

# The Next Wave of Al Agents in Business

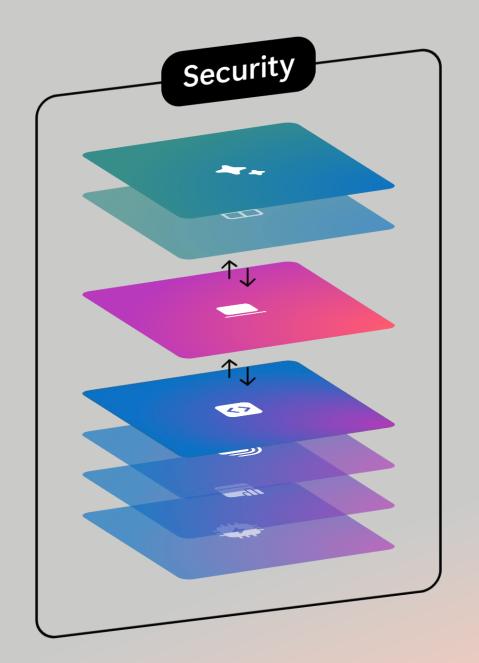
Speaker: Syed Safdar Hussain Head of Cloud Solution Architecture & CoE



## Every app will be reinvented with Al

New apps will be built that weren't possible before





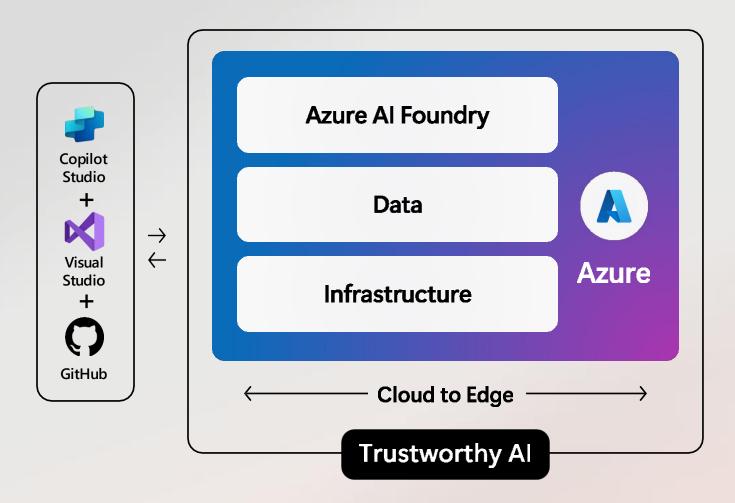
# Copilot

## **Copilot devices**

# **Copilot & Al stack**



## **Copilot & AI stack**





# **A New Frontier Agentic Al**

# **Agentic Al**

use AI to automate and execute business processes, nearly autonomously

Al agents can plan, adapt to new information, and execute tasks independently, making them capable of handling complex, dynamic environments, decisionmaking and autonomous action. They can learn from feedback, adjust their strategies, and operate with minimal human oversight.



Reason over a provided business process

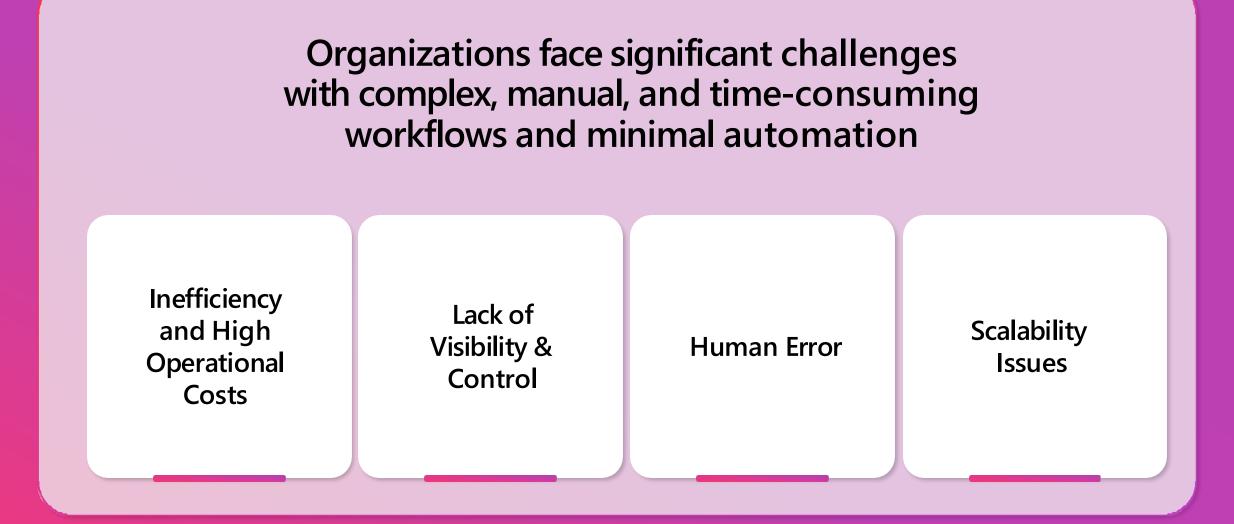


Retrieve context to complete the process

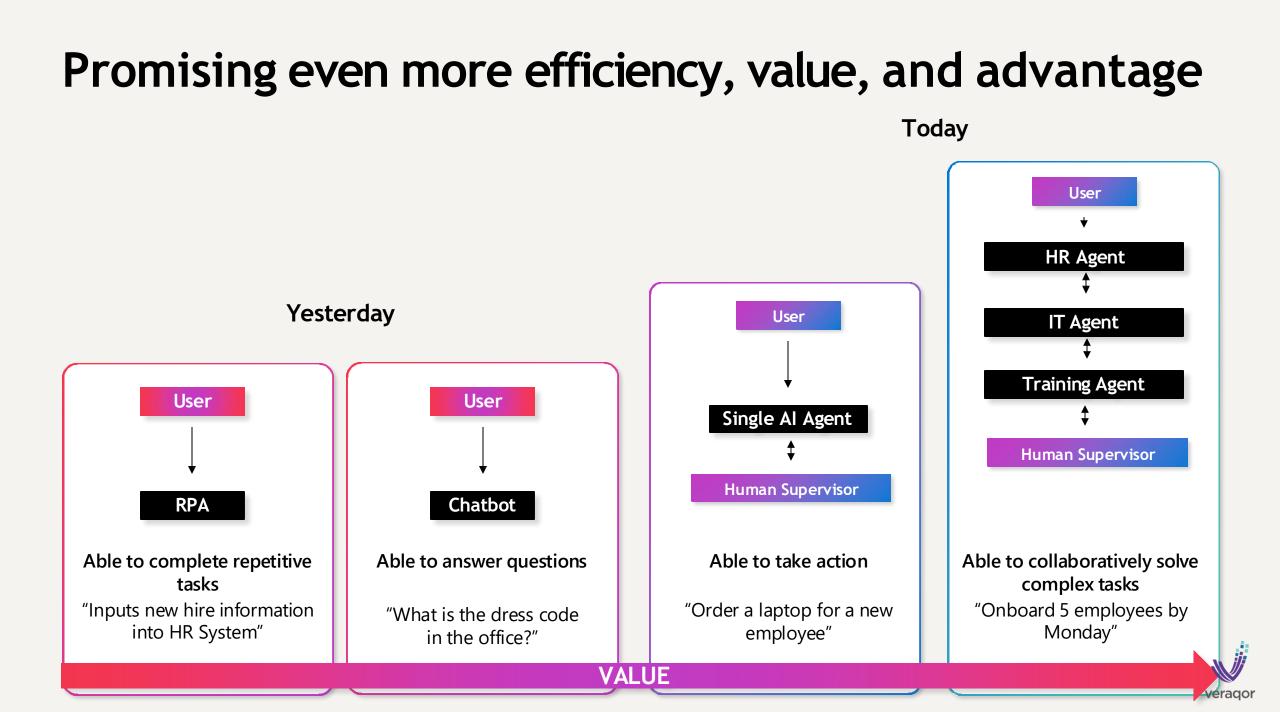


Perform an action for the end-user



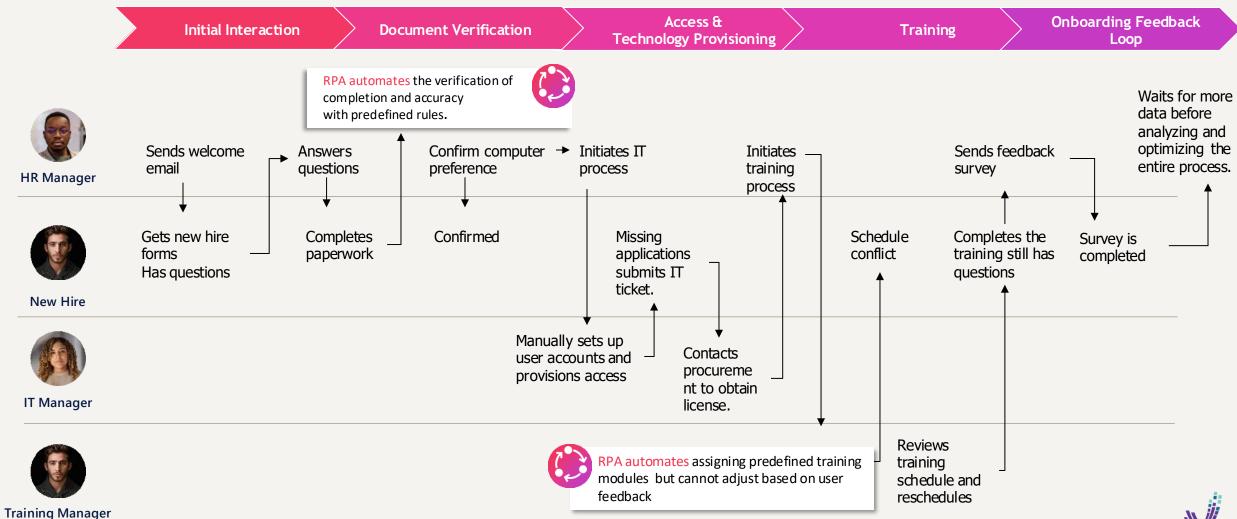






## TODAY

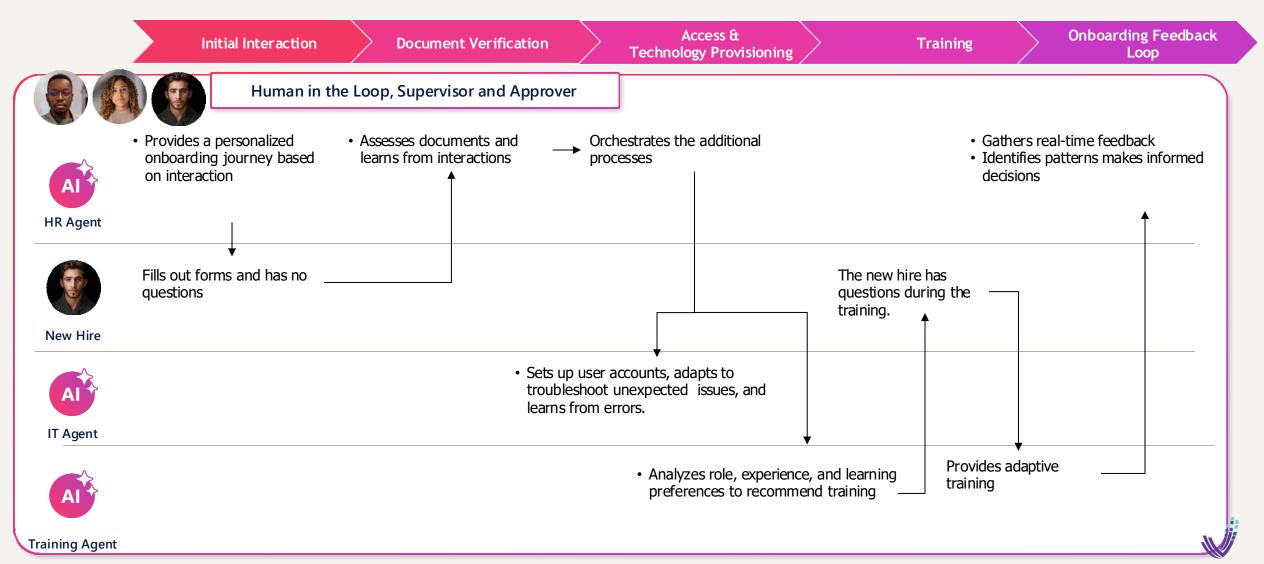
Existing solutions can only automate very specific tasks that have clear inputs and outputs



veragor

# TOMORROW

With AI Agents, these steps can be automated for the first time, while keeping human in the loop.



# Key Applications of Al Agents

# This cutting-edge shift promises to ignite a fresh wave of productivity and efficiency

Automation is undergoing a thrilling transformation with autonomous agents now being designed to proactively plan actions and perform tasks for users.

**Intelligent Decision Making** 

**Increased Productivity** 

**Enhanced Efficiency** 

**Cost Reduction** 

#### **Key Use Cases Across Industries**

## Al Agent workflows can enhance efficiency, accuracy, and customer satisfaction

### Travel Booking & Expense Management

- Assists employees in booking business trips
- Integrates with Tripadvisor, Outlook, and SharePoint
- Books via Teams chat or email
- Uses OCR to gather receipts
- Automates expense report submission and tracking

#### **Employee Onboarding**

- Personalized onboarding assistant for new hires
- Uses LLMs grounded in HR data from SharePoint
- Provide relevant training materials
- Schedule orientations and set up software accounts
- Monitor task completion and ensure efficient onboarding

#### Personalized Customer Support

- Diagnoses issues by referencing history and product manuals
- Provides tailored solutions or escalates through automated workflows
- Creates tickets and schedules follow-ups
- Updates CRM records, enhancing future support

#### Data Analytics and Reporting

- Analytics data from data lake and data warehouse
- Responds to user requests in natural language
- Generates insights, visualization, and sends via Teams or email
- Automates data handling for real-time, effortless decision-making



# Challenges in Development with AI Agents

# Successfully developing AI Agents requires 4 primary considerations

Knowledge	Actions	Security	Evaluation
Providing agents with the right context	Giving them the tools to complete their tasks	Ensuring they only have access to what they should	Ensuring they complete their tasks correctly



# Organizations need platforms that enable rapid development of performant, secure AI Agents

#### **Current Frameworks**

Lack of integrated tools, insecure data grounding, challenging orchestration

Ineffective deployment of AI across websites, applications, and production environments

Restrictive pre-defined models that are challenging to customize

Security and data privacy risks

#### What's Needed

Connected complex workflow automation grounded by seamless connection to enterprise data

Tools and APIs that seamlessly integrate across enterprise applications

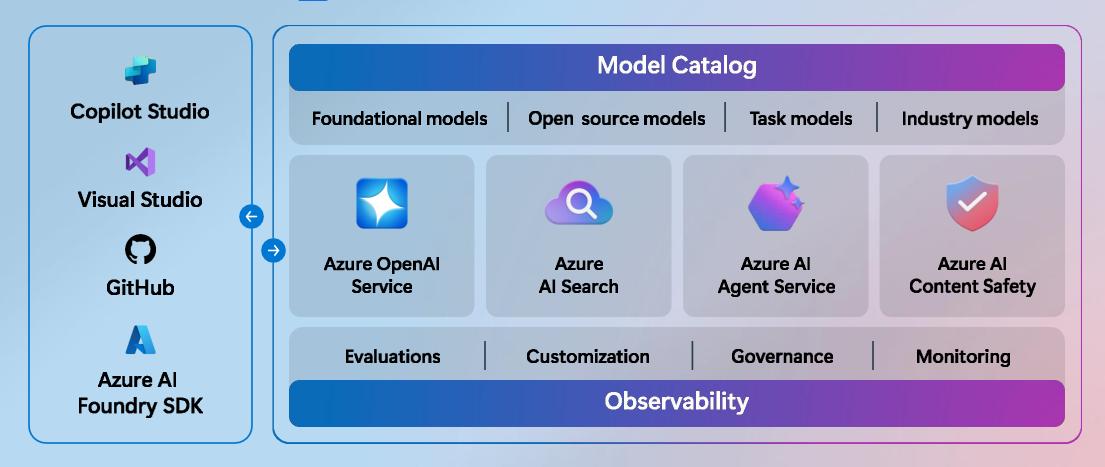
Flexible models that enable processing and integration of information from multiple modalities or types of data

Secure, responsible AI that protects sensitive information and behaves compliantly

veradoi

#### Announcing

# Azure Al Foundry





**Public Preview** 



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

Rapid Development and Automation Extensive Data Connections Flexible Model Selection Enterprise grade Security

Al.Azure.com



# The full package

Built-in enterprise readiness	Extensive Ecosystem of Tools Knowledge Microsoft Fabric* SharePoint*	Model Catalog
BYO-file storage (coming soon)	Grounding with Bing Search     Azure Al Search	Azure OpenAl Service (GPT-40, GPT-40 mini)
BYO-search index	Files (local or Azure Blob) Your own licensed data*	Models-as-a-Service
OBO Authorization Support	File Search Code Interpreter	Llama 3.1-405B-Instruct
Enhanced Observability	Actions   Azure Logic Apps*   OpenAPI 3.0 Specified Tools   Azure Functions	Mistral Large

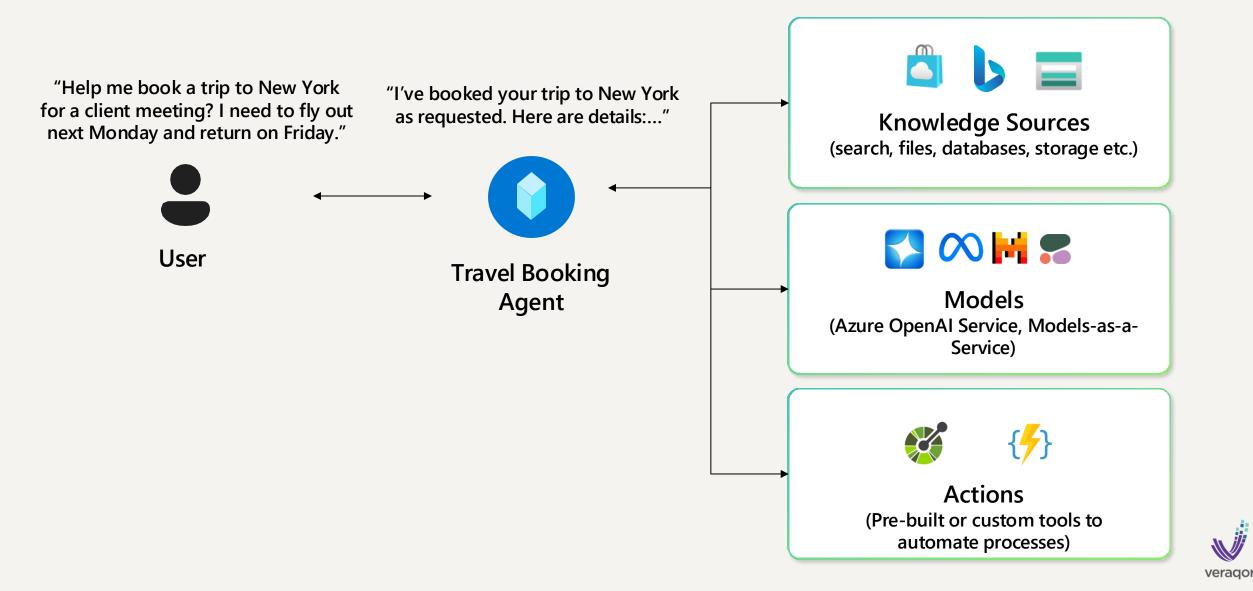
\*Indicates feature is coming soon

veragor

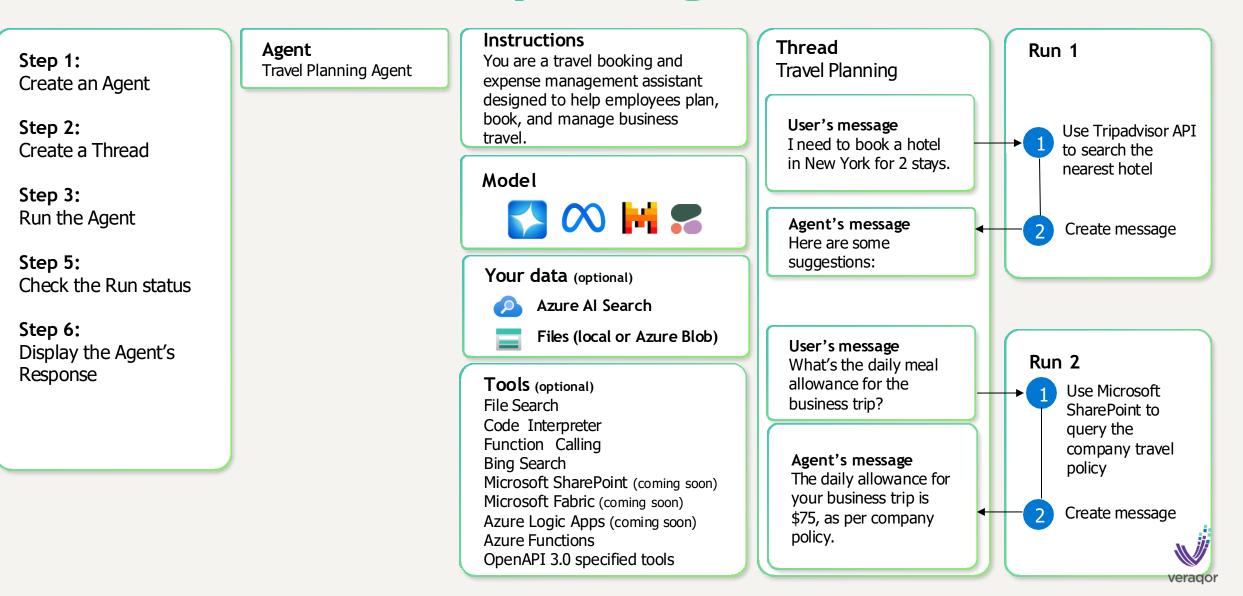
# Introducing

Azure Al Foundry SDK - Agent Service				
Azure OpenAl Assistants API File Search	Model Catalog   Model Catalog   Azure OpenAl Service   (GPT-40, GPT-40 mini)   Models-as-a-Service Llama 3.1-405B-Instruct Mistral Large Cohere-Command-R-Plus	Extensive Ecosystem of Tools         Knowledge         Image: Imag		
		<ul> <li>OpenAPI 3.0 Specified Tools</li> <li>Azure Functions</li> </ul>		
Built-In Enterprise Readiness				
BYO-file storage BYO-search index (coming soon)	x BYO-thread storage BYO-virtual network (coming soon)	OBO Authorization Support Enhanced Observability		

# How does your agent work?



# How does your agent work?



# 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 sto ra ge
  - File Search (API handles chunking, embeddings storage and creation, and implementing vector search)
- Code Interpreter
- Function Calling

#### **Azure AI Agent Service**

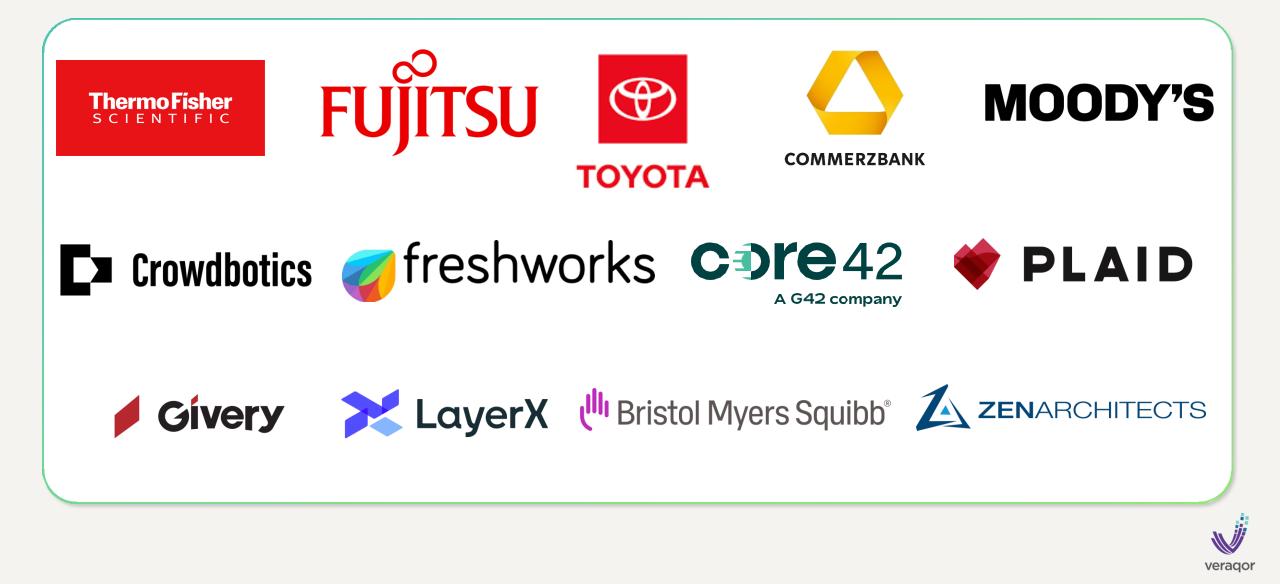
- Assistants API features, plus
- Build with a model of your choice (OpenAI, Llama, Mistral, Cohere...)
- Real time web-grounding with Bing
- Secure grounding on enterprise data in SharePoint and Fabric
- Bring Your Own Licensed data (Tripadvisor)
- Connect to 1400+ data sources and services with Azure Logic Apps
  - Long running, event driven actions with Azure Functions
  - Standardized OpenAPI 3.0 tools
  - Bring Your Own Storage

VS

- Bring Your Own Private Network
- Bring Your Own AI Search Resource
- Limitless scaling with PTUs
- Open Telemetry based tracing

veragor

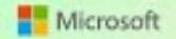
# Customers are innovating with Azure AI Agent Service



## Demo

<ul> <li>Agents playgrou</li> </ul>	na °	
Connected resource	→ New agent Delete	w code
🗹 Clear chat 🔍 Logs	🗄 Thread files 🗸 🔳 JSON response 🛈 0 tokens 🛈	Setup
	New thread started thread_z8eKh6WYDFeb2koBkR0EFmeU	Agent id ()
		Agent name
		Scheduler Agent
		Azure OpenAl resource ①
When is spring break for Pied	mont High School in CA?	Azure AI Search resource ()
		AzureAlSearch
		Deployment * Create new depl

https://youtu.be/ltt7JNd30Ag?si=Y2Cbdx8YZOPFWJzB&t=1506



# Azure Al Agents Service Building intelligent multi-agent systems

Marco Casalaina, Vice Present of Products, Azure Al Mads Bolaris, Groupt Product Manager, Azure Al Lambert Ninteman, Global Lead, GenAl Democratization & Enablement

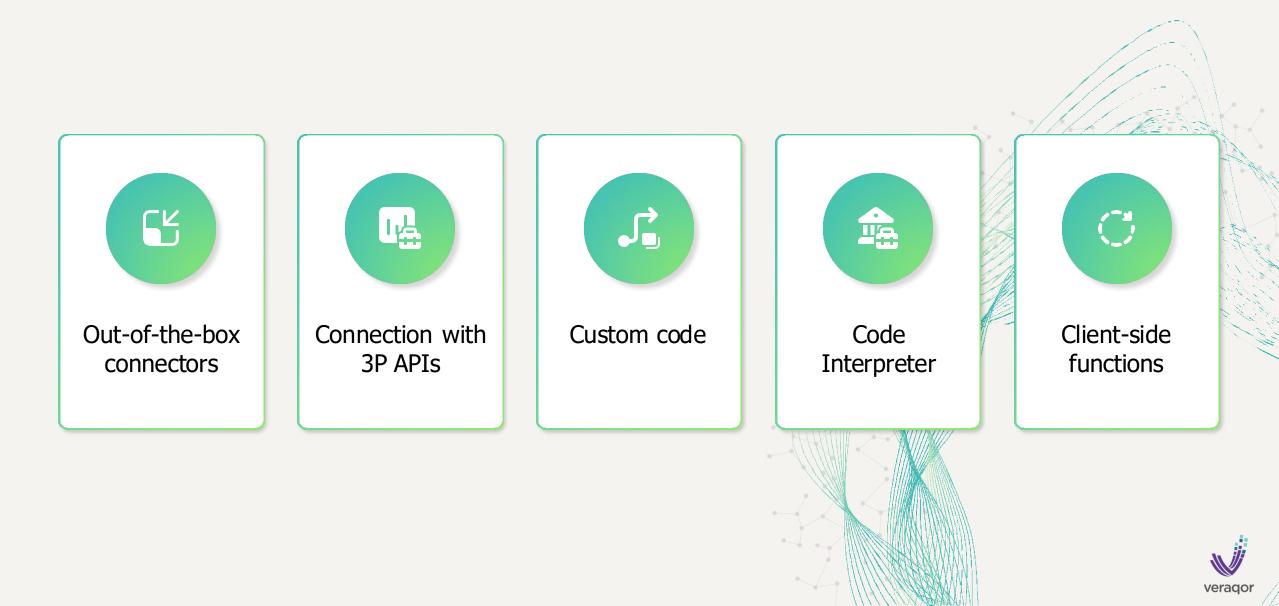
Brian King, GM, Al and Search Platform



# Rapid Development and Automation



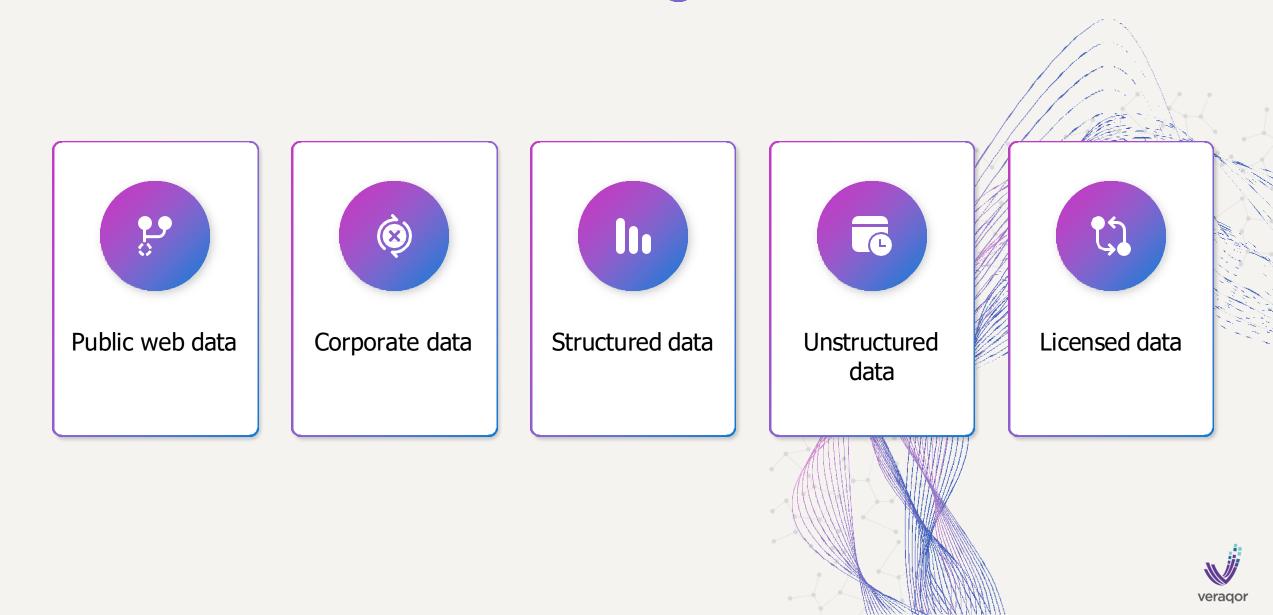
# **Rapid Development and Automation**



# Extensive Data Connections



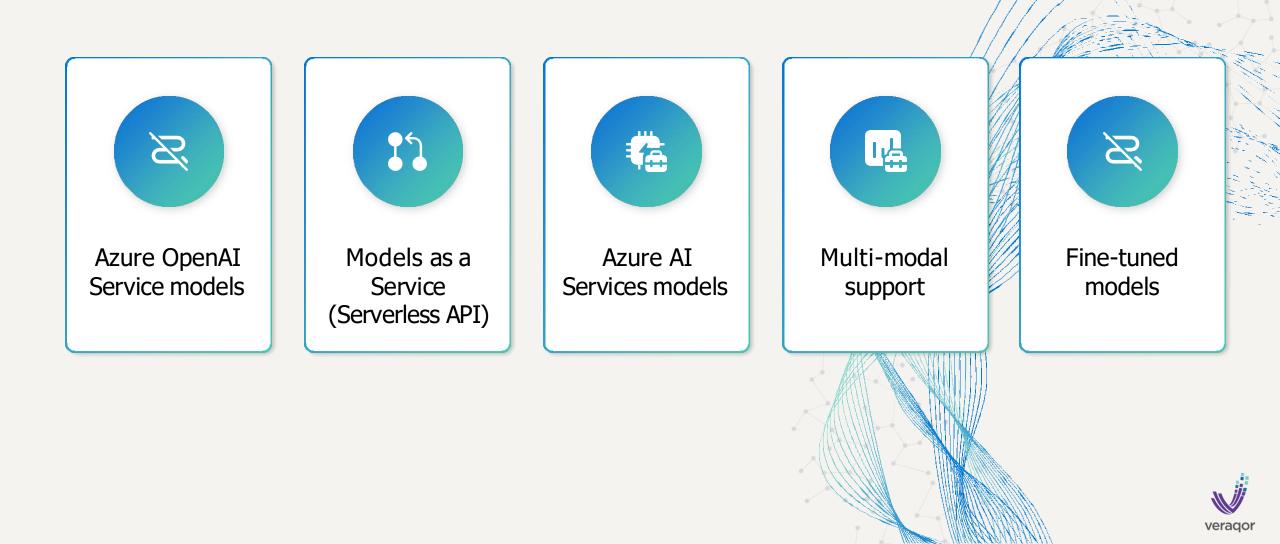
# **Extensive Knowledge Connections**



# Flexible Model Selection



# **Flexible Model Selection**



# Leverage Models as a Service (MaaS)

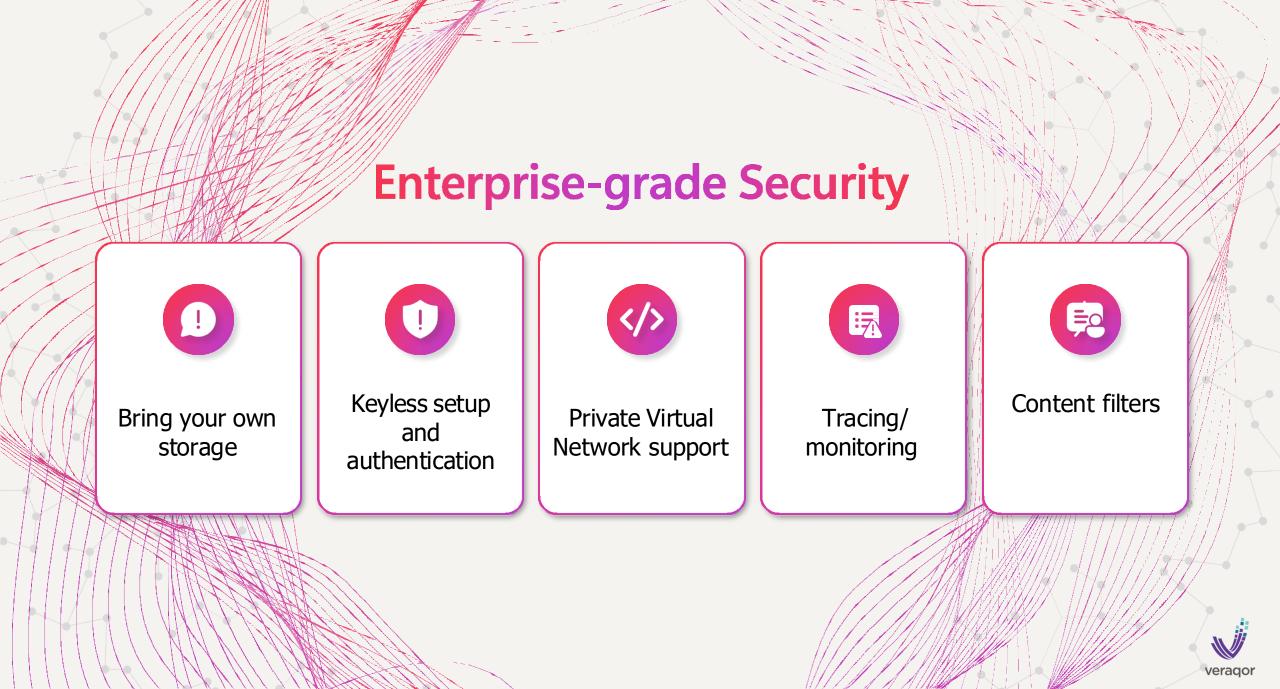
- Support a variety of models from Azure AI Foundry model catalog to power your agent's reasoning. The supported models include:
  - GPT-40
  - GPT-40 mini
  - Llama 3.1
  - Mistral Large
  - Cohere Command R+
- Leverage Model Inference API to easily swap models and compare performance to find the best model for your specific needs.

Œ	Find the right model to b	ouild your custom AI so	lution						⑦ He
Overview	What's new?								
Model catalog									
Playgrounds	tsuzumi is now available!	Announcing BRIA 3.2 Fast	News from Cohere	-	New SLM from Mistr	al 📊	Meta Llama 3.2 mo	odels are here!	Ex
AI Services	tsuzumi is a lightweight, fine-tuneable model with strong Japanese capabilities!	A text-to-image model trained exc on licensed data with legal liability			Ministral 38 provide a co and low-latency solution		Llama 3.2 118 Vision In Vision Instruct are here		The
Id and customize	Check out models	Check out models	Check out mode		Check out models			Is Read blog 🖸	
Code PREVIEW				Read blog E				in the biog to	
Fine-tuning Prompt flow				•• >					
ess and improve	New model benchmarks available no								
Tracing PREVIEW		w in model catalog o model catalog for easier navigation. Co	many handhmadic access models and d	tagets available in the indust	outo access which one a	ante conce husineses		+ Quality	
indening manan	model delicinitarias are integrated inte	o model catalog for easier havigation, co	inpare benchinarks across models and d	casets available in the indust	ry to assess which one h	leets your business			1000
Evaluation	scenario.						(C) Madata to a	Safety "	· compare
		v model benchmarks are scored					Passie take	Safety -	• •
Safety + security		v model benchmarks are scored 🕻					Passiler tasks	salety	
Safety + security assets	Compare with benchmarks How		nos tudo — 🚝 Ene tudos tudos -	-			Pegalar tasks 14 taunt	Cost Salety	
Safety + security assets  Models + endpoints			nce tasks 💛 ा≣ Fine-tuning tasks 🔻	₩ Licenses V			Pegalar tasks 14 taunt	salety	e models
Safety + security assets Models + endpoints Data + indexes	Compare with benchmarks How		nce tasks $\checkmark$ $\cong$ Fine-tuning tasks $\checkmark$	₩ Licenses →			Pegalar tasks 14 taunt	Cost W Compare	
Evaluation Safety + security assets ^ Models + endpoints Data + indexes Web apps	Compare with benchmarks How	D Deployment options ↓ I≣ Infer				🕶 ol-mini	Hadri Mar Hanne Start Start Mar Mar Mar Mar Mar Mar Mar Mar Mar Mar	Cost Cost Cost Compare More	
Safety + security assets Models + endpoints Data + indexes	Compare with benchmarks How			■ Licenses → ■ Licenses → ■ Cat-preview Chat completic	n	ol-mini Chat completion	Hadri Mar Hanne Start Start Mar Mar Mar Mar Mar Mar Mar Mar Mar Mar	Cost W Compare	
Safety + security assets Models + endpoints Data + indexes	Compare with benchmarks How Collections V IE Industry V C Search Search Search	▷ Deployment options ∨ III Infer	gpt-35-turbo (]) Chat completion	Chat completion				Cost Compare Cost Mon gpt-4o-mini 1 Chat completion	e models
Safety + security assets Models + endpoints Data + indexes	Compare with benchmarks How	D Deployment options ↓ I≣ Infer				Chat completion		Cost Cost Cost Compare More	
Safety + security assets Models + endpoints Data + indexes	Compare with benchmarks How Collections Collections III Industry Collections Collections III Industry Collection Collections III Industry Collection Collec	<ul> <li>▷ Deployment options ∨ III Inference</li> <li>○ gpt-4 ① Chat completion</li> <li>○ gpt-4-32k ① Chat completion</li> </ul>	gpt-35-turbo ()         Chat completion         gpt-35-turbo-instruct         Chat completion	Chat completic	<b>16k</b>	dall-e-3 Text to image		Cost Compare Cost Compare Mor Chat completion dall-e-2 Fect to image	
Safety + security assets Models + endpoints Data + indexes	Compare with benchmarks How Collections V IE Industry V C Search Search Search	▷ Deployment options ∨ III Infer	gpt-35-turbo (]) Chat completion	Chat completion	<b>16k</b>	Call-e-3 Text to image		Cost Compare Cost Mon gpt-4o-mini 1 Chat completion	



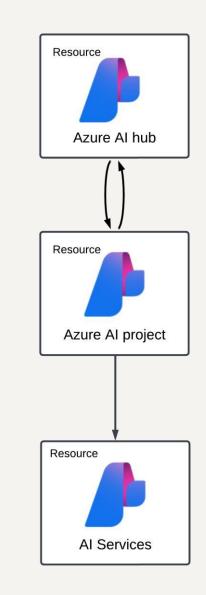
# Enterprise-grade Security





# **Basic agent setup**

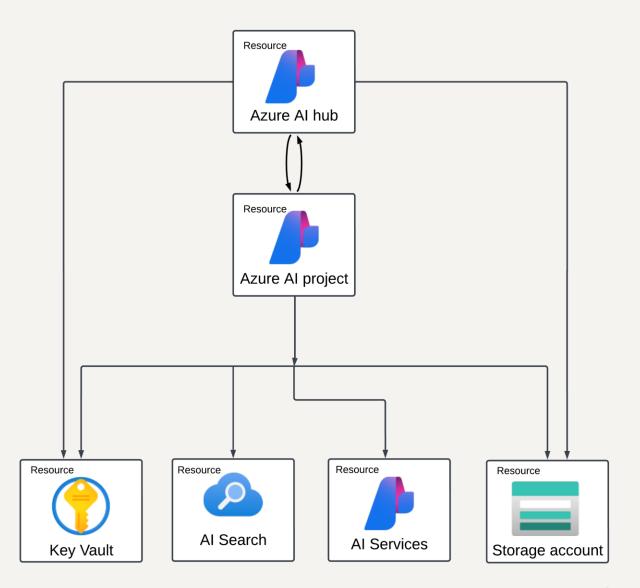
- Overview: All agents created in this project use multitenant search and storage resources fully managed by Microsoft. You don't have visibility or control over these underlying Azure resources.
- Required customer resources:
  - AI hub
  - AI project
  - AI Services/AOAI
- How to use: Deploy the basic setup template
- Tool Implications:
  - File search and code interpreter
    - Uploaded files get stored in Microsoft managed storage
    - Vector stores get created using a Microsoft managed search resource
  - Azure AI Search tool is not supported
  - Azure Blob Storage with file search is not supported





# Standard agent setup

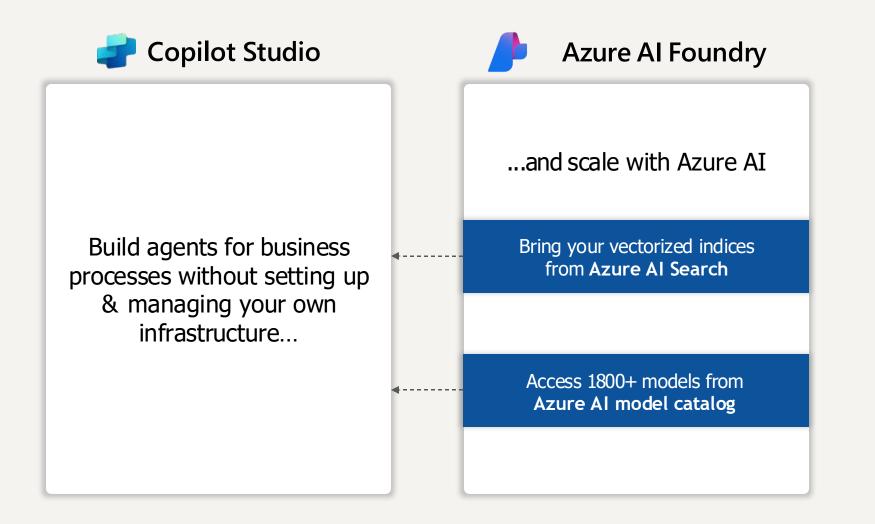
- Overview: Agents use customer-owned, single-tenant search and storage resources. With this setup, you have full control and visibility over these resources, but you incur costs based on your usage.
- Required customer resources:
  - AI hub
  - AI project
  - AI Services/AOAI
  - Key Vault
  - Storage account
  - Azure AI Search
- How to use: Deploy the standard setup template
- Tool Implications:
  - File search and code interpreter
    - Uploaded files get stored in your Azure Blob Storage account
    - Vector stores get created using your Azure
       AI Search resource





# **Start building agents with Copilot Studio...**

Build, deploy, realize value and prototype faster for simple scenarios



veragor



Single-agent

## Deploy agents with Azure Al Foundry



Managed agent micro-services Multi-agent

2

# Orchestrate them together with AutoGen and Semantic Kernel



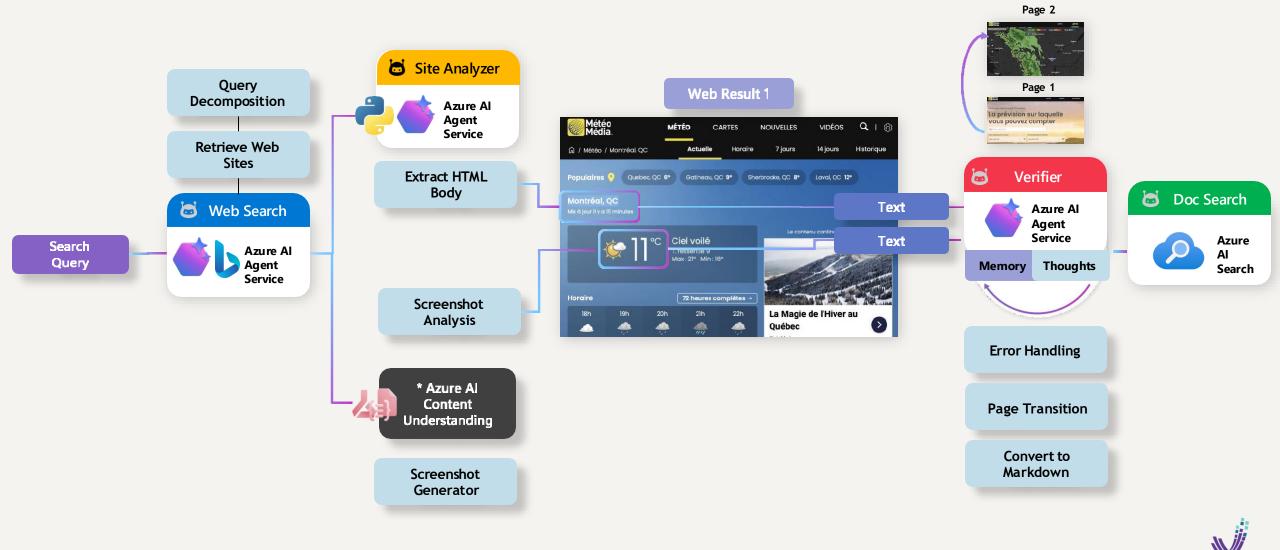
State-of-the-art research SDK Production-ready and stable SDK

Ideation

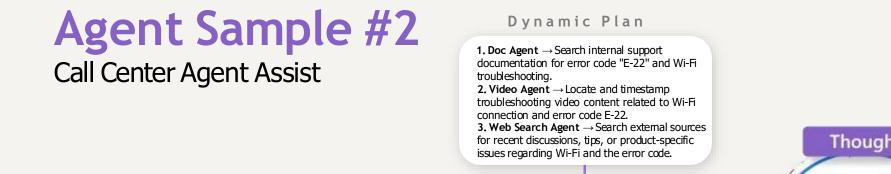
Production



## Agent Sample #1 Web Search and Extraction



veragor



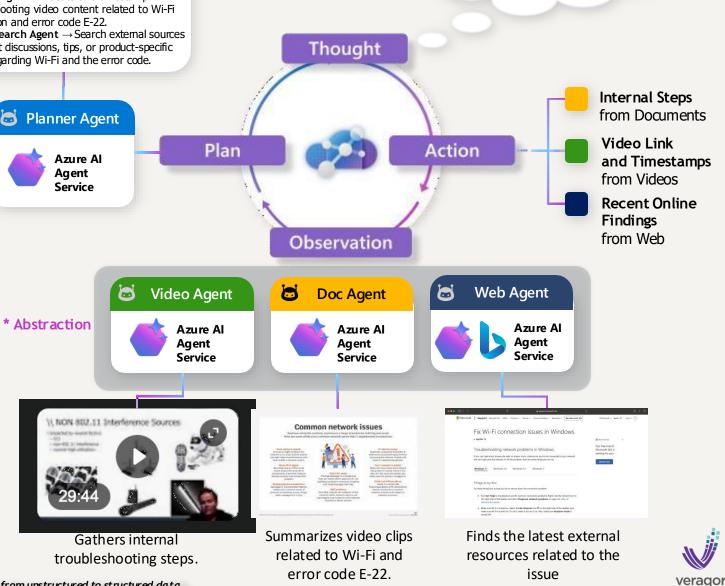
{Wi-Fi}

{Error: E-22}

{ Thermostat X10 }

**Query Parsing** 

LLM



Thought based on observation

Did I find the relevant data to assist the

agent?

\* We can use the Azure AI Content Understanding service to extract images and videos, and transform them from unstructured to structured data

Customer Query

My SmartHome thermostat isn't

can I fix this?

connecting to Wi-Fi,

and I keep seeing an

error code E-22. How

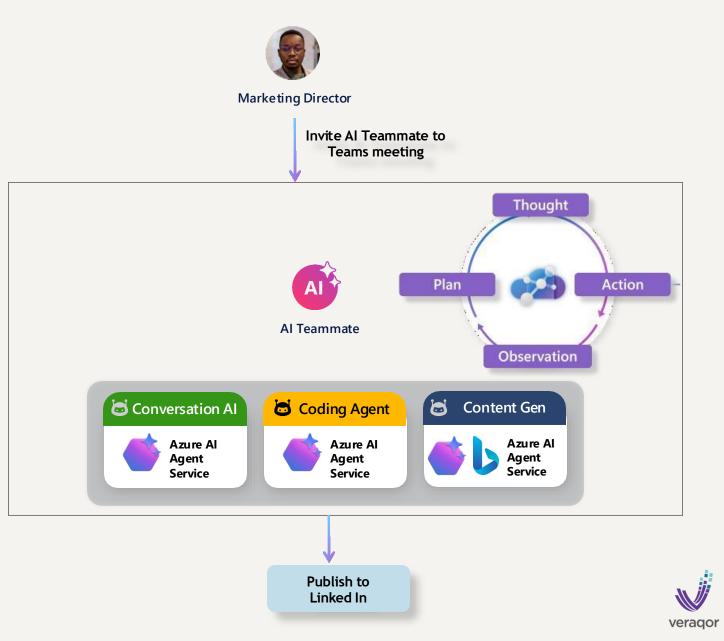
# Agent Sample #3

AI Teammates powered by multi-agents

#### **Create an AI Teammate**



#### **Collaborate with AI Teammate**



## How We Can Help

Briefing	Free 1-Hour Consultation Unlock the potential of AI-powered agents to transform business operations with our 1-hour free briefing session. This session provides decision-makers and technical stakeholders with insights into how AI agents can enhance efficiency, automate workflows, and drive business value across industries.
Assessment	One-Week Engagement   Fixed Cost \$5,000 Gain in-depth insights into the feasibility, readiness, and impact of AI-powered agents for your business with our 1-week AI Agent Assessment. This structured assessment identifies opportunities, evaluates technical requirements, and provides a strategic roadmap for AI agent adoption within your organization.
Proof of Concept	2-4 Weeks Engagement   Fixed Cost \$10,000 Accelerate AI adoption with a 2 to 4-week Proof of Concept (PoC) engagement that demonstrates the feasibility and impact of AI-powered agents in your business environment. This structured PoC validates AI agent capabilities, ensures seamless integration, and delivers a working prototype aligned with your specific use case.
Implementation	<b>6-8 Weeks Engagement   Fixed Cost \$30,000</b> Achieve full-scale deployment of AI-powered agents with our 6 to 8-week Implementation Solution. This engagement ensures end-to-end AI agent integration, automation, and optimization for real-world business applications, delivering a production-ready solution that drives efficiency and innovation.

veragor

#### What Can We Do For You



#### Business Outcome Strategy Workshop & Roadmap

Understanding the complexities and best practices to get you where you want to be. A complimentary 2-hour advisory session to assess your current state and provide the guidance you need.

#### Deeper Solution Session & Briefing

See what Veraqor can do for you! An hour-long envisioning workshop explores the impact of digital transformation and innovation to help customers with visionsetting, strategy, roadmaps, and organizational alignment.

#### Architectural Design Sessions

This custom 2-hour session focuses on your technical solution objectives and aligns them with specific components of solutions to help you not only meet your goals but also capitalize on them.

#### Industry Best Practices, Tools & Frameworks

This custom 1-hour session focuses on guidance of best practices working directly with our specialized architects on innovative new capabilities and wellarchitected framework.

Complimentary Offers for Customers