Microsoft

05 Abschluss: Whats coming next?

Recommended "Golden Path" Architecture

BYO resources

- **Cosmos DB** Thread storage
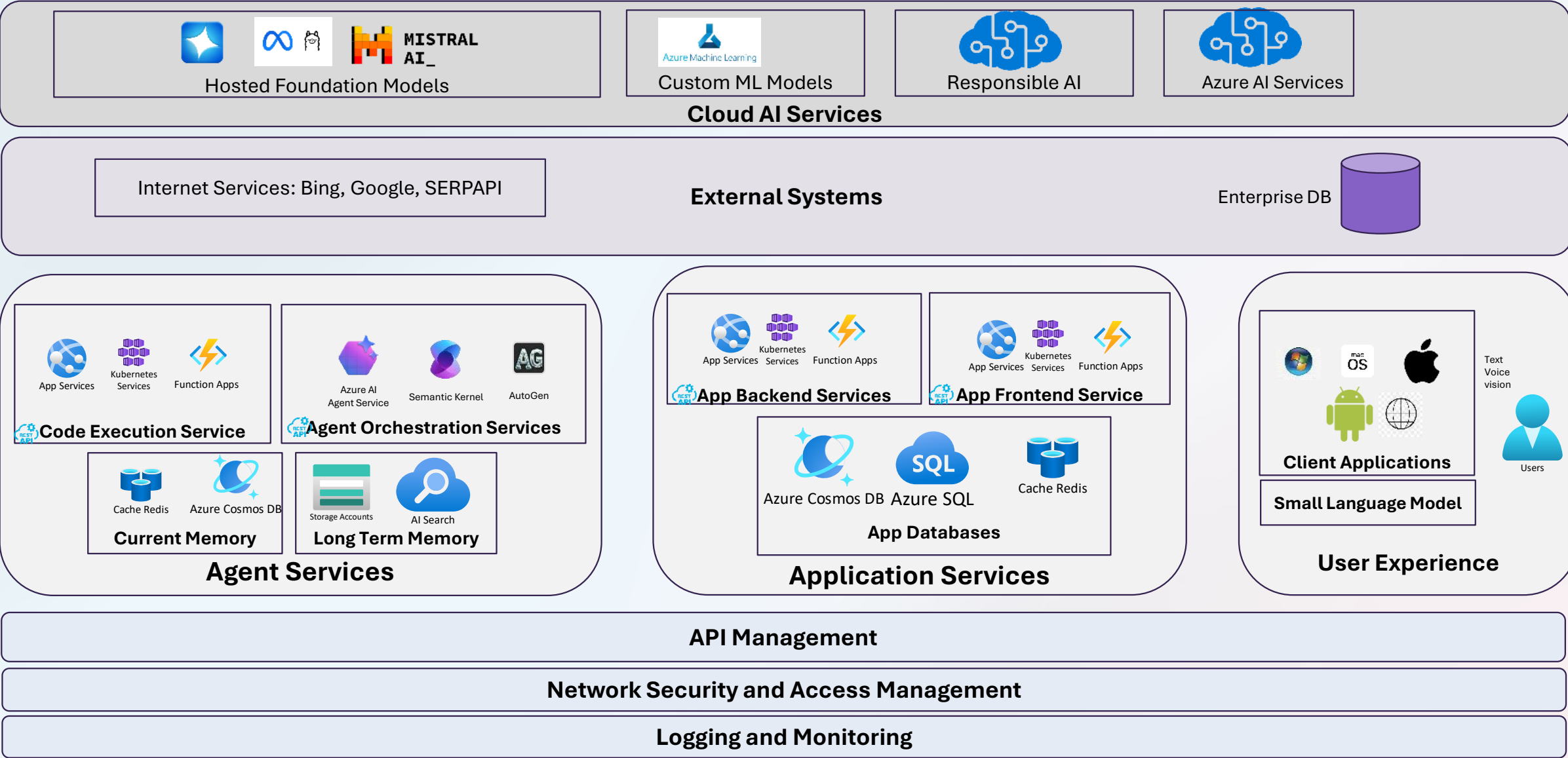- **Key vault** Connections
- **Azure Storage** File storage
- **Azure AI Search** File search index

Azure AI Foundry
Standard setup

Built-in AI tools	Agents	Models
File search		
Code interpreter		

Bot Service Channels

AI tool resources

- Azure AI Search
- Grounding with Bing Search
- Logic Apps
- Azure Functions

Semantic Kernel
Multi-agent orchestrator

Azure Container Apps

- Fabric
- SharePoint
- **External APIs** OpenAPI specs
- MCP servers
- A2A Servers

Agentic AI System Architecture Logical View

Hosted Foundation Models	AI Safety Services	AI Services	Hosted Custom AI/ML Models

External Systems

Agent Orchestration Service

Policy

Memory

Tools Calling	Code execution environment

Multi-Agent orchestration

Application Platform

Front end

Backend

Database

Infrastructure Platform

User Experience (Text, Voice, Video)
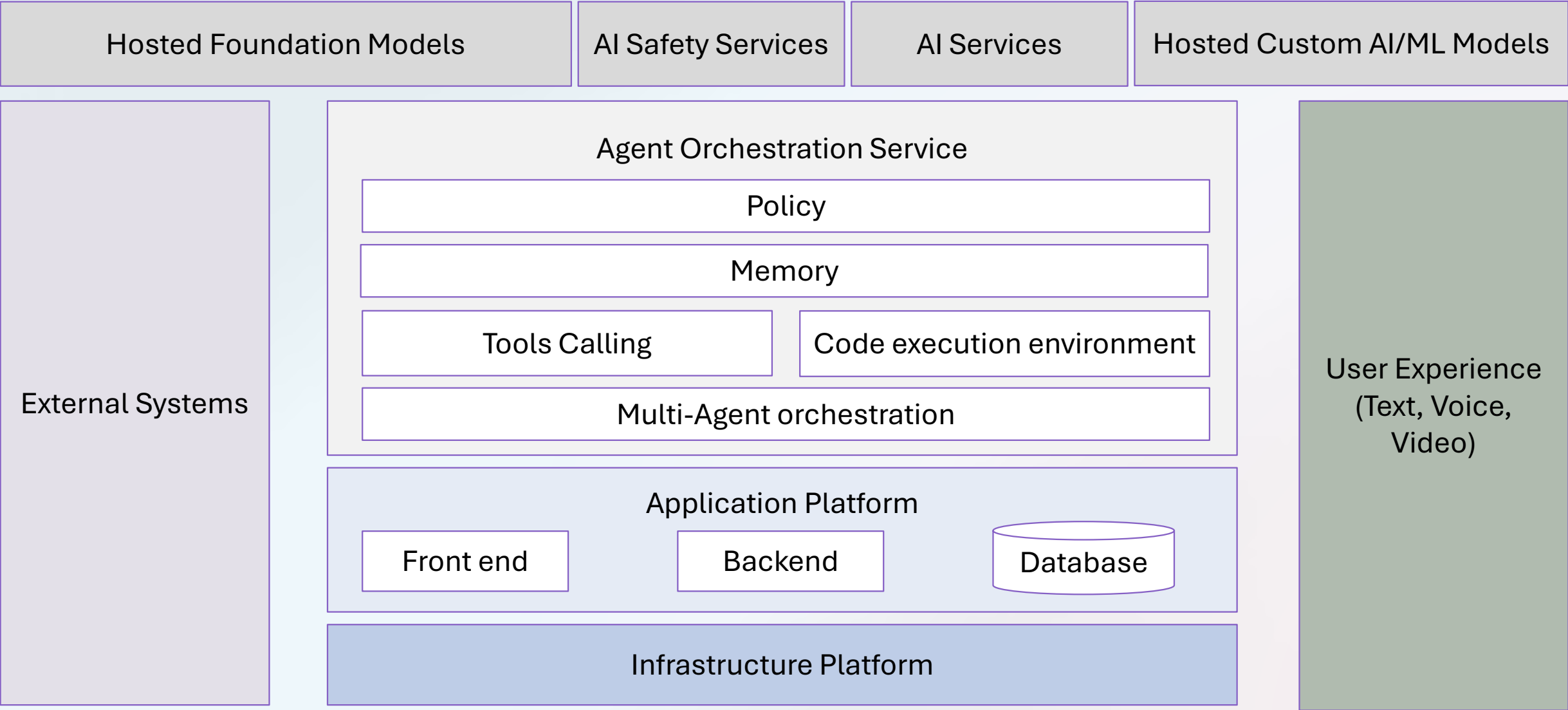
System Architecture: Component View

Cloud AI Services
- Hosted Foundation Models
- MISTRAL AI_
- Custom ML Models (Azure Machine Learning)
- Responsible AI
- Azure AI Services

External Systems
- Internet Services: Bing, Google, SERPAPI
- Enterprise DB

Agent Services
- Code Execution Service (App Services, Kubernetes Services, Function Apps)
- Agent Orchestration Services (Azure AI Agent Service, Semantic Kernel, AutoGen)
- Current Memory (Cache Redis, Azure Cosmos DB)
- Long Term Memory (Storage Accounts, AI Search)

Application Services
- App Backend Services (App Services, Kubernetes Services, Function Apps)
- App Frontend Service (App Services, Kubernetes Services, Function Apps)
- App Databases (Azure Cosmos DB, Azure SQL, Cache Redis)

User Experience
- Client Applications
- Small Language Model
- Users
- Text Voice vision

API Management

Network Security and Access Management

Logging and Monitoring

Entra IDs for agents created with Azure AI Foundry

Manage AI agent access to data and resources.
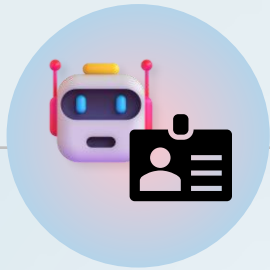
Built-in identifications for AI agents.
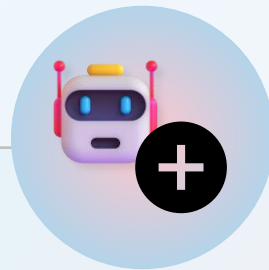
Embedded librariess & configurations.
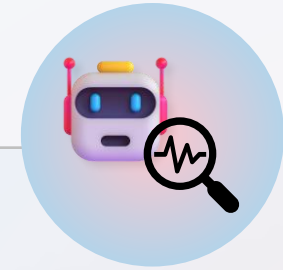
Audit agent actions

Microsoft 365 Agents Toolkit

Publish Azure AI Foundry powered agents to Microsoft 365 Copilot, Teams, custom apps, web, and more.

- Bring Copilot capabilities to your agents with Microsoft Copilot APIs
- Build once and publish to multiple (15+) channels
- Publish to Microsoft 365 Copilot, the UI for AI
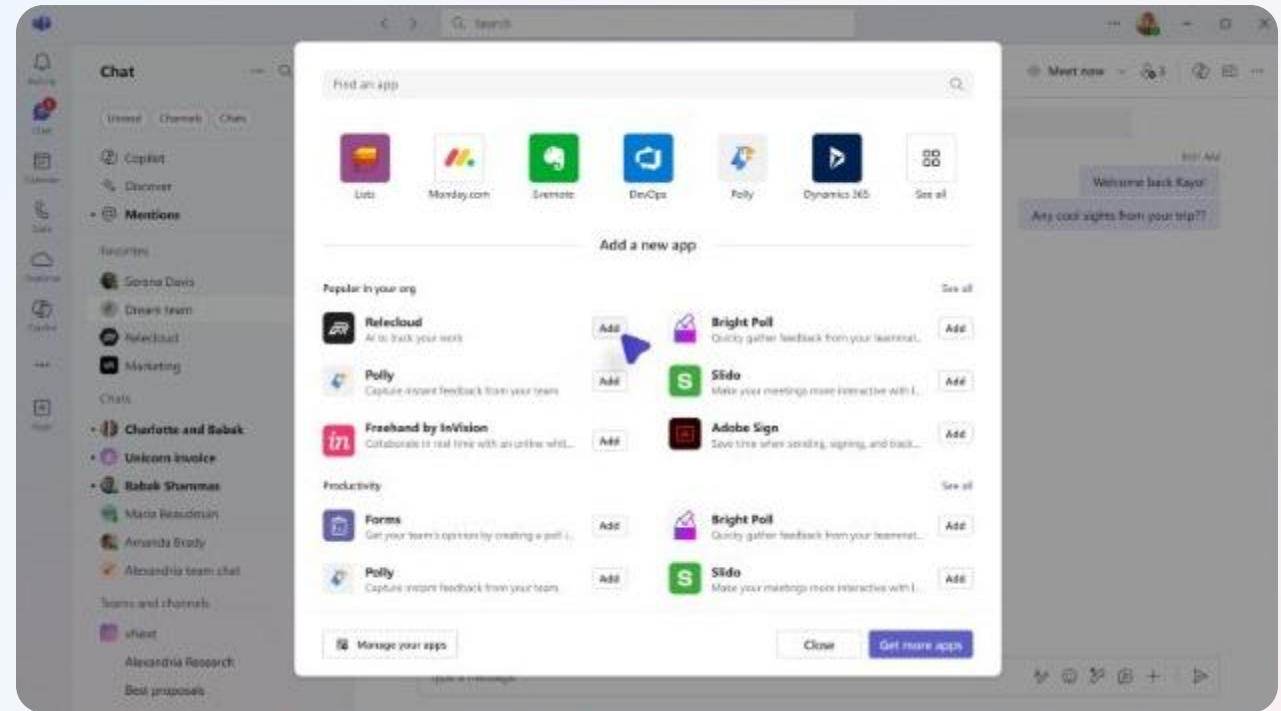
Azure AI Foundry

New | Preview

Introducing Voice Live API

Redefining agent communication with advanced, real-time voice

| Use the foundation model of your choice | Unparalleled speech recognition and synthesis capabilities | Fully customizable voices and avatars |

aka.ms/voice-live

Voice-enabled AI Agents: transforming customer engagement with Azure AI Speech | Microsoft Community Hub

Multi-agent workflow automation

Multi-agent workflow automation

Workflow specification is done via **VS Code AI Foundry Extension** or **SDK**

Defined workflows run **on managed AI Foundry platform** and can be accessed via **one single API**
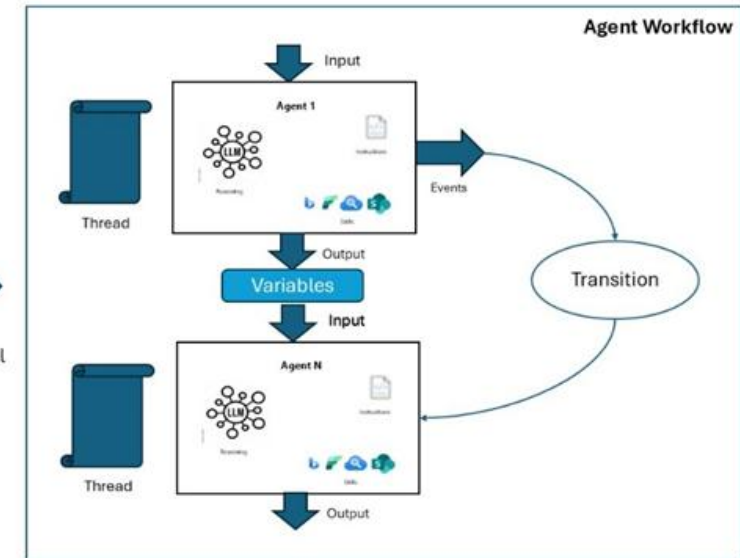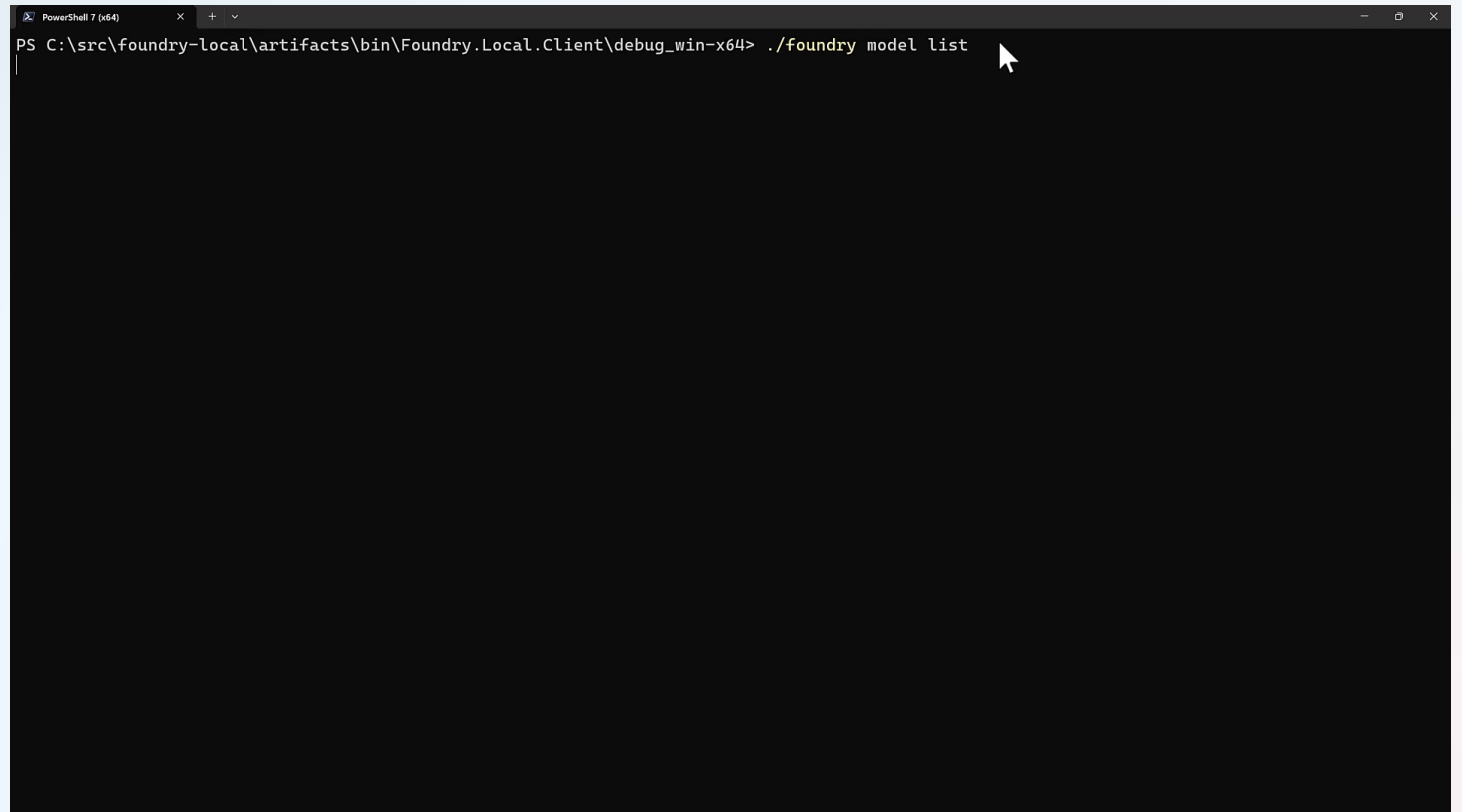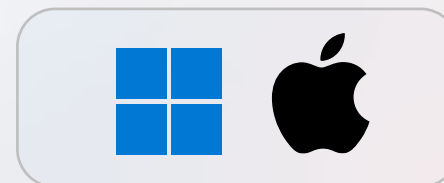
Foundry Local

Comprehensive support for managing models, agents, and their components on Windows and MacOS devices.

- Accelerate on-device inferencing

- Ship local AI apps, not just explore

- Bring the power of Azure AI Foundry locally

```
PS C:\src\foundry-local\artifacts\bin\Foundry.Local.Client\debug_win-x64> ./foundry model list
```

Microsoft

Thank You!

Feedback and Q&A