Microsoft Azure

CLOUD CHAMPION

Intro to Azure AI Agent Service

February 2025

Introduction – Guy Gregory

Partner Solution Architect,
EMEA Partner Team (GPS)

Aligned to Azure AI

Focus on Channel Partner

20+ years of Microsoft partner experience

https://linkedin.com/in/guygregory

Agenda

01

Introducing Azure AI Agent Service

02

Demo 1: Azure AI Agent Service in the Foundry portal UI experience

03

Ecosystem of tools

04

Demo 2: Contoso Enterprise Agent

05

What next? Multi-agent orchestration

06

Useful links

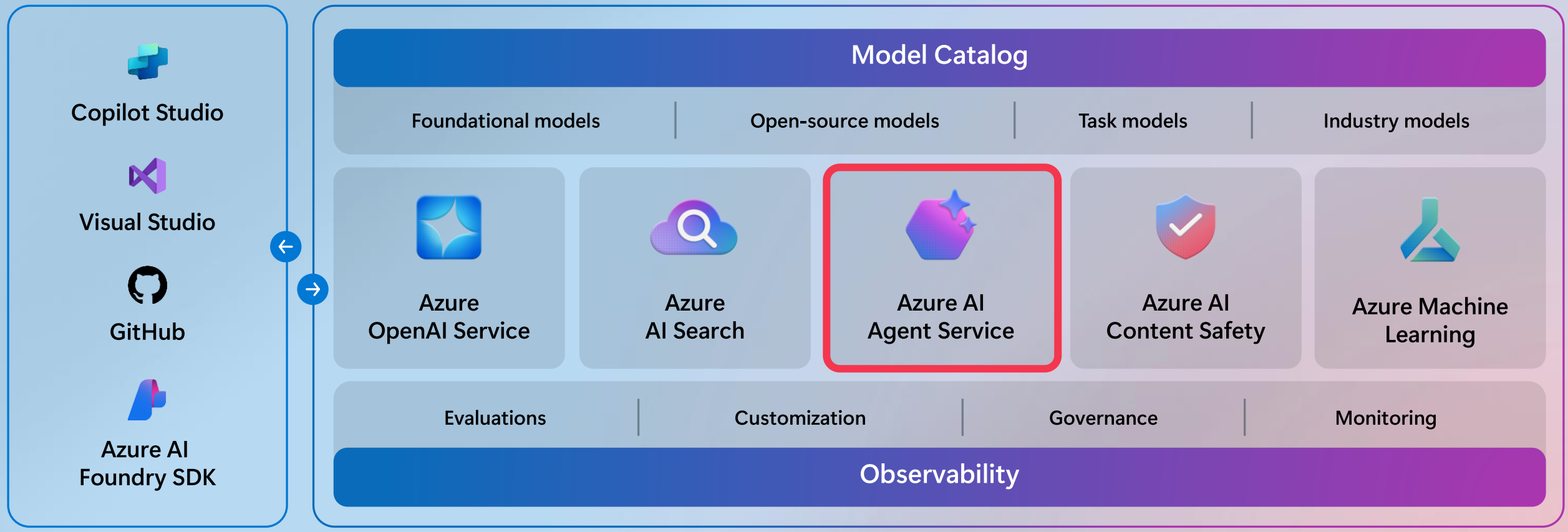
07

Q&A

Ground rules / Disclaimers!

- Azure AI Agent Service is in Preview, and is subject to change
- Parts of the service are not available yet
- My experience is mainly based on 1 month+ of access via the SDK
- Treat this as a ‘first look’ introduction. Objective: from zero to (one) agent.
- Unless it’s mentioned in this deck, I can’t answer detailed roadmap questions - I am not in the product team / engineering
- I don’t have any pricing information to share
- I will demo some code. Some of it might even work(!)
- All of the demo repos are publicly available and can be replicated by anyone with an Azure subscription
- Please ask questions in the Q&A – I’ll address at the end, time-permitting

Azure AI Foundry



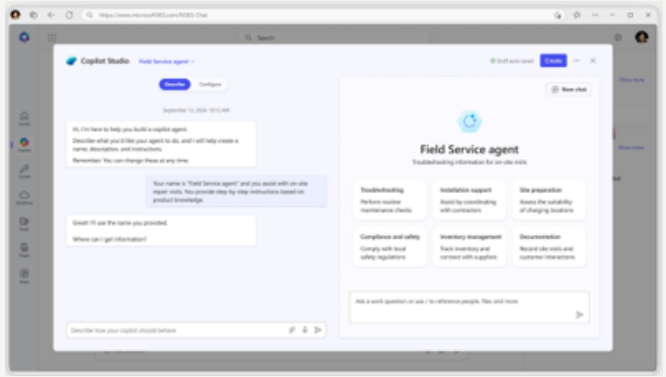
Public Preview

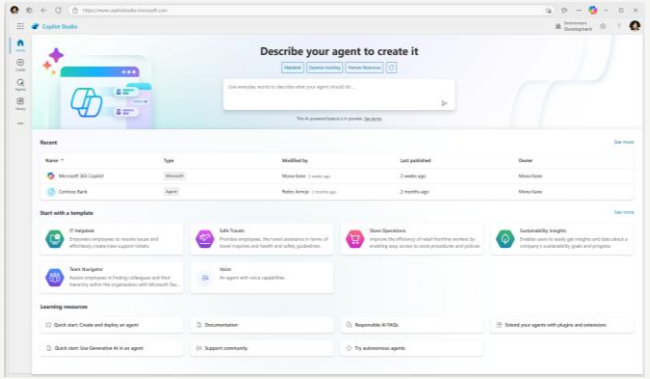
Azure AI Agent Service

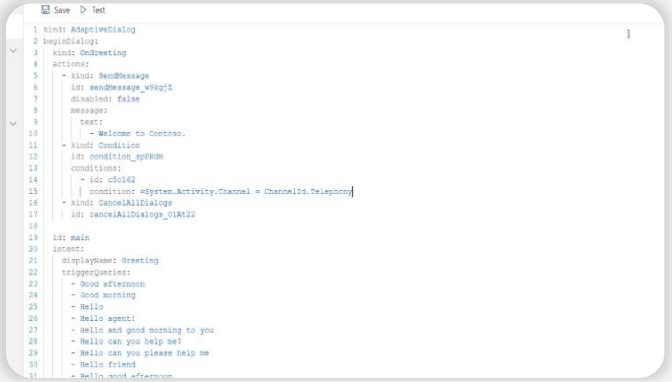
Empower developers to securely build, deploy, and scale AI agents with ease

ai.azure.com

Azure AI Agent Service or Copilot Studio?







Natural Language — Code first

Agent builder

For End Users

Copilot Studio

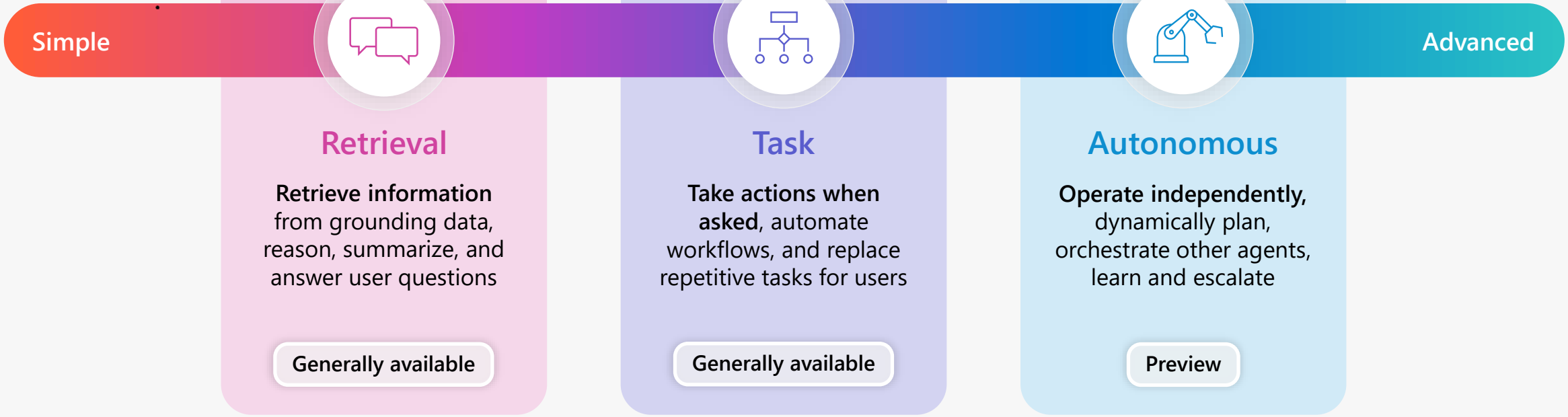
For Makers

Azure AI Foundry+ Visual Studio/GitHub

Powered by Azure AI Agent Service

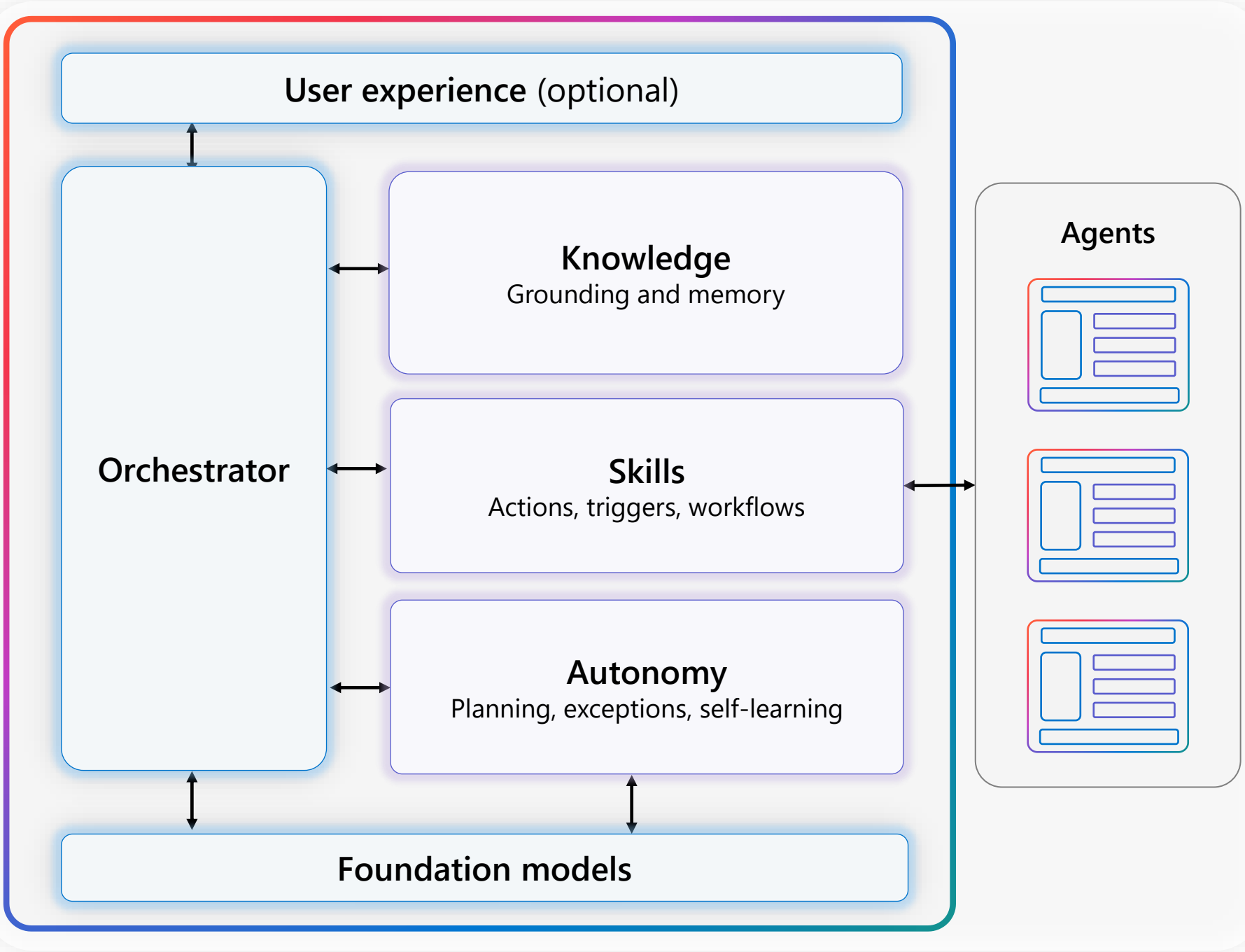
For Developers

What are AI Agents?



Agents vary in levels of complexity and capabilities depending on your need

Agent architecture



Demo 1 – Azure AI Foundry portal UI experience

https://ai.azure.com

Demo file (nifty_500_quarterly_results.csv):
https://aka.ms/cloudchampion/agents/demo1

Chat Completions API vs Assistants API vs Agents

Chat Completions API

- Lightweight and powerful
- Inherently stateless

VS

Assistants API

- Build with OpenAI models
- Stateful (inbuilt conversation state management)
- Access persistent threads
- Automatic management of the model's context window
- Access files in several formats
- Utilized Microsoft-managed storage
- File Search (API handles chunking, embeddings storage and creation, and implementing vector search)
- Code Interpreter
- Function Calling

VS

Azure AI Agent Service

- **Assistants API features, plus**
- Build with a model of your choice (OpenAI, Llama, Mistral, Cohere…)
- Real time web-grounding with Bing
- Secure grounding on enterprise data in SharePoint and Fabric
- Bring Your Own Licensed data (Tripadvisor)
- Connect to 1400+ data sources and services with Azure Logic Apps
- Long running, event driven actions with Azure Functions
- Standardized OpenAPI 3.0 tools
- Bring Your Own Storage
- Bring Your Own Private Network
- Bring Your Own AI Search Resource
- Limitless scaling with PTUs
- Open Telemetry based tracing

The full package

Azure AI Foundry SDK – Agent Service

Built-in enterprise readiness

BYO-file storage (coming soon)

BYO-search index

OBO Authorization Support

Enhanced Observability

Extensive Ecosystem of Tools

Knowledge

- Microsoft Fabric*
- SharePoint*
- Grounding with Bing Search
- Azure AI Search
- Files (local or Azure Blob)
- Your own licensed data*
- File Search
- Code Interpreter

Actions

- Azure Logic Apps*
- OpenAPI 3.0 Specified Tools
- Azure Functions

Model Catalog

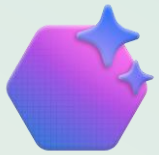
Azure OpenAI Service
(GPT-4o, GPT-4o mini)

Models-as-a-Service

- Llama 3.1-405B-Instruct
- Mistral Large
- Cohere-Command-R-Plus

*Indicates feature is coming soon

Public Preview



Azure AI Agent Service

Empower developers to securely build, deploy, and scale AI agents with ease

Rapid Development and Automation	Extensive Data Connections	Flexible Model Selection	Enterprise Readiness

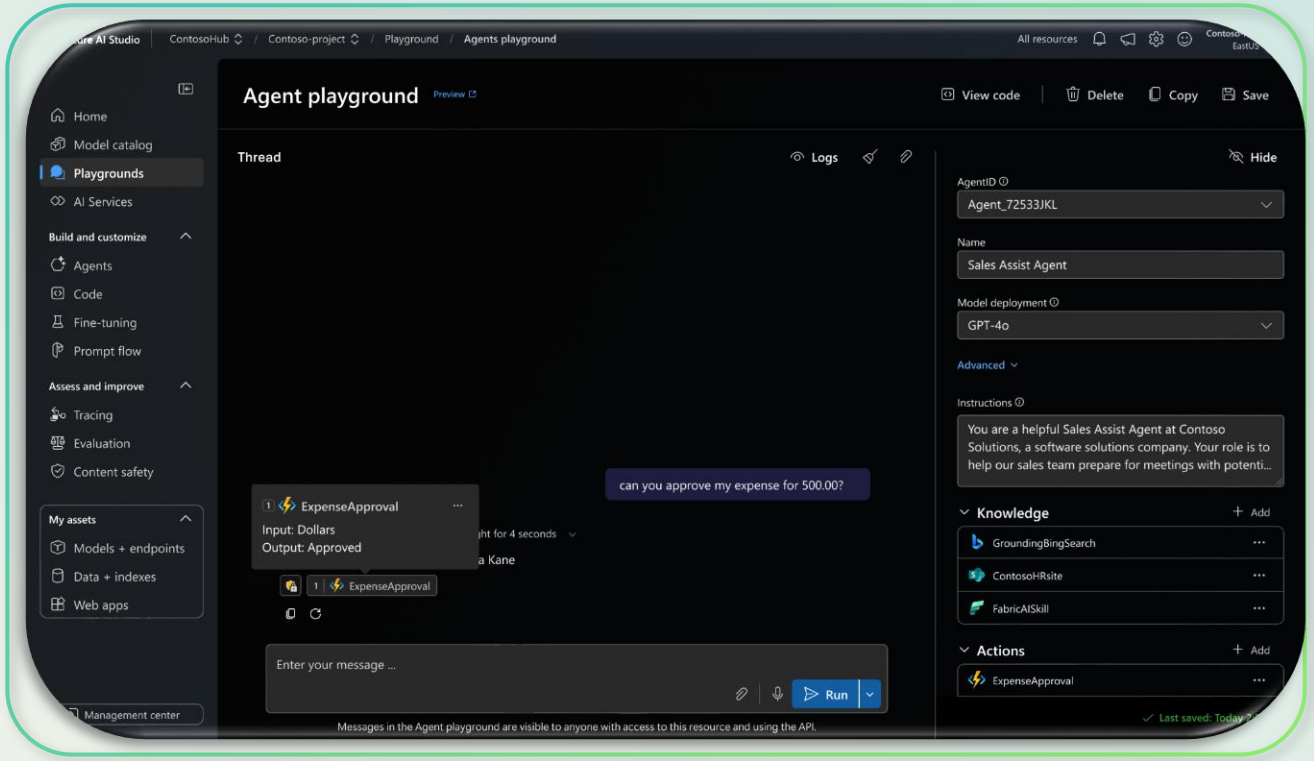
ai.azure.com

Rapid Development and Automation

Implement stateless or stateful code-based actions with Azure Functions

Automation

- Enable your agent to perform external interactions with the world, such as calling APIs or asynchronously sending and waiting for events.
- Azure Functions and Azure Durable Actions enable you to execute serverless code for **synchronous, asynchronous, long-running, and event-driven actions** such as approving invoices with **human-in-the-loop**, monitor an end-to-end product supply chain over long periods of time, and many more.

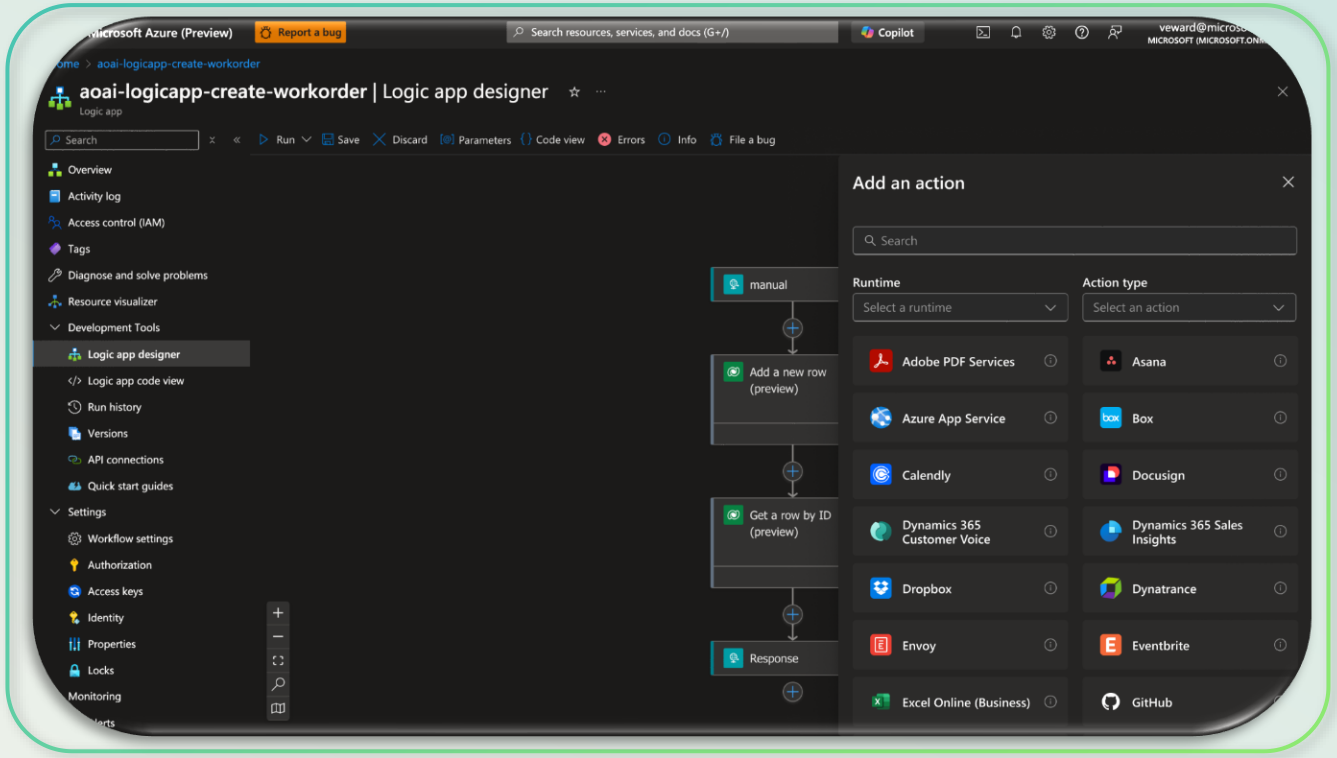


Playground experience in Azure AI Foundry portal is coming soon in January

Automation

1,400 built-in connectors through Azure Logic Apps (coming soon)

- Create and run automated workflows with Azure Logic Apps with little to no code using the visual designer and selecting from prebuilt operations.
- Simply define the business logic for your workflow to connect your agent to external systems, tools, and APIs.
- Microsoft connectors include Microsoft products such as Azure App Service, Dynamics365 Customer Voice, Microsoft Teams, M365 Excel
- Non-Microsoft connectors include enterprise services such as MongoDB, Dropbox, Jira, Gmail, Twilio, SAP, Stripe, ServiceNow and many more.

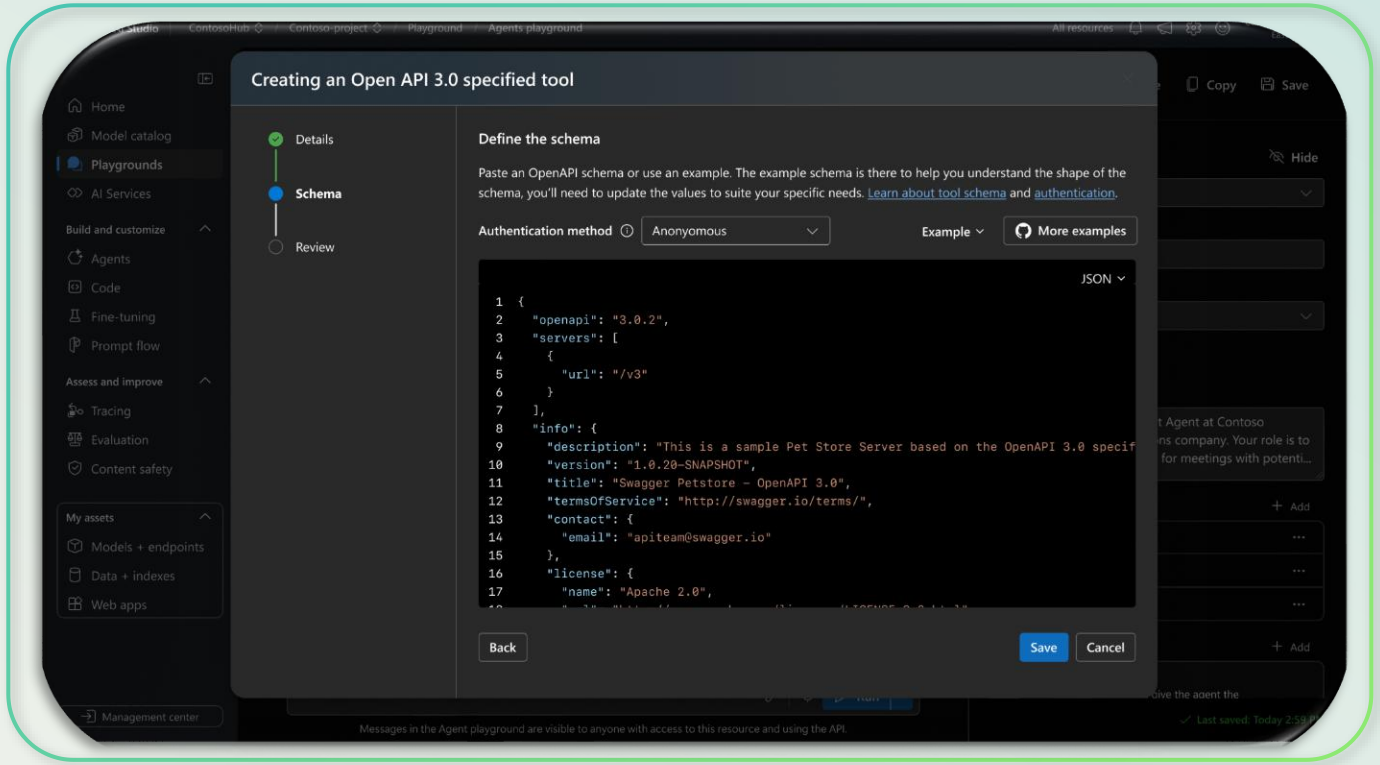


Playground experience in Azure AI Foundry portal is coming soon in January

Build standardized custom tools with OpenAPI 3.0 specifications

Automation

- Connect your AI agent to an external API using an OpenAPI 3.0 specified tool, allowing for scalable interoperability with various applications
- Build an organization-side tools library specified with the OpenAPI 3.0 standard
- Enable your custom tools to authenticate access and connections with managed identities (Microsoft Entra ID) for added security, making it ideal for integrating with existing infrastructure or web services



Playground experience in Azure AI Foundry portal is coming soon in January

Azure AI Agent SDK - OpenAPI 3.x example

Powered by CLOUD CHAMPION

Ask a question:

Give me some links to security training in Spanish

Submit

Response:

Here are some links to security training in Spanish:

1. Conceptos de Seguridad para Pymes - En esta sesión repasaremos lo de M365 enfocado al segmento Pymes.
2. Ciberseguridad: el eslabón más débil - Repasaremos diferentes compe cliente al que podemos dirigirnos para hacer ese upsell.
3. Acelera tu negocio aumentando la seguridad en tus clientes con Micro Microsoft 365 E5 Security - Aprende a aumentar la seguridad en tus cl
4. Acelera la innovación y mejora la seguridad de tu código con GitHub - completo de tus soluciones y las herramientas de GitHub.
5. Seguridad en la nube - Veremos los controles de seguridad en Azure c Trust.
6. Servicio de analítica de seguridad (SIEM/SOAR) - Aprende a posiciona
7. Pregunta cualquier duda sobre cómo vender las soluciones de seguridad de Microsoft. Consulta con especialistas en ventas de Seguridad de Microsoft.
8. ¿Cómo vende Microsoft sus soluciones de seguridad? - Completa este learning path para conseguir la insignia de Security Solutions Sales.
9. Seguridad en la nube: 5 principios básicos que cualquier empresa debe seguir - Webinar que explica cinco principios básicos para proteger el entorno en la nube.
10. Campaña de Modernización de SQL (3/4): Descubra cómo mejorar la seguridad de bases de datos en un entorno de Microsoft Azure - Aprende sobre mejorar la seguridad de bases de datos en Microsoft Azure.

These resources should be helpful for anyone looking to enhance their knowledge of security in a Spanish-speaking context.

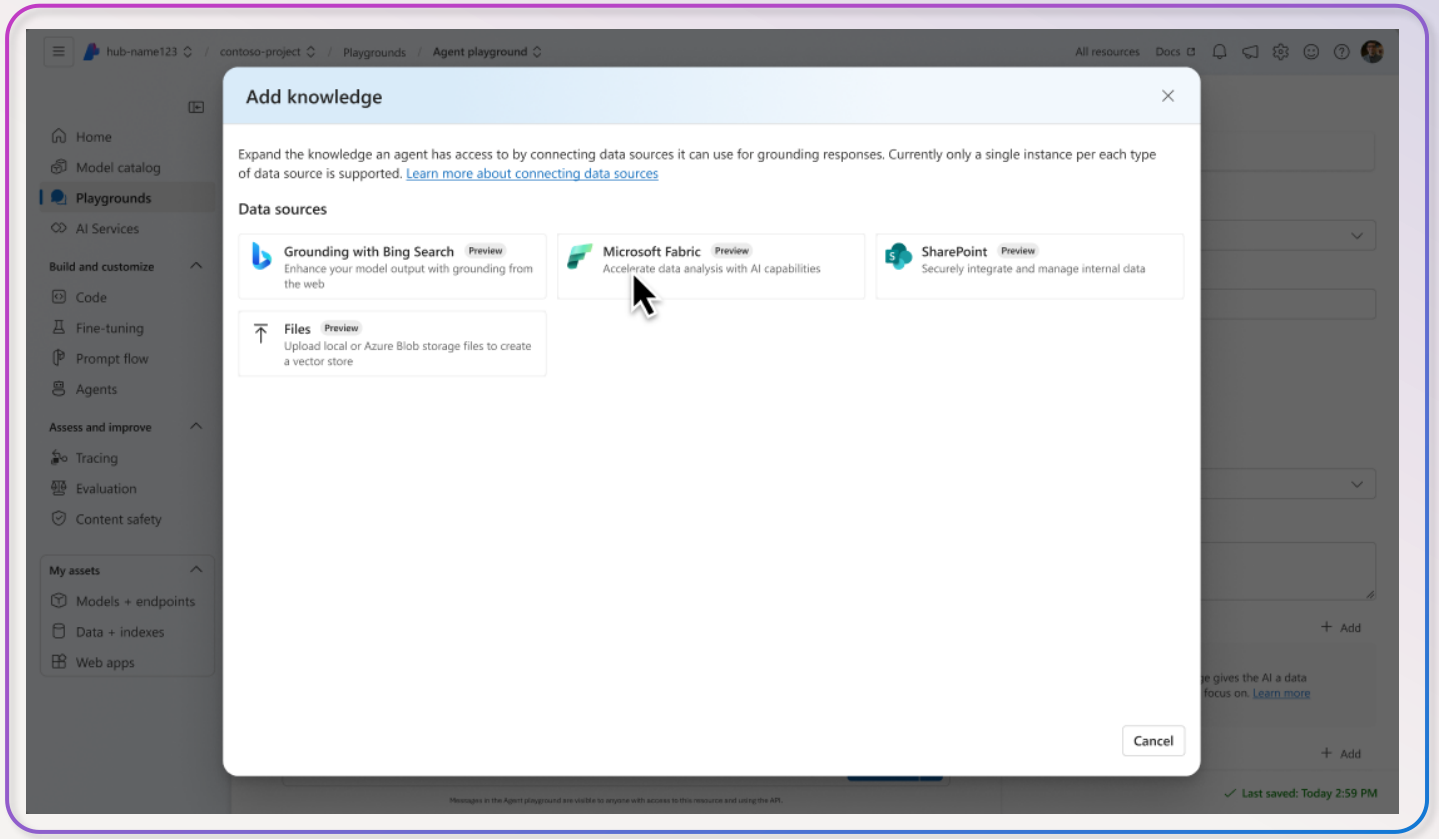
Extensive Data Connections

Ground your AI agents with enterprise knowledge

Offers a rich ecosystem of enterprise knowledge sources to ground your agent's outputs

- Access real-time web search results with Microsoft Bing
- Integrate with SharePoint to retrieve and use internal documents and files (coming soon)
- Leverage built-in AI capabilities in Microsoft Fabric to retrieve structured data insights (coming soon)
- Bring your own vector and hybrid search with Azure AI Search
- Utilize licensed data sources like Tripadvisor to enhance responses (coming soon)

Data Connections

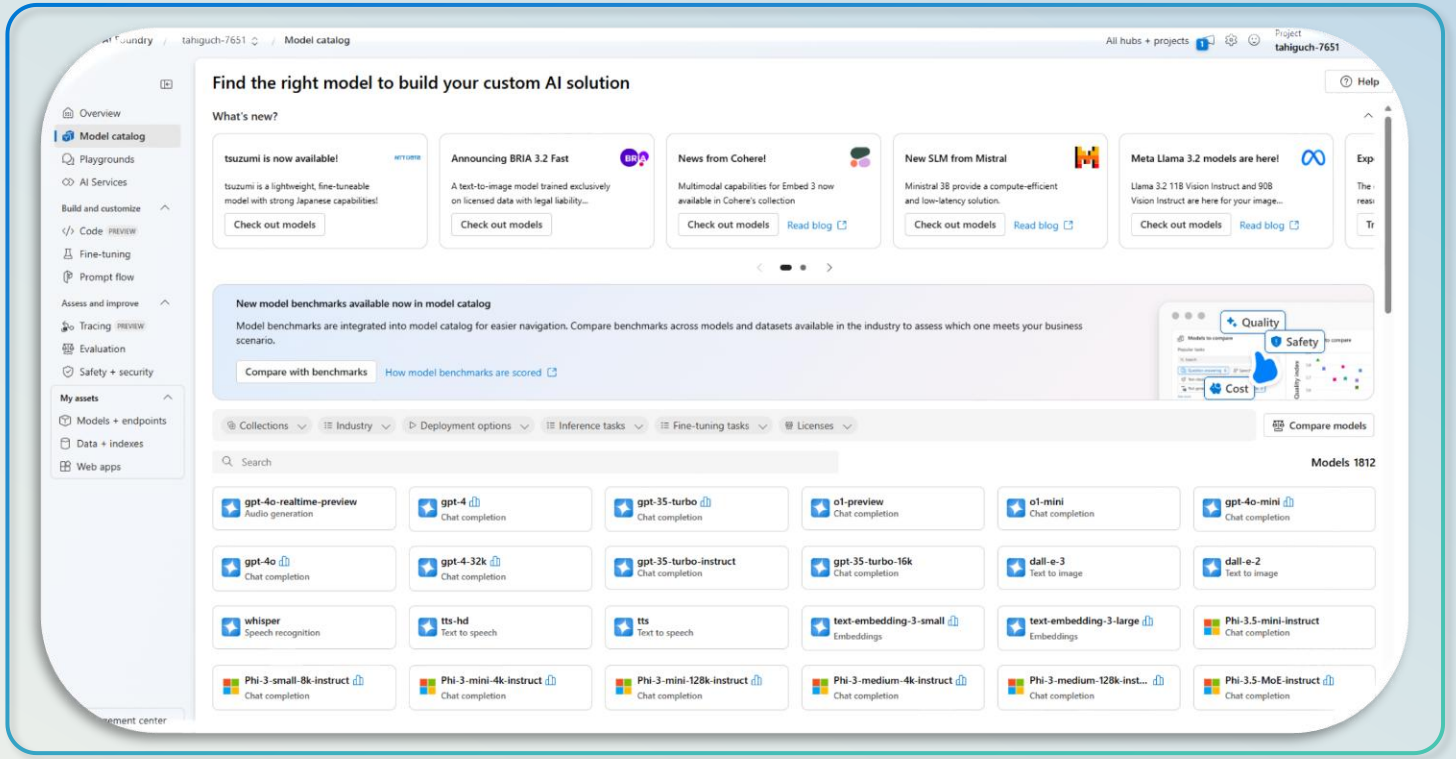


Flexible Model Selection

Leverage Models as a Service (MaaS)

- Support a variety of models from Azure AI Foundry model catalog to power your agent's reasoning. The supported models include:
  - GPT-4o
  - GPT-4o mini
  - Llama 3.1
  - Mistral Large
  - Cohere Command R+
- Leverage Model Inference API to easily swap models and compare performance to find the best model for your specific needs.

Flexible Model Selection



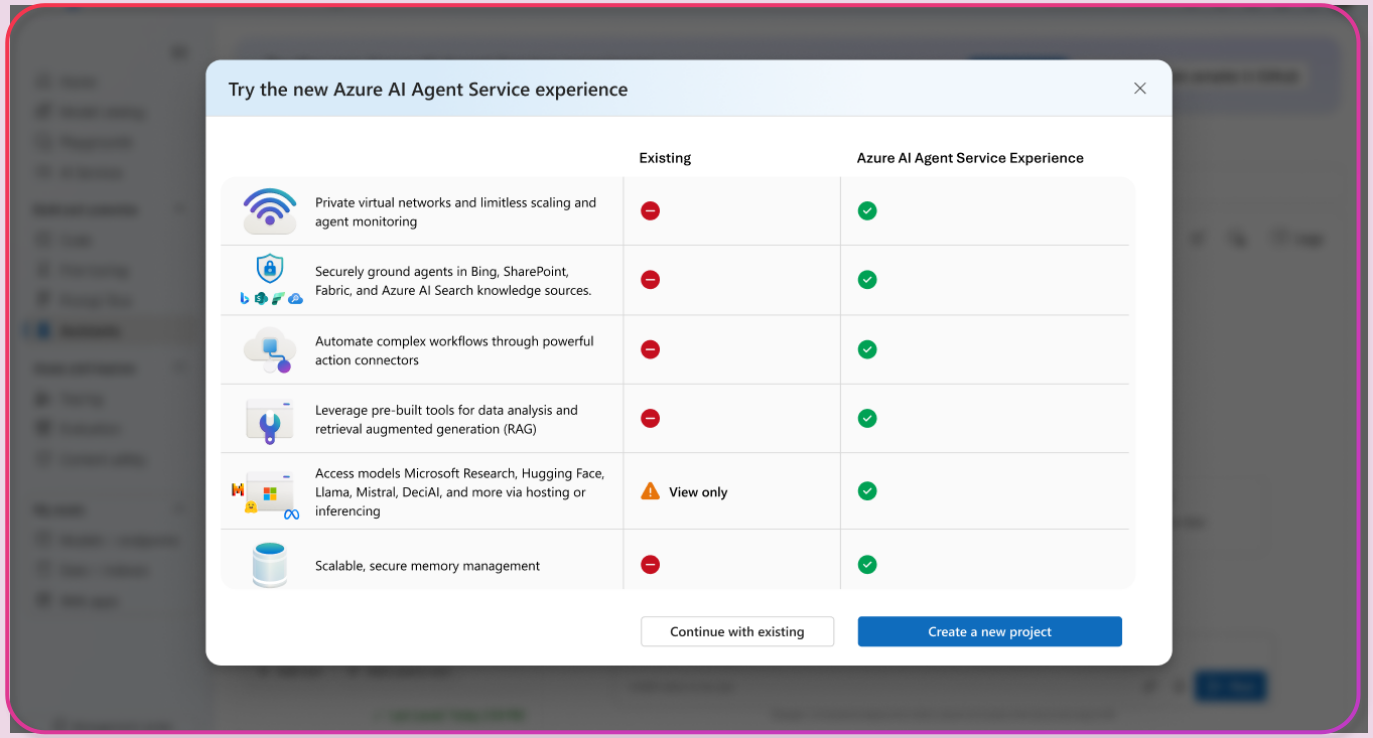
Enterprise Readiness

Ensure data privacy and compliance

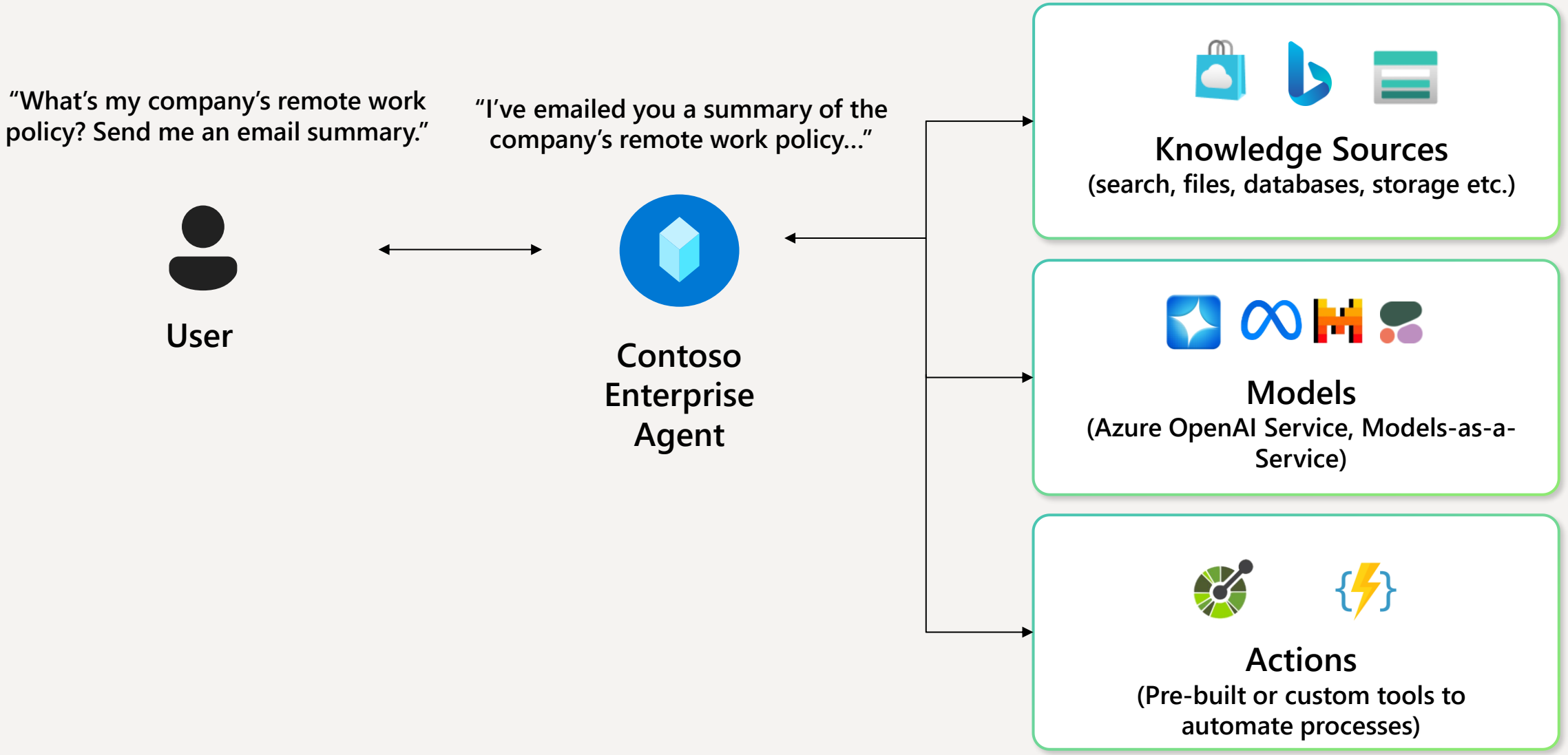
Security

Easily build enterprise-grade agents that are secure, scalable, and performant

- Securely retrieve data and execute actions inside your private virtual network (coming soon)
- Easily configure and authenticate your agents with keyless setup and on-behalf-of (OBO) authentication
- Ensure data privacy and compliance with no public egress for your traffic (Limited Private Preview)
- Gain insights into your agent's performance with tracing and monitoring (coming soon)
- Support Azure OpenAI Service's built-in content filters and prompt injection attack mitigations



Contoso enterprise agent demo



Contoso enterprise agent demo

Step 1:
Create an Agent

Step 2:
Create a Thread

Step 3:
Run the Agent

Step 5:
Check the Run status

Step 6:
Display the Agent's Response

Agent
Contoso Enterprise Agent

Instructions
You are a helpful enterprise assistant at Contoso. You have access to the following tools, actions, and data…

Model

Your data (optional)

Azure AI Search (optional)

Files (local or Azure Blob)

Tools
File Search (HR Documents)
Bing Search
Weather API
Stock Price lookup
Twilio SMS
Email (mocked)
User lookup (mocked)

Thread
Remote work policy

User's message
Summarise the remote work policy

Agent's message
Here is the summary:

User's message
Email it to me

Agent's message
I have emailed you the summary

Run 1

1. Use File Search to summarise the company's remote work policy
2. Create message

Run 2

1. Use the email tool to send the summary
2. Create message

Demo 2 – Contoso Enterprise Agent

Demo repo:
https://aka.ms/cloudchampion/agents/demo2

Contoso Enterprise Agent - Step by step

01

Obtain an Azure subscription

02

Install Python

Install VS Code

03

Install VS Code Extensions for Python, Jupyter

04

Clone the GitHub repo

05

Create and activate a virtual environment (venv)

06

Install required modules

pip install -r requirements.txt

07

Create your Azure AI Foundry Hub and Project

08

Configure your .env file (inc. Project Connection String)

09

Authenticate using az login

10

Run the Jupyter notebook and open the Gradio site

11

Optional: Deploy Grounding with Bing Search, and configure your Connection

12

Optional: Request an API key from OpenWeather (free)

1

Single-agent

Deploy agents with **Azure AI Foundry**

Managed agent micro-services

2

Multi-agent

Orchestrate them together with **AutoGen** and **Semantic Kernel**

State-of-the-art research SDK

Production-ready and stable SDK

Ideation → **Production**

Sponsored by Microsoft

WEBINAR 19 March | 10:00 - 11:00 AM CET

Introduction to AI Agents and Multi-Agent Systems with Semantic Kernel

Speaker

Henri Schulte

AI Partner Solution Architect

Microsoft

CLOUD CHAMPION

https://aka.ms/cloudchampion/multiagents

Get started with Azure AI Agents

Cloud Champion resources:

Downloads these slides:
https://aka.ms/IntroToAgentService

Watch the recording on demand:
https://aka.ms/cloudchampion/agents

Demo repos:
https://aka.ms/cloudchampion/agents/demo1
https://aka.ms/cloudchampion/agents/demo2

Multi-Agent Systems with SK - 19th March
https://aka.ms/cloudchampion/multiagents

Agent Service on Microsoft Learn Docs
https://aka.ms/MicrosoftLearn_Agents

Watch the Microsoft Ignite sessions on-demand

- Streamlining customer service with AI-powered agents (THR608)
- Building intelligent multi-agent systems with Azure AI (BRK102)

Read the Tech Community announcement

Microsoft Azure

CLOUD CHAMPION

Thank you