Building a GenAI-Powered RAG App Integrated with Microsoft Teams

By Natarajan Ramasamy

Building a GenAI-Powered RAG App Integrated with Microsoft Teams – From Concept to Real MVP

Overview

In this post, I want to share an exciting journey I recently completed — building a practical **GenAI MVP using the Retrieval-Augmented Generation (RAG) pattern**, integrated directly into **Microsoft Teams**, powered by **Azure AI Search** and **Azure OpenAI**. This was more than just a tech experiment — it's a real, scalable architecture that can power intelligent, enterprise-ready conversational AI tools.

This article is for:

- AI enthusiasts and developers looking to explore GenAI in practice
- Enterprise architects interested in Microsoft ecosystem integrations
- Anyone curious about building smarter internal tools using ChatGPT-like models

-

Table of Contents

1. What is Generative AI?
2. What is RAG (Retrieval-Augmented Generation)?
3. Why RAG is Better than Pure LLM
4. Understanding Azure AI Search
5. A Primer on Azure OpenAI Service
6. Why Microsoft Teams Integration Matters
7. The Architecture of My MVP Project
8. Step-by-Step: Building Your Own RAG+Teams App
9. Common Challenges and How I Solved Them
10. Future Plans and How You Can Get Involved

What is Generative AI?

Generative AI refers to AI models that can **generate new content** — text, images, code, music — from learned patterns in massive datasets. Technologies like GPT-4, DALL·E, and Codex are famous examples. These models don’t just classify or predict; they create.

Example use cases:

- Summarizing documents
- Writing emails or code
- Creating product descriptions
- Automating chatbot conversations

What is RAG?

RAG (Retrieval-Augmented Generation) is a pattern that combines:

- **Retrieval**: Find relevant documents using a search engine (e.g., vector search).
- **Generation**: Feed those documents into a Generative AI model to produce accurate, context-rich responses.

RAG bridges the gap between **LLM hallucinations** and **enterprise-grade information accuracy**.

Why RAG is Better than Pure LLMs

LLMs alone don't “know” your enterprise data unless you fine-tune them — a costly and inflexible process.

RAG provides:

Real-time access to current data
Explainability via citations
No need to fine-tune the LLM
Dynamic context adaptation

Understanding Azure AI Search

Azure AI Search is Microsoft's cloud-based search-as-a-service platform, supporting:

- Full-text search
- Vector search for semantic similarity
- Cognitive skills to extract and index metadata from documents

In this project, we used Azure AI Search to:

- Ingest PDFs, DOCXs, and web content
- Generate semantic embeddings
- Enable fast retrieval for the RAG flow

Azure OpenAI Service

Azure provides enterprise access to OpenAI models (e.g., GPT-4) via REST APIs. Benefits:

- Regionally hosted in Azure (compliance and data security)
- Authentication via Azure AD
- Integrated with other Azure services

We use these models to **generate natural language answers** from retrieved documents.

Why Microsoft Teams Integration Matters

Most enterprises live inside Microsoft Teams.

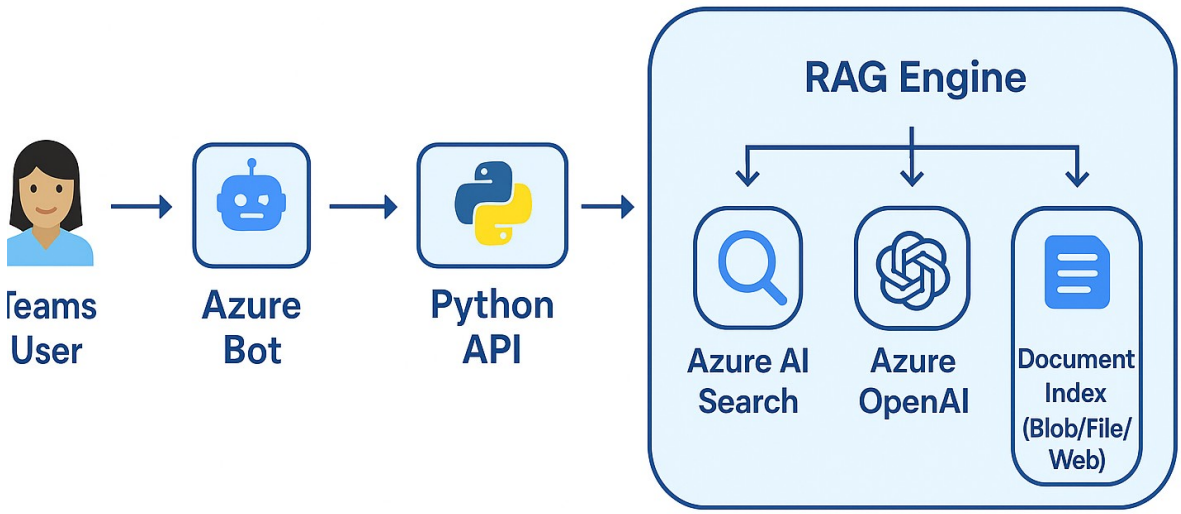
Imagine if your internal AI assistant was:

- Available inside Teams
- Could answer employee queries using internal documents
- Could escalate to humans or services when needed

That's what we built. A fully functional AI bot within Teams — answering real business questions.

Architecture of My MVP Project

Here’s the high-level design of our solution:



Key components:

- Microsoft Bot Framework (via Azure Bot Service)
- Python backend with `/chat` and `/ask` endpoints
- Azure AI Search for document retrieval
- Azure OpenAI for language generation
- Microsoft Teams as the user interface

How You Can Build This Yourself (Hands-On)

Here's a simplified guide. You can do this in a week with a basic Azure subscription:

Step 1: Set Up Azure Resources

- Azure AI Search instance
- Azure OpenAI
- Azure Bot Channel registration
- Azure Blob Storage (for document ingestion)

Step 2: Prepare the Codebase

Use the official GitHub sample from Microsoft:
[Natarajans GitHub Project - RAG with Teams]

Modify it to:

- Customize indexing pipeline
- Add support for your document types
- Secure your backend APIs

Step 3: Create the Python Backend

- Flask or Quart app
- Routes for `/ask`, `/chat`, `/api/messages`
- JWT authentication
- RAG orchestration logic (retrieval → prompt → generation)

Step 4: Deploy to Azure VM or App Service

- Use `systemd` for service management
- HTTPS via NGINX + SSL

Step 5: Connect Teams to Your Bot

- Register Teams channel in Azure Bot Channel
- Test via Bot Emulator and live Teams chat
- Implement token validation for Teams authentication

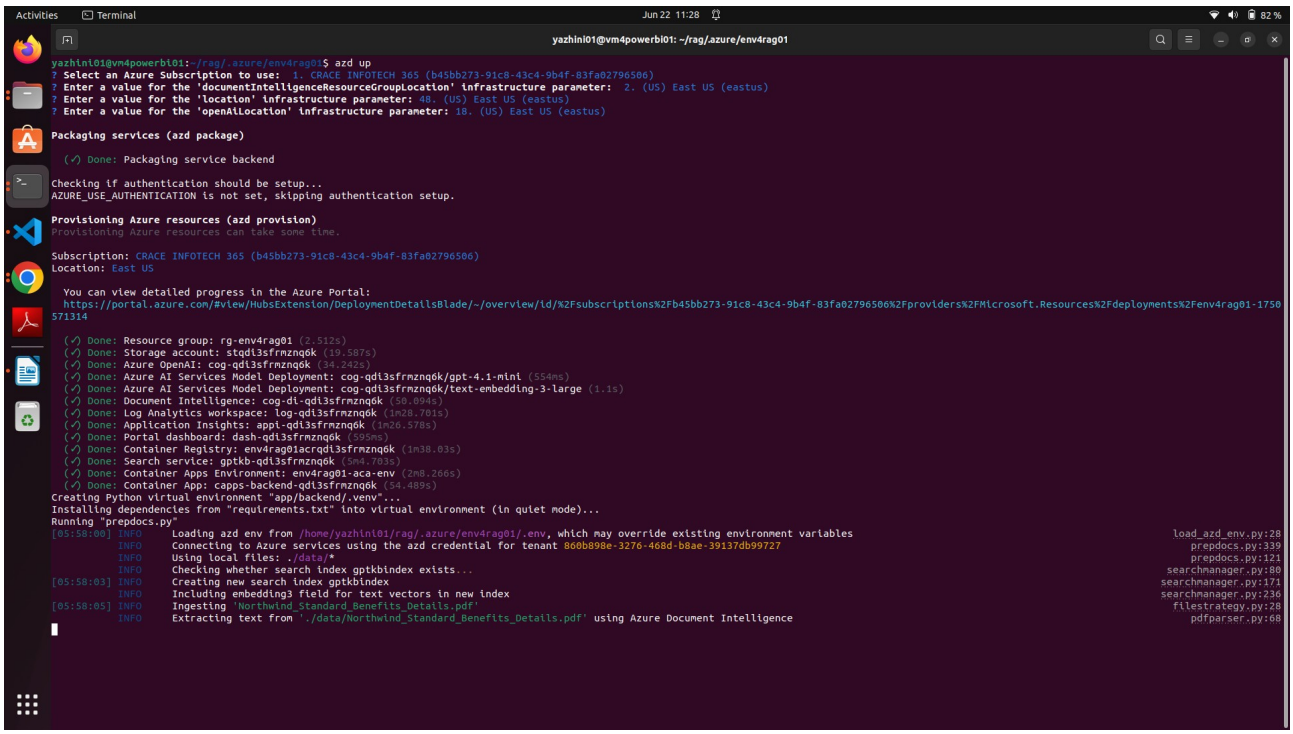
Common Challenges I Faced

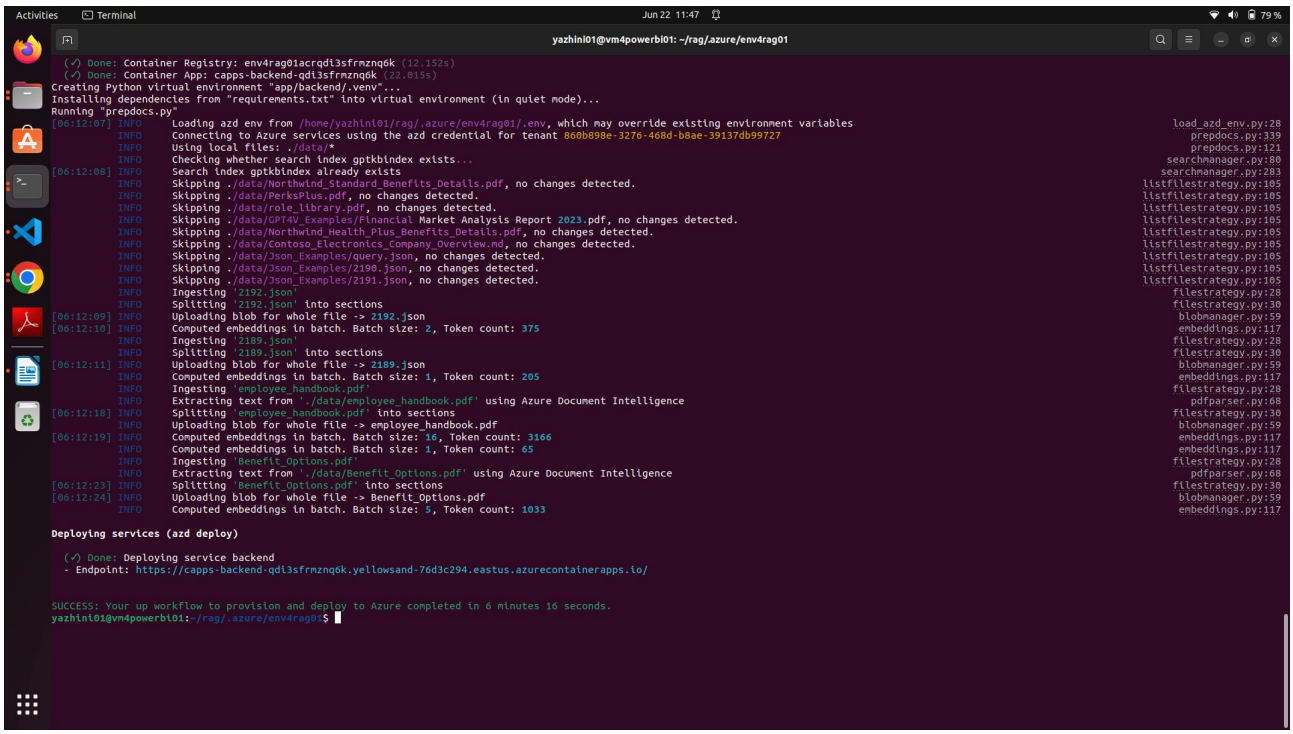
- Original Microsoft Github project document guides us to install many components which are not required immediately or mandatory
- For example Simple low cost Linux VM is enough to install this Web application but default scripts will create expensive App service plans for either function apps or container apps which are not required in initial MVP stage.
- Default scripts creates Managed Identity and assigns RBAC roles which requires Owner Permissions in Subscription level , But using API keys we can run this application without Owner Permissions.
- For Documents Ingestion , Documents Intelligence service is good but expensive, but we can manage with python parsing.
- By default, only web chat is available . Exposing this web application in publich internet is risky and requires extra precautions to protect our corporate assets and Customer Information. So i have created Teams channel which ensure that our web applicaiton is only accessible by corporate employess only
- To enable Teams integration , we need azure bot service , ie backend for Teams bot. For that i have adapted and reused the existing Web API backend.
- Local Dev server deployments consistently throws errors in all deployments which can be cured by installing python setuptools package in addtion.
- Current Code base can ingest data from the documents uploaded only. But that is not enough. In addtion to documents , organizations may have their Knowledge / Data in Confluence Pages, Sharepoints sites, and other web sites. So we need web scrapting . I have tested it with Atlassian Confluence Sites , but still the code requires thorough testing and enhancement so it is not integrated as of now.
- In addtion to Documents Upload, Crawling Corporate Web sites both makes more corporate knowledge repository availalbe offline in the AI Search index, but they are not enough , as soon they may become outdated. So planning to include real time AI Agent to retrieve data from the corporate web sites on demand.
- This web application is like chatGPT, which can be thought of having a virtual SME to help our corporate employee. This Virtual SME has brain (LLM) , can hear what we ask and can speak the relevant answer for our query. Ie it is just a Question Answer Machine . But we can also add two more senses like Seeing the input and outputs and doing ie executing tasks on behalf of us is a great enhancement we can get with these kind of Virtual SMEs / Assistant , they are called AI Agents and I am planning to include them here.

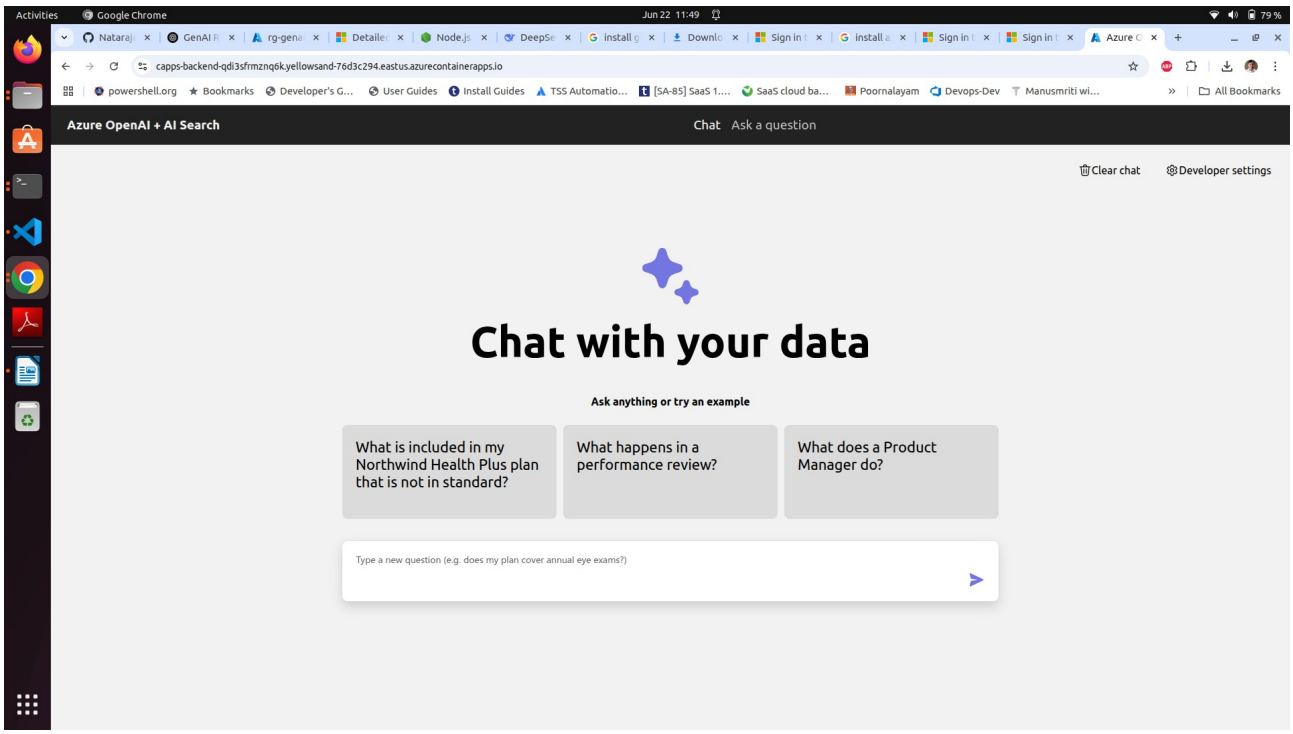
Quick glance over hands-on

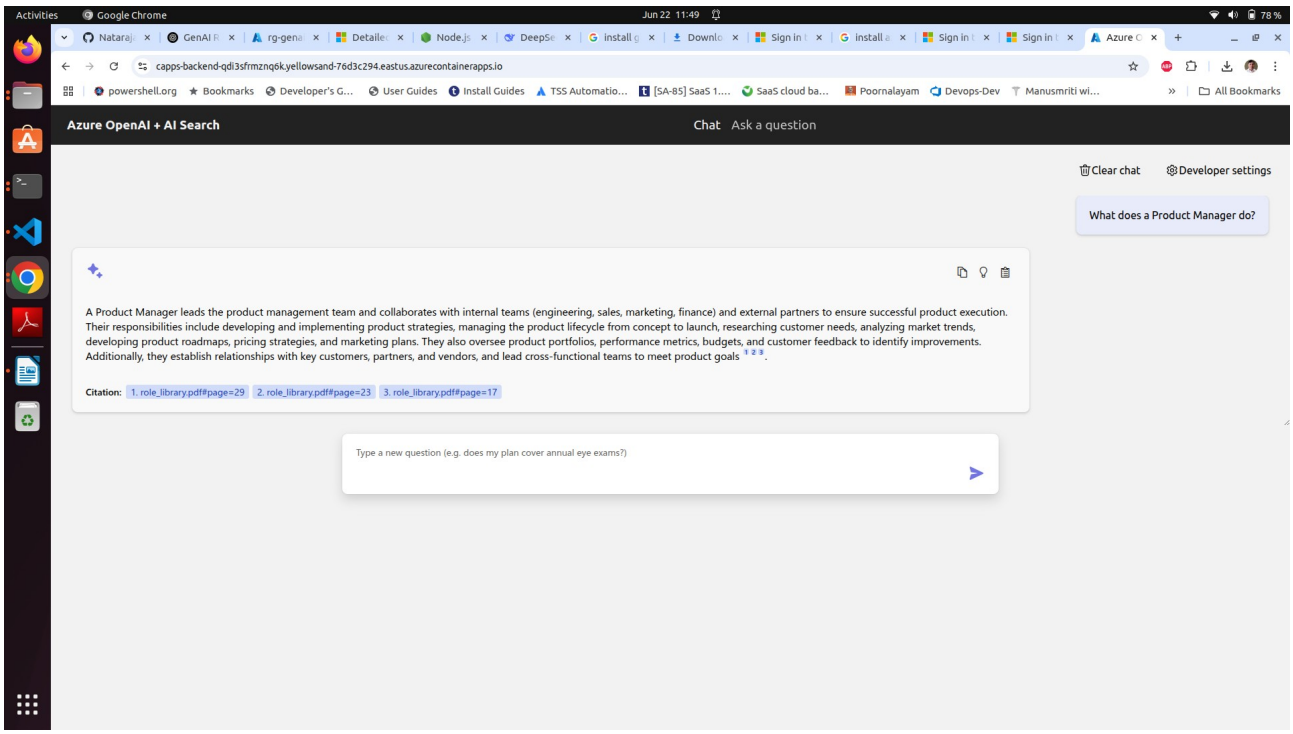
(Complete fully functional and tested source code is available in github)

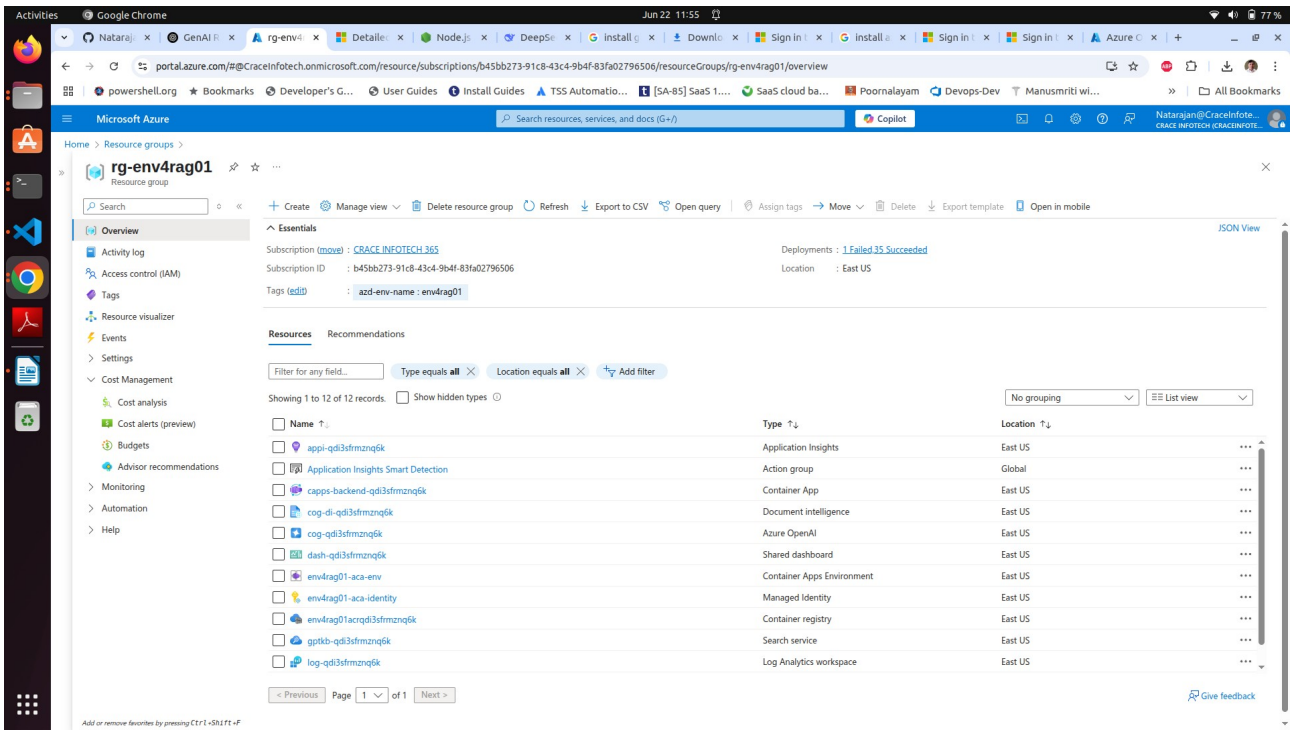
```
yazhini01@vm4powerbi01:~/rag$ azd version
azd version 1.17.1 (commit a20d045e2de92b98589422a1c26821597fc7b5d7)
yazhini01@vm4powerbi01:~/rag$ python --version
Python 3.12.3
yazhini01@vm4powerbi01:~/rag$ node -v
v22.16.0
yazhini01@vm4powerbi01:~/rag$ git -v
git version 2.43.0
yazhini01@vm4powerbi01:~/rag$ az --version
azure-cli                         2.74.0

core                              2.74.0
telemetry                          1.1.0

Dependencies:
msal                              1.32.3
azure-mgmt-resource               23.3.0

Python location '/opt/az/bin/python3'
Config directory '/home/yazhini01/.azure'
Extensions directory '/home/yazhini01/.azure/cliextensions'

Python (Linux) 3.12.10 (main, May 27 2025, 09:12:37) [GCC 13.3.0]

Legal docs and information: aka.ms/AzureCliLegal


Your CLI is up-to-date.
yazhini01@vm4powerbi01:~/rag$
```

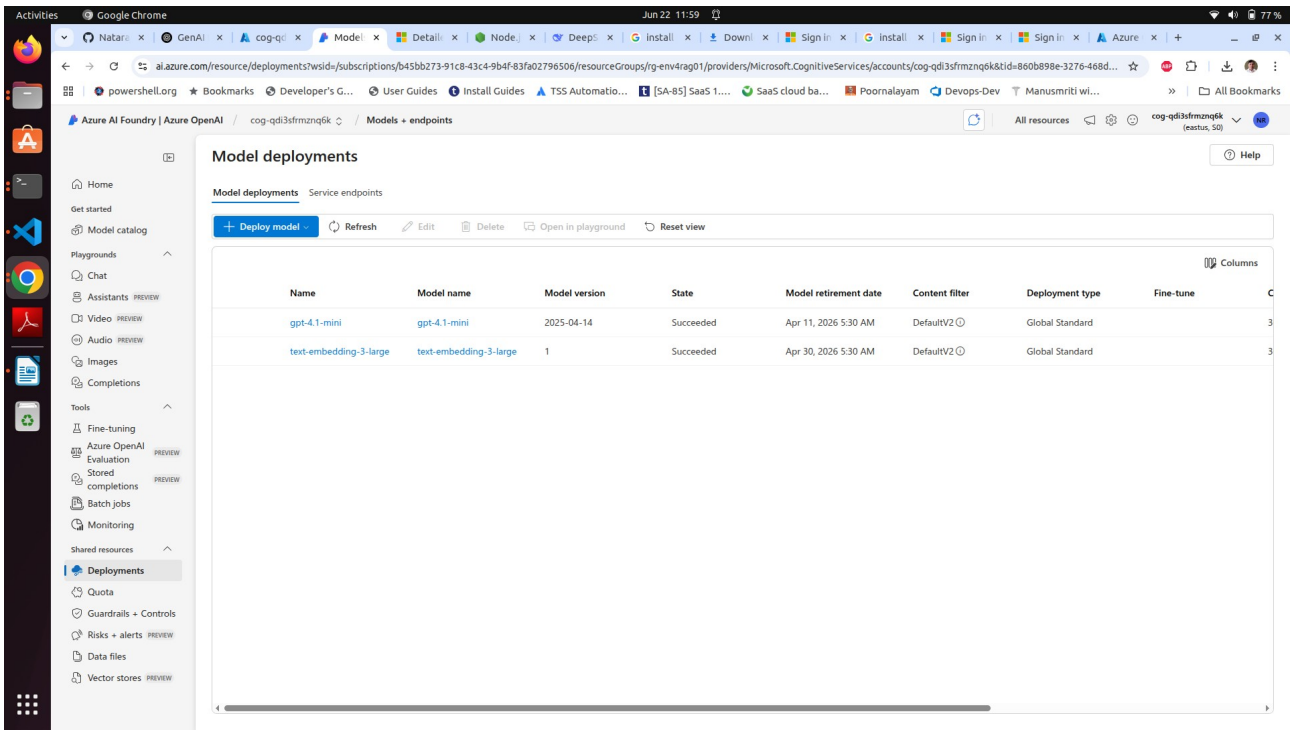

Activities
Terminal
Jun 22 11:28
82 %
yazhini01@vm4powerbi01: ~/rag/.azure/env4rag01
yazhini01@vm4powerbi01:~/rag/.azure/env4rag01$ azd up
? Select an Azure Subscription to use: 1. CRACE INFOTECH 365 (b45bb273-91c8-43c4-9b4f-83fa02796506)
? Enter a value for the 'documentIntelligenceResourceGroupLocation' infrastructure parameter: 2. (US) East US (eastus)
? Enter a value for the 'location' infrastructure parameter: 48. (US) East US (eastus)
? Enter a value for the 'openAiLocation' infrastructure parameter: 18. (US) East US (eastus)
Packaging services (azd package)
(✓) Done: Packaging service backend
Checking if authentication should be setup...
AZURE_USE_AUTHENTICATION is not set, skipping authentication setup.
Provisioning Azure resources (azd provision)
Provisioning Azure resources can take some time.
Subscription: CRACE INFOTECH 365 (b45bb273-91c8-43c4-9b4f-83fa02796506)
Location: East US
You can view detailed progress in the Azure Portal:
https://portal.azure.com/#view/HubsExtension/DeploymentDetailsBlade/~/overview/id/%2Fsubscriptions%2Fb45bb273-91c8-43c4-9b4f-83fa02796506%2Fproviders%2FMicrosoft.Resources%2Fdeployments%2Fenv4rag01-1750571314
(✓) Done: Resource group: rg-env4rag01 (2.512s)
(✓) Done: Storage account: stqdi3sfrmznq6k (19.587s)
(✓) Done: Azure OpenAI: cog-qdi3sfrmznq6k (34.242s)
(✓) Done: Azure AI Services Model Deployment: cog-qdi3sfrmznq6k/gpt-4.1-mini (554ms)
(✓) Done: Azure AI Services Model Deployment: cog-qdi3sfrmznq6k/text-embedding-3-large (1.1s)
(✓) Done: Document Intelligence: cog-di-qdi3sfrmznq6k (50.094s)
(✓) Done: Log Analytics workspace: log-qdi3sfrmznq6k (1m28.701s)
(✓) Done: Application Insights: appi-qdi3sfrmznq6k (1m26.578s)
(✓) Done: Portal dashboard: dash-qdi3sfrmznq6k (595ms)
(✓) Done: Container Registry: env4rag01acrqdi3sfrmznq6k (1m38.03s)
(✓) Done: Search service: gptkb-qdi3sfrmznq6k (5m4.703s)
(✓) Done: Container Apps Environment: env4rag01-aca-env (2m8.266s)
(✓) Done: Container App: capps-backend-qdi3sfrmznq6k (54.489s)
Creating Python virtual environment "app/backend/.venv"...
Installing dependencies from "requirements.txt" into virtual environment (in quiet mode)...
Running "prepdocs.py"
[05:58:00] INFO Loading azd env from /home/yazhini01/rag/.azure/env4rag01/.env, which may override existing environment variables load_azd_env.py:28
INFO Connecting to Azure services using the azd credential for tenant 860b898e-3276-468d-b8ae-39137db99727 prepdocs.py:339
INFO Using local files: ./data/* prepdocs.py:121
INFO Checking whether search index gptkbindex exists... searchmanager.py:80
[05:58:03] INFO Creating new search index gptkbindex searchmanager.py:171
INFO Including embedding3 field for text vectors in new index searchmanager.py:236
[05:58:05] INFO Ingesting 'Northwind_Standard_Benefits_Details.pdf' filestrategy.py:28
INFO Extracting text from './data/Northwind_Standard_Benefits_Details.pdf' using Azure Document Intelligence pdfparser.py:68

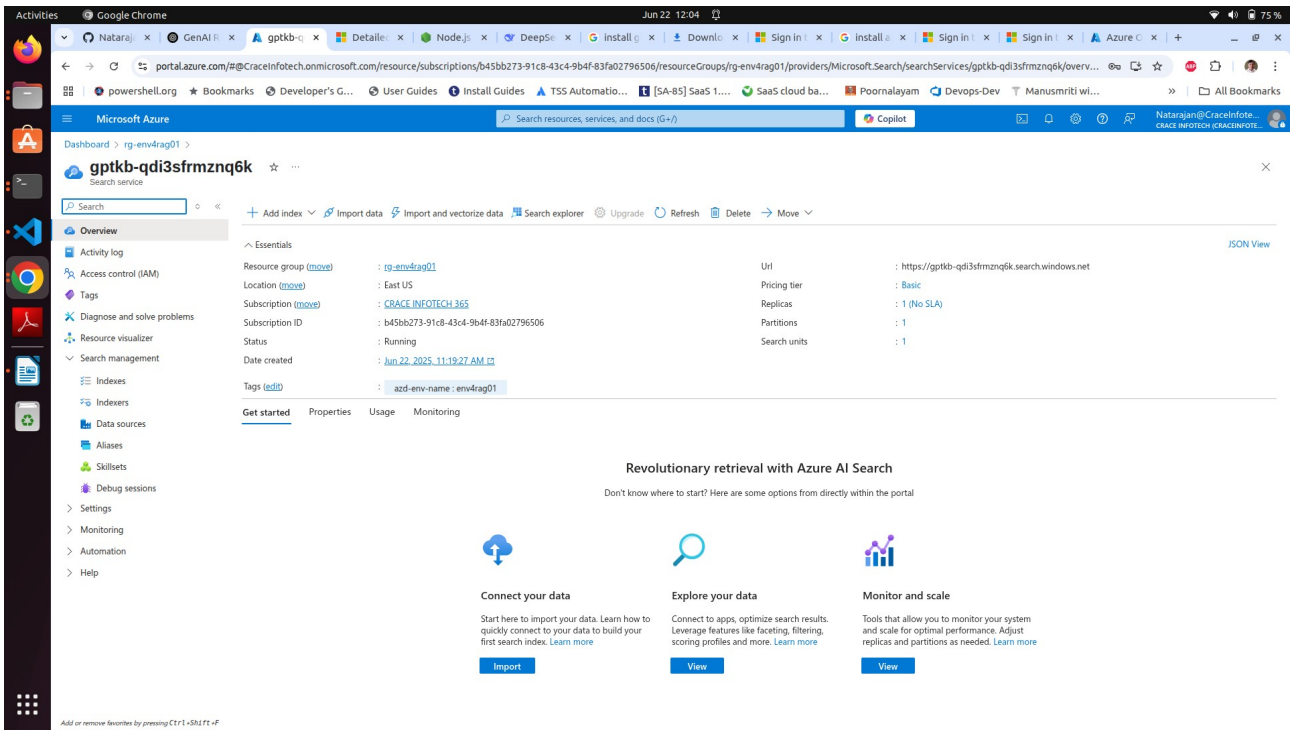

Activities
Terminal
Jun 22 11:47
79 %
yazhini01@vm4powerbi01: ~/rag/.azure/env4rag01
(✓) Done: Container Registry: env4rag01acrqdi3sfrmznq6k (12.152s)
(✓) Done: Container App: capps-backend-qdi3sfrmznq6k (22.015s)
Creating Python virtual environment "app/backend/.venv"...
Installing dependencies from "requirements.txt" into virtual environment (in quiet mode)...
Running "prepdocs.py"
[06:12:07] INFO Loading azd env from /home/yazhini01/rag/.azure/env4rag01/.env, which may override existing environment variables load_azd_env.py:28
INFO Connecting to Azure services using the azd credential for tenant 860b898e-3276-468d-b8ae-39137db99727 prepdocs.py:339
INFO Using local files: ./data/* prepdocs.py:121
INFO Checking whether search index gptkbindex exists... searchmanager.py:80
[06:12:08] INFO Search index gptkbindex already exists searchmanager.py:283
INFO Skipping ./data/Northwind_Standard_Benefits_Details.pdf, no changes detected. listfilestrategy.py:105
INFO Skipping ./data/PerksPlus.pdf, no changes detected. listfilestrategy.py:105
INFO Skipping ./data/role_library.pdf, no changes detected. listfilestrategy.py:105
INFO Skipping ./data/GPT4V_Examples/Financial Market Analysis Report 2023.pdf, no changes detected. listfilestrategy.py:105
INFO Skipping ./data/Northwind_Health_Plus_Benefits_Details.pdf, no changes detected. listfilestrategy.py:105
INFO Skipping ./data/Contoso_Electronics_Company_Overview.md, no changes detected. listfilestrategy.py:105
INFO Skipping ./data/Json_Examples/query.json, no changes detected. listfilestrategy.py:105
INFO Skipping ./data/Json_Examples/2190.json, no changes detected. listfilestrategy.py:105
INFO Skipping ./data/Json_Examples/2191.json, no changes detected. listfilestrategy.py:105
INFO Ingesting '2192.json' filestrategy.py:28
INFO Splitting '2192.json' into sections filestrategy.py:30
[06:12:09] INFO Uploading blob for whole file -> 2192.json blobmanager.py:59
[06:12:10] INFO Computed embeddings in batch. Batch size: 2, Token count: 375 embeddings.py:117
INFO Ingesting '2189.json' filestrategy.py:28
INFO Splitting '2189.json' into sections filestrategy.py:30
[06:12:11] INFO Uploading blob for whole file -> 2189.json blobmanager.py:59
INFO Computed embeddings in batch. Batch size: 1, Token count: 205 embeddings.py:117
INFO Ingesting 'employee_handbook.pdf' filestrategy.py:28
INFO Extracting text from './data/employee_handbook.pdf' using Azure Document Intelligence pdfparser.py:68
[06:12:18] INFO Splitting 'employee_handbook.pdf' into sections filestrategy.py:30
INFO Uploading blob for whole file -> employee_handbook.pdf blobmanager.py:59
[06:12:19] INFO Computed embeddings in batch. Batch size: 16, Token count: 3166 embeddings.py:117
INFO Computed embeddings in batch. Batch size: 1, Token count: 65 embeddings.py:117
INFO Ingesting 'Benefit_Options.pdf' filestrategy.py:28
INFO Extracting text from './data/Benefit_Options.pdf' using Azure Document Intelligence pdfparser.py:68
[06:12:23] INFO Splitting 'Benefit_Options.pdf' into sections filestrategy.py:30
[06:12:24] INFO Uploading blob for whole file -> Benefit_Options.pdf blobmanager.py:59
INFO Computed embeddings in batch. Batch size: 8, Token count: 1033 embeddings.py:117
Deploying services (azd deploy)
(✓) Done: Deploying service backend
- Endpoint: https://capps-backend-qdi3sfrmznq6k.yellowsand-76d3c294.eastus.azurecontainerapps.io/
SUCCESS: Your up workflow to provision and deploy to Azure completed in 6 minutes 16 seconds.
yazhini01@vm4powerbi01:~/rag/.azure/env4rag01$

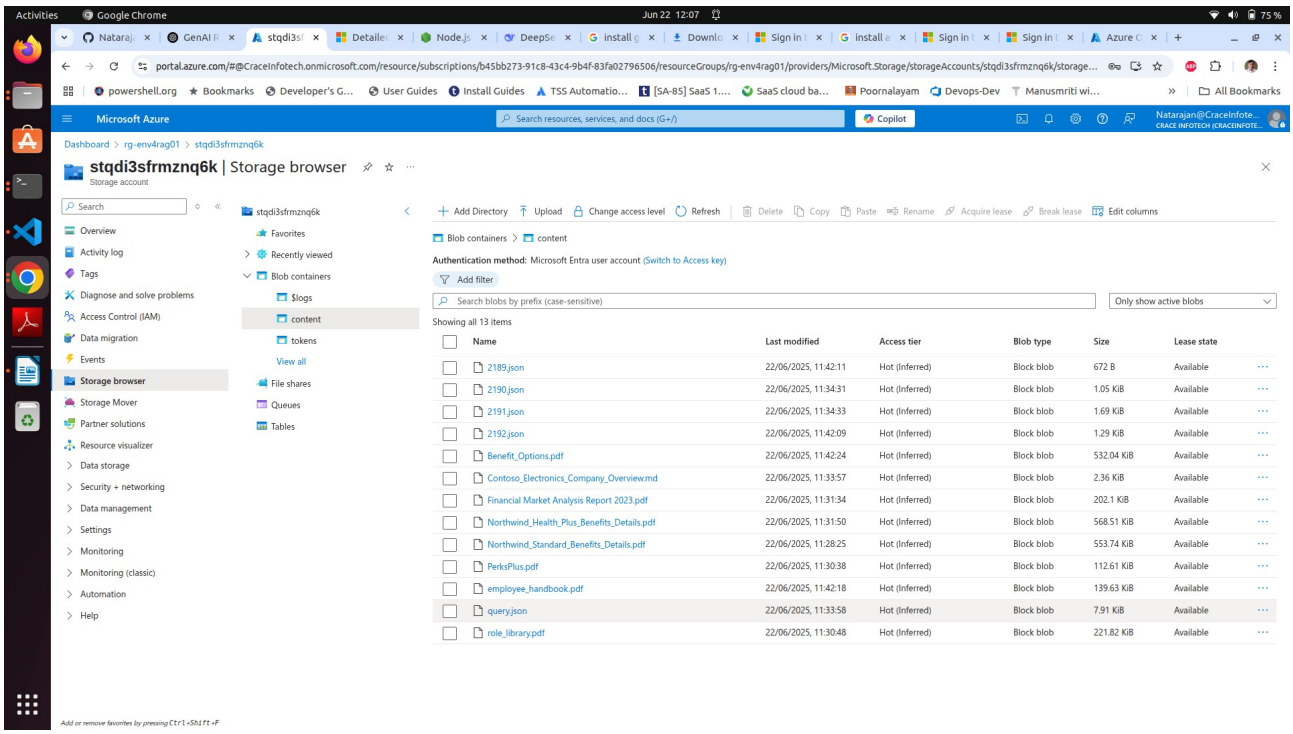

Activities
Google Chrome
Jun 22 11:49
79 %
capps-backend-qdi3sfrmznq6k.yellowsand-76d3c294.eastus.azurecontainerapps.io
Azure OpenAI + AI Search
Chat
Ask a question
Clear chat
Developer settings
Chat with your data
Ask anything or try an example
What is included in my Northwind Health Plus plan that is not in standard?
What happens in a performance review?
What does a Product Manager do?
Type a new question (e.g. does my plan cover annual eye exams?)

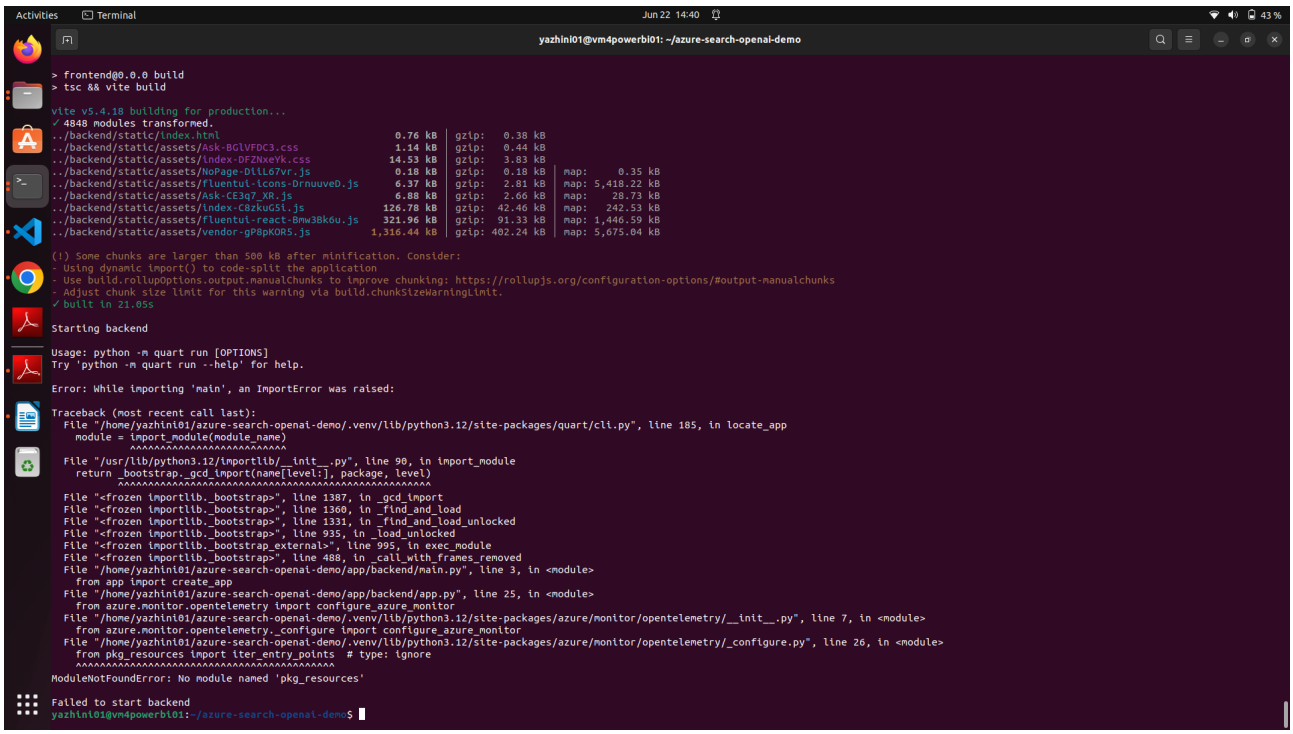

Activities
Google Chrome
Jun 22 11:49
78 %
capps-backend-qdi3sfrmznq6k.yellowsand-76d3c294.eastus.azurecontainerapps.io
Azure OpenAI + AI Search
Chat
Ask a question
Clear chat
Developer settings
What does a Product Manager do?
A Product Manager leads the product management team and collaborates with internal teams (engineering, sales, marketing, finance) and external partners to ensure successful product execution. Their responsibilities include developing and implementing product strategies, managing the product lifecycle from concept to launch, researching customer needs, analyzing market trends, developing product roadmaps, pricing strategies, and marketing plans. They also oversee product portfolios, performance metrics, budgets, and customer feedback to identify improvements. Additionally, they establish relationships with key customers, partners, and vendors, and lead cross-functional teams to meet product goals [1] [2] [3].
Citation:
1. role_library.pdf#page=29
2. role_library.pdf#page=23
3. role_library.pdf#page=17
Type a new question (e.g. does my plan cover annual eye exams?)

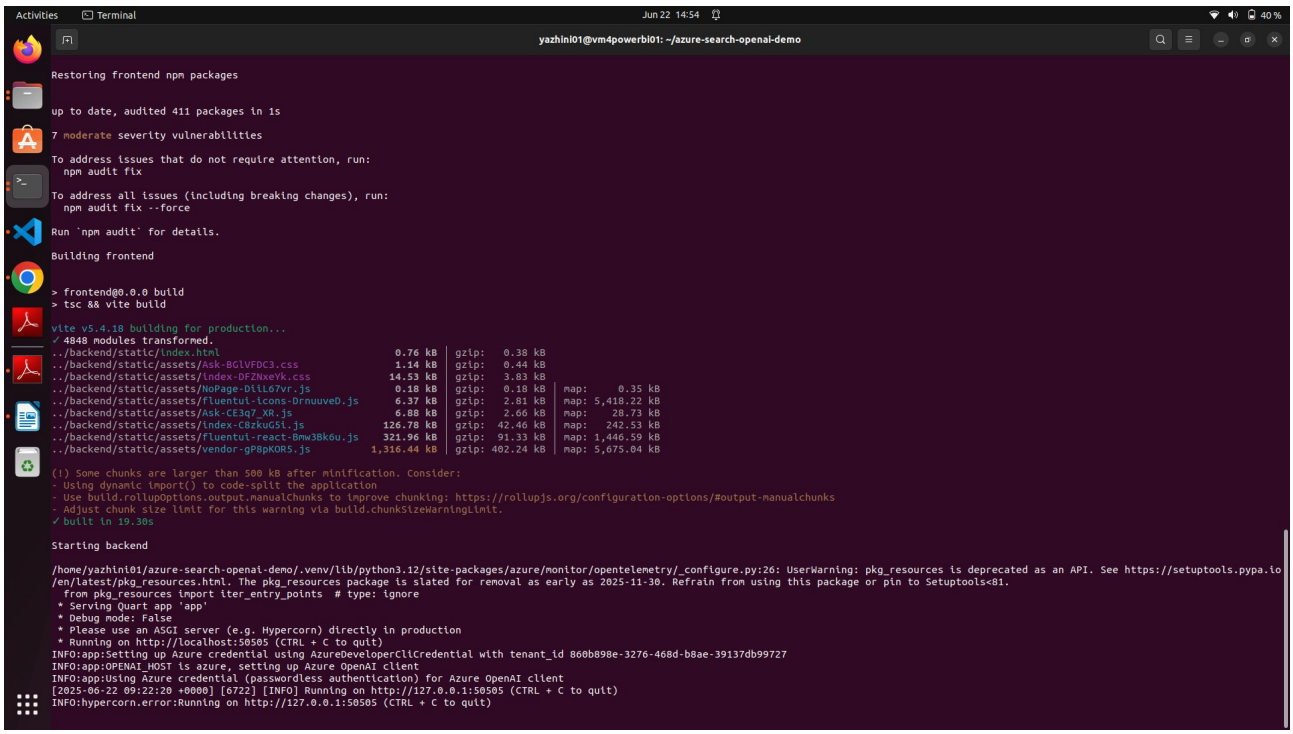

Activities
Google Chrome
Jun 22 11:55
77 %
portal.azure.com/#@CraceInfotech.onmicrosoft.com/resource/subscriptions/b45bb273-91c8-43c4-9b4f-83fa02796506/resourceGroups/rg-env4rag01/overview
Microsoft Azure
Search resources, services, and docs (G+/)
Copilot
Home > Resource groups >
rg-env4rag01
Resource group
Search
Overview
Activity log
Access control (IAM)
Tags
Resource visualizer
Events
Settings
Cost Management
Cost analysis
Cost alerts (preview)
Budgets
Advisor recommendations
Monitoring
Automation
Help
Create
Manage view
Delete resource group
Refresh
Export to CSV
Open query
Assign tags
Move
Delete
Export template
Open in mobile
Essentials
JSON View
Subscription (move) : CRACE INFOTECH 365
Subscription ID : b45bb273-91c8-43c4-9b4f-83fa02796506
Tags (edit) : azd-env-name : env4rag01
Deployments : 1 Failed,35 Succeeded
Location : East US
Resources
Recommendations
Filter for any field...
Type equals all
Location equals all
Add filter
Showing 1 to 12 of 12 records.
Show hidden types
No grouping
List view
Name
Type
Location
appi-qdi3sfrmznq6k Application Insights East US
Application Insights Smart Detection Action group Global
capps-backend-qdi3sfrmznq6k Container App East US
cog-di-qdi3sfrmznq6k Document intelligence East US
cog-qdi3sfrmznq6k Azure OpenAI East US
dash-qdi3sfrmznq6k Shared dashboard East US
env4rag01-aca-env Container Apps Environment East US
env4rag01-aca-identity Managed Identity East US
env4rag01acrqdi3sfrmznq6k Container registry East US
gptkb-qdi3sfrmznq6k Search service East US
log-qdi3sfrmznq6k Log Analytics workspace East US
< Previous
Page 1 of 1
Next >
Give feedback
Add or remove favorites by pressing Ctrl+Shift+F

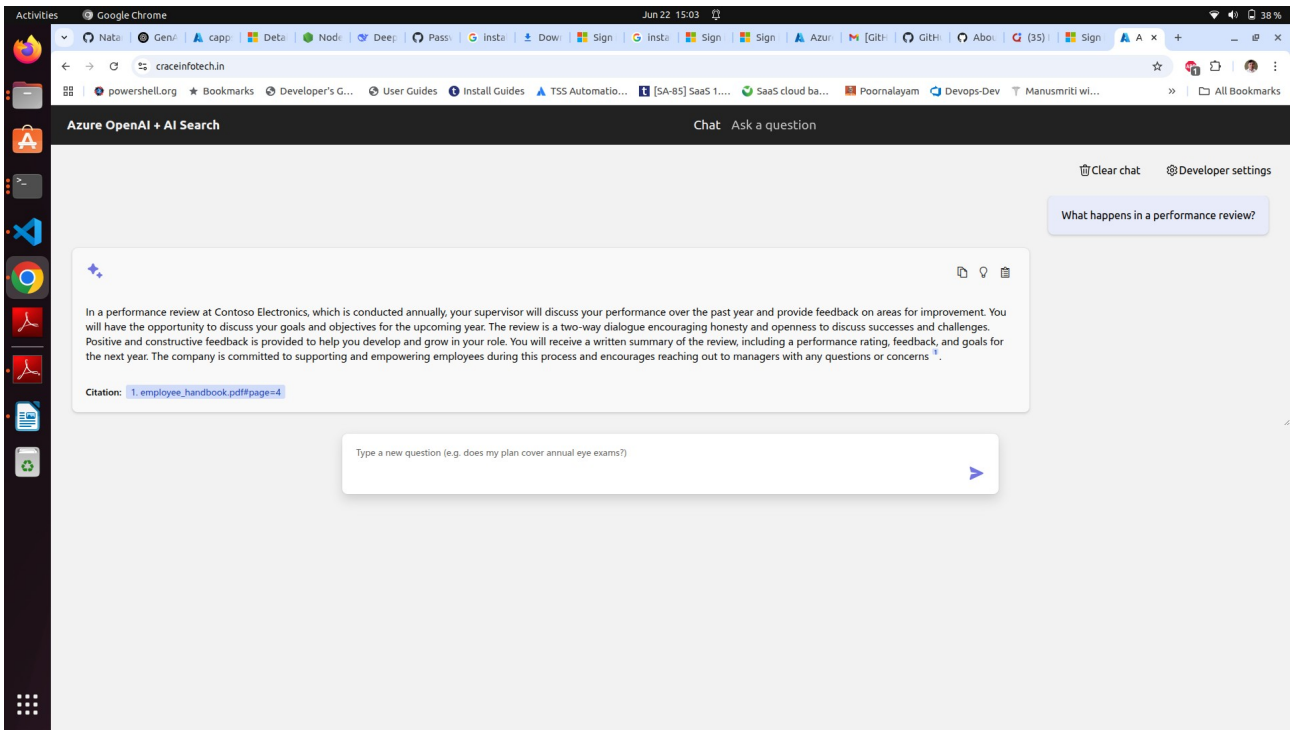

Activities
Google Chrome
Jun 22 11:59
77 %
ai.azure.com/resource/deployments?wsid=/subscriptions/b45bb273-91c8-43c4-9b4f-83fa02796506/resourceGroups/rg-env4rag01/providers/Microsoft.CognitiveServices/accounts/cog-qdi3sfrmznq6k&tid=860b898e-3276-468d...
powershell.org
Bookmarks
Developer's G...
User Guides
Install Guides
TSS Automatio...
[SA-85] SaaS 1....
SaaS cloud ba...
Poornalayam
Devops-Dev
Manusmriti wi...
All Bookmarks
Azure AI Foundry | Azure OpenAI
cog-qdi3sfrmznq6k
Models + endpoints
All resources
cog-qdi3sfrmznq6k
(eastus, S0)
Home
Get started
Model catalog
Playgrounds
Chat
Assistants PREVIEW
Video PREVIEW
Audio PREVIEW
Images
Completions
Tools
Fine-tuning
Azure OpenAI Evaluation PREVIEW
Stored completions PREVIEW
Batch jobs
Monitoring
Shared resources
Deployments
Quota
Guardrails + Controls
Risks + alerts PREVIEW
Data files
Vector stores PREVIEW
Model deployments
Help
Model deployments
Service endpoints
Deploy model
Refresh
Edit
Delete
Open in playground
Reset view
Columns
Name
Model name
Model version
State
Model retirement date
Content filter
Deployment type
Fine-tune
gpt-4.1-mini
gpt-4.1-mini
2025-04-14
Succeeded
Apr 11, 2026 5:30 AM
DefaultV2
Global Standard
text-embedding-3-large
text-embedding-3-large
1
Succeeded
Apr 30, 2026 5:30 AM
DefaultV2
Global Standard

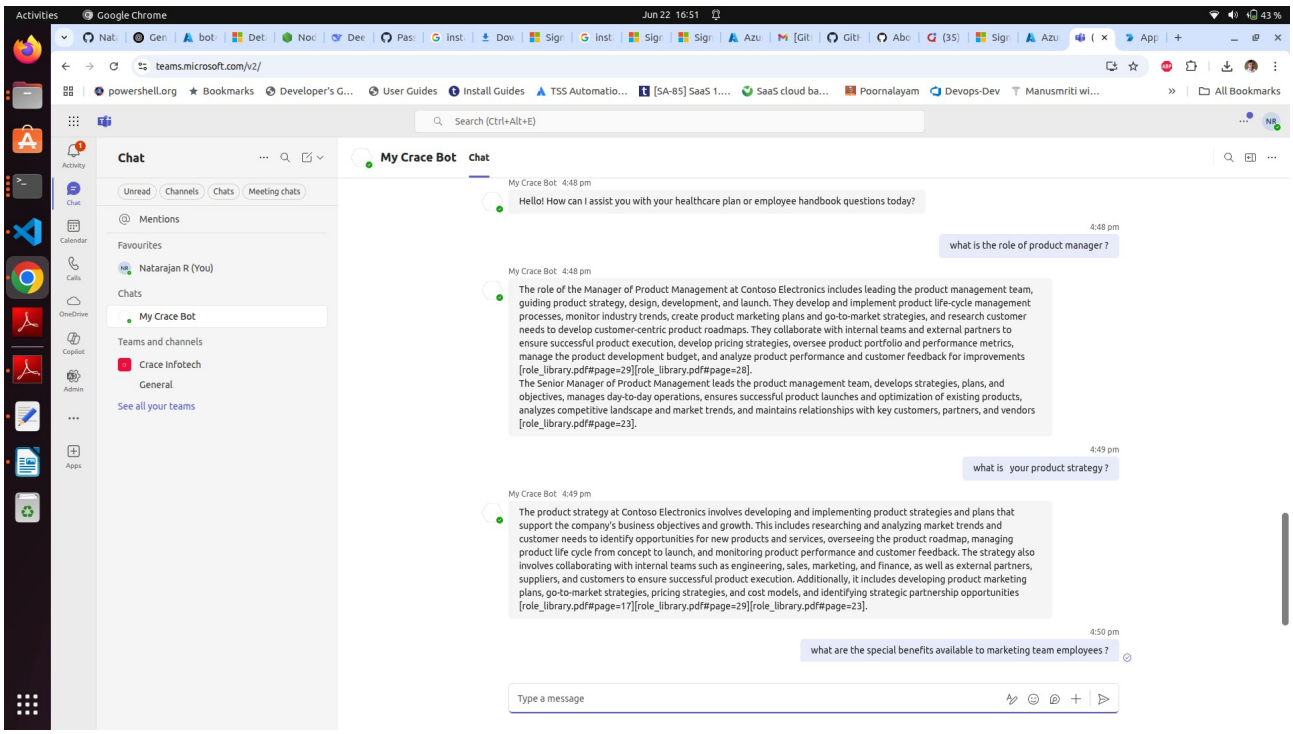

Activities
Google Chrome
Jun 22 12:04
75 %
portal.azure.com/#@CraceInfotech.onmicrosoft.com/resource/subscriptions/b45bb273-91c8-43c4-9b4f-83fa02796506/resourceGroups/rg-env4rag01/providers/Microsoft.Search/searchServices/gptkb-qdi3sfrmznq6k/overv...
Microsoft Azure
Search resources, services, and docs (G+/)
Copilot
Natarajan@CraceInfote...
CRACE INFOTECH (CRACEINFOTE...
Dashboard > rg-env4rag01 >
gptkb-qdi3sfrmznq6k
Search service
Search
Overview
Activity log
Access control (IAM)
Tags
Diagnose and solve problems
Resource visualizer
Search management
Indexes
Indexers
Data sources
Aliases
Skillsets
Debug sessions
Settings
Monitoring
Automation
Help
Add index
Import data
Import and vectorize data
Search explorer
Upgrade
Refresh
Delete
Move
Essentials
JSON View
Resource group (move) : rg-env4rag01
Location (move) : East US
Subscription (move) : CRACE INFOTECH 365
Subscription ID : b45bb273-91c8-43c4-9b4f-83fa02796506
Status : Running
Date created : Jun 22, 2025, 11:19:27 AM
Tags (edit) : azd-env-name : env4rag01
Url : https://gptkb-qdi3sfrmznq6k.search.windows.net
Pricing tier : Basic
Replicas : 1 (No SLA)
Partitions : 1
Search units : 1
Get started
Properties
Usage
Monitoring
Revolutionary retrieval with Azure AI Search
Don't know where to start? Here are some options from directly within the portal
Connect your data
Start here to import your data. Learn how to quickly connect to your data to build your first search index. Learn more
Import
Explore your data
Connect to apps, optimize search results. Leverage features like faceting, filtering, scoring profiles and more. Learn more
View
Monitor and scale
Tools that allow you to monitor your system and scale for optimal performance. Adjust replicas and partitions as needed. Learn more
View
Add or remove favorites by pressing Ctrl+Shift+F

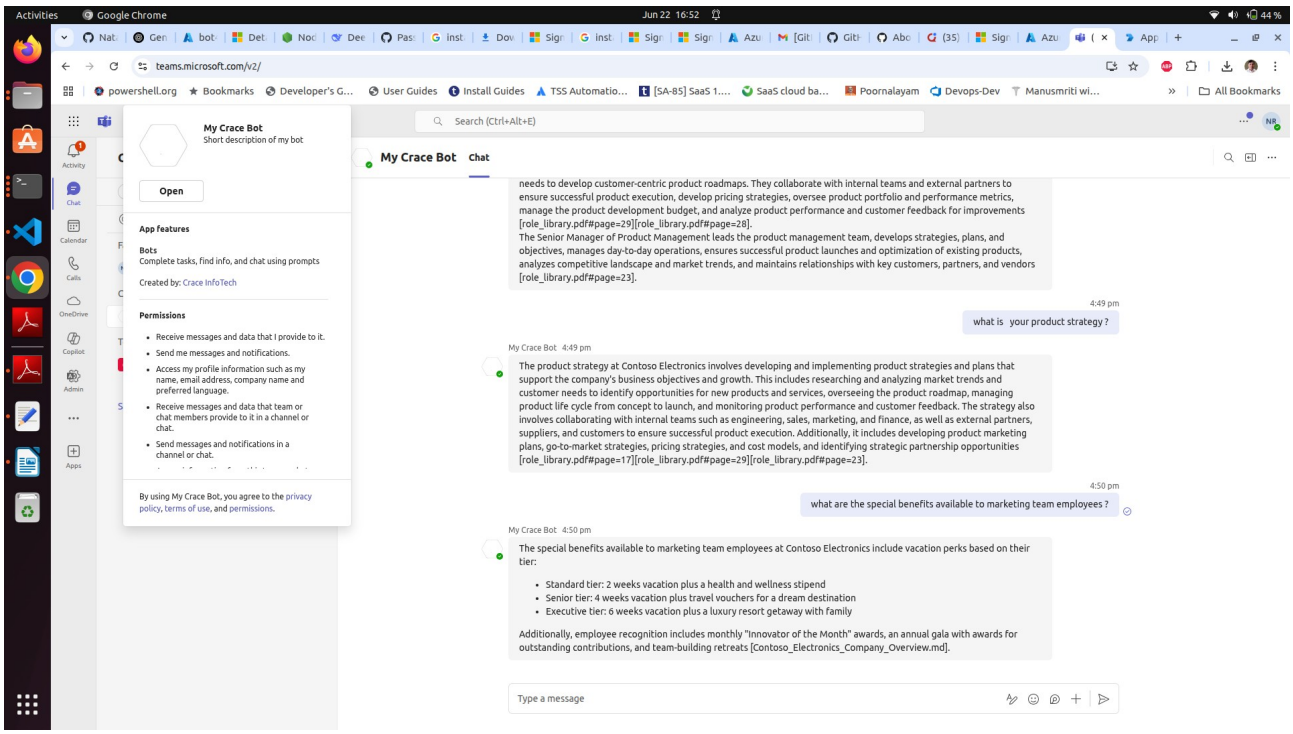

Activities
Google Chrome
Jun 22 12:07
75 %
portal.azure.com/#@CraceInfotech.onmicrosoft.com/resource/subscriptions/b45bb273-91c8-43c4-9b4f-83fa02796506/resourceGroups/rg-env4rag01/providers/Microsoft.Storage/storageAccounts/stqdi3sfrmznq6k/storage...
Microsoft Azure
Search resources, services, and docs (G+/)
Copilot
Natarajan@CraceInfote...
CRACE INFOTECH (CRACEINFOTE...
Dashboard > rg-env4rag01 > stqdi3sfrmznq6k
stqdi3sfrmznq6k | Storage browser
Storage account
Search
Overview
Activity log
Tags
Diagnose and solve problems
Access Control (IAM)
Data migration
Events
Storage browser
Storage Mover
Partner solutions
Resource visualizer
Data storage
Security + networking
Data management
Settings
Monitoring
Monitoring (classic)
Automation
Help
Add or remove favorites by pressing Ctrl+Shift+F
stqdi3sfrmznq6k
Favorites
Recently viewed
Blob containers
$logs
content
tokens
View all
File shares
Queues
Tables
Add Directory
Upload
Change access level
Refresh
Delete
Copy
Paste
Rename
Acquire lease
Break lease
Edit columns
Blob containers > content
Authentication method: Microsoft Entra user account (Switch to Access key)
Add filter
Search blobs by prefix (case-sensitive)
Only show active blobs
Showing all 13 items
Name
Last modified
Access tier
Blob type
Size
Lease state
2189.json 22/06/2025, 11:42:11 Hot (Inferred) Block blob 672 B Available
2190.json 22/06/2025, 11:34:31 Hot (Inferred) Block blob 1.05 KiB Available
2191.json 22/06/2025, 11:34:33 Hot (Inferred) Block blob 1.69 KiB Available
2192.json 22/06/2025, 11:42:09 Hot (Inferred) Block blob 1.29 KiB Available
Benefit_Options.pdf 22/06/2025, 11:42:24 Hot (Inferred) Block blob 532.04 KiB Available
Contoso_Electronics_Company_Overview.md 22/06/2025, 11:33:57 Hot (Inferred) Block blob 2.36 KiB Available
Financial Market Analysis Report 2023.pdf 22/06/2025, 11:31:34 Hot (Inferred) Block blob 202.1 KiB Available
Northwind_Health_Plus_Benefits_Details.pdf 22/06/2025, 11:31:50 Hot (Inferred) Block blob 568.51 KiB Available
Northwind_Standard_Benefits_Details.pdf 22/06/2025, 11:28:25 Hot (Inferred) Block blob 553.74 KiB Available
PerksPlus.pdf 22/06/2025, 11:30:38 Hot (Inferred) Block blob 112.61 KiB Available
employee_handbook.pdf 22/06/2025, 11:42:18 Hot (Inferred) Block blob 139.63 KiB Available
query.json 22/06/2025, 11:33:58 Hot (Inferred) Block blob 7.91 KiB Available
role_library.pdf 22/06/2025, 11:30:48 Hot (Inferred) Block blob 221.82 KiB Available


Activities
Terminal
Jun 22 14:40
43 %
yazhini01@vm4powerbi01: ~/azure-search-openai-demo
> frontend@0.0.0 build
> tsc && vite build
vite v5.4.18 building for production...
✓ 4848 modules transformed.
../backend/static/index.html 0.76 kB | gzip: 0.38 kB
../backend/static/assets/Ask-BGlVFDC3.css 1.14 kB | gzip: 0.44 kB
../backend/static/assets/index-DFZNxeYk.css 14.53 kB | gzip: 3.83 kB
../backend/static/assets/NoPage-DlLG7vr.js 0.18 kB | gzip: 0.18 kB | map: 0.35 kB
../backend/static/assets/fluentui-icons-DrnuuveD.js 6.37 kB | gzip: 2.81 kB | map: 5,418.22 kB
../backend/static/assets/Ask-CE3q7_XR.js 6.88 kB | gzip: 2.66 kB | map: 28.73 kB
../backend/static/assets/index-C8zkuG5i.js 126.78 kB | gzip: 42.46 kB | map: 242.53 kB
../backend/static/assets/fluentui-react-Bmw3Bk6u.js 321.96 kB | gzip: 91.33 kB | map: 1,446.59 kB
../backend/static/assets/vendor-gP8pKOR5.js 1,316.44 kB | gzip: 402.24 kB | map: 5,675.04 kB
(!) Some chunks are larger than 500 kB after minification. Consider:
- Using dynamic import() to code-split the application
- Use build.rollupOptions.output.manualChunks to improve chunking: https://rollupjs.org/configuration-options/#output-manualchunks
- Adjust chunk size limit for this warning via build.chunkSizeWarningLimit.
✓ built in 21.05s
Starting backend
Usage: python -m quart run [OPTIONS]
Try 'python -m quart run --help' for help.
Error: While importing 'main', an ImportError was raised:
Traceback (most recent call last):
File "/home/yazhini01/azure-search-openai-demo/.venv/lib/python3.12/site-packages/quart/cli.py", line 185, in locate_app
module = import_module(module_name)
File "/usr/lib/python3.12/importlib/__init__.py", line 90, in import_module
return _bootstrap._gcd_import(name[level:], package, level)
File "<frozen importlib._bootstrap>", line 1387, in _gcd_import
File "<frozen importlib._bootstrap>", line 1360, in _find_and_load
File "<frozen importlib._bootstrap>", line 1331, in _find_and_load_unlocked
File "<frozen importlib._bootstrap>", line 935, in _load_unlocked
File "<frozen importlib._bootstrap_external>", line 995, in exec_module
File "<frozen importlib._bootstrap>", line 488, in _call_with_frames_removed
File "/home/yazhini01/azure-search-openai-demo/app/backend/main.py", line 3, in <module>
from app import create_app
File "/home/yazhini01/azure-search-openai-demo/app/backend/app.py", line 25, in <module>
from azure.monitor.opentelemetry import configure_azure_monitor
File "/home/yazhini01/azure-search-openai-demo/.venv/lib/python3.12/site-packages/azure/monitor/opentelemetry/__init__.py", line 7, in <module>
from azure.monitor.opentelemetry._configure import configure_azure_monitor
File "/home/yazhini01/azure-search-openai-demo/.venv/lib/python3.12/site-packages/azure/monitor/opentelemetry/_configure.py", line 26, in <module>
from pkg_resources import iter_entry_points # type: ignore
ModuleNotFoundError: No module named 'pkg_resources'
Failed to start backend
yazhini01@vm4powerbi01:~/azure-search-openai-demo$


Activities
Terminal
Jun 22 14:54
40 %
yazhini01@vm4powerbi01: ~/azure-search-openai-demo
Restoring frontend npm packages
up to date, audited 411 packages in 1s
7 moderate severity vulnerabilities
To address issues that do not require attention, run:
npm audit fix
To address all issues (including breaking changes), run:
npm audit fix --force
Run `npm audit` for details.
Building frontend
> frontend@0.0.0 build
> tsc && vite build
vite v5.4.18 building for production...
✓ 4848 modules transformed.
../backend/static/index.html 0.76 kB | gzip: 0.38 kB
../backend/static/assets/Ask-BGlVFDC3.css 1.14 kB | gzip: 0.44 kB
../backend/static/assets/index-DFZNxeYk.css 14.53 kB | gzip: 3.83 kB
../backend/static/assets/NoPage-DiLG7vr.js 0.18 kB | gzip: 0.18 kB | map: 0.35 kB
../backend/static/assets/fluentui-icons-DrnuuveD.js 6.37 kB | gzip: 2.81 kB | map: 5,418.22 kB
../backend/static/assets/Ask-CE3q7_X8.js 6.88 kB | gzip: 2.66 kB | map: 28.73 kB
../backend/static/assets/index-C8zkuG5i.js 126.78 kB | gzip: 42.46 kB | map: 242.53 kB
../backend/static/assets/fluentui-react-Bmw3Bk6u.js 321.96 kB | gzip: 91.33 kB | map: 1,446.59 kB
../backend/static/assets/vendor-gP8pKOR5.js 1,316.44 kB | gzip: 402.24 kB | map: 5,675.04 kB
(!) Some chunks are larger than 500 kB after minification. Consider:
- Using dynamic import() to code-split the application
- Use build.rollupOptions.output.manualChunks to improve chunking: https://rollupjs.org/configuration-options/#output-manualchunks
- Adjust chunk size limit for this warning via build.chunkSizeWarningLimit.
✓ built in 19.30s
Starting backend
/home/yazhini01/azure-search-openai-demo/.venv/lib/python3.12/site-packages/azure/monitor/opentelemetry/_configure.py:26: UserWarning: pkg_resources is deprecated as an API. See https://setuptools.pypa.io/en/latest/pkg_resources.html. The pkg_resources package is slated for removal as early as 2025-11-30. Refrain from using this package or pin to Setuptools<81.
from pkg_resources import iter_entry_points # type: ignore
* Serving Quart app 'app'
* Debug mode: False
* Please use an ASGI server (e.g. Hypercorn) directly in production
* Running on http://localhost:50505 (CTRL + C to quit)
INFO:app:Setting up Azure credential using AzureDeveloperCliCredential with tenant_id 860b898e-3276-468d-b8ae-39137db99727
INFO:app:OPENAI_HOST is azure, setting up Azure OpenAI client
INFO:app:Using Azure credential (passwordless authentication) for Azure OpenAI client
[2025-06-22 09:22:20 +0000] [6722] [INFO] Running on http://127.0.0.1:50505 (CTRL + C to quit)
INFO:hypercorn.error:Running on http://127.0.0.1:50505 (CTRL + C to quit)


Activities
Google Chrome
Jun 22 15:03
38 %
craceinfotech.in
powershell.org
Bookmarks
Developer's G...
User Guides
Install Guides
TSS Automatio...
[SA-85] SaaS 1....
SaaS cloud ba...
Poornalayam
Devops-Dev
Manusmriti wi...
All Bookmarks
Azure OpenAI + AI Search
Chat
Ask a question
Clear chat
Developer settings
What happens in a performance review?
In a performance review at Contoso Electronics, which is conducted annually, your supervisor will discuss your performance over the past year and provide feedback on areas for improvement. You will have the opportunity to discuss your goals and objectives for the upcoming year. The review is a two-way dialogue encouraging honesty and openness to discuss successes and challenges. Positive and constructive feedback is provided to help you develop and grow in your role. You will receive a written summary of the review, including a performance rating, feedback, and goals for the next year. The company is committed to supporting and empowering employees during this process and encourages reaching out to managers with any questions or concerns [1].
Citation:
1. employee_handbook.pdf#page=4
Type a new question (e.g. does my plan cover annual eye exams?)


Activities
Google Chrome
Jun 22 16:51
teams.microsoft.com/v2/
Search (Ctrl+Alt+E)
Chat
Unread
Channels
Chats
Meeting chats
Mentions
Favourites
Natarajan R (You)
Chats
My Crace Bot
Teams and channels
Crace Infotech
General
See all your teams
My Crace Bot
Chat
My Crace Bot 4:48 pm
Hello! How can I assist you with your healthcare plan or employee handbook questions today?
4:48 pm
what is the role of product manager ?
My Crace Bot 4:48 pm
The role of the Manager of Product Management at Contoso Electronics includes leading the product management team, guiding product strategy, design, development, and launch. They develop and implement product life-cycle management processes, monitor industry trends, create product marketing plans and go-to-market strategies, and research customer needs to develop customer-centric product roadmaps. They collaborate with internal teams and external partners to ensure successful product execution, develop pricing strategies, oversee product portfolio and performance metrics, manage the product development budget, and analyze product performance and customer feedback for improvements [role_library.pdf#page=29][role_library.pdf#page=28].
The Senior Manager of Product Management leads the product management team, develops strategies, plans, and objectives, manages day-to-day operations, ensures successful product launches and optimization of existing products, analyzes competitive landscape and market trends, and maintains relationships with key customers, partners, and vendors [role_library.pdf#page=23].
4:49 pm
what is your product strategy ?
My Crace Bot 4:49 pm
The product strategy at Contoso Electronics involves developing and implementing product strategies and plans that support the company's business objectives and growth. This includes researching and analyzing market trends and customer needs to identify opportunities for new products and services, overseeing the product roadmap, managing product life cycle from concept to launch, and monitoring product performance and customer feedback. The strategy also involves collaborating with internal teams such as engineering, sales, marketing, and finance, as well as external partners, suppliers, and customers to ensure successful product execution. Additionally, it includes developing product marketing plans, go-to-market strategies, pricing strategies, and cost models, and identifying strategic partnership opportunities [role_library.pdf#page=17][role_library.pdf#page=29][role_library.pdf#page=23].
4:50 pm
what are the special benefits available to marketing team employees ?
Type a message


Activities
Google Chrome
Jun 22 16:52
teams.microsoft.com/v2/
My Crace Bot
Short description of my bot
Open
App features
Bots
Complete tasks, find info, and chat using prompts
Created by: Crace InfoTech
Permissions
Receive messages and data that I provide to it.
Send me messages and notifications.
Access my profile information such as my name, email address, company name and preferred language.
Receive messages and data that team or chat members provide to it in a channel or chat.
Send messages and notifications in a channel or chat.
By using My Crace Bot, you agree to the privacy policy, terms of use, and permissions.
Search (Ctrl+Alt+E)
My Crace Bot
Chat
needs to develop customer-centric product roadmaps. They collaborate with internal teams and external partners to ensure successful product execution, develop pricing strategies, oversee product portfolio and performance metrics, manage the product development budget, and analyze product performance and customer feedback for improvements [role_library.pdf#page=29][role_library.pdf#page=28].
The Senior Manager of Product Management leads the product management team, develops strategies, plans, and objectives, manages day-to-day operations, ensures successful product launches and optimization of existing products, analyzes competitive landscape and market trends, and maintains relationships with key customers, partners, and vendors [role_library.pdf#page=23].
4:49 pm
what is your product strategy ?
My Crace Bot 4:49 pm
The product strategy at Contoso Electronics involves developing and implementing product strategies and plans that support the company's business objectives and growth. This includes researching and analyzing market trends and customer needs to identify opportunities for new products and services, overseeing the product roadmap, managing product life cycle from concept to launch, and monitoring product performance and customer feedback. The strategy also involves collaborating with internal teams such as engineering, sales, marketing, and finance, as well as external partners, suppliers, and customers to ensure successful product execution. Additionally, it includes developing product marketing plans, go-to-market strategies, pricing strategies, and cost models, and identifying strategic partnership opportunities [role_library.pdf#page=17][role_library.pdf#page=29][role_library.pdf#page=23].
4:50 pm
what are the special benefits available to marketing team employees ?
My Crace Bot 4:50 pm
The special benefits available to marketing team employees at Contoso Electronics include vacation perks based on their tier:
Standard tier: 2 weeks vacation plus a health and wellness stipend
Senior tier: 4 weeks vacation plus travel vouchers for a dream destination
Executive tier: 6 weeks vacation plus a luxury resort getaway with family
Additionally, employee recognition includes monthly "Innovator of the Month" awards, an annual gala with awards for outstanding contributions, and team-building retreats [Contoso_Electronics_Company_Overview.md].
Type a message

1 What’s Next and How You Can Join Me

I plan to:

- Extend this to ingest Data from many sources
- Add action-based prompts (form filling, KB updates) along with AI Agents as Knowledge Tools
- Build plug-and-play modules for different enterprise domains
- In Electronics, if anything can be converted into electrical signal, we can automate them. Ie Simple Sensors gives us extensive power in making things Digital. But Sensors are somewhat expensive, but the same effect, ie converting domain variables into computer signals is very easy with LLMs. Thus we have enoromous virtual sensor capability with the LLMs. It increases our Automation capability also in the same scale.
- I started my career in Electronics and learned Microprocessors programming (from 4bit, 8 bit, to current 64 bit). So I can see through what is happening inside Microprocessors and CPUs (All of data and flags) . Now I can extend the same knowledge and experience in understandaing LLM and thier peripherals interactions clearly.

If you:

- Want to try this on your own and need guidance
- Have internal use cases that could benefit from an AI assistant
- Want to collaborate on open-source contributions

Feel free to **comment, DM me, or connect**. I’d love to help you build your GenAI journey.

Let’s Connect!

If this post resonated with you or sparked new ideas, let’s connect. I’m always open to conversations around:

- Generative AI & RAG
- Enterprise automation
- Azure AI ecosystem
- Intelligent bots & assistants

THANK YOU