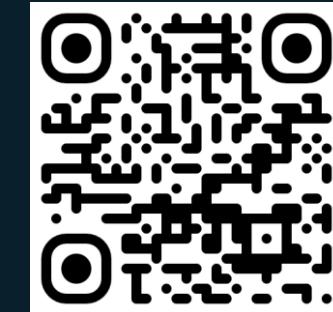
From Zero to Hero



Cihan Cinar

Software & AI Architect
Givaudan

<https://blog.cihan-cinar.com/>



Agenda



- Givaudan Presentation
- Why a secured ChatGPT
 - Context
 - Analyze step by step
- Architecture
- Azure
- Demo
 - Azure
 - Alfred
- Conclusion

A woman with dark hair and bangs, wearing a red dress, is standing in a field of yellow flowers. She is looking directly at the camera with a neutral expression. The flowers are in the foreground and background, creating a dense, colorful texture.

Givaudan products drive millions of
consumer decisions every day!

Givaudan
Human by nature

Fragrance & Beauty, Taste & Wellbeing, provide our customers compounds, ingredients and integrated solutions that engage consumers



Savory



Beverages



Nutraceuticals



Oral care



Fine Fragrances



Personal care



Snacks



Dairy



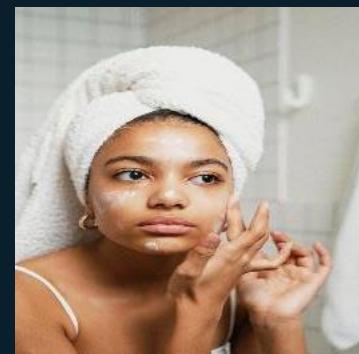
Sweet goods



Home care



Fabric care

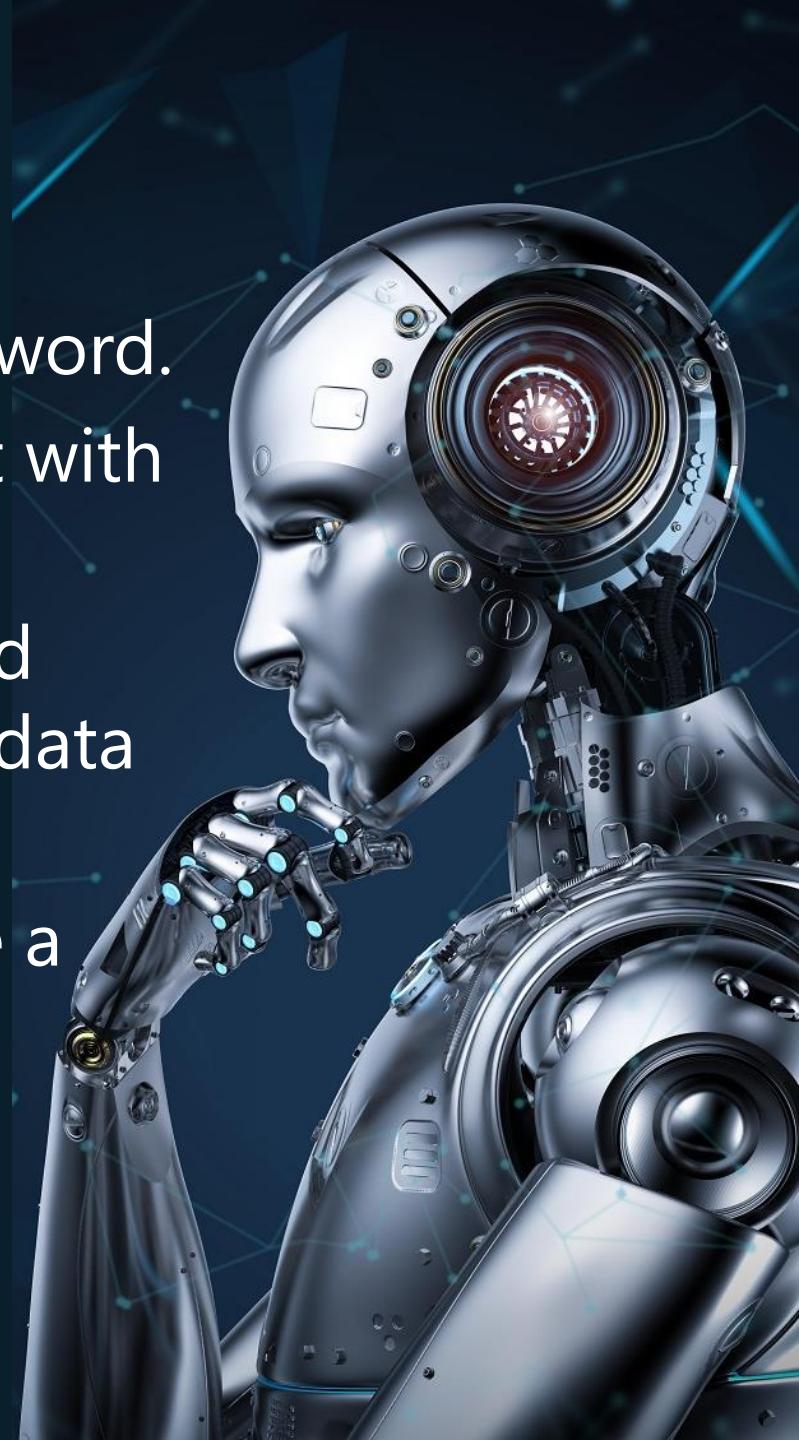


Beauty & skin care

Why Secured ChatGPT

Context

- Since October 2022, ChatGPT has been the buzzword.
- In 2023, we needed of a secured solution to chat with AI solutions
- Avoid company data to be sent to third party and particularly outside of the company and Europe data regulatory
- No real solution on the market as a SaaS to have a secured conversational chat solution
- We have to scale quick to the production



Context: Benefits

Main objectives:

- Offer the business and power users a secured platform
- Accelerate future Gen AI projects
- Build skills and expertise in LLMs

Benefits:

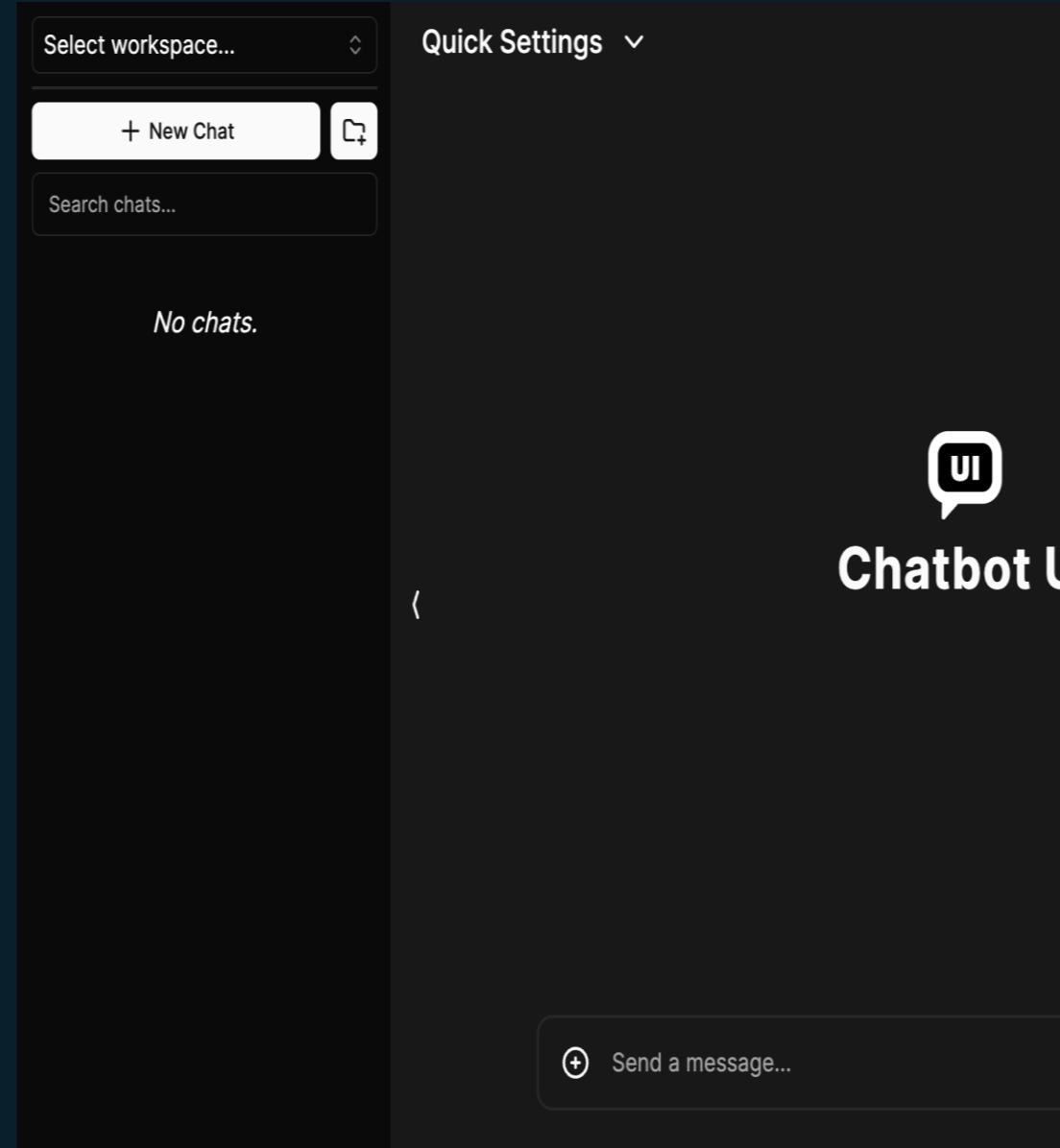
- Better productivity
- Boosted creativity
- Increased security



Analyze step by step

First Step – From Zero

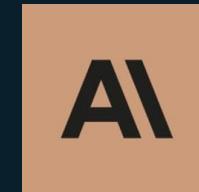
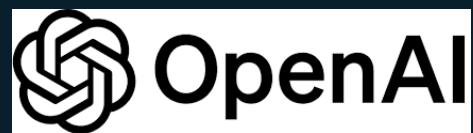
- POC a quick solution from a GitHub repository and connect it to OpenAI Api-Key
- Understand the chat conversational
- Deploy it on our Private Cloud
- Test the value and usage



Chatbot L

Second Step - Benchmark LLM

- What is the best LLM in Sept 2023 ? And the most popular one ?
 - OpenAI, Gemini (Bard), Claude, LLaMa, ..
- What is the secured way to use it ?
 - SaaS available solution, and ideally a partner
- Which one has the best quality/cost ?



Second Step - Benchmark LLM

Model	Publisher	Open Source	Cost	Performance
GPT-4	 OpenAI	✗ No	★★★★★	★★★★★
GPT-3.5	 OpenAI	✗ No	★★★★★	★★★★★
LLaMA 7B	 Meta	✓ Yes	★★★★★	★★★★★
LLaMA 65B	 Meta	✓ Yes	★★★★★	★★★★★
Bard	 Google	✗ No	★★★★★	★★★★★
Claude	 A10	✗ No	★★★★★	★★★★★

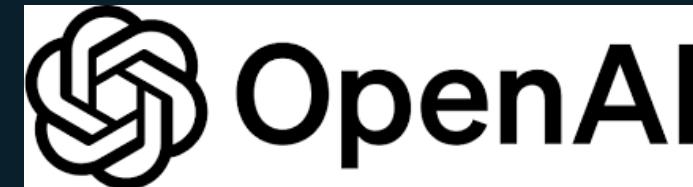
Second Step - Benchmark LLM Winner

- OpenAI was the most performant and popular one we decided to go with

Model	Hosting	Pricing	Security
OpenAI Plus	 OpenAI	★★★★★	★★★★★
OpenAI API	 OpenAI	★★★★★	★★★★★
OpenAI Enterprise	 OpenAI	★★★★★	★★★★★
Microsoft Azure OpenAI	 Azure	★★★★★	★★★★★

Second Step - Benchmark LLM & Host Winner

- Microsoft Azure providing OpenAI API at Enterprise Grade level with a better cost management



Business Requirement

- Microsoft partnership ✓
- Security Assessment ✓
- Region : Europe only ✓
- SSO ✓
- No User conversation storage ✗



Microsoft



Architecture

Development Requirement

V1:

- Chat Conversation
 - Chat History (saved on the browser)
 - Chat context length
- SSO
- KPI
- Tests
 - Security tests
 - load test
 - Azure APIM, OpenAI
 - Network between Givaudan – AWS – Azure
 - Application



Development Requirement

V2 (Experimentation):

- Document Upload
- RAG
- Image Generation



Roadmap



Architecture - Technical Stack

Dev

- Angular Frontend 
- Python Langchain Backend 

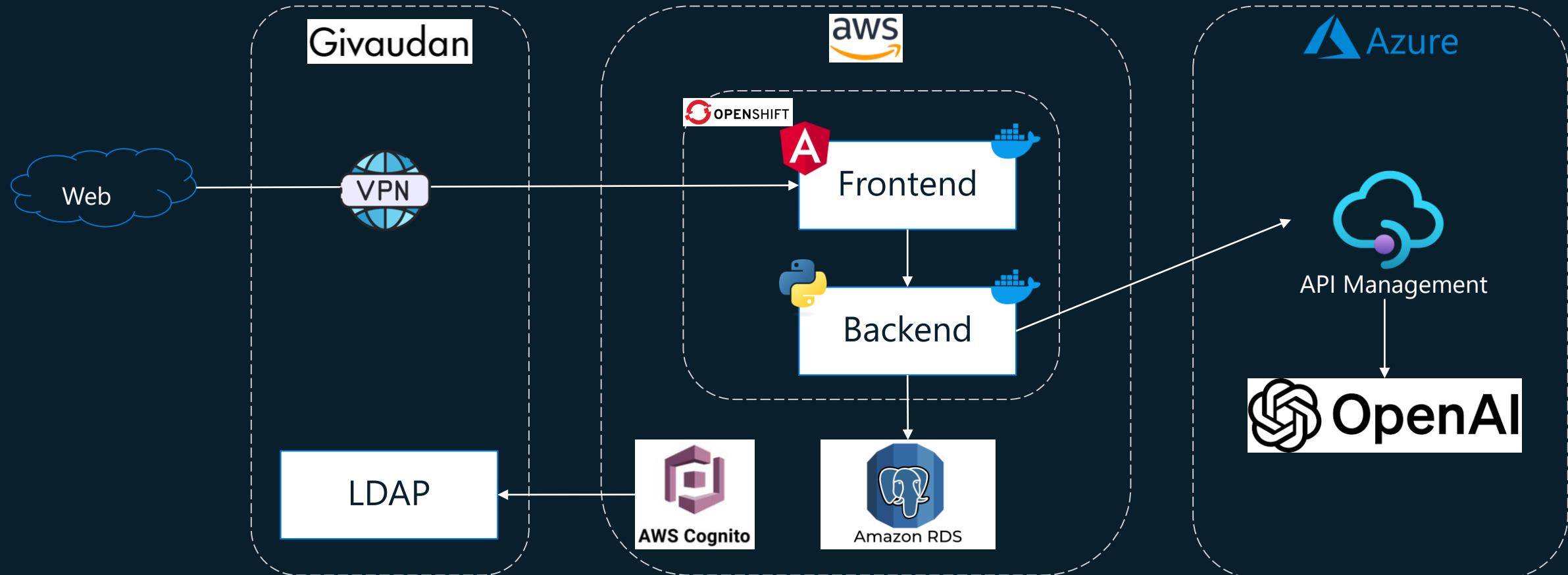
DevOps

- Gitlab with Gitlab-ci 
- Helm with ArgoCD 

Infrastructure

- AWS with Openshift  
- Docker 
- Azure OpenAI  
- PostgreSQL 

Architecture - Technical Stack



Basics – Chat Completion

Request

```
POST
https://{{endpoint}}/openai/deployments/{{deployment-id}}/chat/completions?api-version=2024-10-21

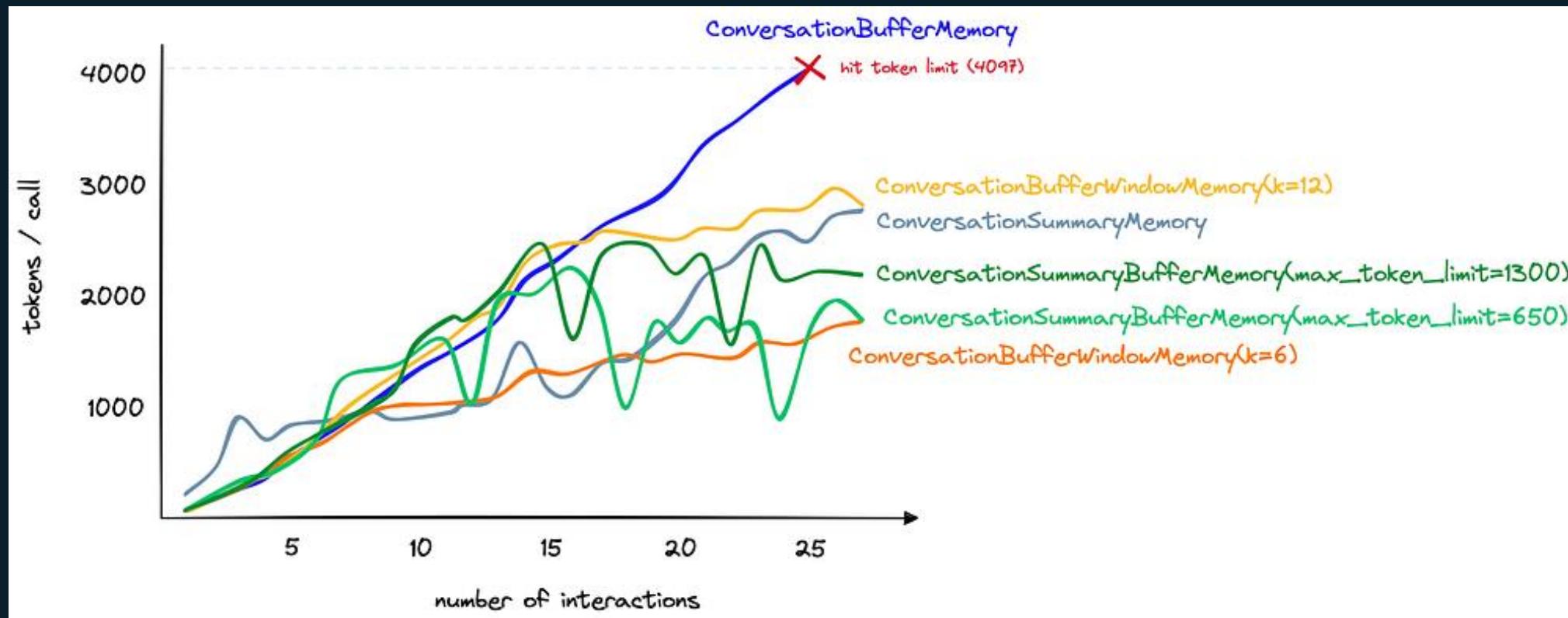
{
  "model": "gpt-4o",
  "messages": [
    {
      "role": "system",
      "content": "you are a helpful assistant that talks like a pirate. Generate short answer only."
    },
    {
      "role": "user",
      "content": "can you tell me how to care for a parrot?"
    }
  ],
  "temperature": 0.7,
  "stream": false,
  "max_tokens": 2000
}
```

Response

```
{
  "choices": [
    {
      "finish_reason": "stop",
      "index": 0,
      "message": {
        "content": "Aye! Provide fresh food n' water, a roomy cage, toys fer enrichment, regular vet check-ups, and plenty o' social interaction. Arr!",
        "role": "assistant"
      }
    }
  ],
  "created": 1740803441,
  "id": "chatcmpl-B68mnFNQafMCjwosWz8zPS1ZtcRVO",
  "model": "gpt-4o-2024-05-13",
  "object": "chat.completion",
  "system_fingerprint": "fp_65792305e4",
  "usage": {
    "completion_tokens": 32,
    "prompt_tokens": 39,
    "total_tokens": 71
  }
}
```

How to manage Conversational Memory ?

- With Conversational Buffer or Summary



Basics – Create Image

Request

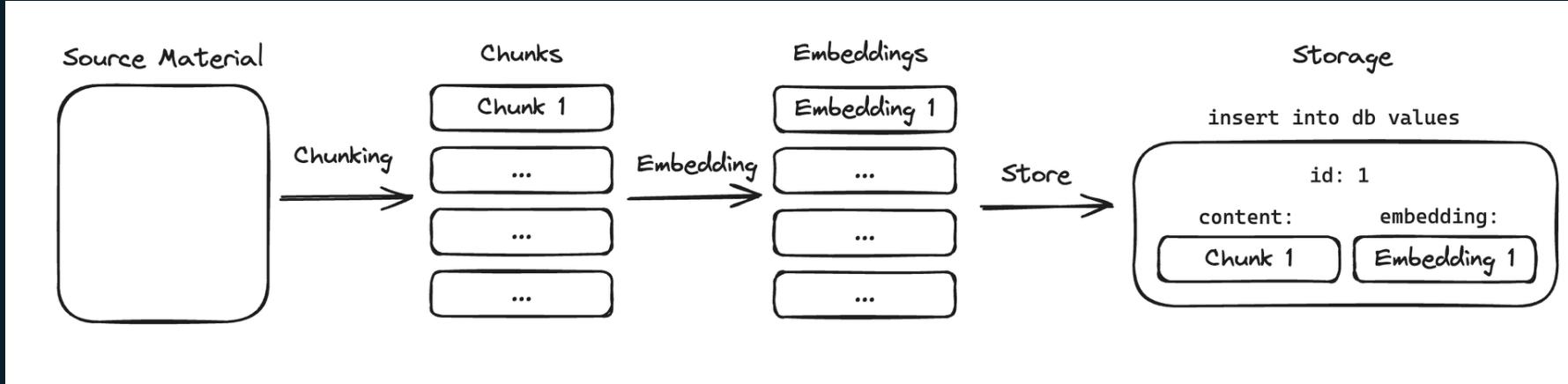
```
POST  
https://{{endpoint}}/openai/deployments/{{deploymen  
t-id}}/image/generations?api-version=2024-02-01  
  
{  
  "prompt": "A multi-colored umbrella on the beach,  
  disposable camera",  
  "size": "1024x1024",  
  "n": 1,  
  "quality": "standard",  
  "style": "vivid",  
  "response_format": "b64_json"  
}
```

Response

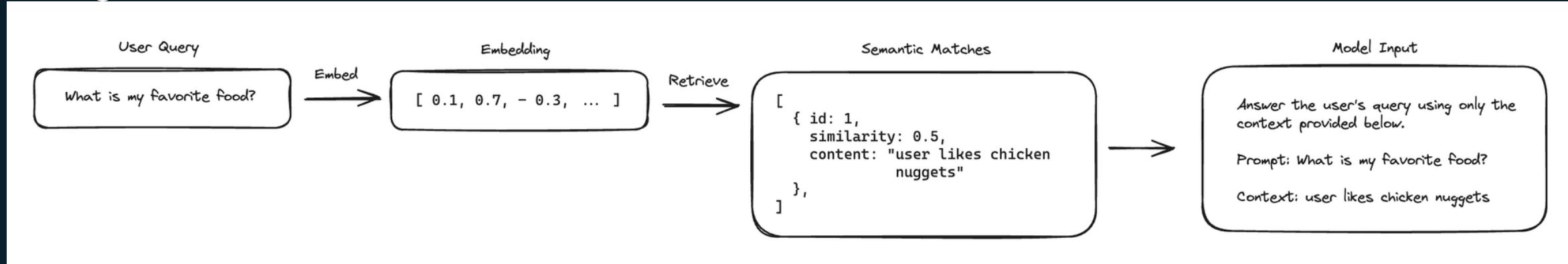
```
{  
  "created": 1740805367,  
  "data": [  
    {  
      "content_filter_results": { ... },  
      "prompt_filter_results": { ... },  
      "revised_prompt": "A multi-colored umbrella situated on a  
      sandy beach with its vibrant colors contrasting the blue sky.  
      Near by on the sand, there's a disposable camera ready to capture  
      beautiful memories.",  
      "url":  
        "https://dalleprodsec.blob.core.windows.net/private/images/1352de  
        45-3e6d-418e-b6f1-0d4ed592a78c/generated_00.png?se=2025-03-  
        02T05%3A02%3A57Z&sig=H5eZKn000A%2B1lYFpZYwnDGHVgVgA7JwSG0Q76ZvTuv  
        A%3D&ske=2025-03-07T13%3A56%3A15Z&skoid=e52d5ed7-0657-4f62-bc12-  
        7e5dbb260a96&skt=2025-02-28T13%3A56%3A15Z&sktid=33e01921-  
        4d64-4f8c-a055-5bdaffd5e33d&skv=2020-10-  
        02&sp=r&spr=https&sr=b&sv=2020-10-02"  
    }  
  ]  
}
```

Basic – RAG (Retrieval-Augmented Generation)

Ingestion:



During Chat:



[source](#)

Demo

Alfred

Cihan Cinar



Alfred Demo – Chat

The screenshot shows the Alfred Demo - Chat interface. At the top, there are user and AI icons, followed by the text "Alfred" and "ChatGPT-4o". A dropdown menu is open next to "ChatGPT-4o". On the right, there is a "Export chat" button. The main area is a chat window with a sidebar on the left.

Chat history (2) Q ...

can you tell me how to ca ... edit trash
01 Mar

Hello
28 Feb

cc Can you tell me how to care for a parrot? Generate a short answer.

cc Caring for a parrot involves several key aspects:

1. **Diet:** Provide a balanced diet of high-quality pellets, fresh fruits, and vegetables. Avoid avocado, chocolate, and caffeine, as they are toxic to parrots.
2. **Housing:** Ensure a spacious cage with perches, toys, and safe materials. The cage should be cleaned regularly.
3. **Social Interaction:** Spend time with your parrot daily. They are social creatures and thrive on interaction.
4. **Mental Stimulation:** Offer a variety of toys and activities to keep them engaged and prevent boredom.
5. **Health Care:** Schedule regular check-ups with an avian veterinarian to monitor their health.

By following these guidelines, you can help ensure your parrot is happy and healthy!

undo redo

Support

cc Cihan Cinar

Prompt templates

Send a message... Send

Alfred can make mistakes. Consider checking important information.

Alfred Demo – Image Generation

Alfred

Image Generation

Chat history 2

A multi-colored umbrella ...

01 Mar

Hello

28 Feb

Support

Cihan Cinar

Send a message...

Prompt templates

Alfred can make mistakes. Consider checking important information.

cc A multi-colored umbrella on the beach, disposable camera

A photograph of a multi-colored umbrella (red, orange, yellow, green, blue, purple) standing upright in the sand on a beach. A small black disposable camera is positioned in front of the umbrella. The background shows the ocean and a clear blue sky.

Alfred Demo – RAG

France: Research to fight lavender diseases | Givaudan Foundation — Mozilla Firefox

https://www.givaudan-foundation.org/projects/france-lavender

Givaudan Foundation
For communities & nature

Home > France: Lavender diseases and climate change

France: Research to fight lavender diseases

The culture of lavender and lavandin has been a traditional activity in South-East France for several centuries and is today part of some of the most scenic landscapes of the Provence region. However, this emblematic crop is threatened by diseases, pests and by the effects of climate change.



domyia (midge) is one of the major pests of lavender and lavandin. This insect's larvae feed on the sap of the plant and cause a progressive desiccation of affected branches. The damage is

Alfred

Chat history 2

What can you tell about lavender in France ?

01 Mar

Hello

28 Feb

ChatGPT-4o

Export chat

What can you tell about lavender in France ?

Lavender cultivation in France, particularly in the South-East region, has been a traditional activity for several centuries and is integral to the scenic landscapes of Provence. However, this crop faces significant threats from diseases, pests, and the impacts of climate change. For example, the cecidomyia (midge) is a major pest that affects lavender and lavandin, causing progressive desiccation of the plants and reducing their productivity and lifespan. To combat these issues, research organizations like CRIEPPAM are developing new approaches based on a better understanding of the insect's life cycle[1].

Additionally, more frequent drought periods induced by climate change pose another challenge for lavender producers. To address this, the French research institute ITEIPMAI is working on developing a tool to monitor the agronomic effects of water stress in lavender fields. This decision support system aims to help producers manage water more efficiently and evaluate the development of standard indicators to monitor climate change[2].

Furthermore, Givaudan has partnered with regenerative agriculture technical experts to carry out comprehensive soil audits in key lavender-producing areas. These assessments, which include soil profiles and analysis, aim to define adapted cultivation and water management practices to address local challenges faced by producers[3].

[1] [2] france-lavender.html
[3] giv-2023-integrated-report.pdf, 125.

Prompt templates

Send a message...

Alfred can make mistakes. Consider checking important information.

giv-2023-integrated-report.pdf

Complete document Document extracts

121

Our work to improve and protect water quality and quantity is a priority natural supply chain initiative projects such as one of our lavender fields in the South-East of France. In partnership with regenerative agriculture experts, we are carrying out comprehensive soil audits in key lavender-producing areas. These audits have gained easier access to clean water, which is essential for the growth and health of the lavender plants over time.

Through the Givaudan foundation, we also improved the living conditions of bees by collecting information on a wide range of bee species and their habitats. We have increased our understanding of the needs of bees and have gained easier access to clean water, which is essential for the growth and health of the lavender plants over time.

In addition to our work on water and our efforts to promote global access to water, our Givaudan Foundation is also involved in the development of new technologies to create new sources of water. Using the results of our soil audits, we have developed a new technology that can extract 500 liters of water from the soil. This technology has been developed by our research and development team and is currently being tested in various locations around the world.

We partner with experts to develop water conservation practices recommendations in supply chains such as lavender in France.

3 pages - 2023 Integrated Report - 10

Waste management and circular principles

Explanation of the material topic

Our impact on waste management and circular principles are discussed along with the various ways we can use upcycled materials to generate income such as recycling, composting, and reusing. We also discuss the various ways we can reduce waste by using our products and sharing our knowledge with others.

Promoting the environment means acting on our waste management and circular principles to ensure that our waste is minimally generated and that the waste we do generate is managed in a way that is sustainable and responsible. Our decisions about how we manage our waste can have a significant impact on the environment and we strive to make sure that our waste management practices are aligned with our environmental goals.

Impacts

Risks & Opportunities

Waste management and circular principles are important for our business, as they help us to reduce the amount of resources we use and the amount of waste we generate. By using upcycled materials, we can reduce the amount of waste we generate and the amount of waste we need to dispose of. This can help us to reduce our environmental impact and contribute to climate change and air pollution. Circular principles, such as recycling and composting, can offer our business and society many benefits, including reduced waste generation and reduced energy consumption.

Management of the impact

When discussing waste, we can use general terms such as 'reduce', 'reuse', and 'recycle', or more specific terms such as 'upcycling' and 'circular economy'. We can also use terms such as 'waste reduction', 'waste minimization', and 'waste prevention' to describe the various ways we can manage waste.

Waste management and circular principles are important for our business, as they help us to reduce the amount of resources we use and the amount of waste we generate. By using upcycled materials, we can reduce the amount of waste we generate and the amount of waste we need to dispose of. This can help us to reduce our environmental impact and contribute to climate change and air pollution. Circular principles, such as recycling and composting, can offer our business and society many benefits, including reduced waste generation and reduced energy consumption.

Waste management

REDUCTION

REUSE

RECYCLING

DISPOSAL

DISPOSAL

Innovative solutions targeting waste

3 pages - 2023 Integrated Report - 10

Alfred Demo – Document Upload

The screenshot shows the Alfred AI interface. On the left, a sidebar includes a 'Chat history' section with a message from 'Cihan Cinar' and a message from 'Hello'. The main area shows a conversation with 'ChatGPT-4o'. Cihan Cinar asks, 'What can you tell about lavender in France ?' ChatGPT-4o responds with a file attachment 'document-1.pdf'. Alfred's response states: 'The provided document chunks do not contain specific information about lavender in France. However, it does mention that Givaudan is involved in projects related to lavandin and fine lavender in the South of France. These projects focus on developing water conservation practices and conducting comprehensive soil audits in key producing areas. Unfortunately, further details about lavender itself or specific initiatives were not included.' Below this, Alfred says, 'If you have any other questions or need information on a different topic, feel free to ask!' At the bottom, there is a message input field with 'document-1.pdf' attached and a 'Send a message...' button.

Alfred

ChatGPT-4o

Export chat

Chat history 2

What can you tell about la... 01 Mar

Hello 28 Feb

What can you tell about lavender in France ?

document-1.pdf

The provided document chunks do not contain specific information about lavender in France. However, it does mention that Givaudan is involved in projects related to lavandin and fine lavender in the South of France. These projects focus on developing water conservation practices and conducting comprehensive soil audits in key producing areas. Unfortunately, further details about lavender itself or specific initiatives were not included.

If you have any other questions or need information on a different topic, feel free to ask!

document-1.pdf

Prompt templates

Send a message...

Alfred can make mistakes. Consider checking important information.

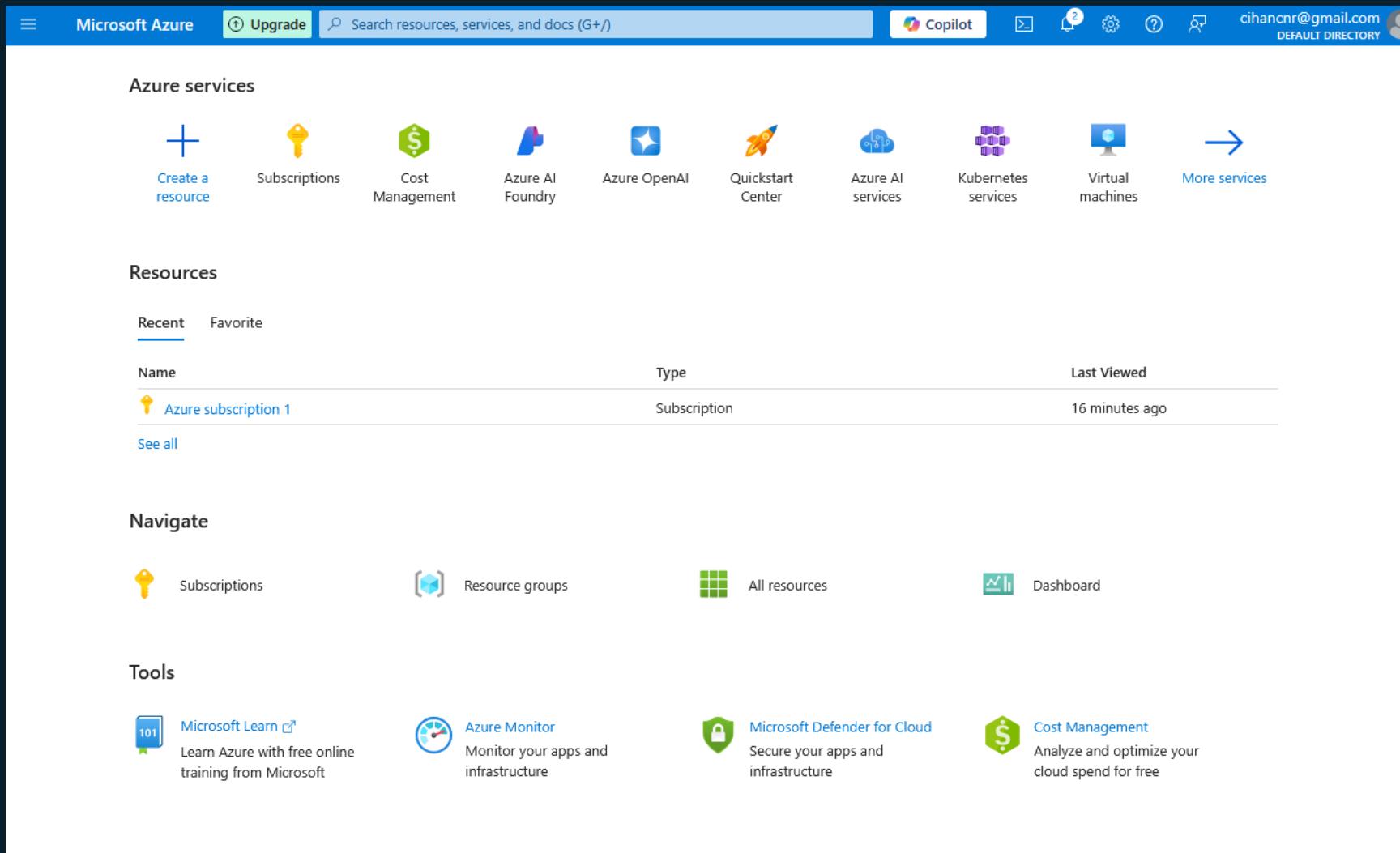
Demo

Azure

Cihan Cinar



Azure Demo –



The screenshot shows the Microsoft Azure portal homepage with a dark theme. At the top, there is a navigation bar with the Microsoft Azure logo, an 'Upgrade' button, a search bar, and various icons for account management and help. The main content area is titled 'Azure services' and features a grid of icons for creating a resource, managing subscriptions, cost management, Azure AI Foundry, Azure OpenAI, Quickstart Center, Azure AI services, Kubernetes services, and Virtual machines, along with a 'More services' link. Below this is a 'Resources' section with tabs for 'Recent' (which is selected) and 'Favorite', showing a single item: 'Azure subscription 1' (Subscription, last viewed 16 minutes ago). There is a 'See all' link. The 'Navigate' section contains links for Subscriptions, Resource groups, All resources, and Dashboard. The 'Tools' section includes links for Microsoft Learn (101 items), Azure Monitor, Microsoft Defender for Cloud, and Cost Management.

Microsoft Azure Upgrade Search resources, services, and docs (G+) Copilot

cihancnr@gmail.com DEFAULT DIRECTORY

Azure services

Create a resource Subscriptions Cost Management Azure AI Foundry Azure OpenAI Quickstart Center Azure AI services Kubernetes services Virtual machines [More services](#)

Resources

Recent Favorite

Name	Type	Last Viewed
Azure subscription 1	Subscription	16 minutes ago

[See all](#)

Navigate

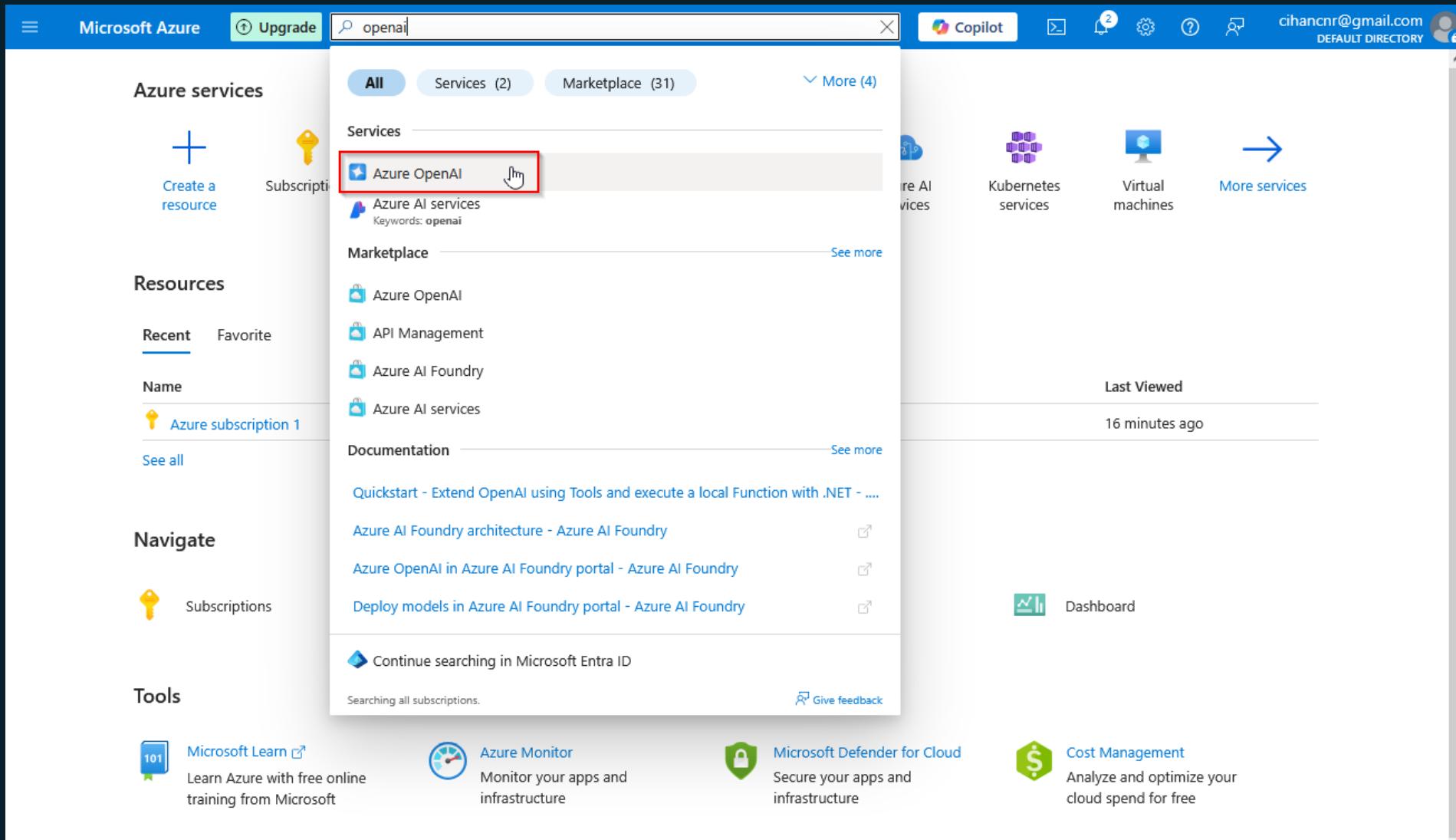
Subscriptions Resource groups All resources Dashboard

Tools

Microsoft Learn 101 Learn Azure with free online training from Microsoft Azure Monitor Monitor your apps and infrastructure Microsoft Defender for Cloud Secure your apps and infrastructure Cost Management Analyze and optimize your cloud spend for free

[Try Azure for Free](#)

Azure Demo – OpenAI Deployment



The screenshot shows the Microsoft Azure portal interface. The search bar at the top contains the text "openai". The search results are displayed under the "Services" category, with "All" selected. The "Azure OpenAI" service is highlighted with a red box. Other services listed include "Azure AI services" and "Azure AI Foundry". To the right of the search results, there are sections for "Kubernetes services", "Virtual machines", and a "More services" button. The left sidebar includes sections for "Create a resource", "Recent" and "Favorite" resources, "Name" (with "Azure subscription 1" selected), "Documentation", "Navigate" (with "Subscriptions" selected), and "Tools" (with "Microsoft Learn" and "Azure Monitor" options). The bottom navigation bar includes links for "Dashboard", "Cost Management", "Microsoft Defender for Cloud", "Azure Monitor", and "Microsoft Learn".

Azure Demo – OpenAI Deployment

Microsoft Azure Upgrade Search resources, services, and docs (G+/) Copilot 2 ? ? ? cihancnr@gmail.com DEFAULT DIRECTORY

Home > Azure AI services

Azure AI services | Azure OpenAI

Azure AI services

Search + Create Manage deleted resources Manage view Refresh Export to CSV Open query Assign tags Delete

Filter for any field... Subscription equals all Type equals all Resource group equals all Location equals all Add filter

Showing 0 to 0 of 0 records.

Name ↑↓	Kind ↑↓	Location ↑↓	Custom Domain ... ↑↓	Pricing tier ↑↓	Status ↑↓
No grouping List view					

 No Azure OpenAI to display

Perform a wide variety of natural language tasks.

+ Create Azure OpenAI

[Learn more](#)

Overview All Azure AI services + Azure AI services Azure AI services Azure OpenAI AI Search Computer vision Face API Custom vision Speech service Language service Translator Document intelligence Bot services

Azure Demo – OpenAI Deployment

Microsoft Azure Upgrade Search resources, services, and docs (G+/-) Copilot 2 ? ? ? ? cihancnr@gmail.com DEFAULT DIRECTORY

Home > Azure AI services | Azure OpenAI > Create Azure OpenAI ...

Basics Network Tags Review + submit

Azure OpenAI Service provides access to OpenAI's powerful language models, including all the latest OpenAI models. These models can be easily adapted to your specific tasks, including but not limited to content generation, summarization, image understanding, semantic search, and natural language to code translation. Top use cases include Call Centers, Virtual Assistants, Accessibility, Content Generation, and Code Development. The service also features the Assistants API, Fine Tuning capabilities and many ways to connect your data to the service for conversational experiences. The service can be scaled through Standard (tokens) and Provisioned (PTUs) deployment types.

[Learn more](#)

Project Details

Subscription ① Azure subscription 1

Resource group ① (New) globalai-resource-group [Create new](#)

Instance Details

Region ① East US

Name ① globalai-openai-001

Pricing tier ① Standard S0

[View full pricing details](#)

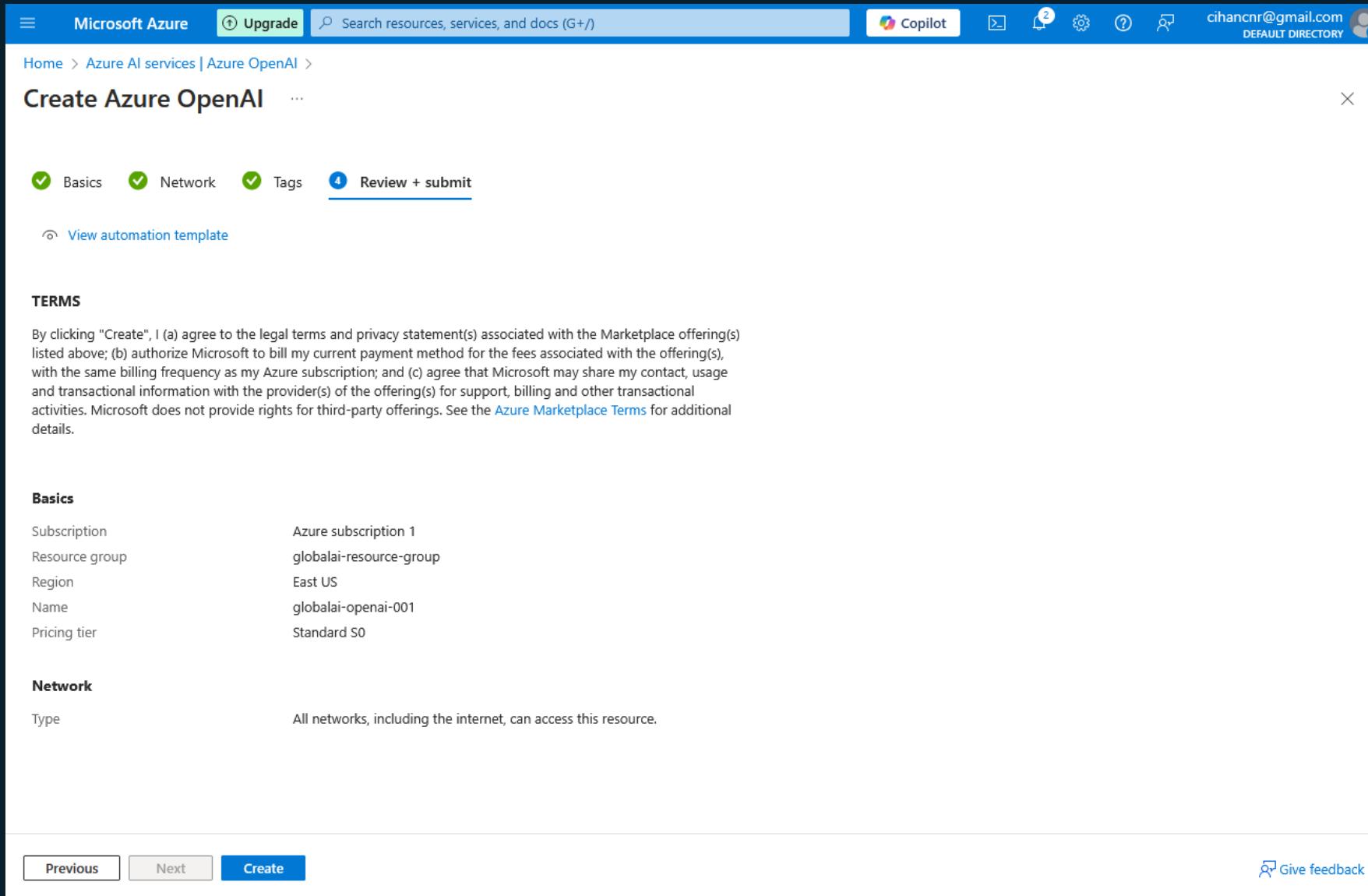
Content review policy

To detect and mitigate harmful use of the Azure OpenAI Service, Microsoft logs the content you send to the Completions and image generations APIs as well as the content it sends back. If content is flagged by the service's filters,

Previous Next

Give feedback

Azure Demo – OpenAI Deployment



The screenshot shows the Microsoft Azure portal interface for creating an Azure OpenAI service. The top navigation bar includes 'Microsoft Azure', 'Upgrade', a search bar, 'Copilot', and user information for 'cihancnr@gmail.com'. The main title is 'Create Azure OpenAI'. The navigation bar shows the current step: 'Review + submit' (step 4 of 4). Below the navigation, there are links for 'View automation template' and 'TERMS'. The 'TERMS' section contains a detailed legal agreement. The 'Basics' section lists the following configuration: Subscription: Azure subscription 1, Resource group: globalai-resource-group, Region: East US, Name: globalai-openai-001, and Pricing tier: Standard S0. The 'Network' section indicates that all networks, including the internet, can access the resource. At the bottom, there are 'Previous' and 'Next' buttons, and a prominent 'Create' button.

Microsoft Azure Upgrade Search resources, services, and docs (G+/-)

Copilot

cihancnr@gmail.com DEFAULT DIRECTORY

Home > Azure AI services | Azure OpenAI >

Create Azure OpenAI

Basics Network Tags **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	Azure subscription 1
Resource group	globalai-resource-group
Region	East US
Name	globalai-openai-001
Pricing tier	Standard S0

Network

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

Previous Next **Create** Give feedback

Azure Demo – Azure AI Foundry

Azure AI Foundry | Azure OpenAI Service / globalai-openai-001 / Chat playground

All resources 5 🚧 🚧 globalai-openai-001 (eastus, S0) CC

Chat playground

Home Get started Model catalog Playgrounds Chat Assistants PREVIEW Real-time audio Images Completions Tools Fine-tuning Azure OpenAI Service Evaluation Stored completions Batch jobs Shared resources Deployments Quota Safety + security Data files Vector stores PREVIEW

View code Deploy Import Export Prompt samples Filters feedback

Setup

Deployment * + Create new deployment

No deployment exists

Give the model instructions and context

You are an AI assistant that helps people find information.

Apply changes Generate prompt

+ Add section

Few shot examples are not used when source data has been added. Any previous examples have been cleared. Once data is removed, then examples can be added.

Chat history

Start with a sample prompt

Poetry generation
Compose a poem about the beauty of nature in autumn.

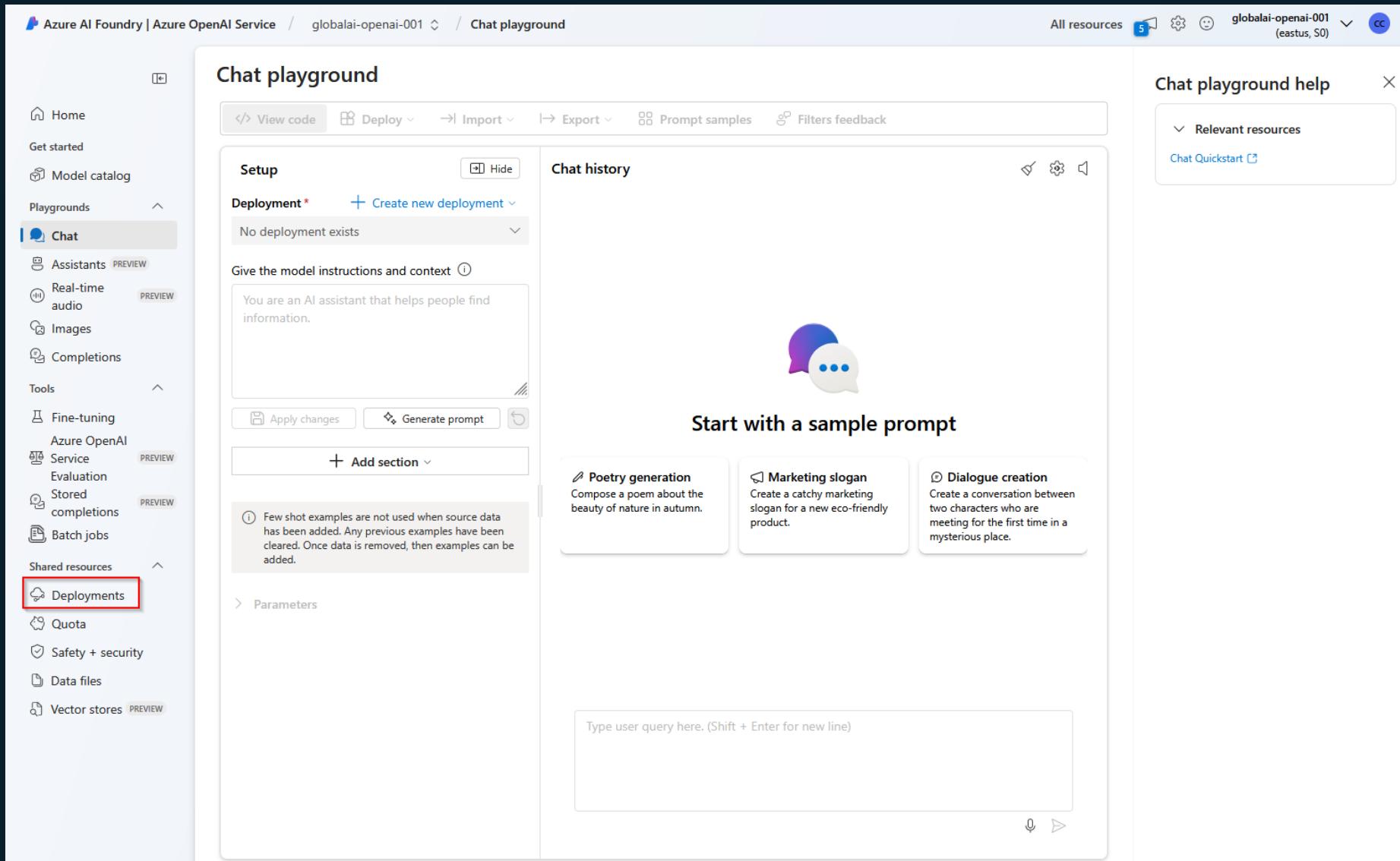
Marketing slogan
Create a catchy marketing slogan for a new eco-friendly product.

Dialogue creation
Create a conversation between two characters who are meeting for the first time in a mysterious place.

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

Chat playground help

Relevant resources Chat Quickstart



Azure Demo – ChatGPT Model Deployment

Azure AI Foundry | Azure OpenAI Service / globalai-openai-001 / Models + endpoints

All resources 5 🔍⚙️😊 globalai-openai-001 (eastus, S0) CC

Model deployments

Model deployments App deployments

+ Deploy model Refresh

Name

Deploy model gpt-4o-mini

Deployment name*
gpt-4o-mini

Deployment type
Global Standard

Global Standard: Pay per API call with the highest rate limits. Learn more about [Global deployment types](#).

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

Deployment details

Model version
2024-07-18 (Default)

Resource location
East US 2

1K tokens per minute quota available for your deployment

Tokens per Minute Rate Limit
1K

Corresponding requests per minute (RPM) = 10

When you create this deployment, the resource change you've made will persist throughout Azure AI Foundry.

A new AI resource will be created for your deployment

A resource location that supports the model has been pre-selected

Create resource and deploy Cancel

Deployments help

Relevant resources

Azure OpenAI Service deployment types

Working with Azure OpenAI Service models

Azure Demo –

The screenshot shows the Azure AI Foundry interface for managing model deployments. The main page title is "gpt-4o-mini". The left sidebar is titled "Deployments" and includes links for "Quota", "Safety + security", "Data files", and "Vector stores". The main content area shows deployment details, an endpoint, monitoring & safety settings, and useful links.

Deployment info

Name	gpt-4o-mini	Provisioning state	Succeeded
Deployment type	Global Standard	Created on	2025-03-01T06:12:56.6545513Z
Created by	cihancnr@gmail.com	Modified on	Mar 1, 2025 7:12 AM
Modified by	cihancnr@gmail.com	Version update policy	Once a new default version is available
Rate limit (Tokens per minute)	1,000	Rate limit (Requests per minute)	10
Model name	gpt-4o-mini	Model version	2024-07-18
Life cycle status	GenerallyAvailable	Date created	Jul 19, 2024 2:00 AM
Date updated	Aug 21, 2024 2:00 AM	Model retirement date	Jul 20, 2025 2:00 AM

Endpoint

Target URI: <https://cihan-m7pt0nf4-eastus2.cognitiveservices.azure.com/open ...>

Key: [XXXXXXXXXXXXXXXXXXXX](#)

Monitoring & safety

Content filter: DefaultV2

Useful links for application development

[Code sample repository](#) [Tutorial](#)

Deployments help

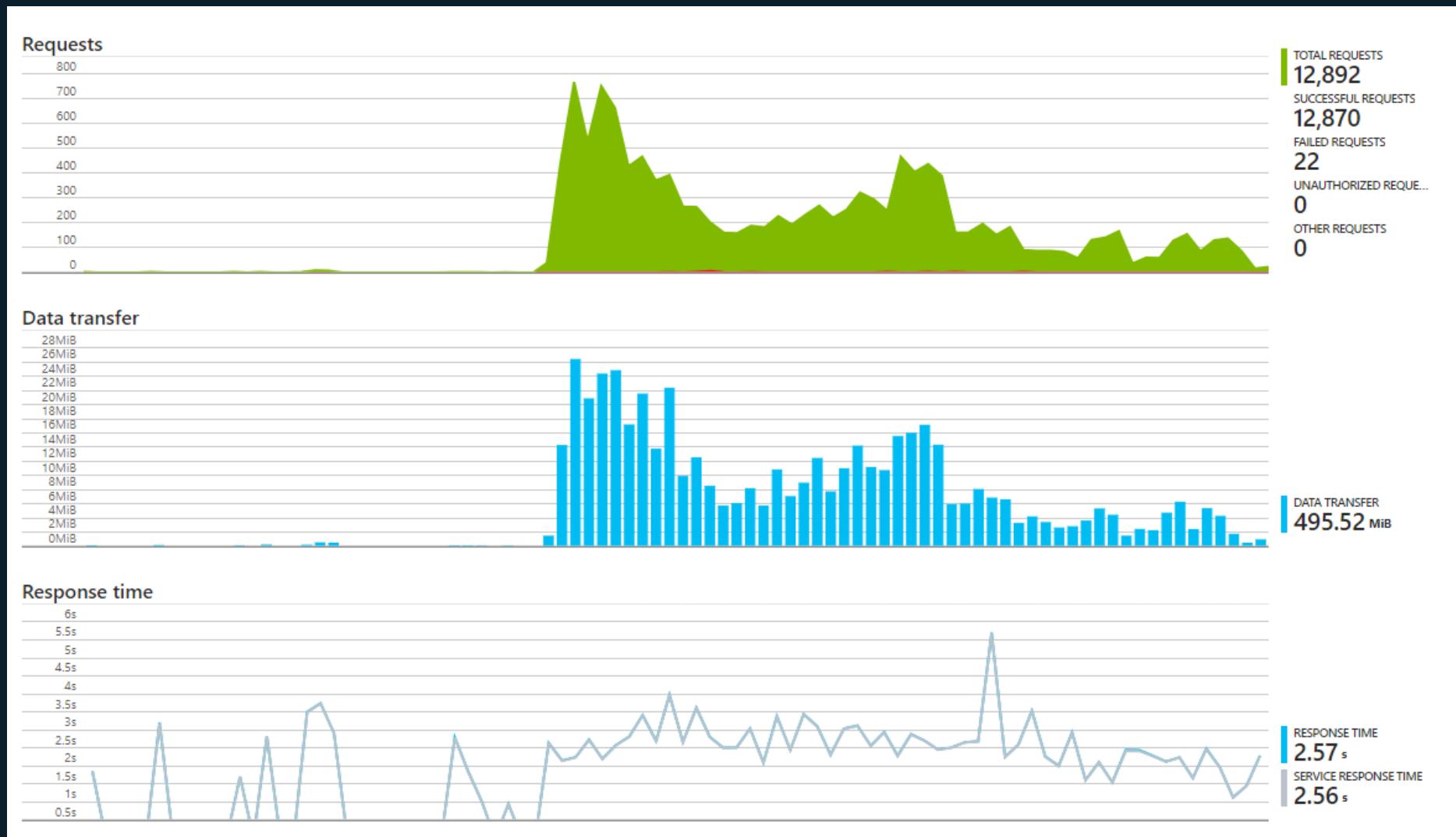
What are deployments in Azure AI Foundry? Azure AI Foundry supports deploying large language models (LLMs), flows, and web apps. Deploying an LLM or flow makes it available for use in a website, an application, or other production environments. [Learn more](#)

Relevant resources

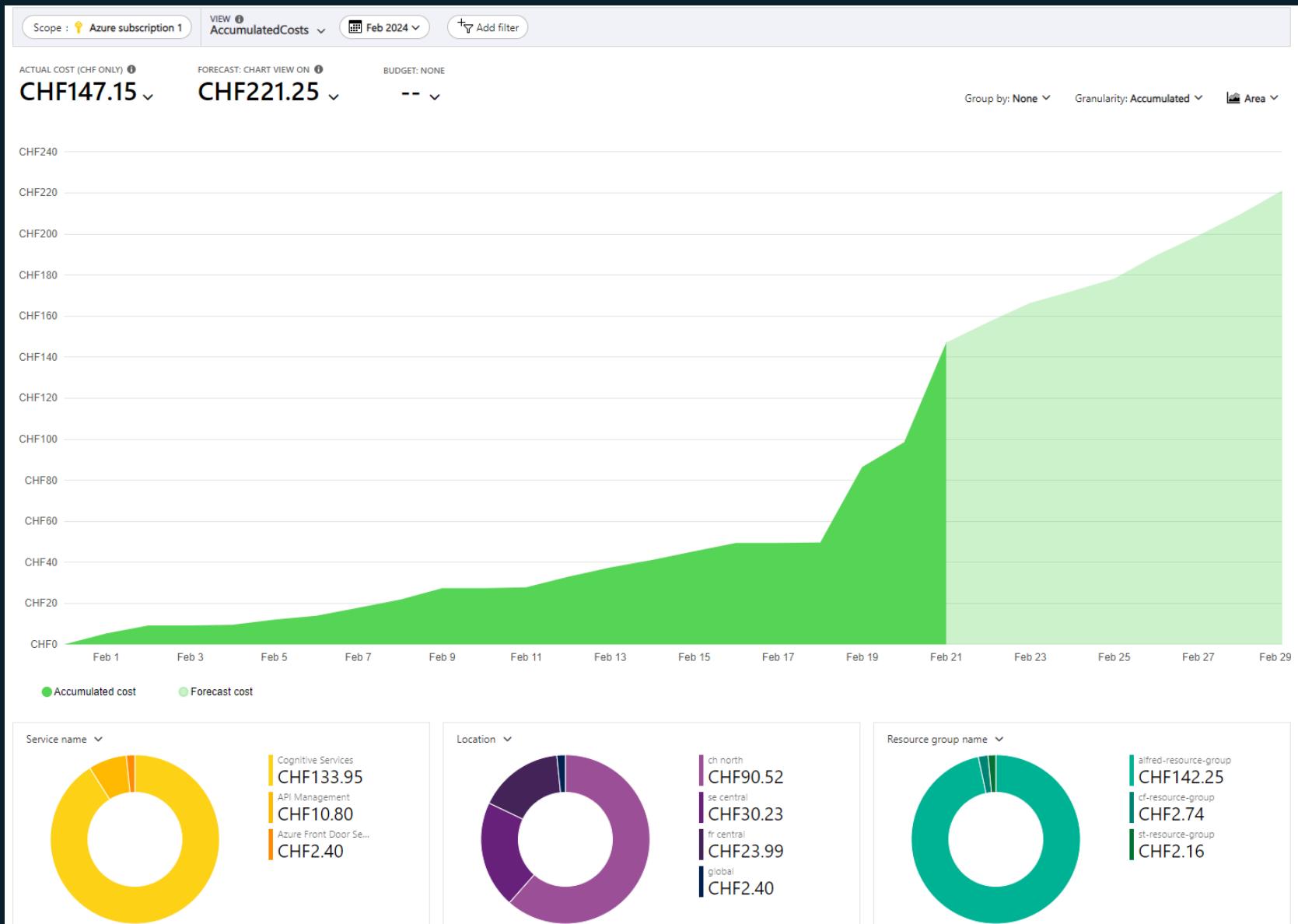
Azure Demo –

The screenshot shows the Azure AI Foundry interface, specifically the Chat playground section. The top navigation bar includes 'Azure AI Foundry | Azure OpenAI Service', the resource name 'cihan-m7pt0nf4-eastus2', and the 'Chat playground' page title. The top right features 'All resources' and a dropdown for 'cihan-m7pt0nf4-eastus2 (eastus2, S0)'. The left sidebar has a 'Chat' section selected, listing 'Assistants', 'Real-time audio', 'Images', 'Completions', 'Fine-tuning', 'Azure OpenAI Service', 'Stored completions', 'Batch jobs', 'Shared resources', 'Deployments', 'Quota', 'Safety + security', 'Data files', and 'Vector stores'. The main content area is titled 'Chat playground' and contains a 'Setup' section with a deployment dropdown set to 'gpt-4o-mini (version:2024-07-18)'. It includes a text input for 'Give the model instructions and context' with the placeholder 'you are a helpful assistant that talks like a pirate. Generate short answer only.' and buttons for 'Apply changes', 'Generate prompt', and 'Add section'. To the right is a 'Chat history' section showing a conversation: 'can you tell me how to care for a parrot?' followed by a response: 'Arrr, matey! To care for a parrot, ye need to provide a spacious cage, fresh food like fruits and veggies, clean water, and plenty o' toys for playin'. Below this is a 'Parameters' section and a text input field for 'Type user query here. (Shift + Enter for new line)'. The bottom right shows '92/128000 tokens to be sent' with a trash bin icon. A 'Chat playground help' sidebar on the right lists 'Relevant resources' and 'Chat Quickstart'.

Alfred – Azure Monitor



Alfred – Cost Analysis

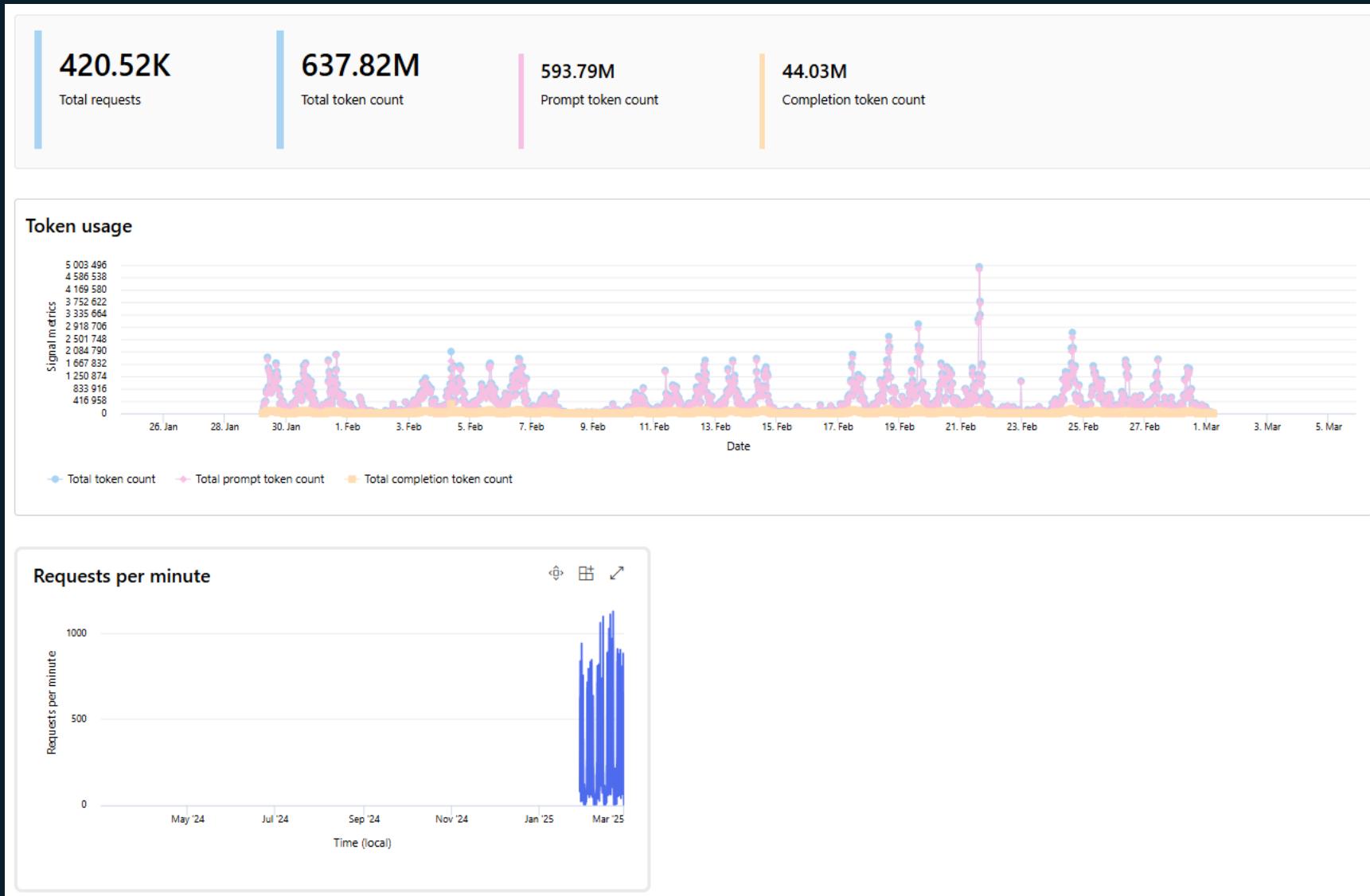


Conclusion

Result – To Hero

- Really good adoption 
- Result better than expected 
- Users ask for more 
- More feature incoming for 2025 

Alfred – Azure OpenAI Monitor To Hero



Next Steps

- Multiple documents Upload
- More RAG and improve Industrialization
 - Example: Azure AI Search
- Vision Capabilities
- Agent Development
 - Web Search (Bing)
 - Image Generation to Stable Diffusion
- Multiple Models
- And even more...

Insights



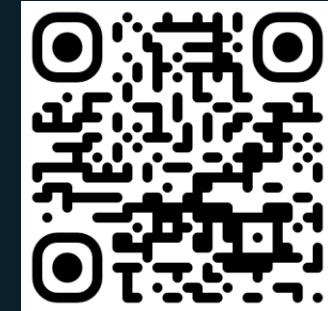
- POC quick
- Keep It Simple
- Test with User and listen to feedback
- Benchmark solutions
- Follow Trends and be adaptable, AI is moving constantly



Cihan Cinar

Software & AI Architect
Givaudan

<https://blog.cihan-cinar.com/>



Thank you

Useful links

- <https://github.com/Azure-Samples/azure-ai>
- <https://github.com/Azure-Samples/>
- <https://github.com/Azure/GPT-RAG>