

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

Cihan Cinar



Sponsors



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.

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



Centrilogic

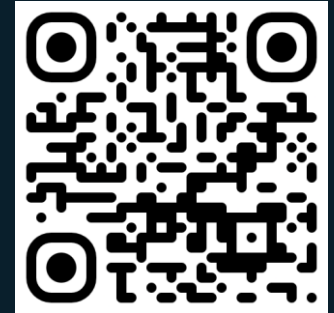
Alfred Givaudan's Internal ChatGPT Journey From Zero to Hero



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





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



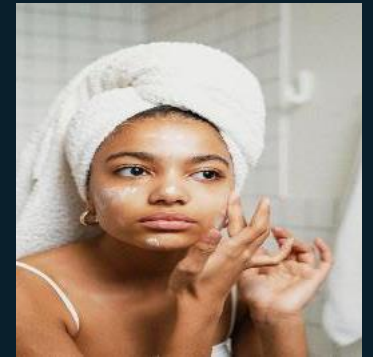
Sweet goods



Home care



Fabric care

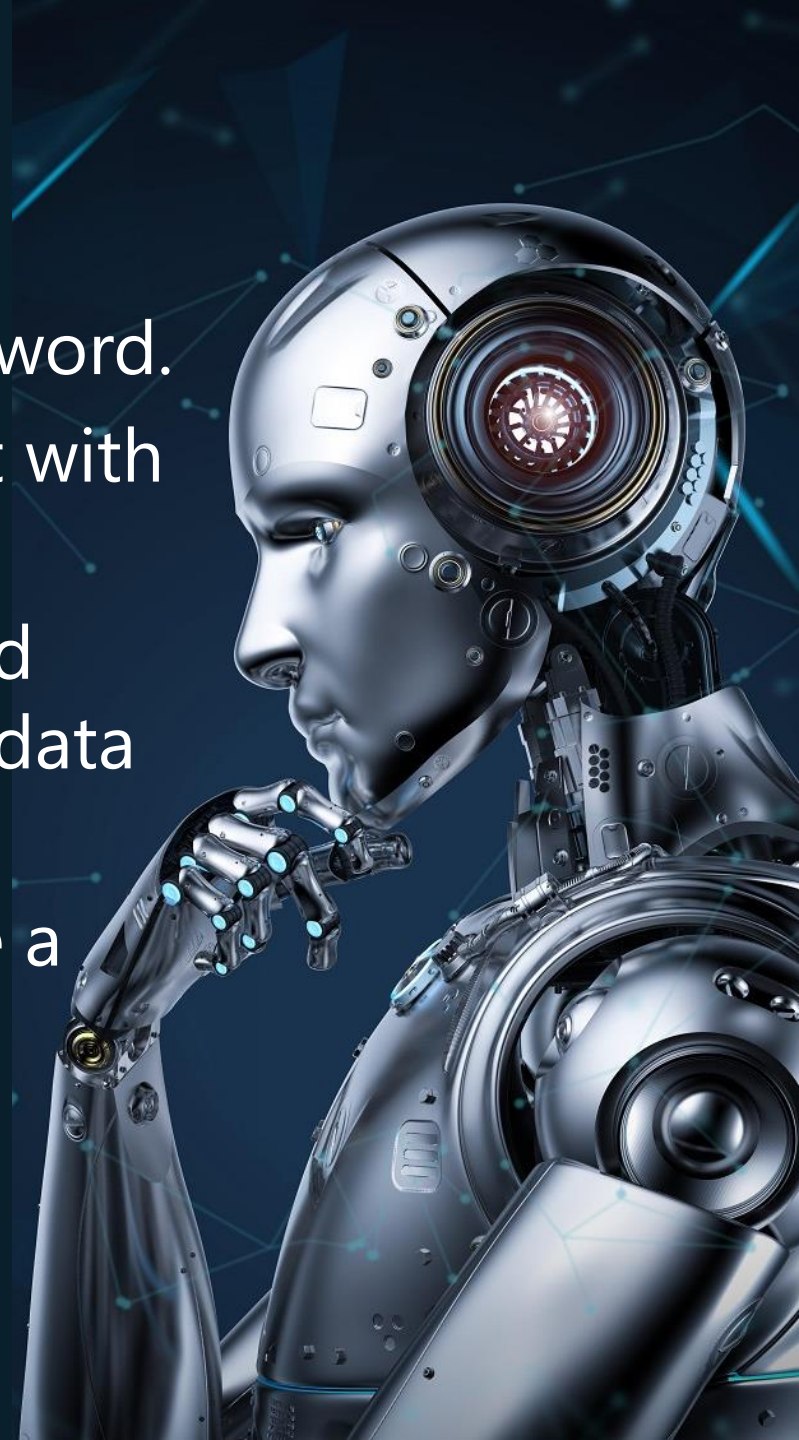


Beauty & skin care

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:

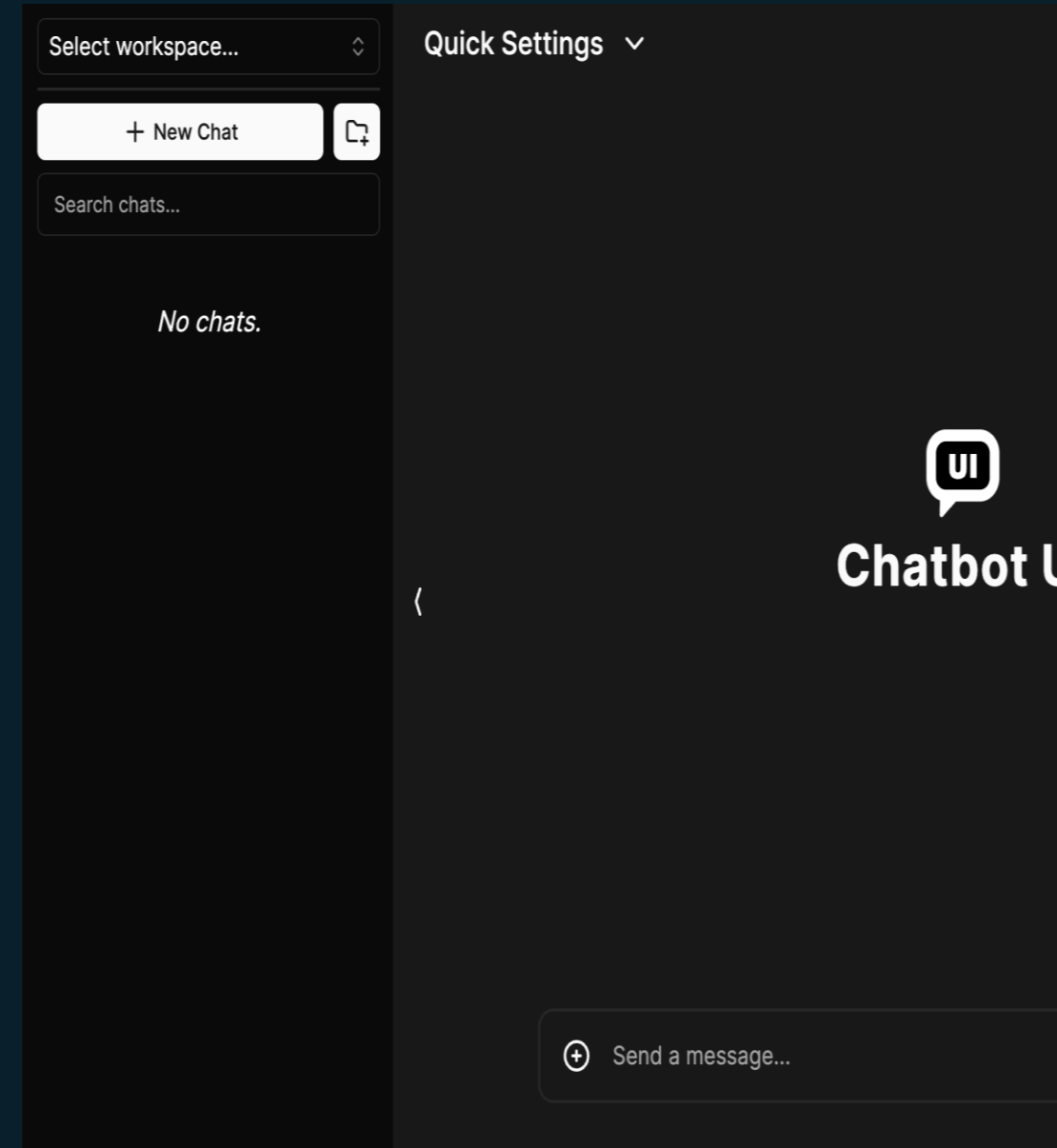
- Better productivity
- Boosted creativity
- Increased security



Analyze step by step

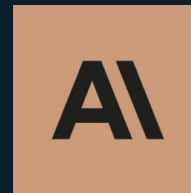
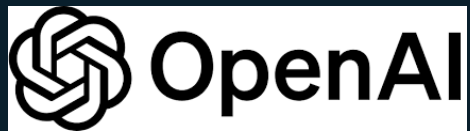
First Step – From Zero

- 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









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	 Anthropic	✗ No	★ ★ ★ ★ ★	★ ★ ★ ★ ★

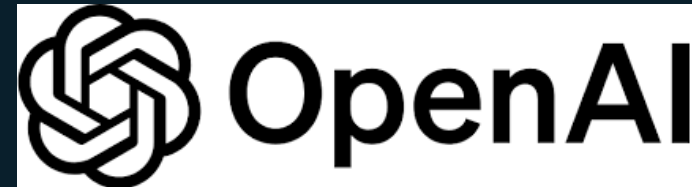
Second Step - Benchmark LLM Winner

- 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 ✗



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 

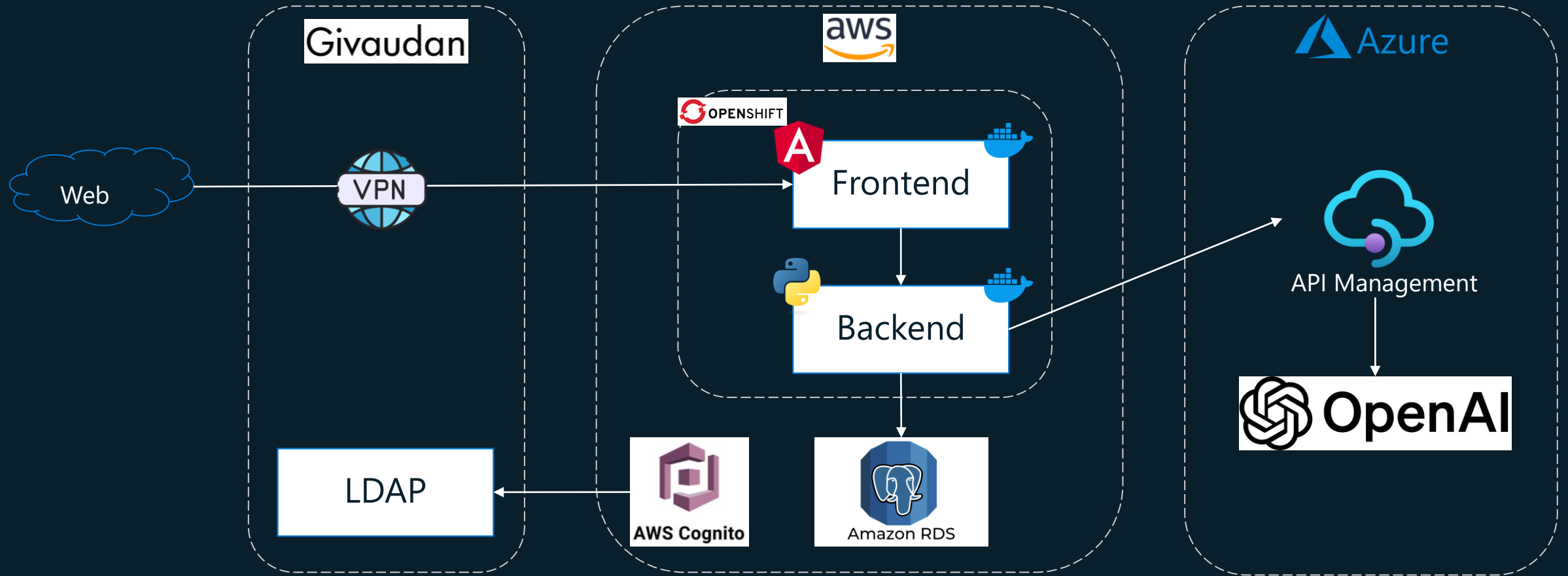
DevOps

- 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

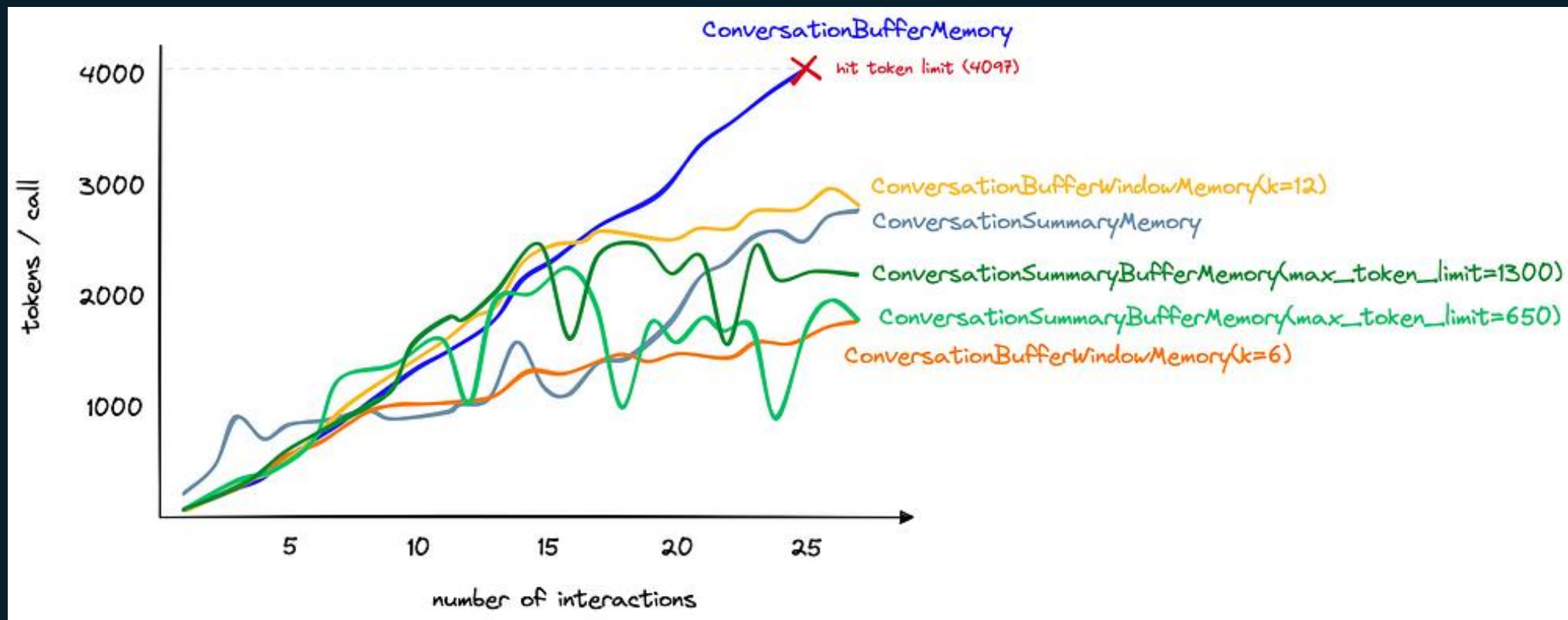
{
  "model": "gpt-4o",
  "messages": [
    {
      "role": "system",
      "content": "you are a helpful assistant that talks like a pirate. Generate short answer only."
    },
    {
      "role": "user",
      "content": "can you tell me how to care for a parrot?"
    }
  ],
  "temperature": 0.7,
  "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-B68mnFNQafMCjwosWz8zPS1ZtcRV0",
  "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/{deployment-id}/image/generations?api-version=2024-02-01
```

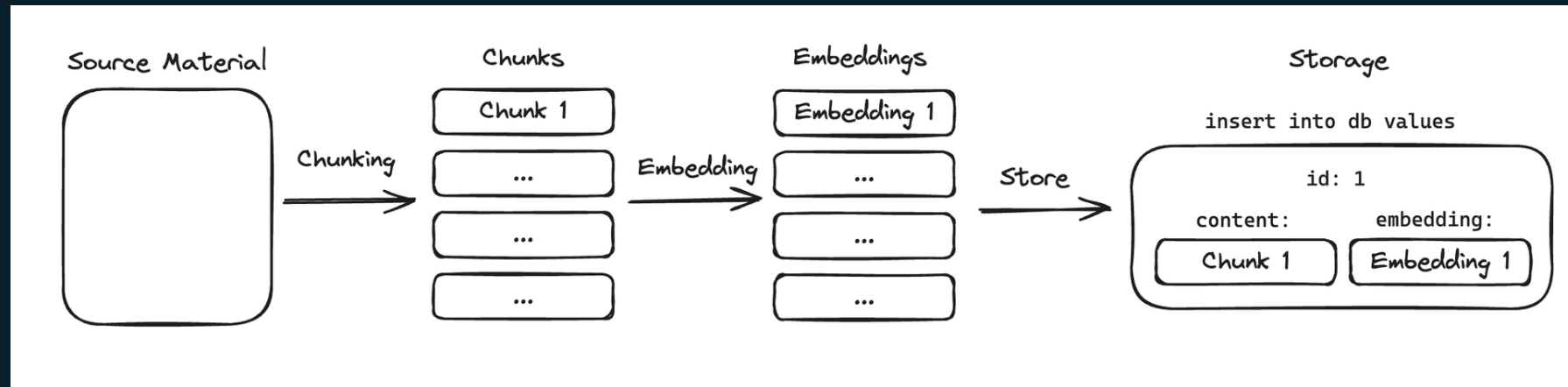
```
{
  "prompt": "A multi-colored umbrella on the beach, disposable camera",
  "size": "1024x1024",
  "n": 1,
  "quality": "standard",
  "style": "vivid",
  "response_format": "b64_json"
}
```

Response

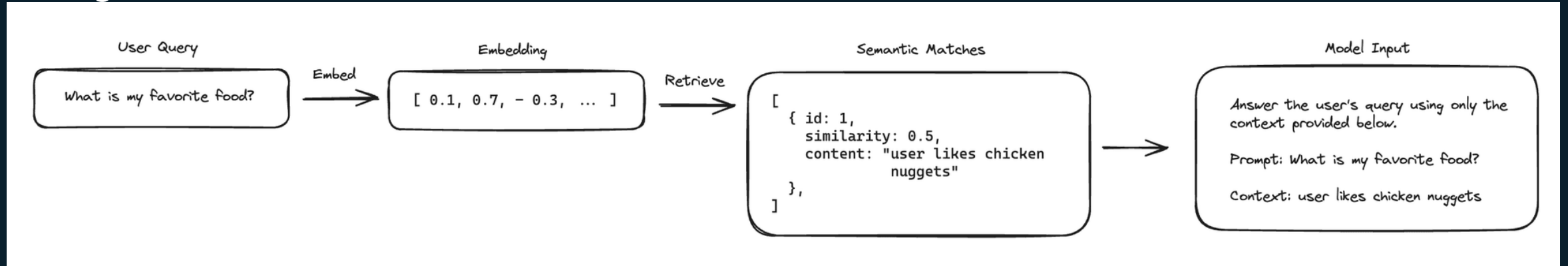
```
{
  "created": 1740805367,
  "data": [
    {
      "content_filter_results": { ... },
      "prompt_filter_results": { ... },
      "revised_prompt": "A multi-colored umbrella situated on a sandy beach with its vibrant colors contrasting the blue sky. Nearby on the sand, there's a disposable camera ready to capture beautiful memories.",
      "url": "https://dalleprodsec.blob.core.windows.net/private/images/1352de45-3e6d-418e-b6f1-0d4ed592a78c/generated_00.png?se=2025-03-02T05%3A02%3A57Z&sig=H5eZKn000A%2B1lYFpZYwnDGhVgVgA7JwSG0Q76ZvTuvA%3D&ske=2025-03-07T13%3A56%3A15Z&skoid=e52d5ed7-0657-4f62-bc12-7e5dbb260a96&sks=b&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:



Demo


Alfred



Cihan Cinar





Alfred Demo – Chat

The screenshot displays the Alfred Chat application interface. On the left is a sidebar with a 'Chat history' section containing two items: 'can you tell me how to ca ...' dated '01 Mar' and 'Hello' dated '28 Feb'. The main chat area at the top shows a user prompt: 'Can you tell me how to care for a parrot? Generate a short answer.' Below this is the AI response, which starts with 'Caring for a parrot involves several key aspects:' followed by a numbered list of five points: 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. The response concludes with 'By following these guidelines, you can help ensure your parrot is happy and healthy!'. At the bottom of the main chat area is a text input field with the placeholder 'Send a message...' and a 'Send' button. The footer of the interface includes a 'Support' link, a user profile for 'Cihan Cinar', and a 'Prompt templates' section.

Alfred 



Chat history 2  


can you tell me how to ca ...  


01 Mar


Hello

28 Feb

 ChatGPT-4o 



 Export chat



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



 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!

  Prompt templates

 Send a message... 

Alfred can make mistakes. Consider checking important information.

Alfred Demo – Image Generation

Alfred

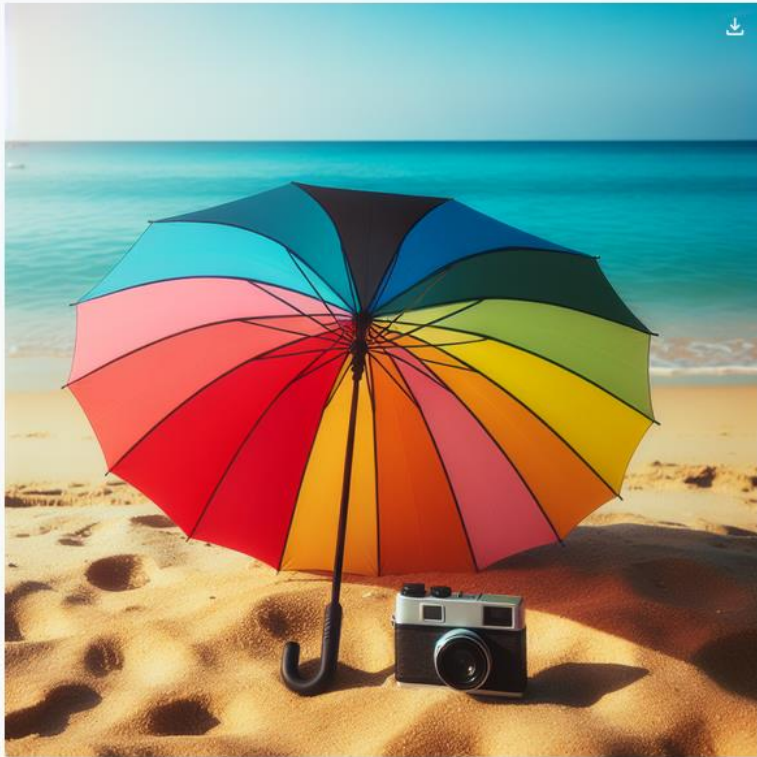
Chat history 2

A multi-colored umbrella ...
01 Mar

Hello
28 Feb

Image Generation

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



A vibrant, multi-colored umbrella with segments in shades of blue, green, yellow, orange, and red is open on a sandy beach. A black disposable camera sits on the sand in front of the umbrella. The background shows the ocean and a clear blue sky.

Send a message...

Support

Cihan Cinar

Prompt templates

Alfred can make mistakes. Consider checking important information.

Alfred Demo – RAG

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

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



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.



cecidomyia (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 la...

01 Mar

Hello

28 Feb

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.

Support

Cihan Cinar

Alfred can make mistakes. Consider checking important information.

giv-2023-integrated-report.pdf

Complete document Document extracts

121

Research value creation Impact and environmental performance

Our work to improve and protect water quality and quantity in our priority natural supply chains includes projects such as one in Provence and the lavender in the South of France. In partnership with regenerative agriculture technical experts, we carried out comprehensive soil audits in key producing areas. The baseline assessment laid down in our plan sheet in 2022 and includes soil profiles, visual inspections of soil structure, and soil analysis. These assessments will help to accurately characterize the soil, its structure, its physico-chemical characteristics, and its water holding and releasing capacity. The baseline assessment will allow for the definition of adapted cultivation and water management practices to address the local challenges faced by the producers.

In alignment with our water policy and our efforts to promote global water security in 2024, in 2023, our Other Natural Fibres, Textiles and Consumer Goods divisions to create new sensory units. Using the regional plants to create everything from walls to floors to the water involved. Meanwhile, the working facility benefited 200 pupils from the French Lavender School, Secondary and Higher Secondary School who had suffered from a lack of access to such activities facilities.

Through the Givaudan foundation, we also improved the living conditions of beekeepers collecting honey in our fields by partnering with UEF to install drinking water infrastructure including water wells and reservoirs. As a result, 1,200 people have gained easier access to their water, especially by reducing the time they need to get to a water source and to maintain and ensure the good functioning of the water points over time.

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

Waste management and circular principles

Exploration of the material topic

Givaudan's impacts on waste management and circular principles are discussed along the whole value chain. Our upstream activities generate waste such as surface agricultural activities and packaging from purchased goods, and our own production and packaging activities create additional waste.

Waste management and circular principles that support us in minimizing the extraction of resources and the generation of waste, in ensuring waste prevention and optimizing environmental, economic and societal value throughout the life cycle of our products are critical elements to long our purpose and showing our love for nature in everything we do.

Protecting the environment means acting on waste and our goal is to continually reduce the waste generated by our business activities including product and process development, manufacturing, formulation, packaging, marketing and distribution. Our decisions about what materials to use also have consequences in terms of waste and we strive to reuse or recycle materials where possible.

If waste is not managed properly, it can result in adverse impacts on the health of employees and people living in surrounding areas, harm environmental quality and reduce the overall cost of operations. Poorly managed waste contributes to climate change and air pollution, and directly affects many ecosystems and biodiversity. Finally, if industry waste can affect our health and

Wellbeing in many ways, e.g., methane gas can contribute to climate change, air pollution can reduce air quality in the atmosphere. Then water sources can be contaminated, crops can grow in contaminated soil, and fish might lose their lives, subsequently ending up in our food. Our intensive management of waste contributes positively to the environment and to people's health by avoiding these effects.

Reducing and preventing natural resources to create new products makes a significant amount of sense and by reducing and recycling waste, we conserve natural resources and reduce energy consumption. Using recycled materials for products instead of new resources results in the same end product with a significantly decreasing energy consumption, reduction of our use of plastic is also beneficial.

Investing resources in circularity and supporting green opportunities to use recycled materials offers the best way to dispose of, reducing our impact on the environment and driving innovation in the same line, reducing waste leads to increased production costs which in turn need for recycling and disposal.

Management of the impact

Waste management

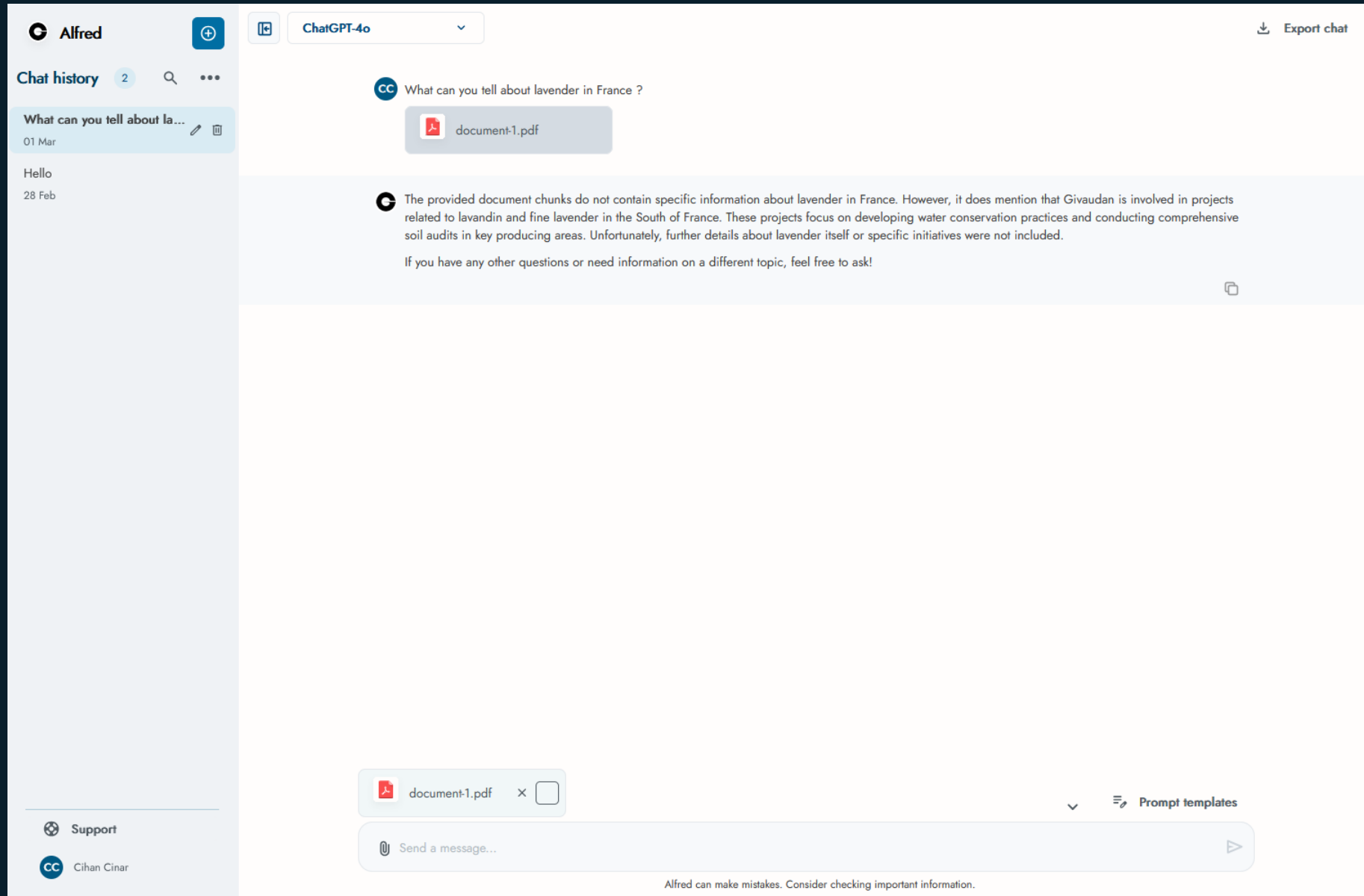
REDUCTION
REUSE
RECYCLE
RECOVERY
Preparation for reuse
Recycle
DISPOSAL
Incineration with and without energy recovery
Landfill

Waste management

Auto

Innovative solutions targeting waste

Alfred Demo – Document Upload



Demo

Azure

Cihan Cinar



Azure Demo –

The screenshot shows the Microsoft Azure portal interface. At the top is a blue header bar with the Microsoft Azure logo, an 'Upgrade' button, a search bar, and a Copilot button. On the right side of the header, there are notification and settings icons, and a user profile for 'cihancnr@gmail.com' with the role 'DEFAULT DIRECTORY'.

Below the header, the main content area is divided into several sections:

- Azure services:** A row of icons for various services: 'Create a resource' (plus icon), 'Subscriptions' (key icon), 'Cost Management' (dollar sign icon), 'Azure AI Foundry' (AI icon), 'Azure OpenAI' (star icon), 'Quickstart Center' (rocket icon), 'Azure AI services' (cloud icon), 'Kubernetes services' (Kubernetes icon), 'Virtual machines' (server icon), and 'More services' (arrow icon).
- Resources:** A section with tabs for 'Recent' and 'Favorite'. Below the tabs is a table with three columns: 'Name', 'Type', and 'Last Viewed'.

Name	Type	Last Viewed
Azure subscription 1	Subscription	16 minutes ago

Below the table is a 'See all' link.
- Navigate:** A row of icons for navigation: 'Subscriptions' (key icon), 'Resource groups' (cube icon), 'All resources' (grid icon), and 'Dashboard' (chart icon).
- Tools:** A row of icons for tools: 'Microsoft Learn' (book icon with '101' badge), 'Azure Monitor' (gauge icon), 'Microsoft Defender for Cloud' (shield icon), and 'Cost Management' (dollar sign icon).

[Try Azure for Free](#)

Azure Demo – OpenAI Deployment

The screenshot displays the Microsoft Azure portal interface. At the top, the header includes the 'Microsoft Azure' logo, an 'Upgrade' button, a search bar containing 'openai', and a 'Copilot' button. The user's profile 'cihancnr@gmail.com' and 'DEFAULT DIRECTORY' are visible on the right. The main content area is divided into several sections: 'Azure services' with a 'Create a resource' button, 'Resources' with 'Recent' and 'Favorite' tabs, 'Navigate' with a 'Subscriptions' link, and 'Tools' with links to 'Microsoft Learn', 'Azure Monitor', 'Microsoft Defender for Cloud', and 'Cost Management'. A search dropdown menu is open, showing results for 'openai'. The 'Services' tab is selected, displaying 'Azure OpenAI' as the top result, which is highlighted with a red box and a hand cursor. Below it, 'Azure AI services' is listed with the keyword 'openai'. The 'Marketplace' tab shows a list of services including 'Azure OpenAI', 'API Management', 'Azure AI Foundry', and 'Azure AI services'. The 'Documentation' section lists several links related to OpenAI deployment. The 'Last Viewed' section shows a timestamp of '16 minutes ago'. The bottom of the page features a row of service tiles: 'Microsoft Learn', 'Azure Monitor', 'Microsoft Defender for Cloud', and 'Cost Management'.

Microsoft Azure Upgrade openai Copilot cihancnr@gmail.com DEFAULT DIRECTORY

Azure services

Create a resource Subscriptions

Resources

Recent Favorite

Name

Azure subscription 1

See all

Navigate

Subscriptions

Tools

Microsoft Learn 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

Search Results for 'openai'

All Services (2) Marketplace (31) More (4)

Services

Azure OpenAI

Azure AI services
Keywords: openai

Marketplace See more

Azure OpenAI

API Management

Azure AI Foundry

Azure AI services

Documentation See more

Quickstart - Extend OpenAI using Tools and execute a local Function with .NET -

Azure AI Foundry architecture - Azure AI Foundry

Azure OpenAI in Azure AI Foundry portal - Azure AI Foundry

Deploy models in Azure AI Foundry portal - Azure AI Foundry

Continue searching in Microsoft Entra ID

Searching all subscriptions. Give feedback

Last Viewed

16 minutes ago

Dashboard

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.

No grouping

List view

Name	Kind	Location	Custom Domain ...	Pricing tier	Status
------	------	----------	-------------------	--------------	--------

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 ...

1 Basics

2 Network

3 Tags

4 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 *

Azure subscription 1

Resource group *

(New) globalai-resource-group

[Create new](#)

Instance Details

Region

East US

Name *

globalai-openai-001

Pricing tier *

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

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

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	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

The screenshot displays the Azure AI Foundry Chat playground interface. The top navigation bar includes the Azure AI Foundry logo, the service name 'Azure OpenAI Service', the resource group 'globalai-openai-001', and the specific playground 'Chat playground'. On the right, it shows 'All resources', a notification icon, settings, a user profile for 'globalai-openai-001 (eastus, S0)', and a Creative Commons license icon.

The left sidebar contains a navigation menu with the following items: Home, Get started, Model catalog, Playgrounds (expanded), Chat (selected), Assistants (PREVIEW), Real-time audio (PREVIEW), Images, Completions, Tools (expanded), Fine-tuning, Azure OpenAI Service (PREVIEW), Evaluation, Stored completions (PREVIEW), Batch jobs, Shared resources (expanded), Deployments (highlighted with a red box), Quota, Safety + security, Data files, and Vector stores (PREVIEW).

The main content area is titled 'Chat playground' and features a toolbar with 'View code', 'Deploy', 'Import', 'Export', 'Prompt samples', and 'Filters feedback'. The interface is divided into three main sections:

- Setup:** Includes a 'Deployment' dropdown (currently showing 'No deployment exists'), a text area for 'Give the model instructions and context' (containing the prompt: 'You are an AI assistant that helps people find information.'), buttons for 'Apply changes' and 'Generate prompt', an 'Add section' button, and a note about few-shot examples.
- Chat history:** Displays a 'Start with a sample prompt' section with three options: '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.), and 'Dialogue creation' (Create a conversation between two characters who are meeting for the first time in a mysterious place.). Below this is a large text input field for the user query, with a hint 'Type user query here. (Shift + Enter for new line)' and a microphone icon.
- Chat playground help:** A sidebar on the right with a 'Relevant resources' section containing a 'Chat Quickstart' link.

Azure Demo – ChatGPT Model Deployment

The screenshot displays the Azure AI Foundry interface for deploying a model. The main header shows 'Azure AI Foundry | Azure OpenAI Service' with a resource identifier 'globalai-openai-001' and the section 'Models + endpoints'. The left sidebar contains navigation options: Home, Get started, Model catalog, Playgrounds, Chat, Assistants, Real-time audio, Images, Completions, Tools, Fine-tuning, Azure OpenAI Service, Evaluation, Stored completions, Batch jobs, Shared resources, Deployments, Quota, Safety + security, Data files, and Vector stores. The central area is titled 'Model deployments' and features a '+ Deploy model' button. A modal dialog titled 'Deploy model gpt-4o-mini' is open, containing the following fields and information:

- Deployment name:** gpt-4o-mini
- Deployment type:** Global Standard (selected from a dropdown). A note explains: '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 (collapsed):**
 - Model version:** 2024-07-18 (Default)
 - Resource location:** East US 2
 - Quota:** 1K tokens per minute quota available for your deployment.
 - Tokens per Minute Rate Limit:** A slider set to 1K. Below it, it states 'Corresponding requests per minute (RPM) = 10'.
 - Warning:** When you create this deployment, the resource change you've made will persist throughout Azure AI Foundry.
 - Information:** A new AI resource will be created for your deployment. A resource location that supports the model has been pre-selected.

At the bottom of the dialog are two buttons: 'Create resource and deploy' and 'Cancel'. On the right side of the interface, there is a 'Deployments help' section with links for 'Relevant resources', 'Azure OpenAI Service deployment types', 'Working with Azure', and 'OpenAI Service models'.

Azure Demo –

The screenshot displays the Azure AI Foundry console interface. The top navigation bar includes the Azure AI Foundry logo, the current subscription 'cihan-m7pt0nf4-eastus2', and the resource group 'Deployments / gpt-4o-mini'. The left sidebar contains a navigation menu with options like Home, Get started, Model catalog, Playgrounds, Chat, Assistants, Real-time audio, Images, Completions, Tools, Fine-tuning, Azure OpenAI Service, Evaluation, Stored completions, Batch jobs, Shared resources, Deployments (selected), Quota, Safety + security, Data files, and Vector stores. The main content area is titled 'gpt-4o-mini' and features tabs for Details, Metrics, Risks & Safety, and Consume. Below the tabs are buttons for 'Open in playground', 'Edit', and 'Delete'. The 'Details' tab is active, showing deployment information in a table format. The table is divided into two columns: 'Deployment info' and 'Endpoint'. The 'Deployment info' column contains details such as 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), and Model retirement date (Jul 20, 2025 2:00 AM). The 'Endpoint' column shows the Target URI (https://cihan-m7pt0nf4-eastus2.cognitiveservices.azure.com/open ...) and a Key (represented by a masked string). Below the table, there are sections for 'Monitoring & safety' (Content filter: DefaultV2) and 'Useful links for application development' (Code sample repository, Tutorial). On the right side of the console, there is a 'Deployments help' panel with a section titled 'What are deployments in Azure AI Foundry?' and a link to 'Relevant resources'.

Azure AI Foundry | Azure OpenAI Service / cihan-m7pt0nf4-eastus2 / Deployments / gpt-4o-mini

All resources 5 cihan-m7pt0nf4-eastus2 (eastus2, S0)

gpt-4o-mini

Details Metrics Risks & Safety Consume

Open in playground Edit Delete

Deployment info		Endpoint	
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

.....

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 displays the Azure AI Foundry Chat playground interface. The top navigation bar includes the Azure AI Foundry logo, the current service (Azure OpenAI Service), the resource group (cihan-m7pt0nf4-eastus2), and the playground name (Chat playground). On the right, it shows 'All resources', a search icon, a settings icon, a user profile (cihan-m7pt0nf4-eastus2), and a subscription status (eastus2, S0).

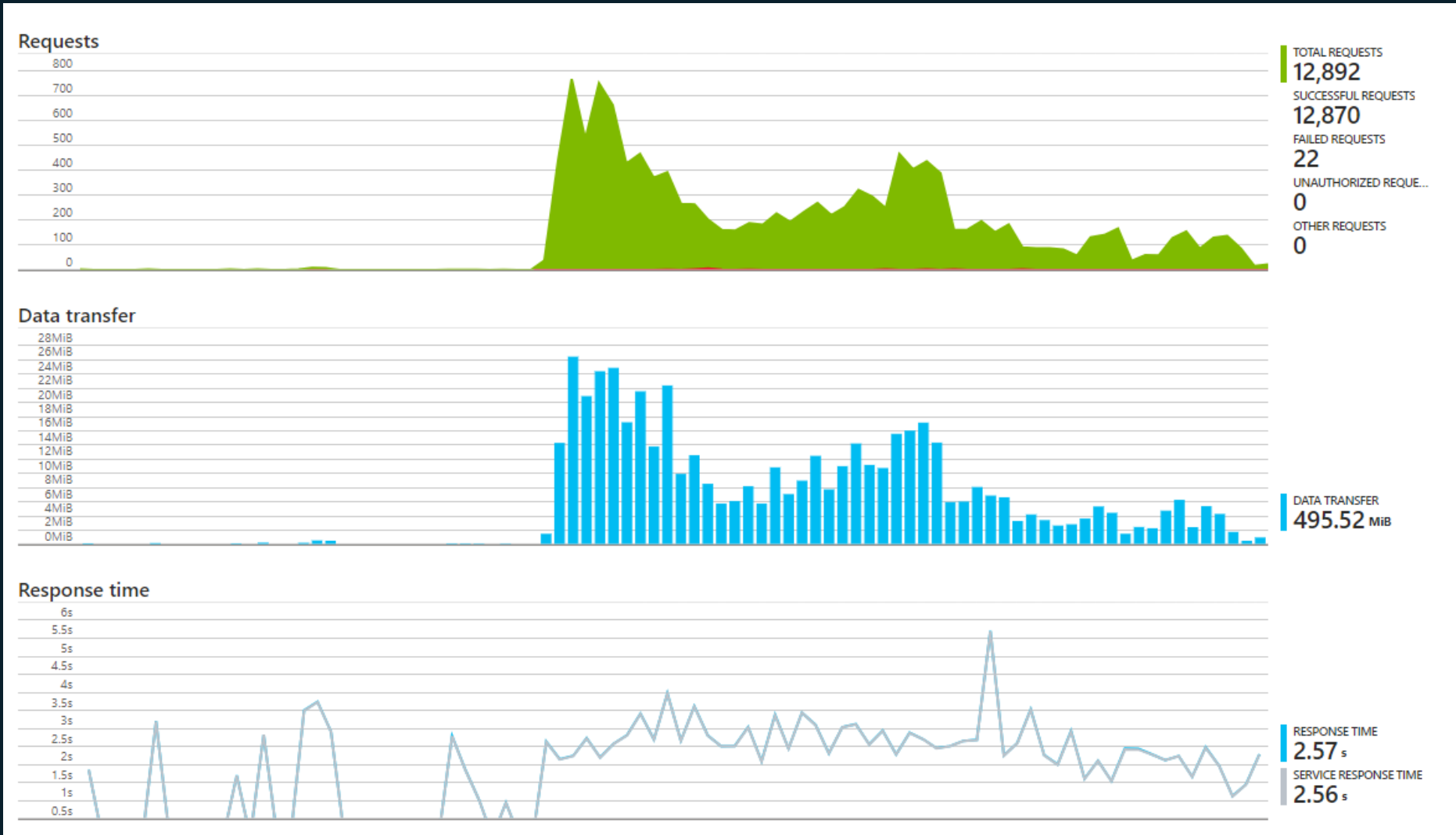
The left sidebar contains a navigation menu with the following items: Home, Get started, Model catalog, Playgrounds (expanded), Chat (selected), Assistants (PREVIEW), Real-time audio (PREVIEW), Images, Completions, Tools, Fine-tuning, Azure OpenAI Service (PREVIEW), Evaluation, Stored completions (PREVIEW), Batch jobs, Shared resources, Deployments, Quota, Safety + security, Data files, and Vector stores (PREVIEW).

The main area is titled 'Chat playground' and features a toolbar with 'View code', 'Deploy', 'Import', 'Export', 'Prompt samples', and 'Filters feedback'. The 'Setup' section on the left includes a 'Deployment' dropdown set to 'gpt-4o-mini (version:2024-07-18)', a 'Create new deployment' link, and a text area for 'Give the model instructions and context' containing the prompt: 'you are a helpful assistant that talks like a pirate. Generate short answer only.' Below this are 'Apply changes' and 'Generate prompt' buttons, and an 'Add section' button.

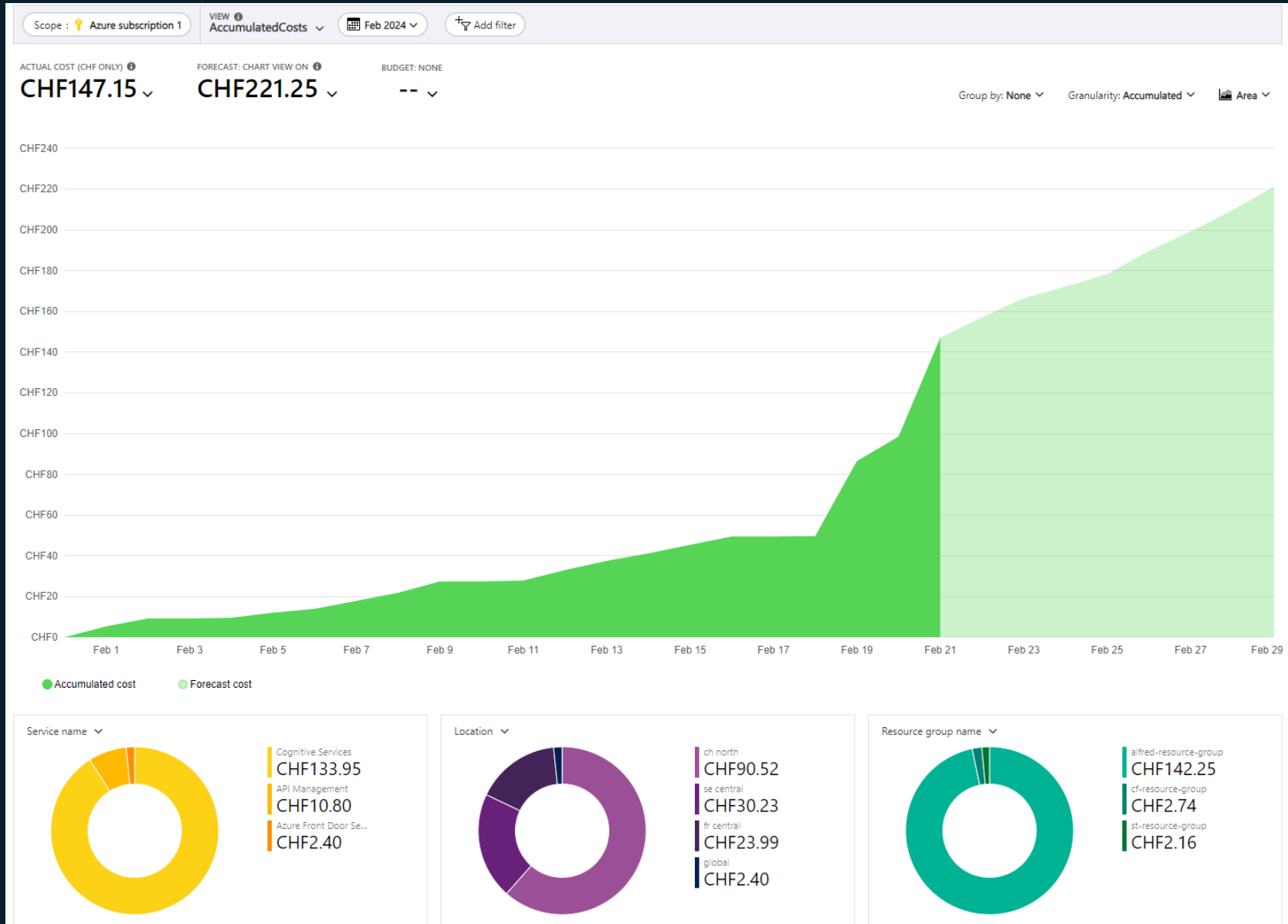
The 'Chat history' section on the right shows a conversation. The user's message is 'can you tell me how to care for a parrot?'. The assistant's response is: 'Arrr, matey! To care fer a parrot, ye need to provide a spacious cage, fresh food like fruits and veggies, clean water, and plenty o' toys fer playin'. Make sure to give 'em attention and social time, too! Avast, keep 'em safe from dangers and give 'em a good vet checkup now and then!'. At the bottom of the chat history is a text input field with the placeholder 'Type user query here. (Shift + Enter for new line)' and a token count '92/128000 tokens to be sent'.

A 'Chat playground help' sidebar on the far right contains a 'Relevant resources' section with a 'Chat Quickstart' link.

Alfred – Azure Monitor







Alfred – Cost Analysis

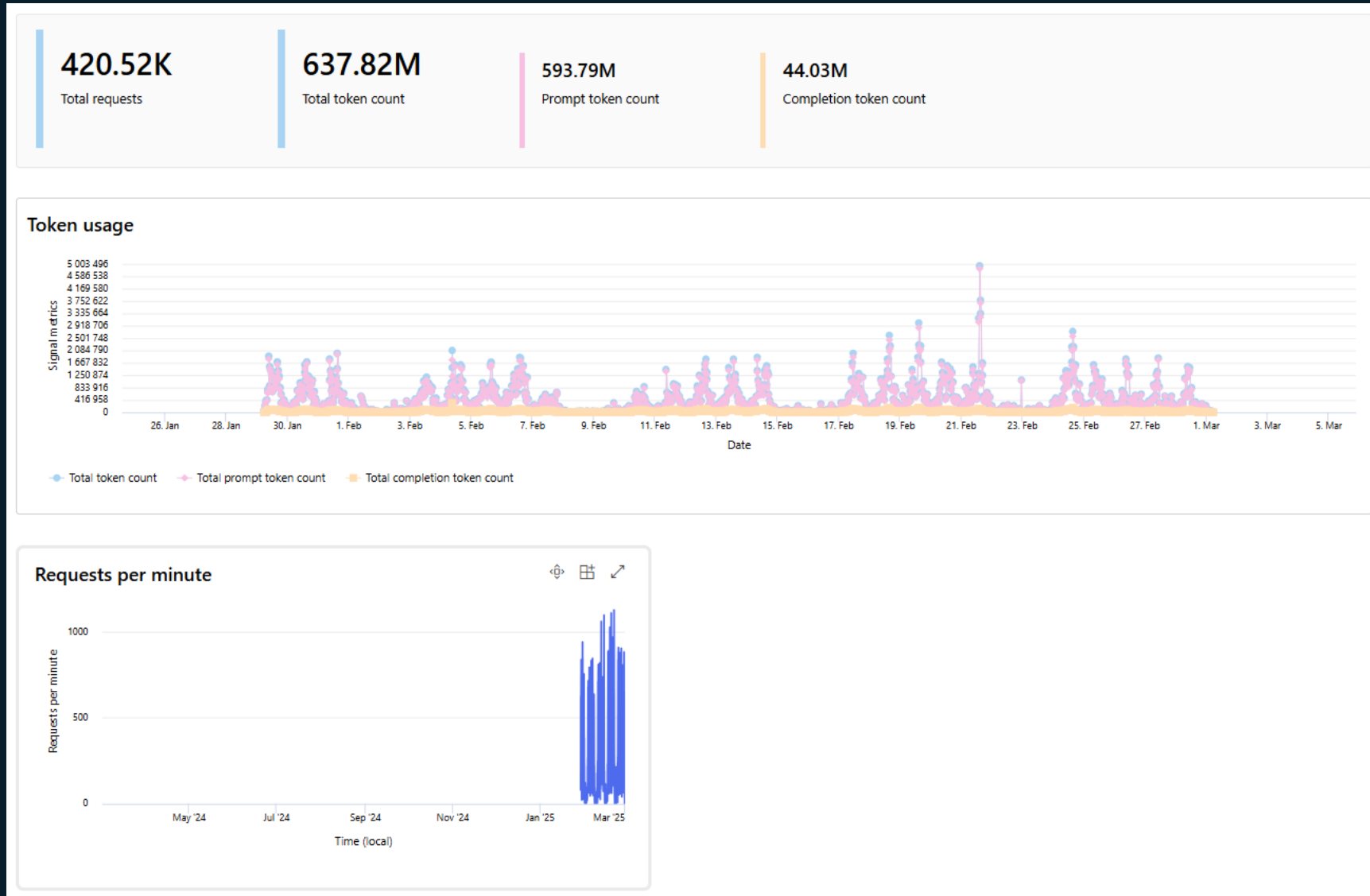


Conclusion

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

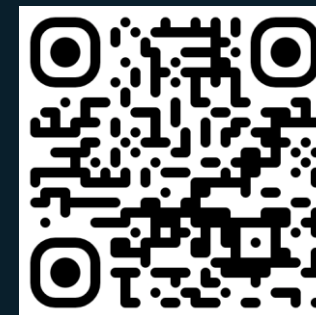


Cihan Cinar

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>