LTIMindtree

Whitepaper

# Data Agents in Microsoft Fabric

# Table of contents

# Executive Summary

Artificial Intelligence (AI), especially Generative AI (GenAI), is everywhere, reshaping our personal and professional worlds. From boosting creativity to supercharging productivity, AI is fast becoming a co-pilot in our daily lives. It's democratizing intelligence through tools like Data Agents, giving users the power to interact with their data using natural language and derive insights effortlessly. From Banking and Financial Services to Healthcare, AI is now woven into every industry and every function.

But unlocking the true value of AI begins with readiness. Organizations need a shift in mindset, one that emphasizes strong data foundations, a culture of experimentation, and domain-specific intelligence. The adoption of specialized agents ensures solutions aren't just smart, they're tailored to industry contexts, delivering meaningful, actionable impact.

# Data, AI, and Gen AI: Strategy and Industry Direction

## Becoming AI-Ready: Building the Foundation

Data, Artificial Intelligence (AI), and Generative AI (Gen AI) are reshaping industries across the board. The explosion of data has opened new possibilities, allowing organizations to gain richer insights and make more informed decisions. AI has emerged as a powerful ally, helping businesses streamline operations, improve customer experiences, and spark innovation through advanced analytics.

Gen AI takes this further by enabling machines to create entirely new content such as text, images, and even music. This unlocks creative applications across domains like marketing, product design, and more.

An effective strategy for leveraging Data, AI, and Gen AI involves three key components:

**Data infrastructure**
Organizations must establish a robust data infrastructure that can handle the volume, variety, and velocity of data generated. This includes investing in data storage, processing, and management solutions.

**AI capability development**
Businesses must develop AI capabilities by adopting machine learning models and algorithms that can analyze data and generate actionable insights. This requires a combination of in-house expertise and collaboration with AI technology providers.

**Gen AI integration**
Integrating Gen AI into the business strategy involves identifying use cases where generative models can add value, such as automating content creation, enhancing customer interactions, and personalizing marketing efforts.

### The Industry Trajectory

The industry direction is clear: companies that tap into the power of Data, AI, and Gen AI will gain a decisive edge. With continuous progress in AI research, increasing availability of data, and growing demand for automation and personalization, the adoption of these technologies is only accelerating.

Sectors such as healthcare, finance, retail, and manufacturing are already reaping the benefits from AI and Gen AI applications. For instance, in healthcare, AI is being used to improve diagnostics and treatment plans, while in finance, it is enhancing fraud detection and risk management. As these technologies continue to evolve, their impact on industry dynamics will only grow, making it imperative for organizations to stay ahead of the curve.

## BI, Copilot, and Agents

Traditional business intelligence (BI) has long helped organizations derive insights and analyze large volumes of data and its KPI in form of reports and dashboards. Predefined and self-service reporting have enabled businesses to track key performance indicators (KPIs), monitor daily monitoring operations, and make informed decisions.

Today, the integration of Gen AI with copilot functionalities is redefining how businesses operate. A compelling example is the marketing copilot, which when paired with Microsoft Fabric and Gen AI enables marketers to gain real-time insights, automate content creation, and optimize campaigns using predictive analytics.

As businesses increasingly rely on data-driven strategies, the synergy between data, AI, and Gen AI is becoming increasingly paramount. Establishing a solid foundation for Gen AI readiness involves technical preparations and strategic alignment. Therefore, companies must ensure that their data infrastructure, AI capabilities, and governance frameworks are robust enough to support the transformative potential of these technologies.

The future of industry will depend on how effectively organizations integrate and harness these technologies. Those who succeed will drive faster innovation and maintain a clear competitive edge in an evolving business landscape.

# Fabric Paradigm

Microsoft Fabric is a unified platform built to streamline and optimize end-to-end business processes. It offers a wide range of capabilities designed to enhance productivity, simplify data management, and drive smarter decisions across the enterprise.

Key capabilities of Fabric include:

**Data integration:** Seamlessly integrates data from multiple sources into a centralized platform, enabling holistic visibility and informed decision-making.

**Data management:** Provides robust tools for data cleansing, validation, governance, and security—ensuring data accuracy and consistency.

**Analytics and reporting:** Offers powerful analytics that support descriptive, diagnostic, predictive, and prescriptive insights, empowering users to make data-backed decisions.

**Automation:** Automates routine processes through workflow orchestration and robotic process automation (RPA), freeing teams from manual, repetitive tasks.

**Collaboration:** Enables cross-functional collaboration with tools for communication, file sharing, and project management, helping teams stay aligned and productive.


**AI and machine learning:** Integrates advanced AI and ML algorithms, including predictive analytics and NLP, to identify trends, detect anomalies, and deliver intelligent recommendations.

**Copilot integration:** The Fabric Copilot acts as a virtual assistant, guiding users through tasks, answering queries, and automating actions. It enhances user experience by offering contextual assistance powered by GenAI.

**Security and compliance:** Ensures end-to-end data protection with encryption, access controls, and audit trails, maintaining compliance with industry regulations.

By integrating these capabilities, Fabric provides a holistic solution that enhances productivity, efficiency, and decision-making across the organization. The inclusion of AI and Copilot further amplifies these benefits by providing intelligent insights and automating routine tasks, allowing users to focus on more strategic activities.

# Fabric and AI – Capabilities

## Copilot

Microsoft Fabric Copilot showcases a myriad of advanced capabilities, transforming the way users interact with data and AI. Under the Fabric paradigm, Copilot integrates into the data management ecosystem, offering intuitive data wrangling and preparation features that facilitate efficient data operations. It empowers data scientists by enhancing machine learning workflows, from model training to deployment, ensuring precision and scalability.

With features like chat with your data, users can ask questions in natural language and receive instant, actionable insights. This conversational interface bridges the gap between complex datasets and user-friendly analysis, accelerating decision-making across roles and functions.

## Data Science and Machine Learning (ML)

Fabric's Data Science and ML capabilities offer a comprehensive suite of tools designed to elevate data analysis and machine learning endeavors. With its seamless integration of advanced analytics and predictive modelling, Fabric empowers data scientists with a comprehensive platform that supports every stage of the data lifecycle. From data ingestion and preprocessing to model development and evaluation, Fabric streamlines workflows that boosts productivity and accuracy.

The platform's robust infrastructure supports scalable machine learning operations, allowing users to deploy models effortlessly and monitor their performance in real time. By leveraging state-of-the-art algorithms and automated machine learning (Auto ML) features, it democratizes access to sophisticated analytics. This, in turn, enables those with limited technical expertise to harness the power of AI and machine learning in their decision-making processes.

## Chat with Your Data – Data Agent

A standout feature of Fabric's AI skillset, the Data Agent harnesses advanced NLP and GenAI algorithms to transform how users interact with data. It simplifies querying by allowing users to "chat" directly with datasets in natural language, no need to understand the underlying data models, queries, or relationships.

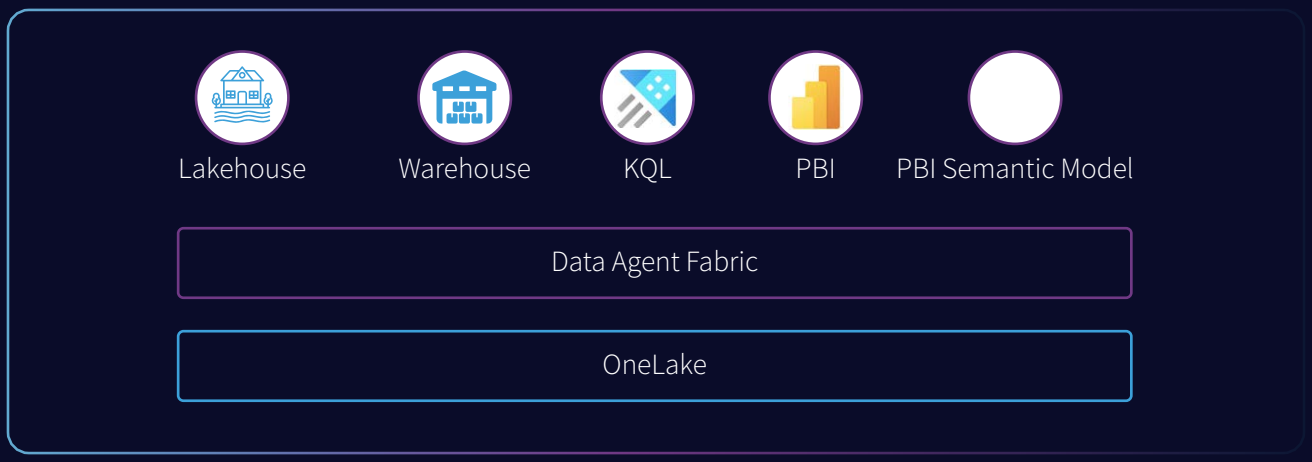
This capability not only improves analytical efficiency but also democratizes access to insights, making them available to both technical and non-technical users. Predictive analytics and automated reporting further empower teams to make informed, strategic decisions across industries.

## What is Data Agent?

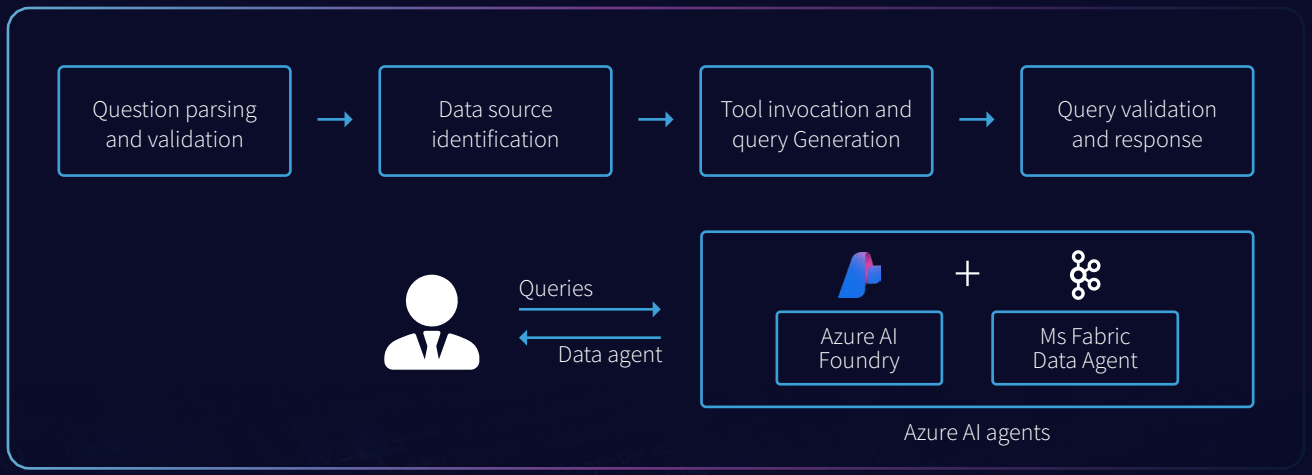
### Overview[1]

Data Agent is a Fabric-native Gen AI capability that enables natural language querying directly on datasets. It's powered by large language models (LLMs) and integrated with OpenAI APIs to offer real-time, conversational access to insights, without requiring knowledge of the dataset structure, schema, or query language.

Users can interact with data across multiple sources, including Data Warehouses, Data Lakehouses, KQL, and Power BI Datasets. Importantly, Data Agent operates with read-only access, ensuring data integrity while providing a seamless and intuitive user experience.

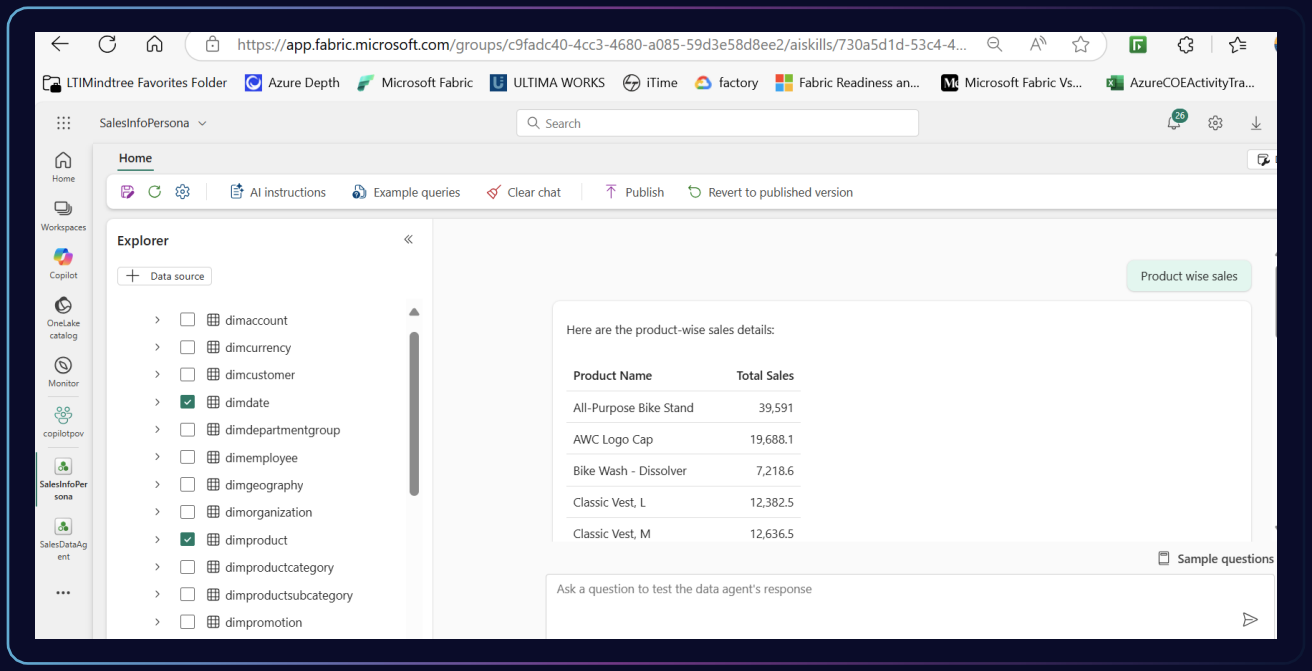



Data Agent overview




Data Agent - Architecture

For creating a data agent, we need to specify the data source, specify the tables, add context and instruction to proceed quickly.




Data Agent – Q&A

## Personas

The Data Agent is purpose-built for non-technical business users, enabling them to extract insights by querying datasets in natural English, without needing technical expertise or support. This democratizes access to data and allows everyday users to harness the power of AI for business decisions.

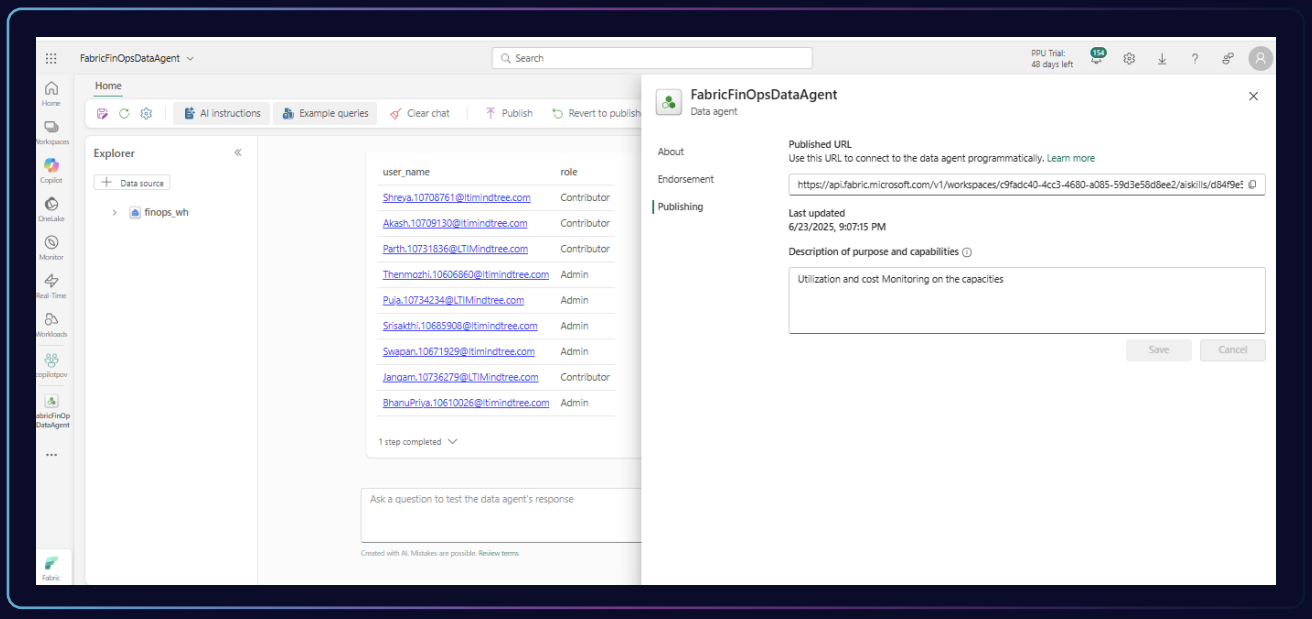
By bridging the gap between complex datasets and human language, the Data Agent introduces a paradigm shift in analytics. Users can interact directly with data using intuitive language, fostering a powerful dialogue between human intuition and machine intelligence. This seamless access unlocks the value of data assets—removing technical barriers and enabling business users to focus on strategy, not syntax.

## Integration

The Fabric Data Agent can be seamlessly integrated into a wide ecosystem, including Azure AI Foundry, Azure Copilot Studio, Microsoft Teams, custom applications, and web-based UX/UI using the Fabric Data Agent endpoints.

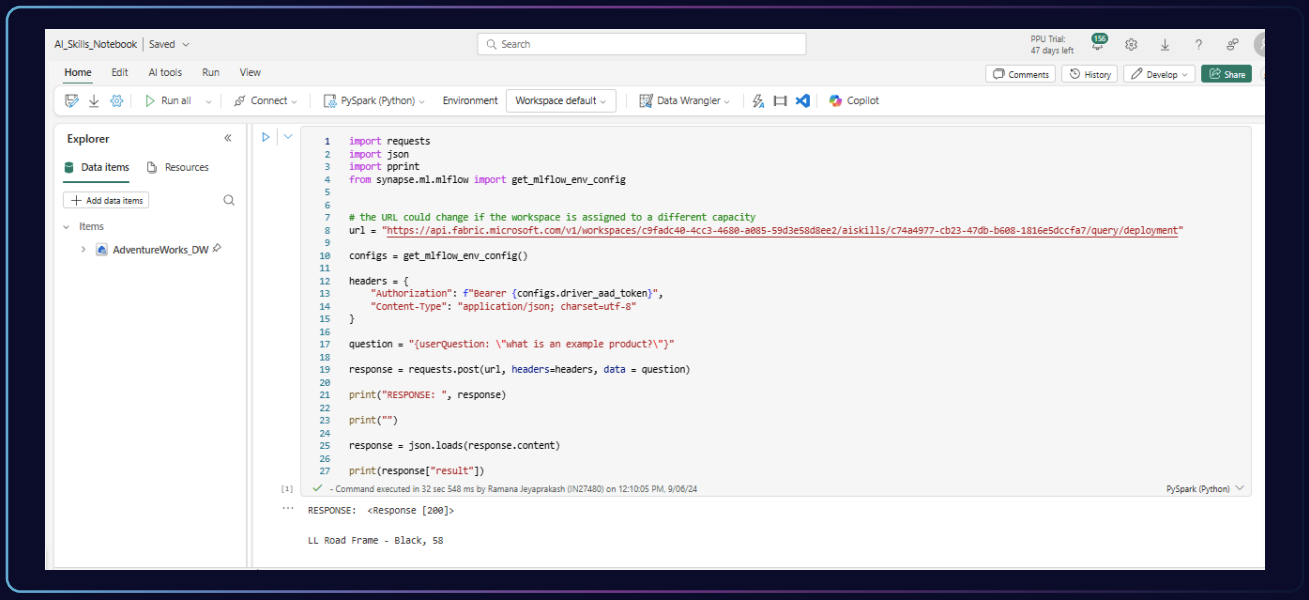
*Integration within the Fabric environment:*

1. Once a data agent is created and published, an API link is generated. This link can be used to run queries directly within Fabric notebooks.




Data Agent Endpoint

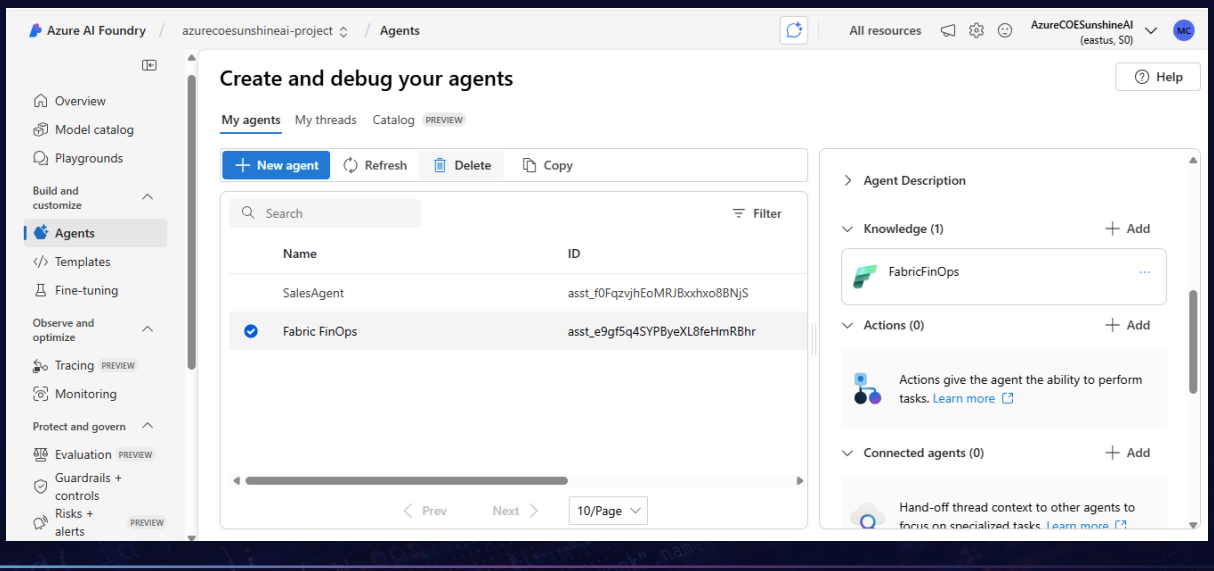
2. These URLs allow automation of workflows by accessing the data agent from within notebooks.



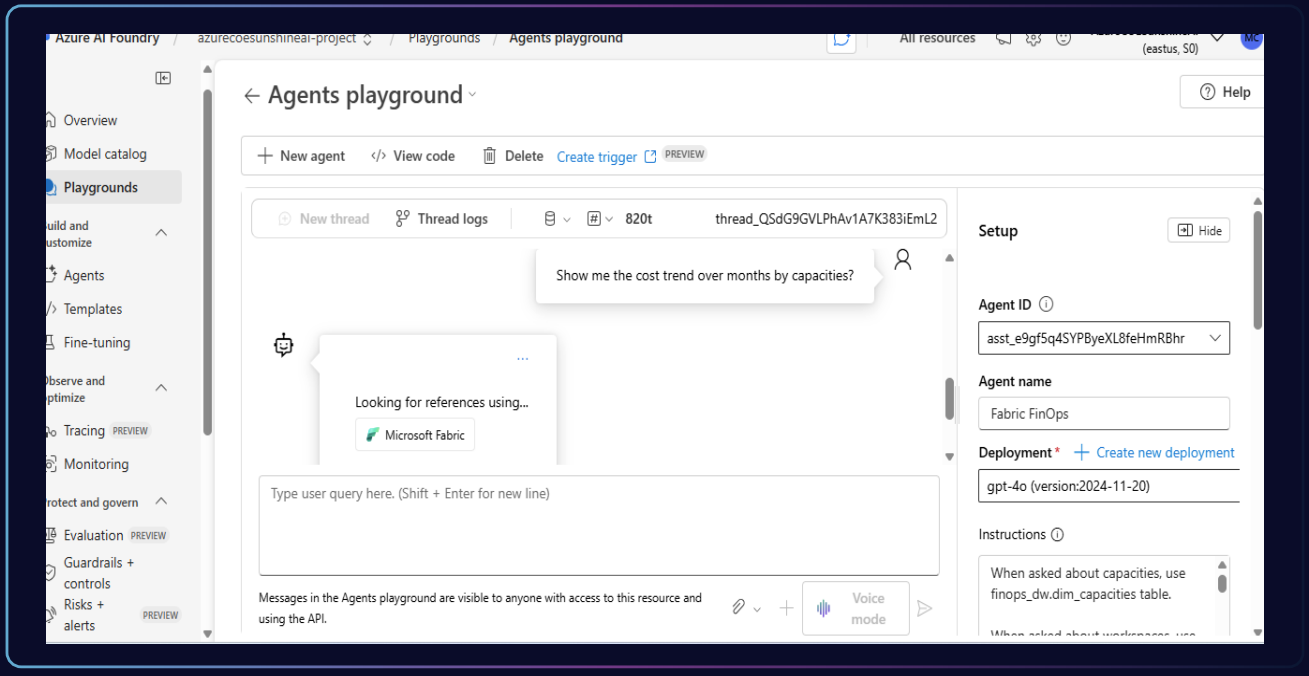

Using Data Agent Endpoint in an Application

*Running the data agent from external web application:*

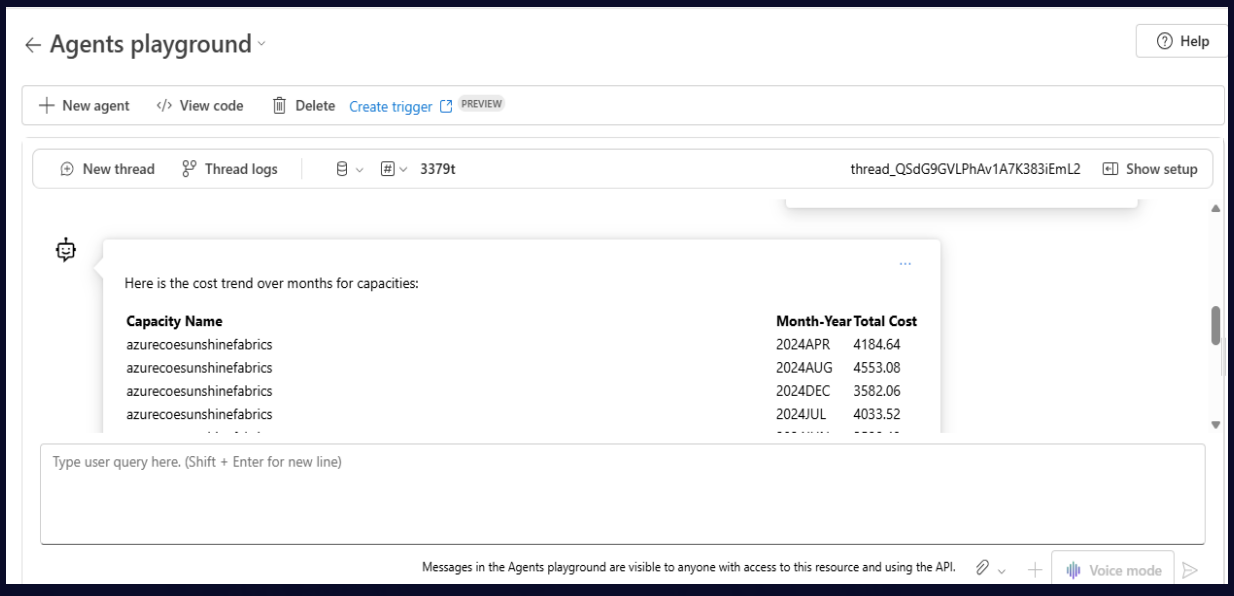
1. Create a data agent in Azure AI Foundry, link it to the Fabric data agent, and begin interacting by simply asking questions in natural language.




Using Data Agent in AI Foundry




Using Data Agent in AI Foundry

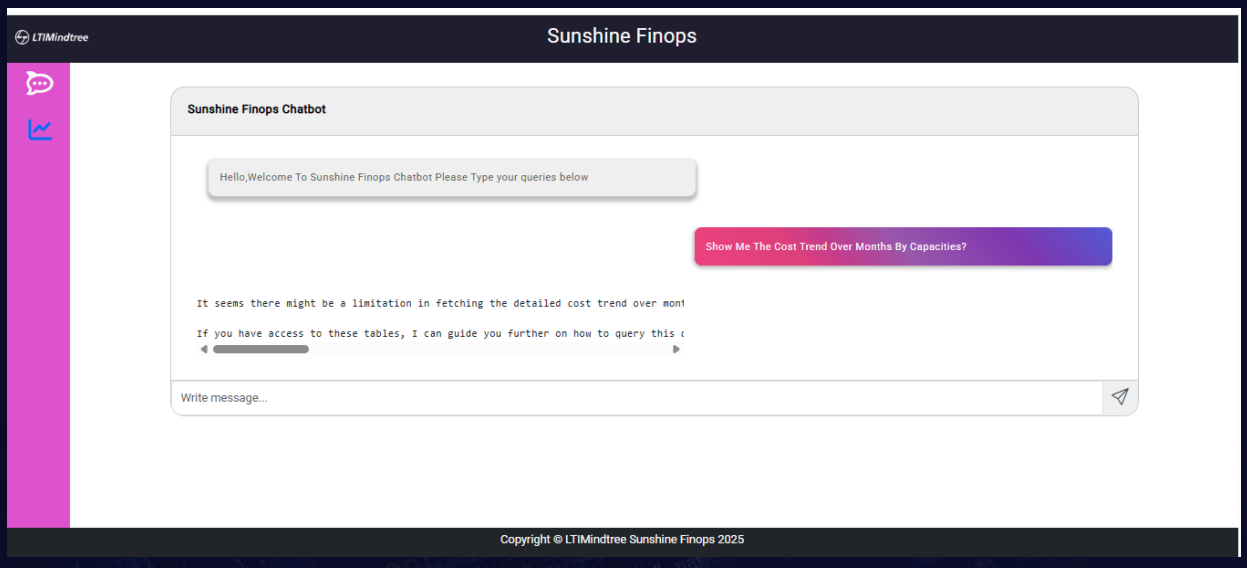



Using Data Agent in AI Foundry

2. Access the Data Agent via REST API and Service Principal (SP)-based authentication. Below are code snippets and examples that demonstrate how to use this in a custom application.

```python
def fetch_query_from_dataagent(user_query):


    credential = ClientSecretCredential(
        tenant_id=settings.AZURE_TENANT_ID,
        client_id=settings.AZURE_CLIENT_ID,
        client_secret=settings.AZURE_CLIENT_SECRET
    )

    agents_client = AgentsClient(
        endpoint=settings.AZURE_PROJECT_ENDPOINT,
        credential=credential,
    )

    with agents_client:
        bikesales_agent = agents_client.get_agent(agent_id='asst_e9gf5q4SYPByeXL8feHmRBhr')
        thread = agents_client.threads.create()
        print(f"Created thread, ID: {thread.id}")

        # full_query = f"{user_query} about {topic}"
        message = agents_client.messages.create(
            thread_id=thread.id,
            role="user",
            content=user_query,
        )
        print(f"Created message, ID: {message.id}")

        try:
            run = agents_client.runs.create_and_process(thread_id=thread.id, agent_id=bikesales_agent.id)
            print(f"Run finished with status: {run.status}")
        except Exception as e:
```

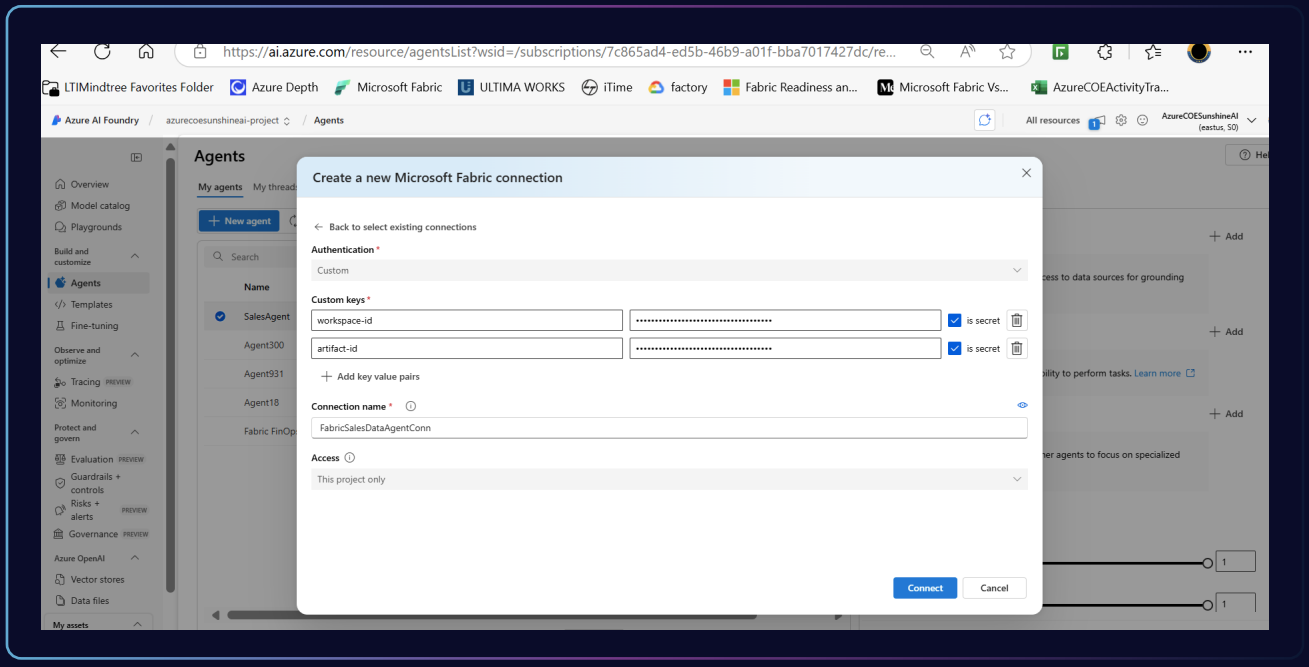
Using Data Agent in Custom Application



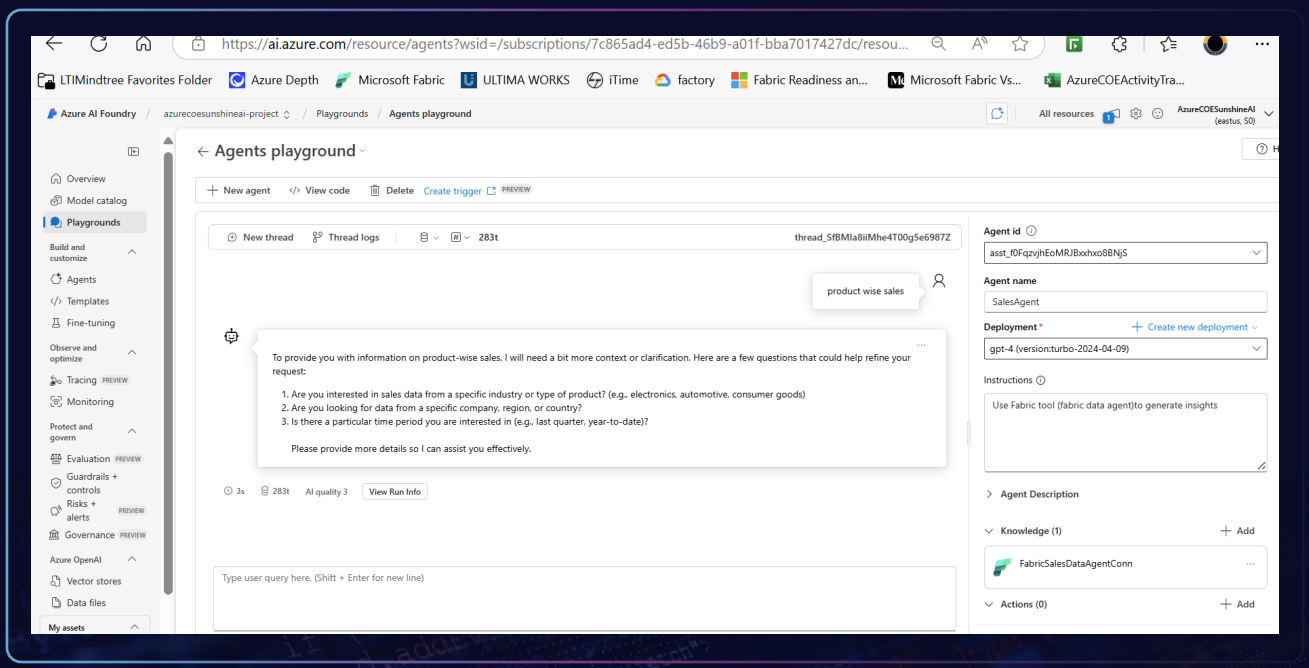

Using Data Agent in Custom Application

Fabric's integration flexibility ensures developers and analysts can embed Data Agent capabilities across multiple interfaces and platforms with ease.

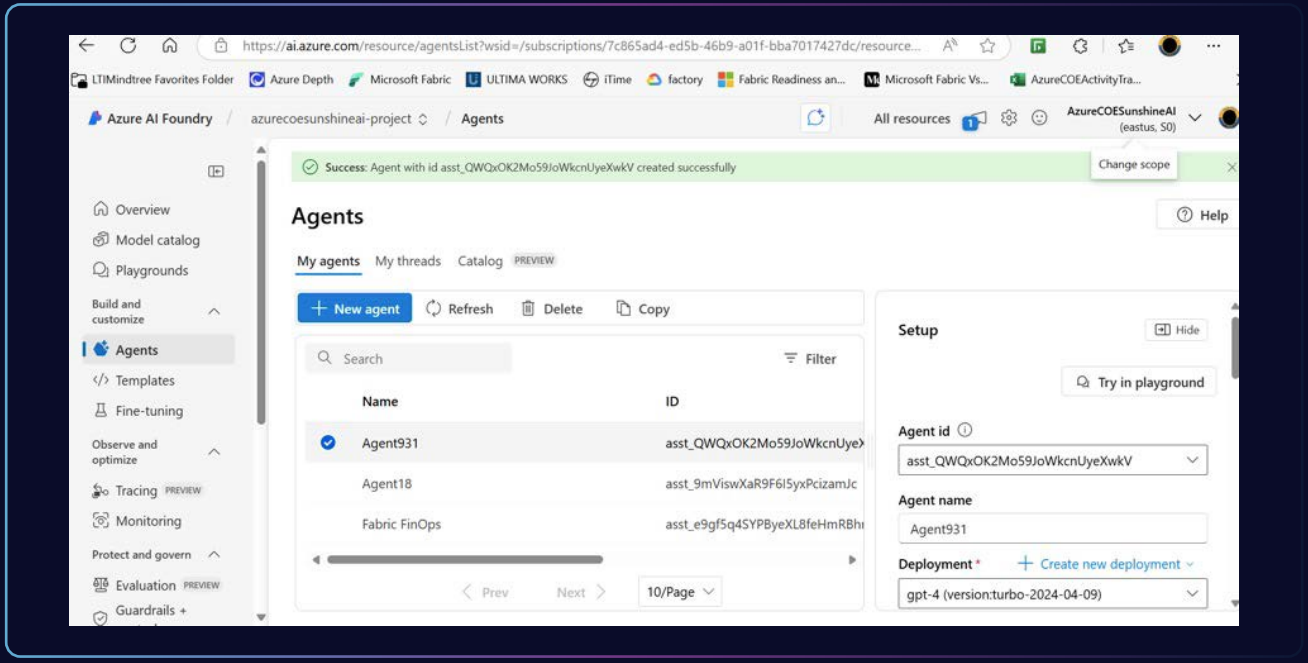
Fabric data agent can easily be integrated into Azure AI Foundry




Using Data Agent in AI Foundry




Using Data Agent in AI Foundry

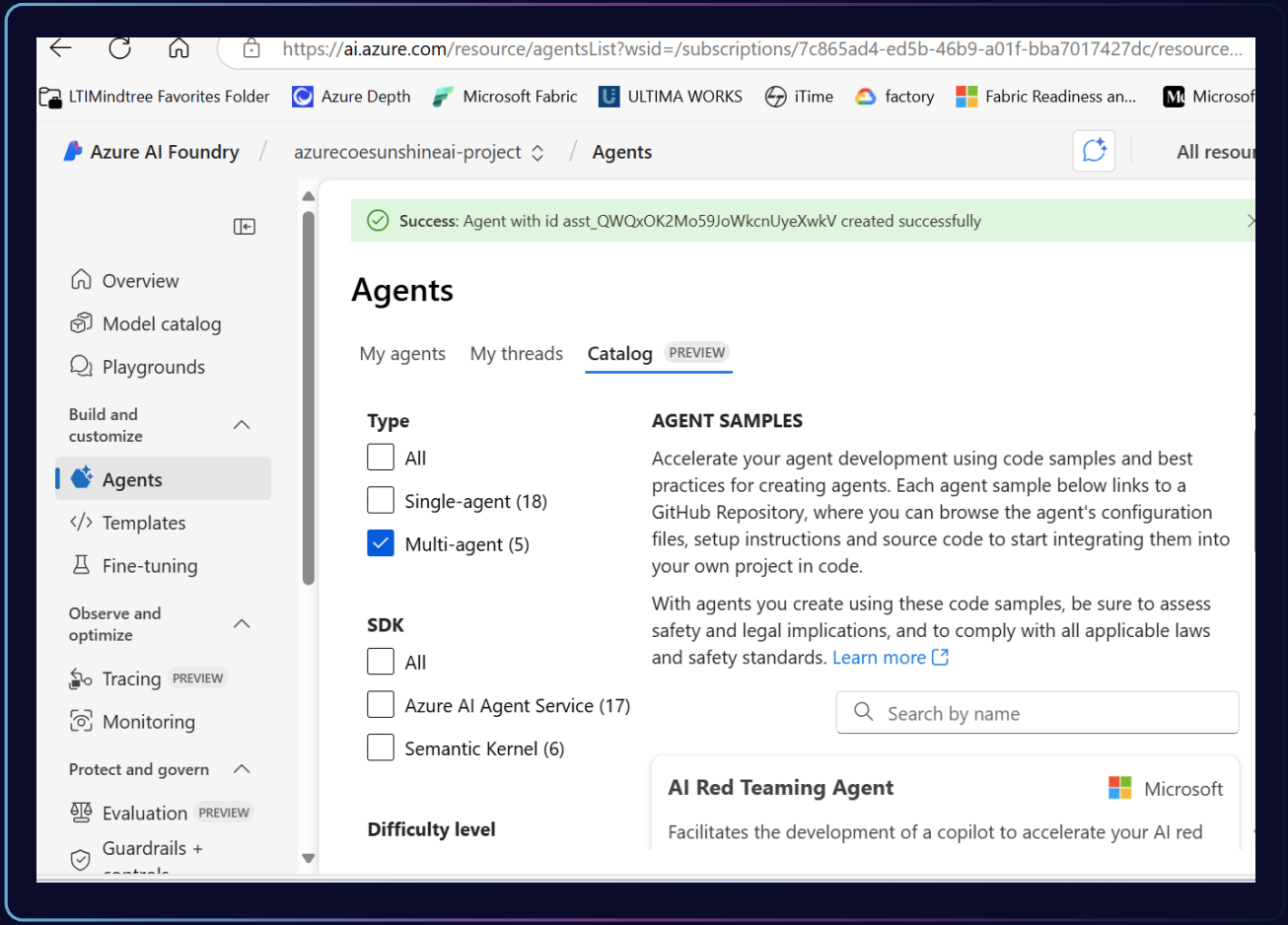



Using Data Agent in AI Foundry

## Specialized Agent

For specific business use cases, organizations may need to build dedicated Data Agents. These agents can be tailored for individual domains or processes, ensuring higher accuracy and relevance in responses.

Multi-agent orchestration can be achieved within Azure AI Foundry or Copilot Studio, enabling collaborative, domain-specific analysis and broader business logic flows.




Multi Agent Integration Scenario

## Security and Governance

Authentication and authorization for Data Agents are governed by the access permissions of the underlying datasets. This ensures data integrity and protects sensitive information based on role-based access.

To enhance governance further, Data Security Posture Management for AI within Microsoft Purview helps manage AI governance for Copilot, AI Agents, and Data Agents, providing centralized oversight and compliance alignment.

Data Agent Security and Governance

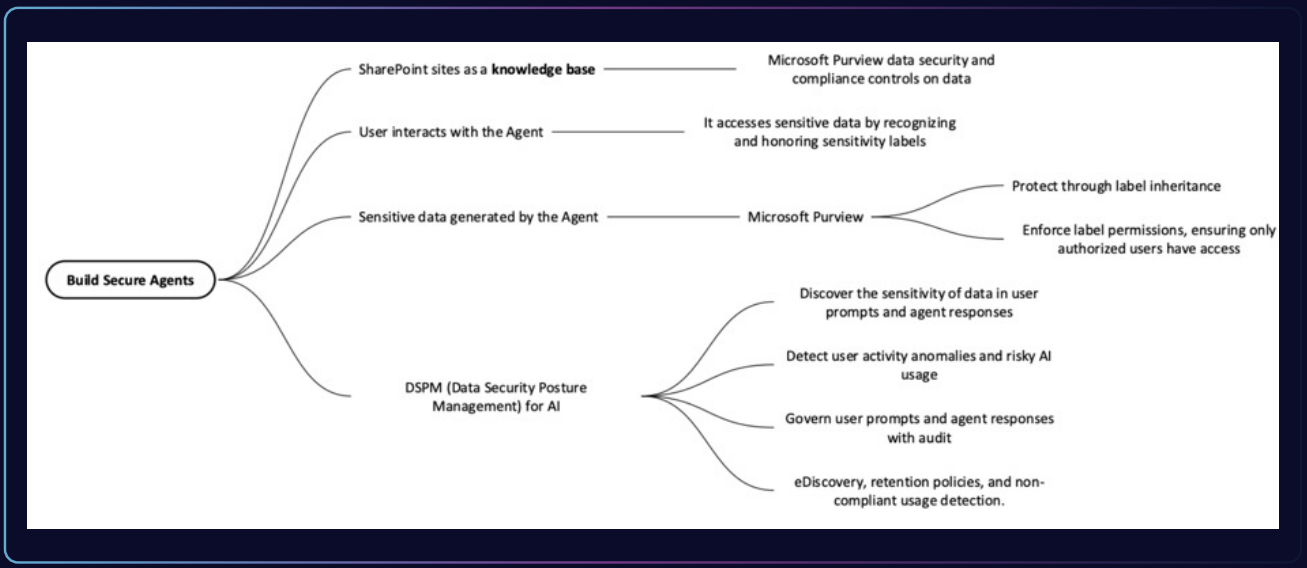


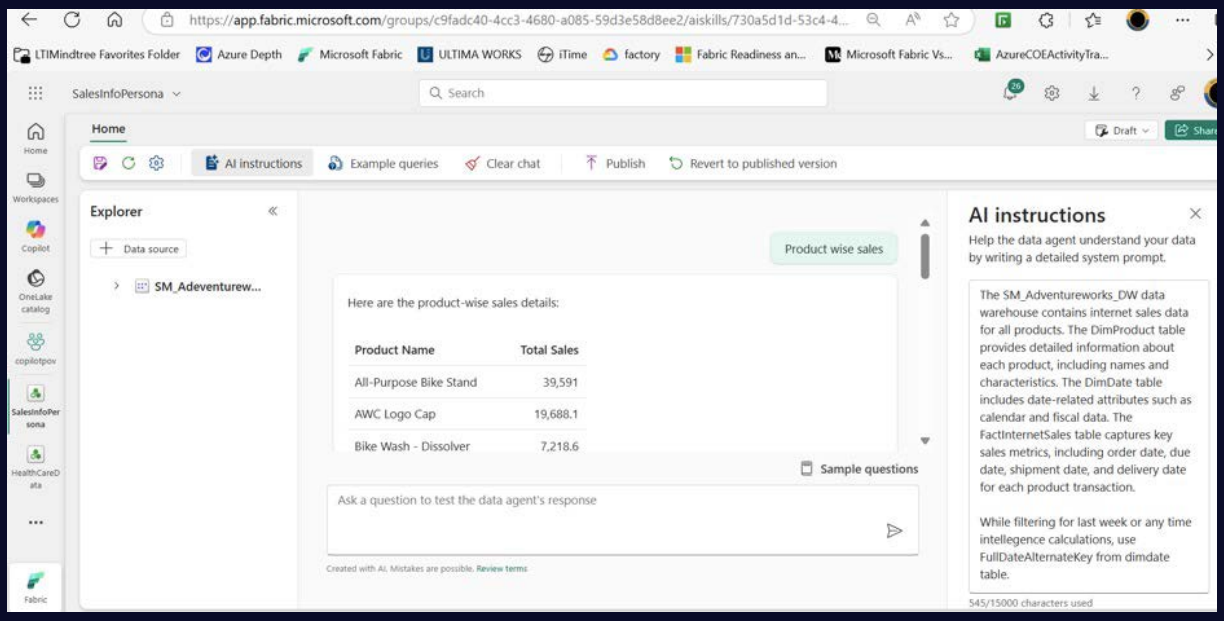

Image Source: Fabric + AI – Data Agents – Microsoft Partner's Co-Sell

## Tuning/Optimization

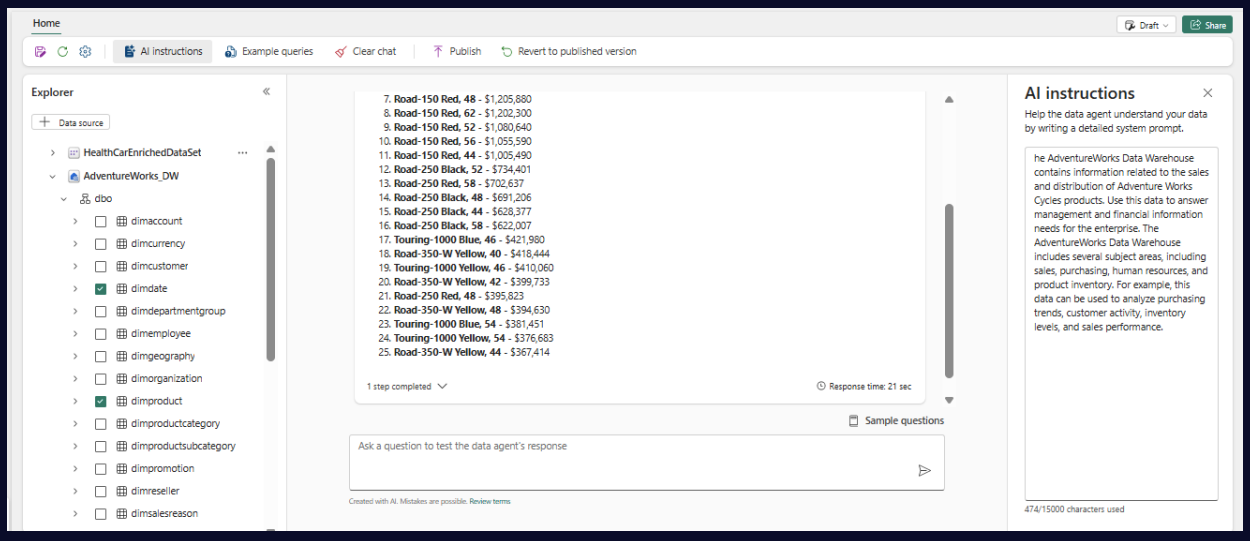
Data agents can be tuned and optimized using organization-specific instructions and contextual examples to improve accuracy and relevance.

With adaptive learning capabilities, the data agent refines its responses over time. Whether analyzing sales trends or customer behavior, it adapts based on feedback, enhancing both precision and contextual understanding. This ongoing feedback loop ensures the data agent evolves in sync with the organization's strategic priorities.

AI instructions allow further customization, defining scope, tone, fallback responses, and preferred terminology.

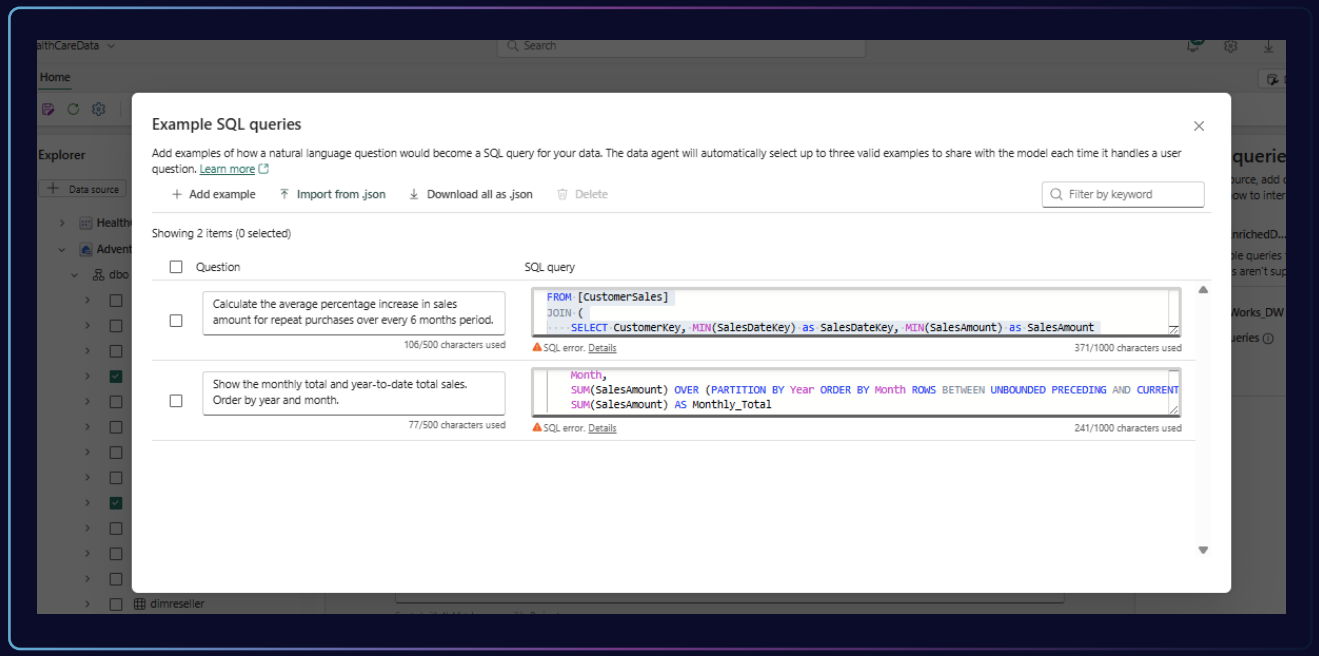



AI Instruction for Data Agent




AI Instruction for Data Agent

Example SQL-formatted queries can also be provided to fine-tune the agent's natural language-to-query translation engine.




Query for Data Agent

### Domain Expert/Validation

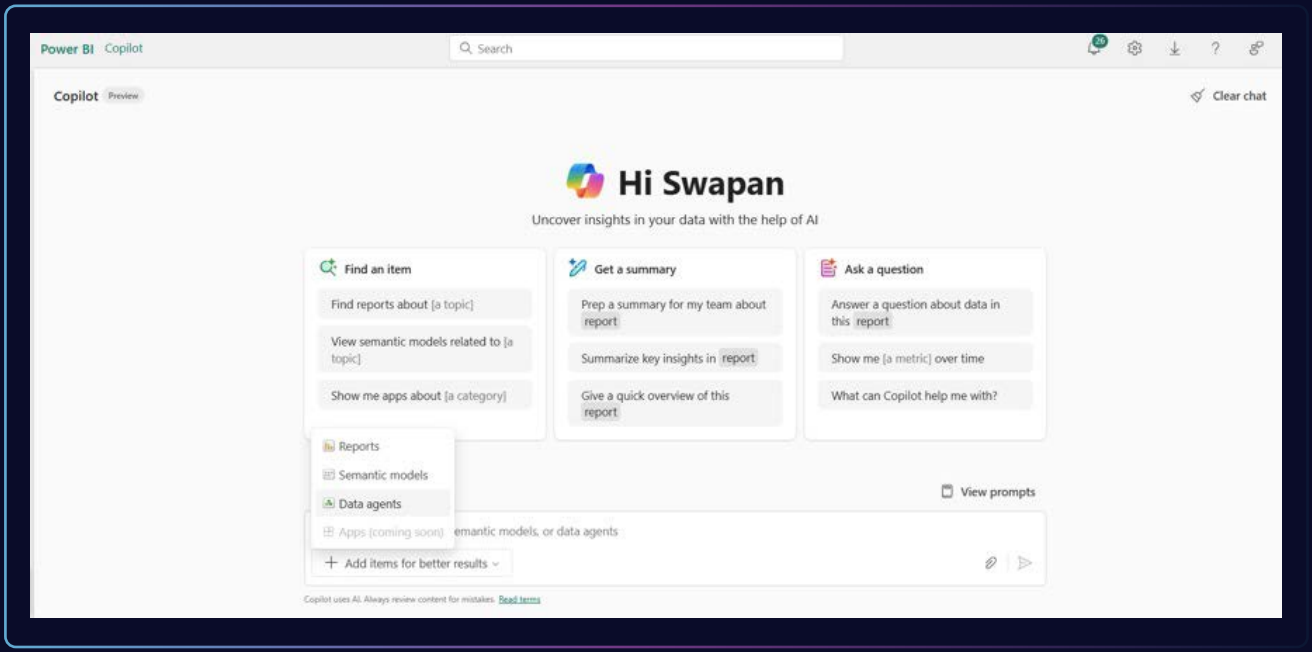
Domain validation is a crucial step in maximizing the value of data agents. Business analysts and subject matter experts play a pivotal role in reviewing and refining query responses.

By analyzing query outputs based on business context, these experts ensure accuracy, consistency, and relevance. Their input helps fine-tune the data agent to deliver better outcomes aligned with domain-specific needs.

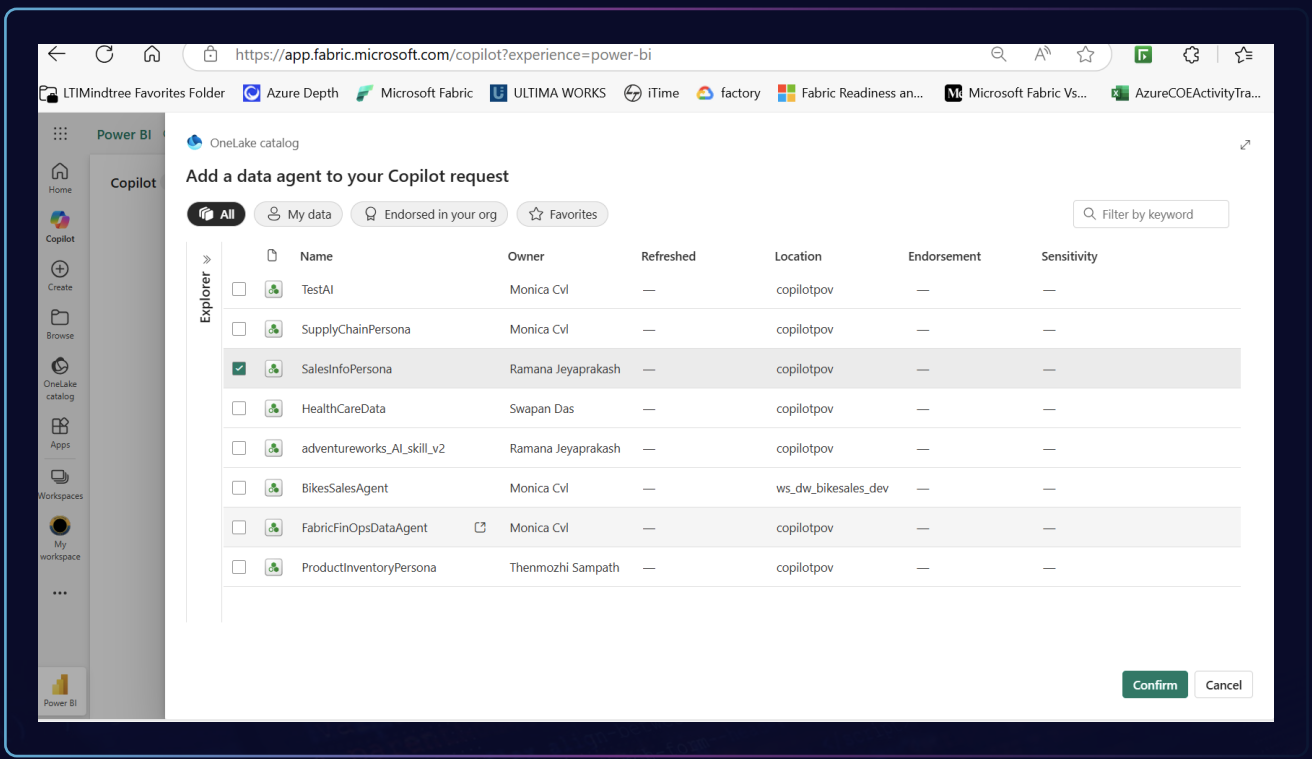
## New Features: Data Agents in Power BI Copilot and Teams[2]

Data Agents are now integrated across multiple Microsoft platforms, including Power BI Copilot, Microsoft 365 Copilot, and custom applications via Data Agent endpoints.

The interplay between AI-driven data agents and tools like Power BI Copilot adds a transformative layer to the analytics landscape. These integrations enable seamless querying, enhancing user experience by delivering precise and actionable insights directly within familiar platforms. By leveraging endpoints, developers can further extend these capabilities into bespoke applications, fostering adaptability across diverse operational environments.

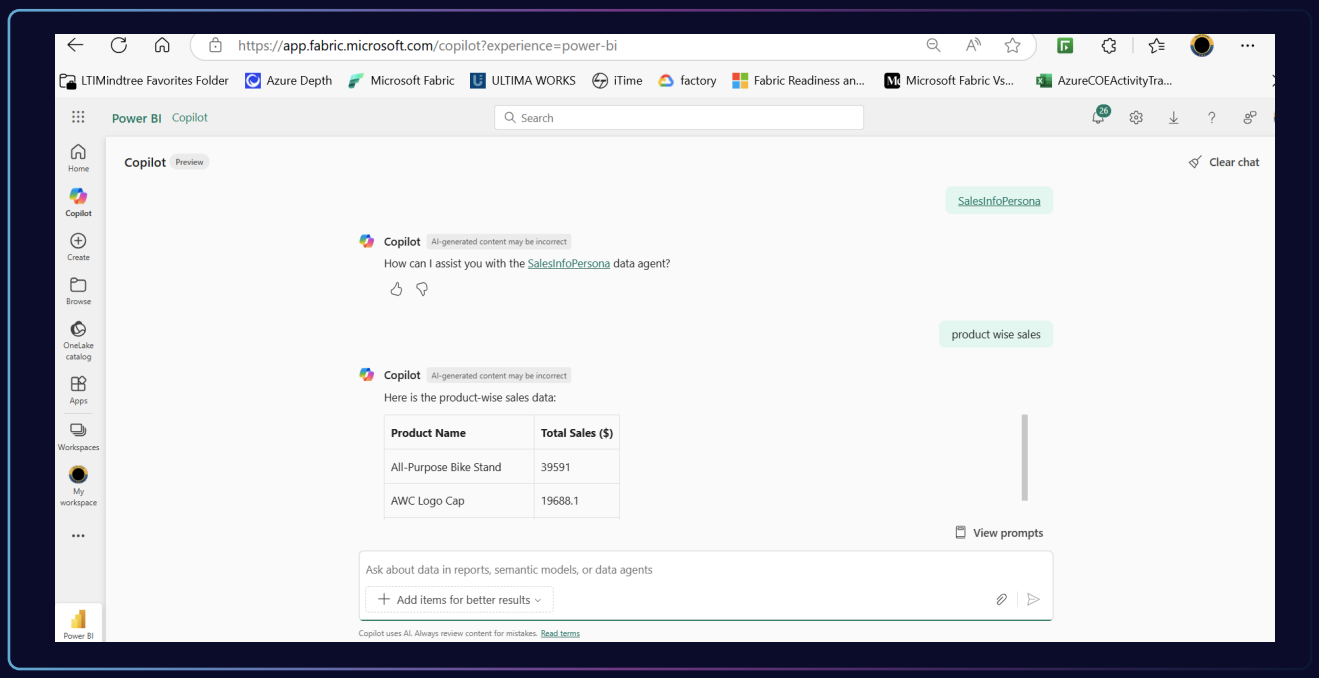



Data Agent as Copilot Knowledge




Data Agent as Copilot Knowledge

Data Agents serve as knowledge sources in Copilot, powering a smarter, more conversational interface for business intelligence and decision support.




Data Agent as Copilot knowledge

# Best Practices[3]

To make the most of data agents, organizations should follow these best practices:

- **Ensure data is AI-ready:** Use descriptive, functional, and meaningful names for tables and fields to improve query performance.
- **Use specialized agents for specific domains:** Create focused agent for specific tasks or domains to enhance accuracy and reduce ambiguity.
- **Minimize data scope:** Include only relevant tables and columns to narrow focus and improve performance.
- Be specific about tasks: Clearly define tasks and expected outputs; also specify responses when data is unavailable.

- **Define business terms, abbreviations, and synonyms:** Specify agent-level instructions to all data sources. Data source instructions can be used to specify how a term is used within a particular dataset.
- **Use natural language cues:** Incorporate SQL-like hints (e.g., LIKE '%...%') in instructions to guide query generation.
- **Give clear, focused instructions:** Specify the agent's function, tone, fallback behavior, and communication style.
- **Provide example queries:** These serve as prompts to help the agent generate correct logic. Vector similarity search identifies the top three most relevant examples and includes them in the agent's prompt to guide its response.

# Conclusion

The Fabric Data Agent empowers business users to interact with their data using natural conversation, without needing dashboards, analytics tools, or technical knowledge.

By leveraging natural language as a user interface, it transforms the way organizations access insights. No complex models, queries, or reports, just simple, powerful, intuitive conversations.

With built-in adaptability, seamless integration, and governance, the Data Agent serves as a bridge between technical complexity and business accessibility. Organizations can now extract strategic value from data—quickly, easily, and intelligently, driven by a GenAI-powered assistant that truly understands business context.

# References

1. *Fabric data agent concepts (preview),* July 8, 2025: https://learn.microsoft.com/en-us/fabric/data-science/concept-data-agent
2. *Extracting deeper insights with Fabric Data Agents in Copilot in Power BI,* Amir Jafari, Joanne Wong, Microsoft Fabric Updates Blog, May 19, 2025: https://blog.fabric.microsoft.com/en-us/blog/extracting-deeper-insights-with-fabric-data-agents-in-copilot-in-power-bi
3. *Best practices for configuring your data agent,* June 18, 2025: https://learn.microsoft.com/en-us/fabric/data-science/data-agent-configuration-best-practices

# About the Authors

## Swapan Kumar Das

Principal – Architecture, Data and Analytics

Swapan brings extensive expertise in technology, leadership, data and application architecture, cloud computing, and big data. With a specialization in data warehousing, business intelligence, analytics, and Azure, he has successfully collaborated with numerous Fortune 500 companies. Swapan is deeply passionate about exploring emerging technologies and advancing research in areas such as data engineering, data quality, data governance, data science, cognitive APIs, quantum computing, AI and ML, generative AI, and agentic AI.

## Prashant Atri

Strategic Advisor – Data & AI, Cross Solutions | Microsoft

As Microsoft's Strategic Advisor – Data & AI, Prashant drives enterprise transformation using AI and analytics. He has led cloud migrations for global pharma clients in regulated environments and works with SIs/ISVs across industries. Passionate about innovation, he champions ideation, frameworks, and community building through impactful knowledge-sharing.

## Monica CVL

Senior Data Engineer, LTIMindtree

Monica is a passionate Data Engineer with expertise in Azure Stack, always eager to explore new technologies and grow her skill set.