



AZURE DAY

Build Intelligent Apps with Microsoft's Copilot stack & Azure OpenAI

Dave Rendon



Microsoft



TECHNOLOGY



Platinum Sponsor



Technical Sponsor



Hello 🖐️

Thank you for joining me today

Dave Rendon

Azure & AI MVP,
Microsoft Certified Trainer

twitter.com/daverndn

linkedin.com/in/daverndn

[Blog.azinsider.net](https://blog.azinsider.net)





Learning Objectives

- You will learn how to leverage the Microsoft Copilot Stack to design and build AI-powered business applications that improve productivity and deliver highly personalized user experiences.
- Imparerai a sfruttare lo stack Microsoft Copilot per progettare e sviluppare applicazioni aziendali potenziate dall'IA, che migliorano la produttività e offrono esperienze utente altamente personalizzate.



Skills and Tools

- **Generative AI:** Understanding fundamentals and applications.
- **Azure OpenAI:** Deploying AI solutions using Azure OpenAI services.
- **Retrieval Augmented Generation (RAG) Patterns:** Improving AI models with RAG.
- **Semantic Kernel:** Using Semantic Kernel for AI-driven tasks.
- **Microsoft Copilot Stack:** Leveraging the Microsoft Copilot stack to build intelligent systems.



Build Intelligent Apps – How?

Large Language Models (LLMs) : AI models trained on vast amounts of text data to understand, generate, and interact using human-like language capabilities.

Azure Open AI: Provides REST API access to OpenAI's powerful language models including o1-preview, o1-mini, GPT-4o, GPT-4o mini, GPT-4 Turbo with Vision, GPT-4, GPT-3.5-Turbo, and Embeddings model series

Semantic Kernel: A framework that enables the integration of advanced language models and custom logic for building AI-driven applications with semantic understanding.

Azure AI Search : A cloud-based search service that provides advanced search capabilities, including full-text search, filtering, and AI-powered insights across various data sources.



Agents

Goal Setting and
Tracking Agent

Investment
Portfolio Manager
(Swarm)

Expense tracking
& Budgeting

User preferences
(Teachable Agent)

Charitable Giving
Advisor

Retirement
Planner (Group
Agent)

Multi-agent Orchestrator



AutoGen



Semantic Kernel

Reasoning engines



Azure
OpenAI



Azure AI
OSS LLMs

Planning Reflection



Assistants API



TaskWeaver

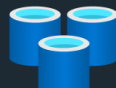
Memory



AI Search
RaG/Vector DB



CosmosDB
Long Term Memory



Redis
Conversation cache



Fabric
Ent. Lakehouse

Runtimes



ACA



AKS



App Service



Functions

Security & Governance



AI Content
Safety Studio



Defender
for Cloud



Entra
MI, AD

Integration



APIM



Logic Apps



Service Bus

LLMOps



PromptFlow
LLM Tracing + Eval



Monitor
AIOps



Workflow

- Azure Open AI and AI Search
- Configure and integrate a sample application with Azure Kubernetes Service (AKS).
- Utilize Azure AI Search
- Expose Azure OpenAI via API Management
- Leverage Semantic Kernel



Prerequisites

Azure OpenAI: oai-copilot

Models:

- CopilotCompletionModel
- CopilotEmbeddingModel

Azure Open AI endpoint/key:

- Key: b06214e1243c4181b0728158734057ca
- Endpoint: <https://oai-copilot.openai.azure.com/>

Azure AI Search URL: <https://acs-copilot.search.windows.net>

Azure AI Search Key: ETd97Me8EuW5brzBnEswvK0veHC9CQieM6p0os78BcAzSeBWFGc3

Cosmos DB name: cosmos-copilot

Cosmos URI: <https://cosmos-copilot.documents.azure.com:443/>

Cosmos Connection String

Storage account: stcopilot002

Endpoint: <https://stcopilot002.blob.core.windows.net/>

APIM URL: <https://springtoys-copilot-apim.azure-api.net>

AOAI APIM Copilot Test: 2407b86ebf544028911c67068bf23609

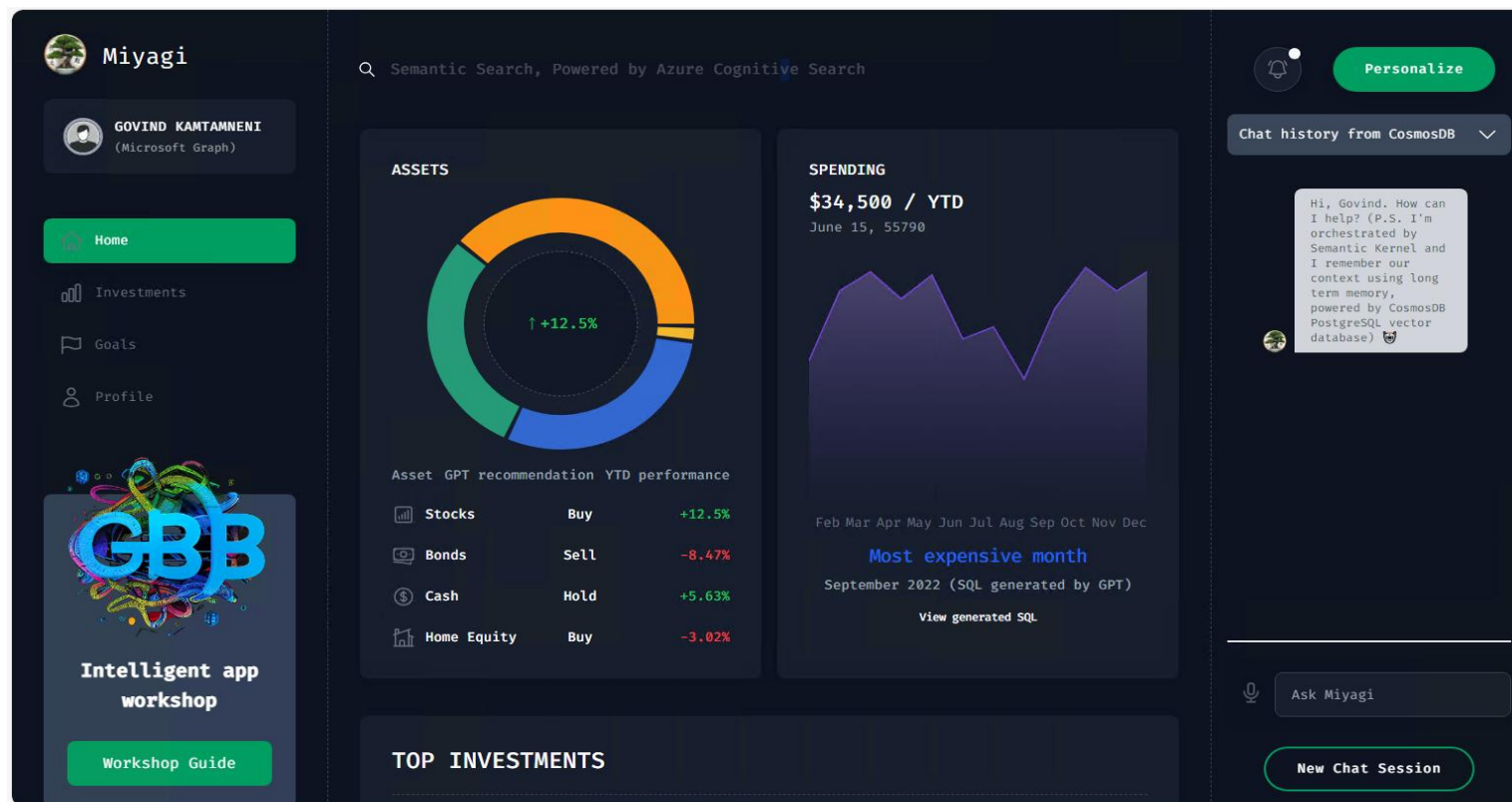
Event Hub Shared Access Policy Connection string



Run the app locally

Project Miyagi leverages Microsoft's Copilot Stack to help you build and deploy smart apps.

- Miyagui UI
- Recommendation service

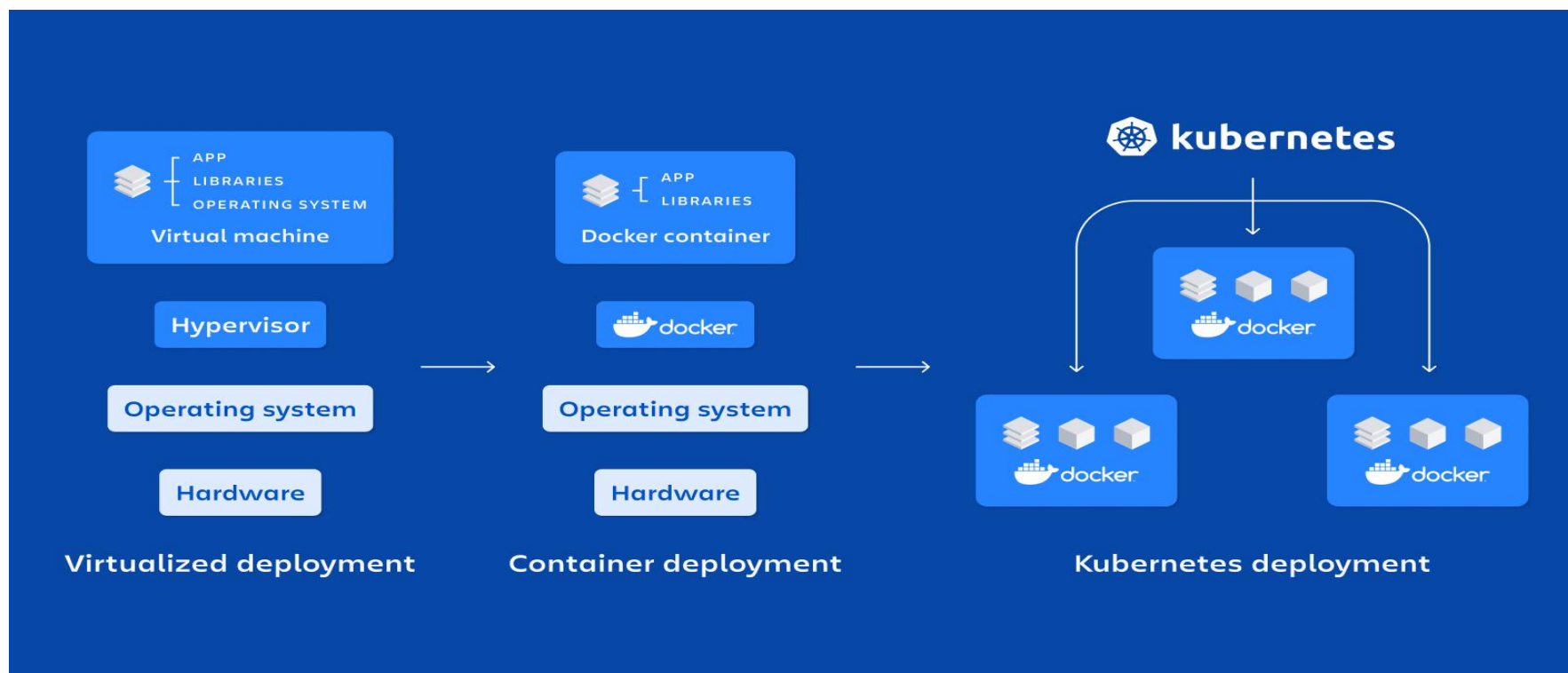




Containerizing the application

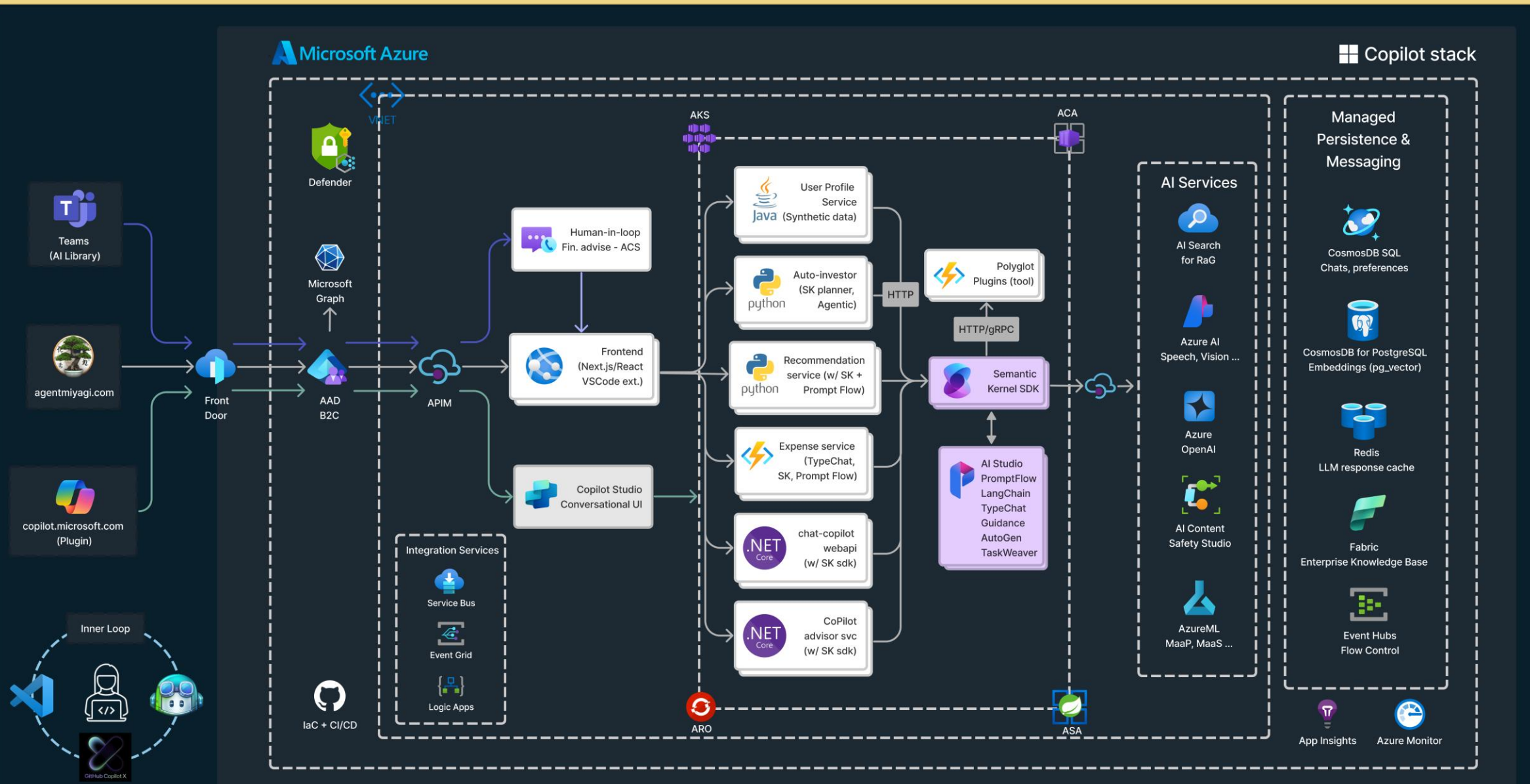
Containerizing Miyagi UI and Recommendation service to Azure Kubernetes Service (AKS):

1. Package these applications into Docker containers
2. Deploying them on AKS for scalable, managed orchestration and efficient operation.





Architecture

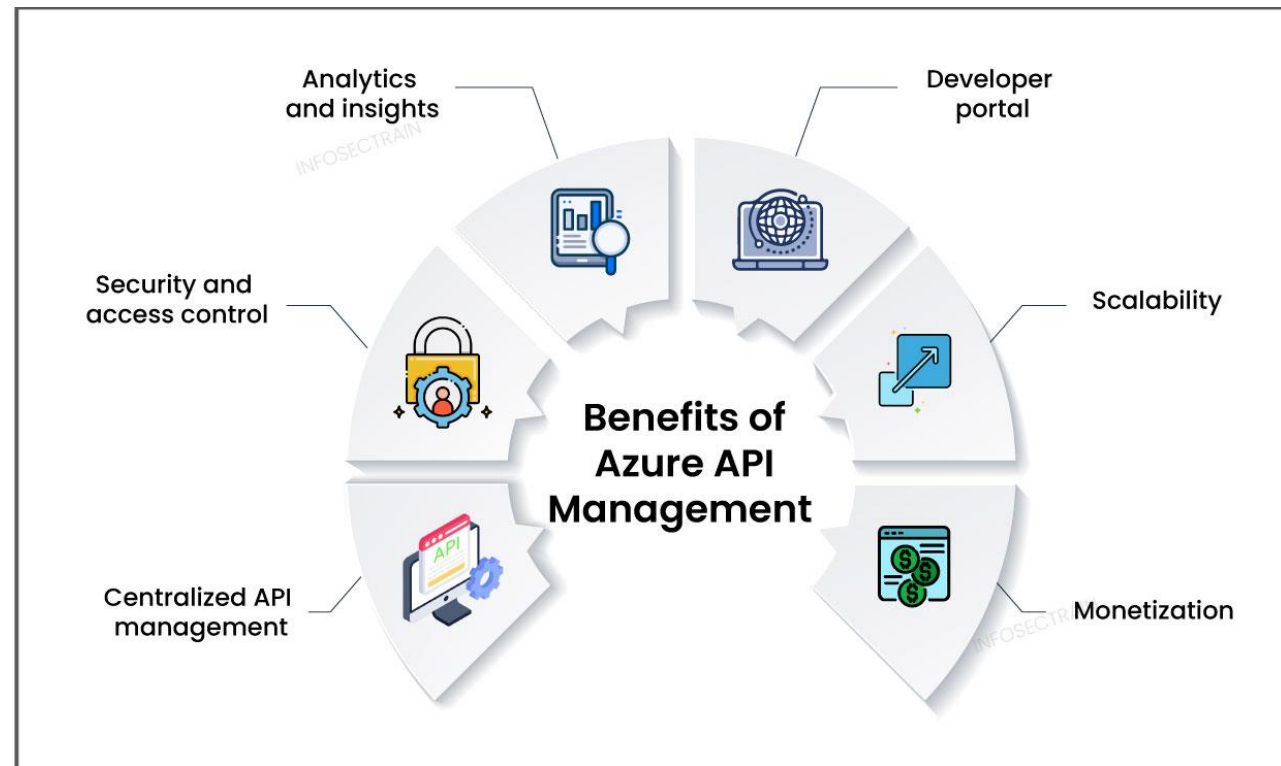




Expose OpenAI using APIM

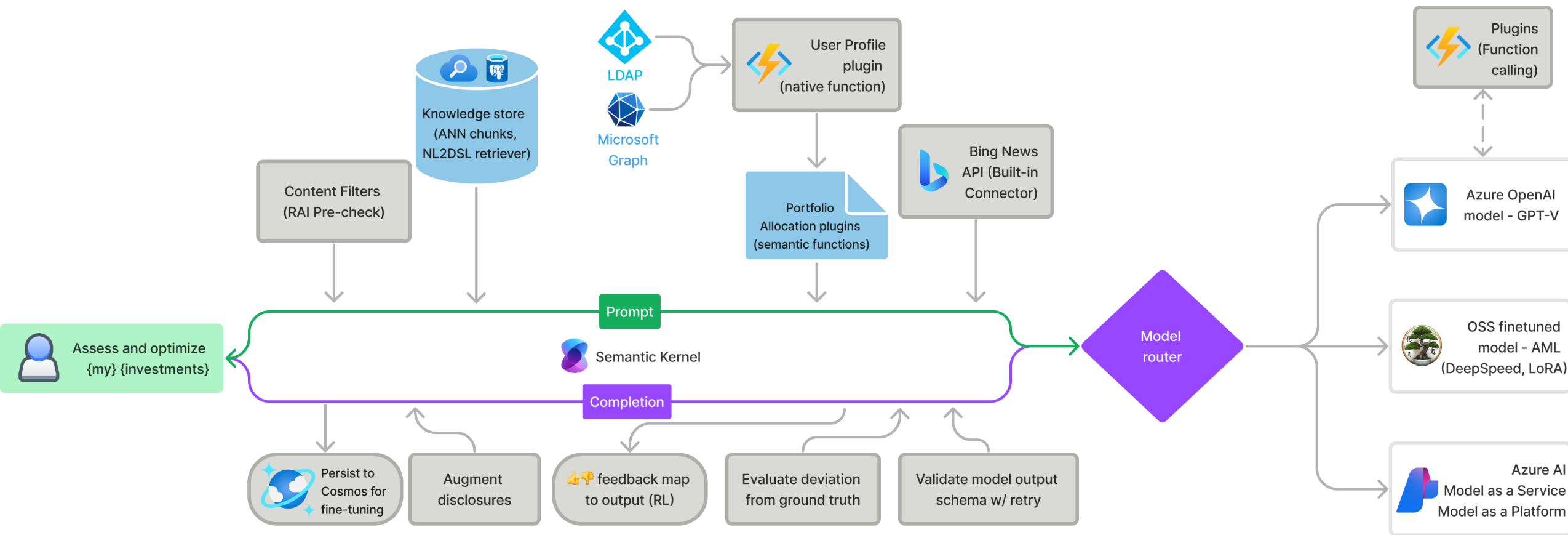
Key features of Azure API Management:

- API Gateway
- Developer Portal
- API Lifecycle Management
- Security and Access Control
- Monitoring and Analytics
- Rate Limiting and Throttling



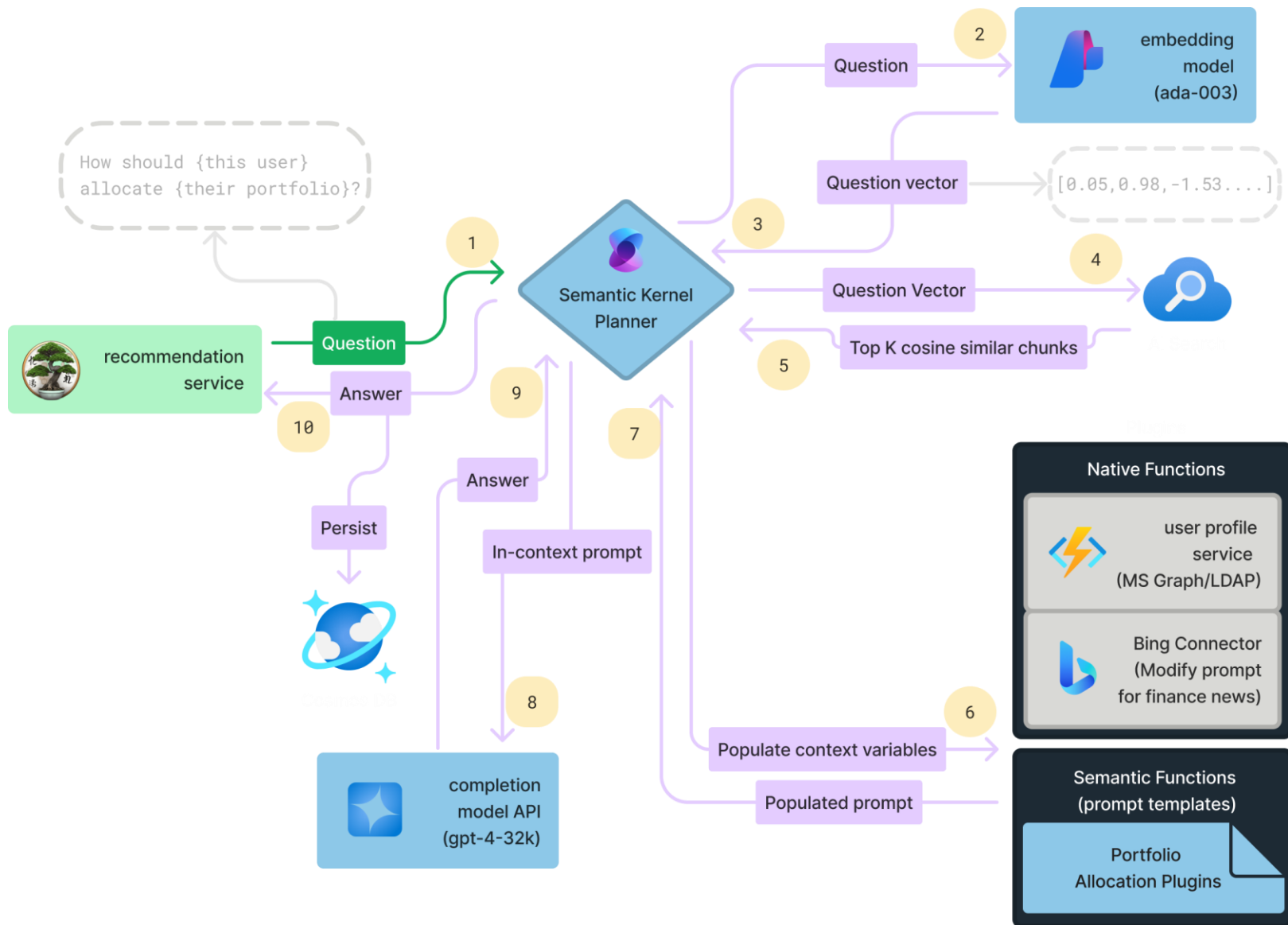


Semantic Kernel – Round trip





Semantic Kernel – Memory Orchestration





Demo - Workflow

Miyagi

GOVIND KAMTAMENI
(Microsoft Graph)

Home

Investments

Goals

Profile

Intelligent app workshop

Workshop Guide

Semantic Search, Powered by Azure Cognitive Search

Asset

Stocks

Bonds

Cash

Home Equity

GPT recommendation

Buy

Sell

Hold

Buy

YTD performance

-8.47%

+5.63%

-3.02%

Semantic Kernel Planner + Skills to orchestrate with GPT-4

TOP INVESTMENTS

| # | Name | Price | 24H Change | GPT Recommendation | 1Y Chart |
|---|--------------------|----------|------------|---|----------|
| 1 | Microsoft MSFT | \$291.24 | 1.06% | BOOYAH! GO BIG AND INVEST IN MSFT - CAPITALIZE ON AI BUZZ | |
| 2 | Accenture ACN | \$282.35 | -2.63% | KEEP UP WITH THE TECH TREND AND ADD ACN TO YOUR PORTFOLIO! | |
| 3 | JPMorgan Chase JPM | \$128.72 | 0.5% | DOLLAR COST AVERAGE OVER THE COMING MONTHS TO CAPITALIZE ON INCREASE DEPOSIT BASE | |
| 4 | PepsiCo PEP | \$165.00 | 0.06% | TAKE ADVANTAGE OF DIVIDENDS FROM PEP AND DIVERSIFY INTELLIGENTLY! | |

EXPENSES

| Date | Category | Description | Amount |
|------|----------|-------------|--------|
| | | | |

Personalize

Chat history from CosmosDB

Hi, Govind. How can I help? (P.S. I remember our context using long term memory via a vector database)

Phase 1: Semantic Kernel Copilot Chat

Work in progress: I will create solutions using Azure OpenAI GPT-4. In this demo, I will utilize Web PubSub, Semantic Kernel, and Qdrant (on AKS/ACA) to orchestrate this chat interaction with GPT-4. Alternatively, wait until Jarvis/AutoGPT matures in order to self-develop without GPT-4.

Phase 2: BYO weights (RAG & RLHF): MiyagiGPT, finetuned by DeepSpeed-Chat

We will need to invest slightly more in RLHF alignment than in prompt engineering!

Ask Miyagi

New Chat Session

GPT recommends to reduce expenses by 5.2% vs. current trend to maintain positive cash flow

Vector Store (private data embeddings) for in-context learning



Thank You!

Feedback



Platinum Sponsor



Technical Sponsor

