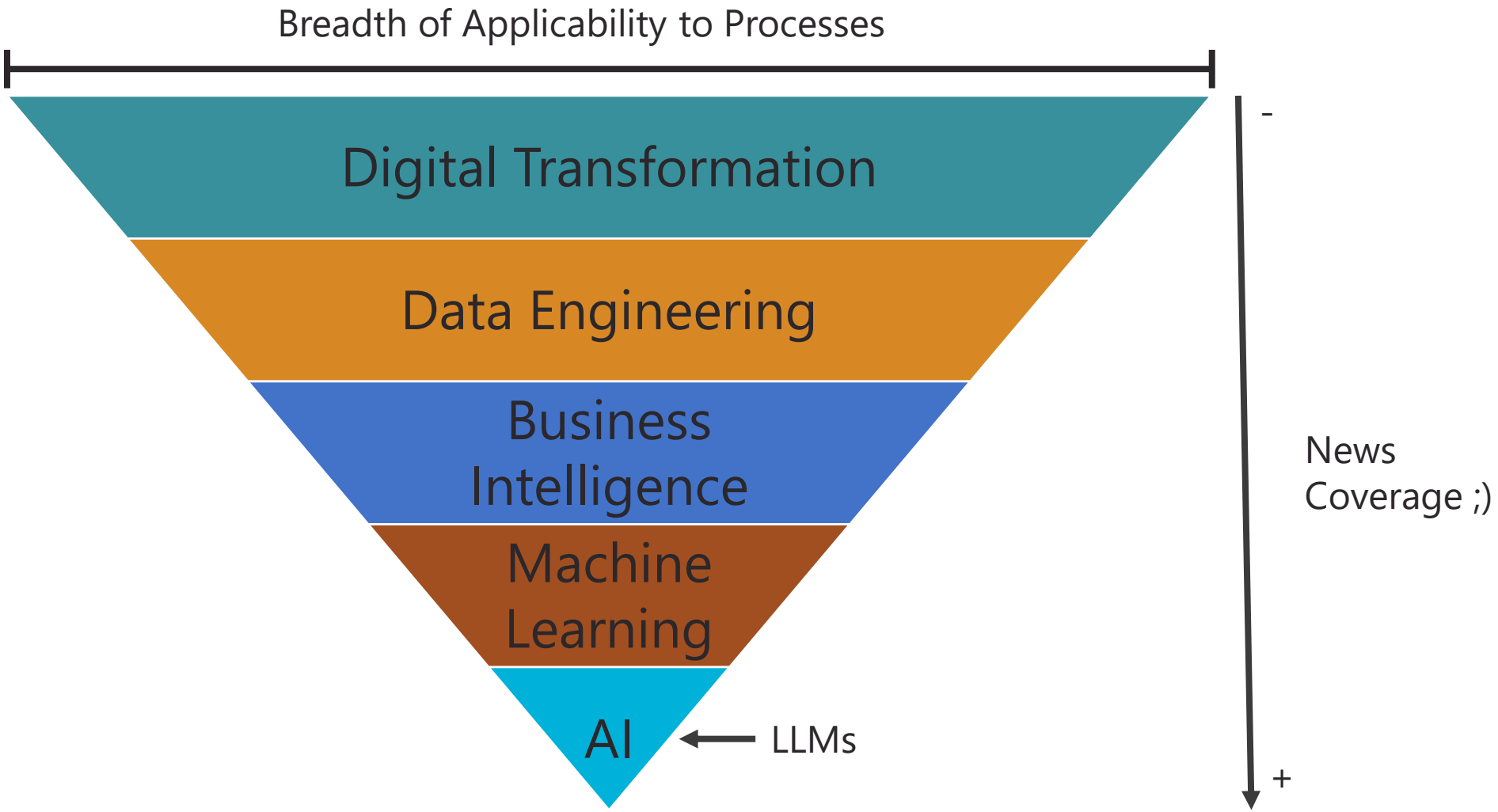
Ventagium
Data Consulting

Fabric Workshop

2025-02-28 Part 2

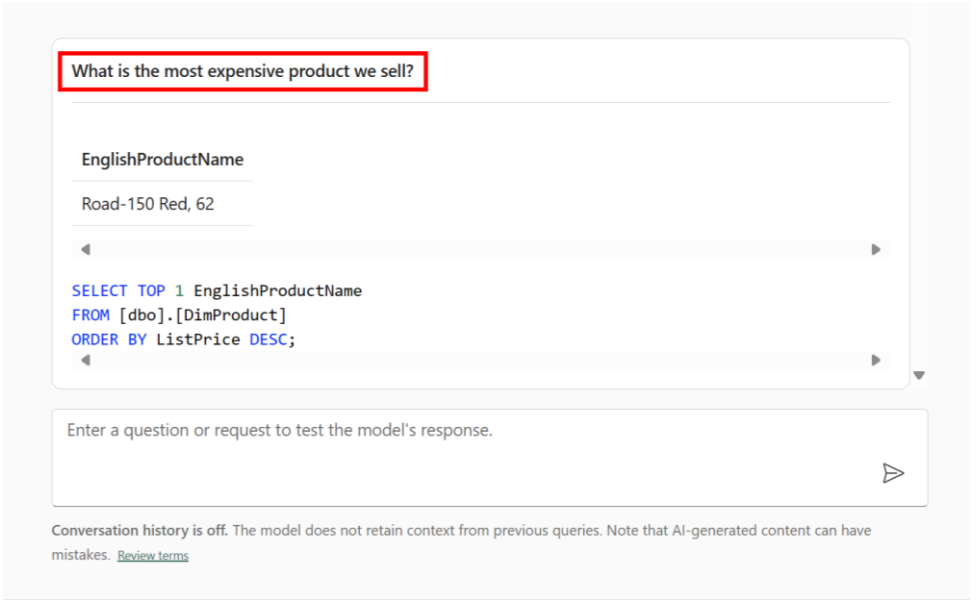
Our thoughts on AI (for Organizations)



AI Skills

What are AI Skills?

- Generative AI system designed to provide insights that answer questions about your data, **as well as the queries to arrive to those insights.**
- Powered by Azure OpenAI.
- Able to interact with data from a Warehouse/Lakehouse, Semantic Models or KQL Databases.
- Additionally trained and tweaked via Example queries and AI instructions.



How AI Skills work

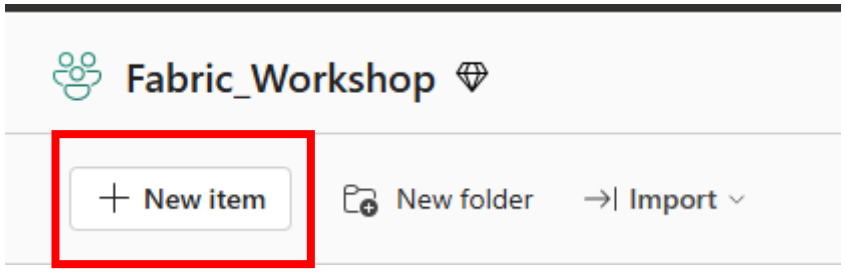
1. System uses credentials to Access the schema of the data source.
2. Prompt is constructed based on:
   A. Question provided by user.
   B. Metadata of the data source(s).
   C. Example queries.
   D. AI Instructions.
3. Query is generated based on Data Source (SQL, DAX or KQL).
4. Response is provided to the user.

Prerequisites.

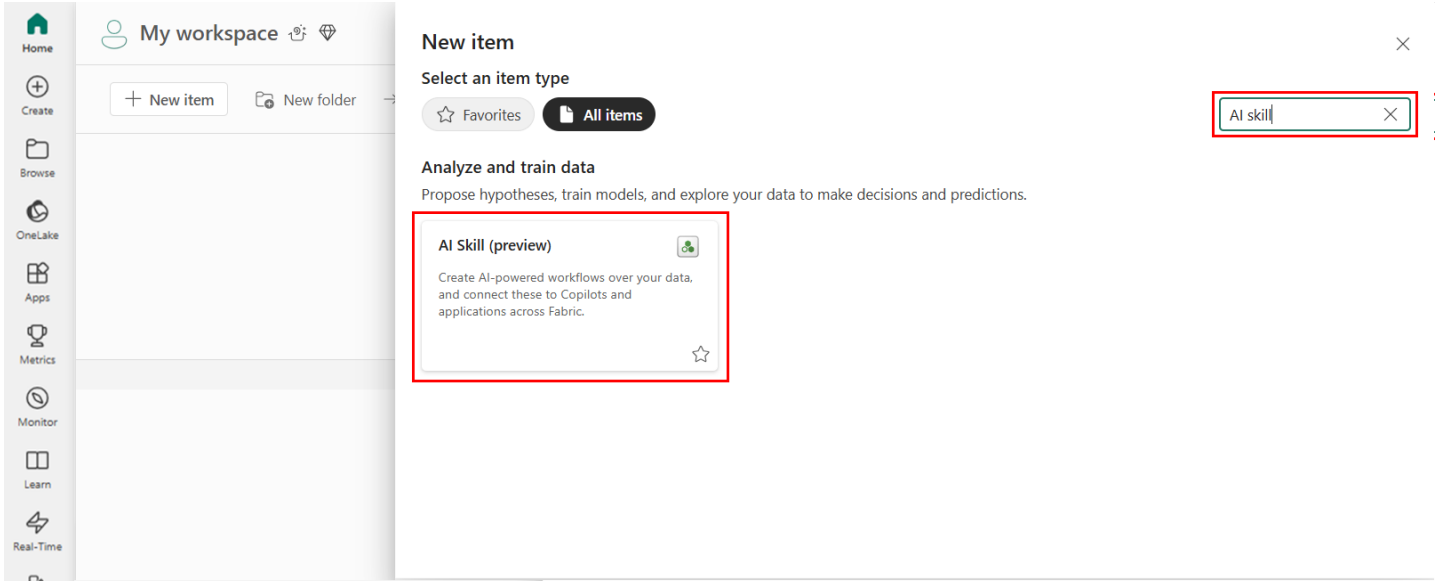
- A paid F64 or higher Fabric capacity resource.
- AI skill tenant switch is enabled.
- Copilot tenant switch is enabled.
- Cross-geo sharing and processing for AI is enabled.
- A Warehouse, Lakehouse, Semantic Model or KQL Database with data.

Getting started in 2 steps…

1. Click on the "**+ New item**" button.

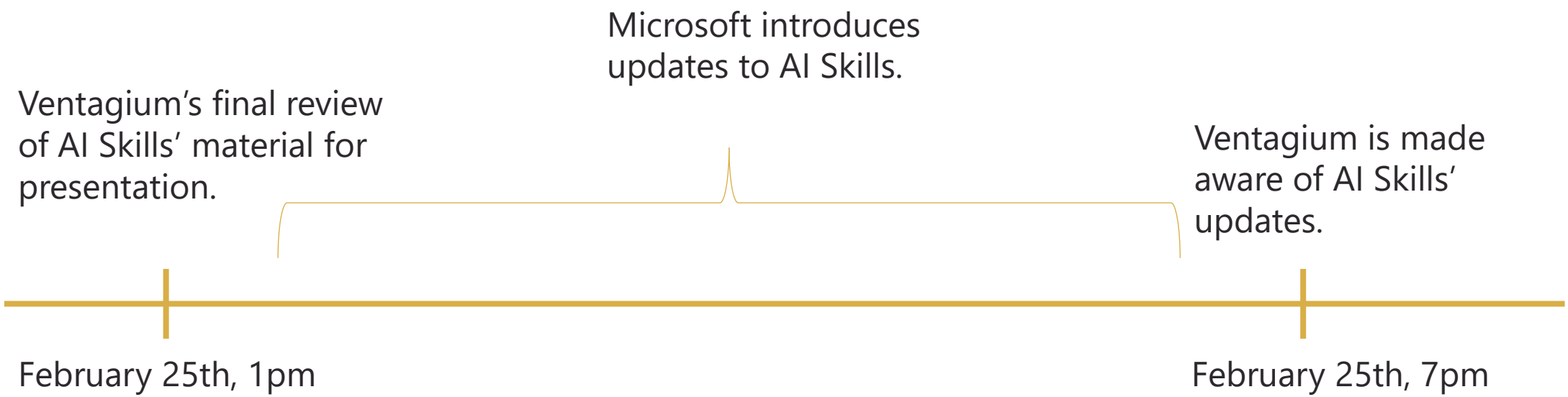


2. Click on **AI Skill** item.

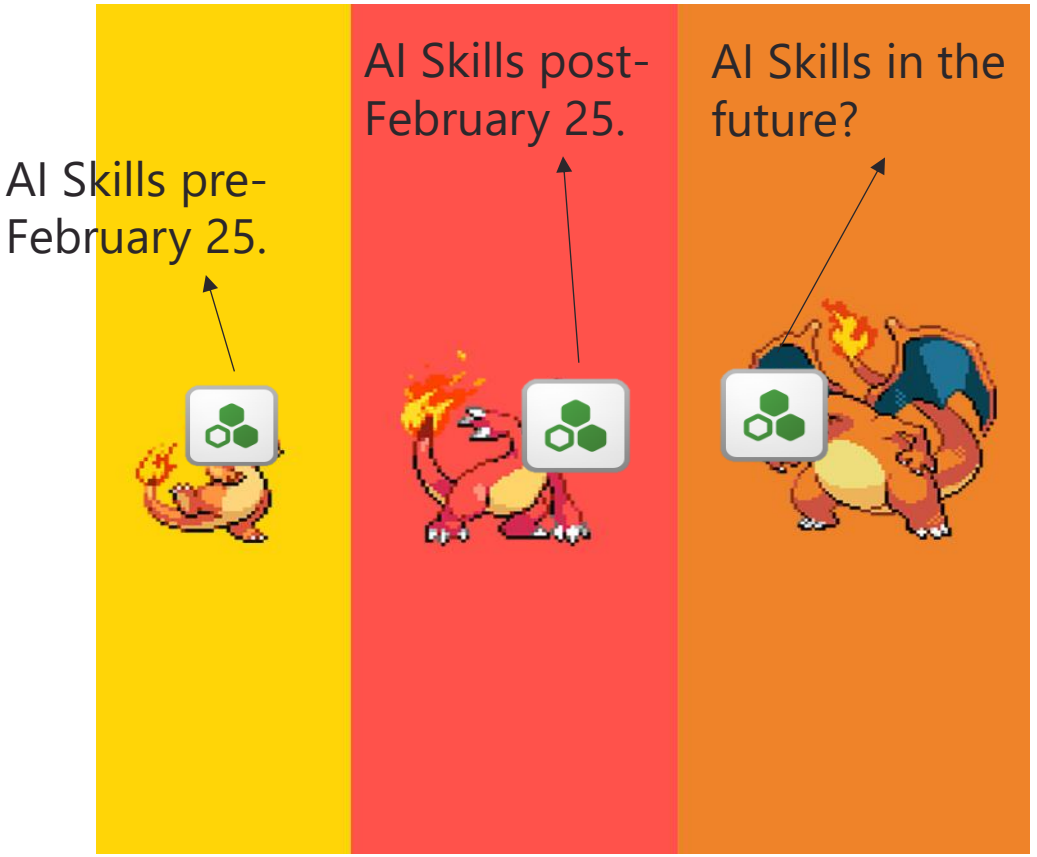


Exercise: AI Skills Tutorial

Could be a joke, but it's actually a true story



AI Skills' evolutionary line



Long list of limitations…

(25 Feb 1 pm)

Limitations

The AI skill is currently in public preview and has limitations. Updates will improve the AI skill over time.

- Generative AI doesn't interpret the results of an executed T-SQL query. It only generates that query.
- The AI skill might return incorrect answers. You should test the AI skill with your colleagues to verify that it answers questions as expected. If it makes mistakes, provide it with more examples and instructions.
- Only T-SQL queries on warehouses and lakehouses are supported.
- The AI skill only generates T-SQL "read" queries. It doesn't generate T-SQL queries that create, update, or delete data.
- The AI skill can only access data that you provide. It only uses the data resource configurations that you provide.
- The AI skill has data access permissions that match the permissions granted to the AI skill questioner. This is true when the AI skill is published to other locations, for example, Copilot for Microsoft 365 or Microsoft Copilot Studio.
- You can't use the AI skill to access unstructured data resources. These resources include .pdf, .docx, or .txt files, for example.
- At this time, you can only select a single warehouse or a single lakehouse.
- The AI skill doesn't support a conversational interface. Every question must be fully self-contained. It doesn't remember earlier questions.
- It blocks non-English language questions or instructions.
- You can't connect the AI skill to Fabric copilots, Microsoft Teams, or other experiences outside of Fabric.
- You can't change the LLM that the AI skill uses.
- The AI skill loses accuracy if you use nondescriptive column names.
- The AI skill loses accuracy if you use large schemas with dozens of tables.
- The AI skill is in a preview status. It has a limited scope and it might have bugs. Because of these considerations, we recommend that you avoid its use in production systems. Also avoid its use for critical decisions.
- Nondescriptive data resource column and table names have a significant, negative impact on generated T-SQL query quality. We recommend the use of descriptive names.
- Use of too many columns and tables might lower AI skill performance.
- The AI skill is currently designed to handle simple queries. Complex queries that require many joins or sophisticated logic tend to have lower reliability.

Even longer list of limitations…

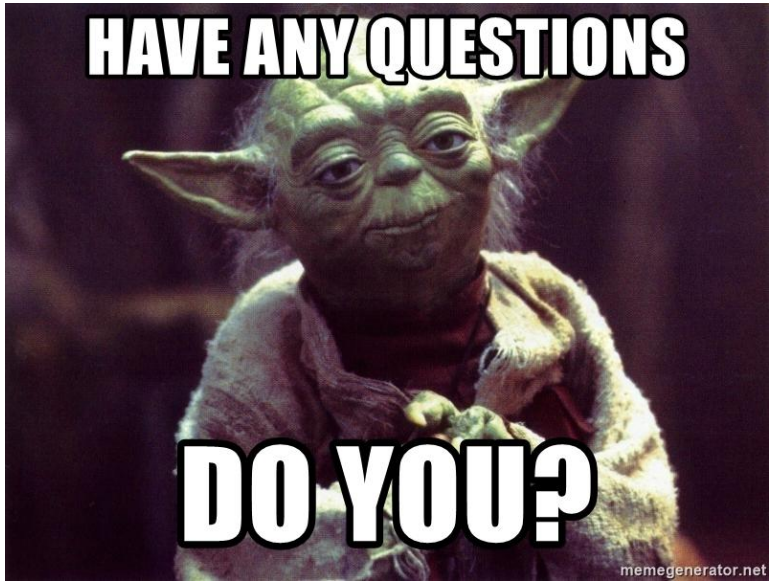
(25 Feb 7 pm)

Limitations

The AI skill is currently in public preview and it has limitations. Updates will improve the AI skill over time.

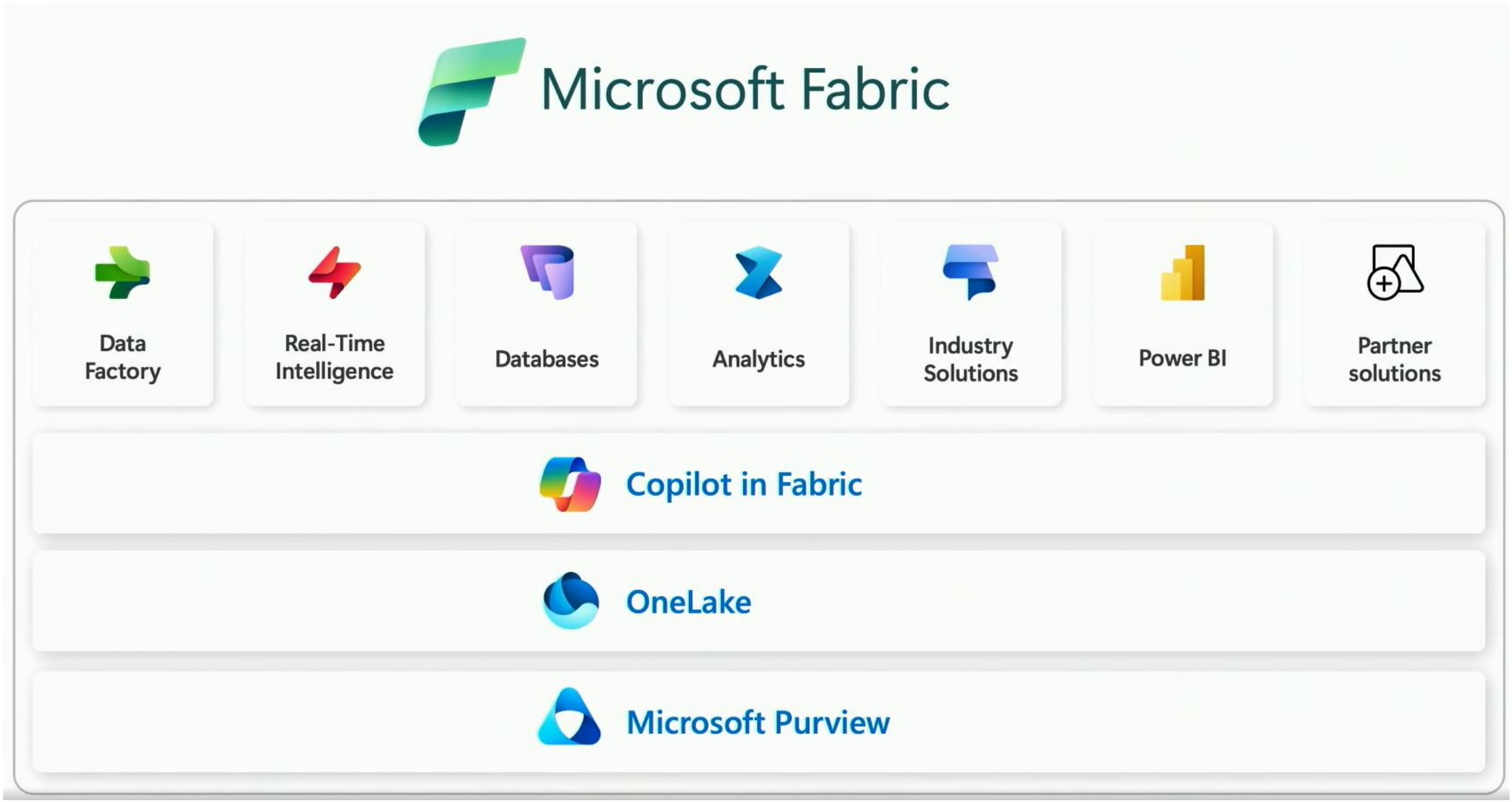
- The AI skill can retrieve data by generating structured queries (SQL, DAX, or KQL) for questions that involve facts, totals, rankings, or filters. However, it can't interpret trends, provide explanations, or analyze underlying causes.
- The AI skill only generates SQL/DAX/KQL "read" queries. It doesn't generate SQL/DAX/KQL queries that create, update, or delete data.
- The AI skill can only access data that you provide. It only uses the data resource configurations that you provide.
- The AI skill has data access permissions that match the permissions granted to the user interacting with the AI skill. This is true when the AI skill is published to other locations-for example, Microsoft Copilot Studio, Azure AI Foundry, and Microsoft Teams.
- You can't add more than five data sources to the AI skill.
- You can't use the AI skill to access unstructured data resources. These resources include .pdf, .docx, or .txt files, for example.
- The AI skill blocks non-English language questions or instructions.
- You can't change the LLM that the AI skill uses.
- You can't add a KQL database as a data source if it has more than 1,000 tables or any table with over 100 columns.
- You can't add a Power BI semantic model as a data source if it contains more than a total of 100 columns and measures.
- The AI skill works best with 25 or fewer tables selected across all data sources.
- Nondescriptive data resource column and table names have a significant, negative impact on generated SQL/DAX/KQL query quality. We recommend the use of descriptive names.
- Use of too many columns and tables might lower AI skill performance.
- The AI skill is currently designed to handle simple queries. Complex queries that require many joins or sophisticated logic tend to have lower reliability.
- If you add a Power BI semantic model as a data source, the AI skill doesn't use any hidden tables, columns, or measures.
- If you previously created an AI skill that used a warehouse as a data source, and the warehouse was located in a workspace that doesn't host that AI skill, you might encounter an error. To resolve this issue, delete the existing data source and add it again.
- To add a Power BI semantic model as a data source for AI skill, you need read/write permissions for that Power BI semantic model. Querying an AI skill that uses a Power BI semantic model also requires that you have read/write permissions for the underlying Power BI semantic model.
- The AI skill might return incorrect answers. You should test the AI skill with your colleagues to verify that it answers questions as expected. If it makes mistakes, provide it with more examples and instructions.
- If you previously created and published an AI skill, and you have used its URL programmatically, the URL will no longer work if you open the AI skill in the AI skill new user interface page. To resolve this, you must republish the AI skill, and use the new URL based on the Assistants API.

Questions?



SQL Database in Fabric

Roberto Hernández
roberto.Hernandez@ventagium.com
https://www.linkedin.com/in/robertohj/


Microsoft Fabric
Data Factory
Real-Time Intelligence
Databases
Analytics
Industry Solutions
Power BI
Partner solutions
Copilot in Fabric
OneLake
Microsoft Purview

SQL database in Fabric

Build AI apps faster and easier than ever

Simple

Autonomous

Optimized for AI

Built upon the familiar SQL Server/Azure SQL Database engine, in Microsoft Fabric

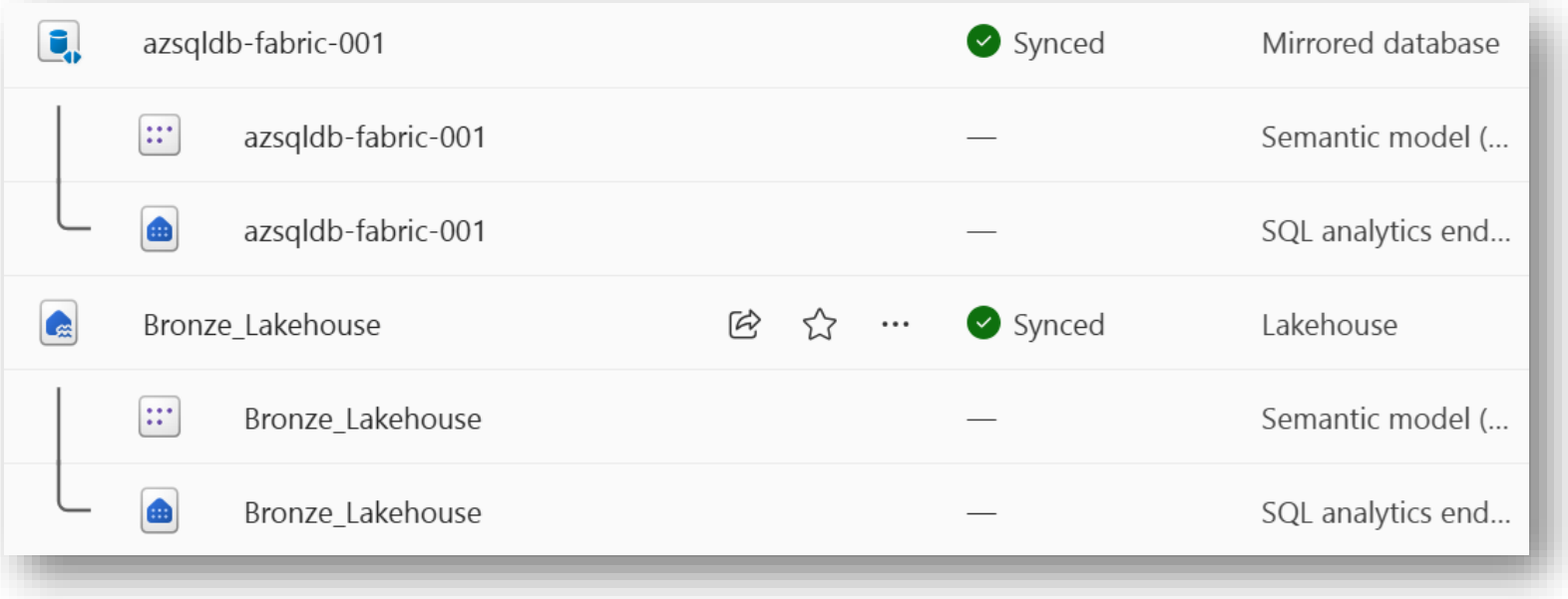
5

SQL Database (Preview)

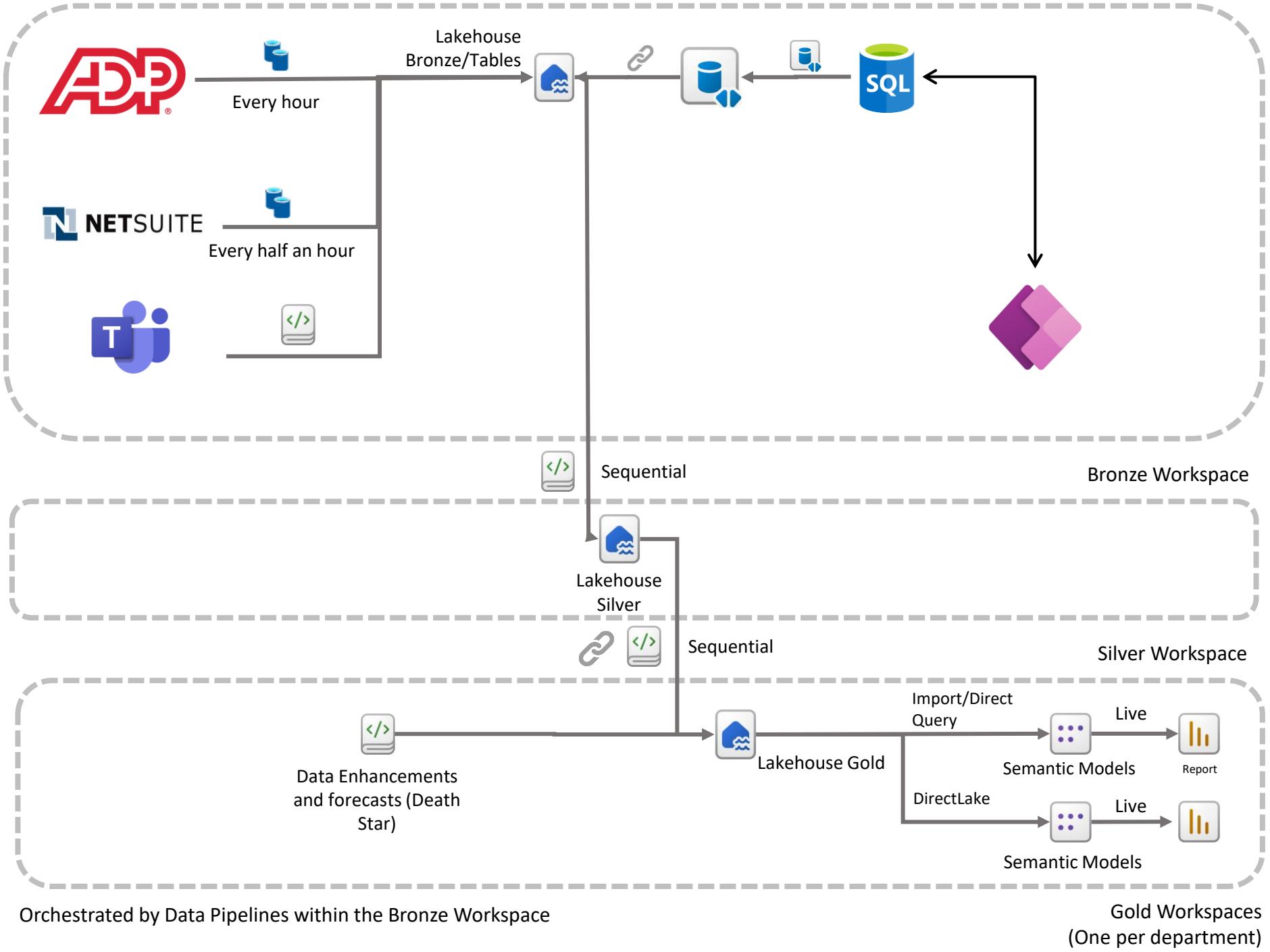
- OLTP (vs OLAP)
- Based on Azure SQL Database
- Automatic provisioning and scaling
- Near real time into OneLake
  - Available with Spark, notebooks, data engineering, reports
- Query editor
  - Fabric Query Editor
  - SQL Server Management Studio
  - Visual Studio Code
- GraphQL APIs

SQL Database – Use case example

- Bring your OLTP data to Fabric: Near Real Time data

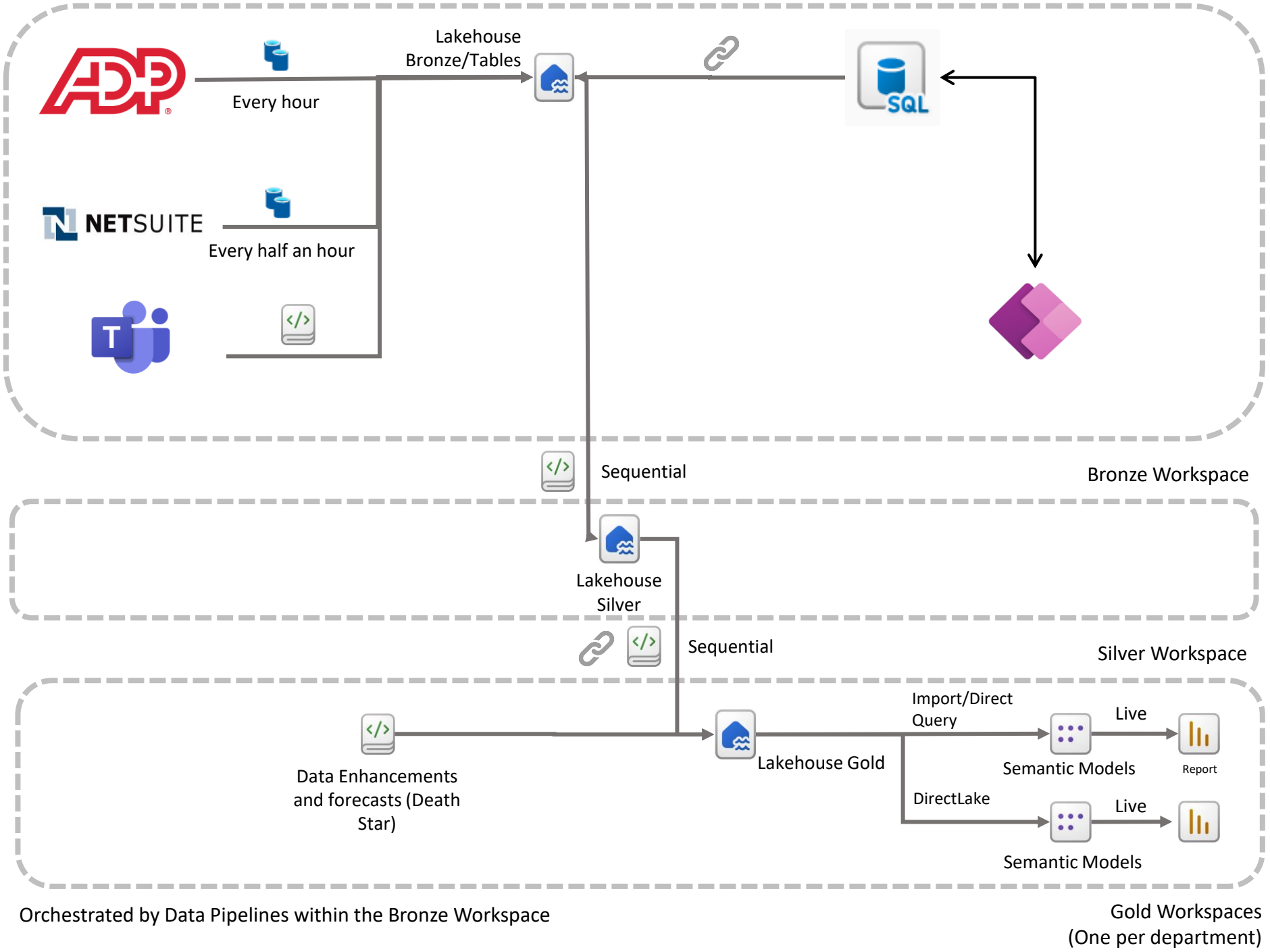


Data Flow
NexaCore



Dataflow Gen2
Warehouse
Lakehouse
Data Pipeline
Report
Semantic Model
Notebook
Copy Data
Shortcuts
Python Script
OneLake
Mirroring
Power Apps

Data Flow
NexaCore



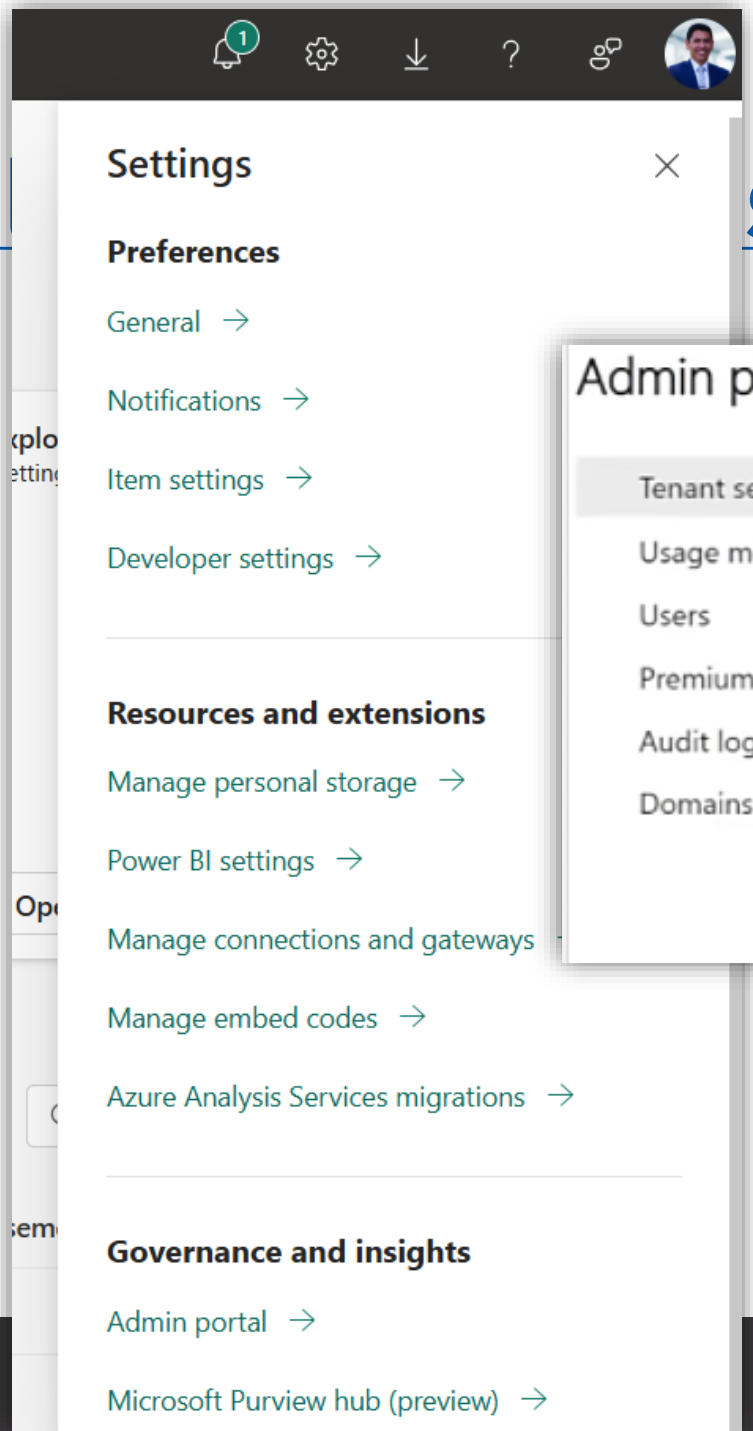
Dataflow Gen2
Warehouse
Lakehouse
Data Pipeline
Report
Semantic Model
Notebook
Copy Data
Shortcuts
Python Script
OneLake
Mirroring
Power Apps

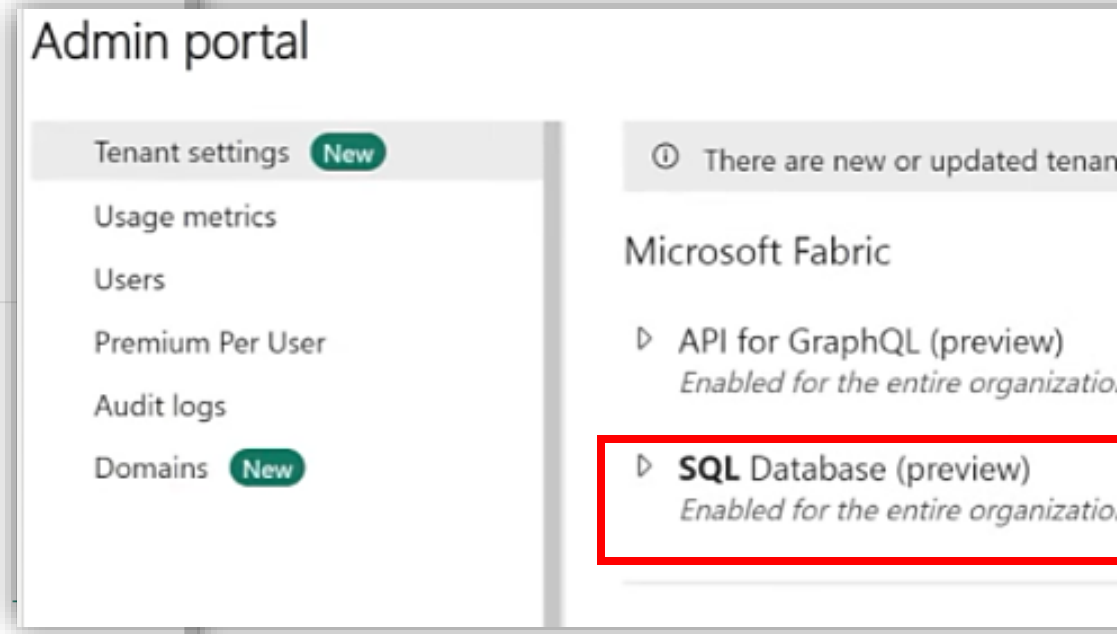
DEMO – SQL Database in Fabric

- Enable SQL Database in Fabric
- 1. Create a SQL DB
- 2. Explore different sections
- 3. Upload/Create data
  - Dataflow
  - Pipeline
- 4. Copilot <- sample query
- 5. API for GraphQL <- notebook (api)
- 6. Connect to an App <- Power App

DEMO – Enab se

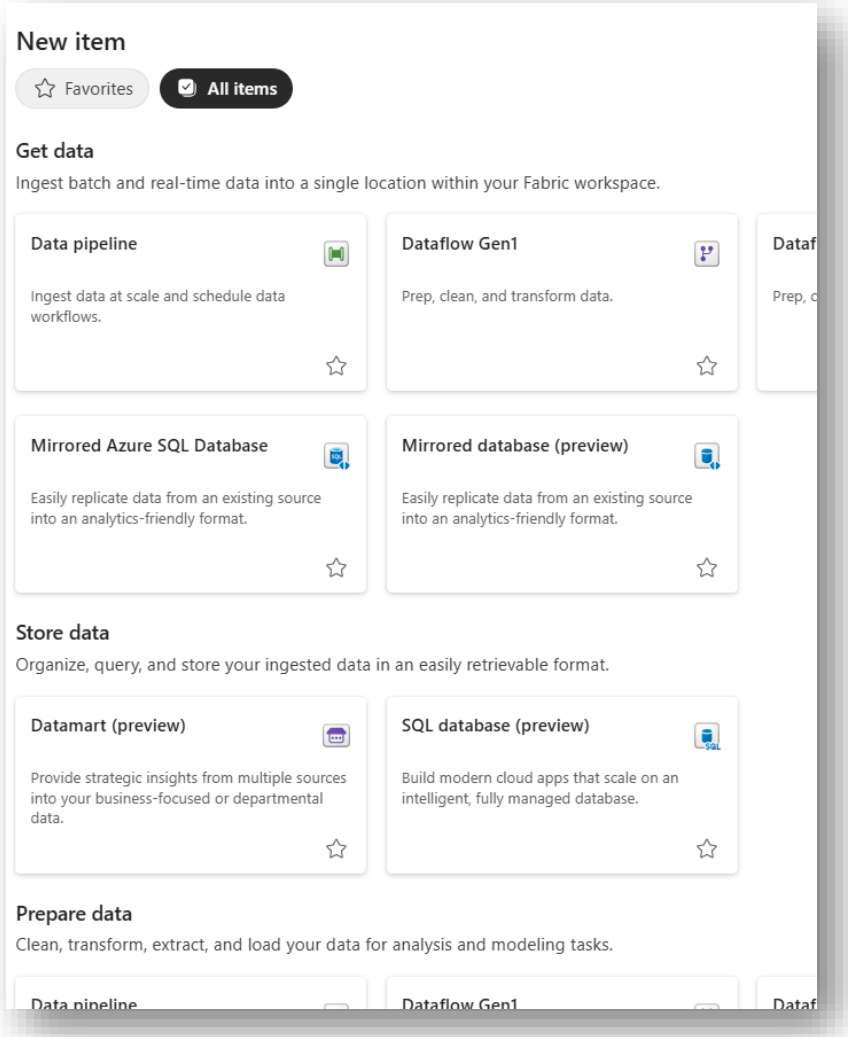
- Admin Portal > tenant settings
  - Enable for your tenant
  - Enable for a capacity
  - Disable

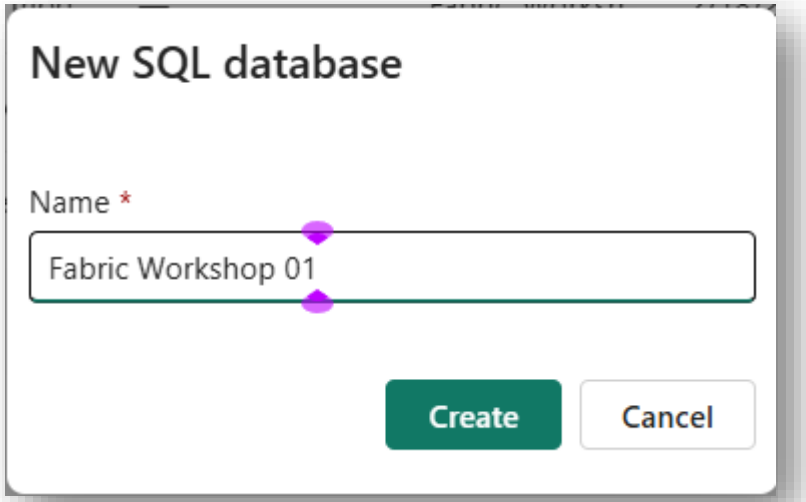




DEMO – Create a SQL Database

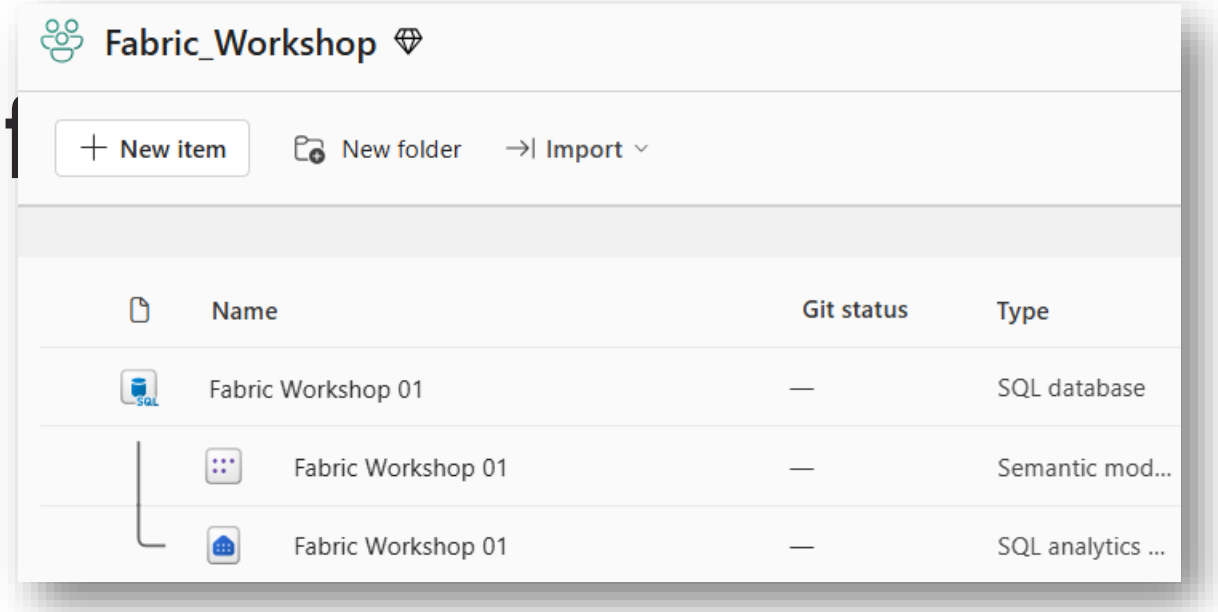
- **1. Create a DB**

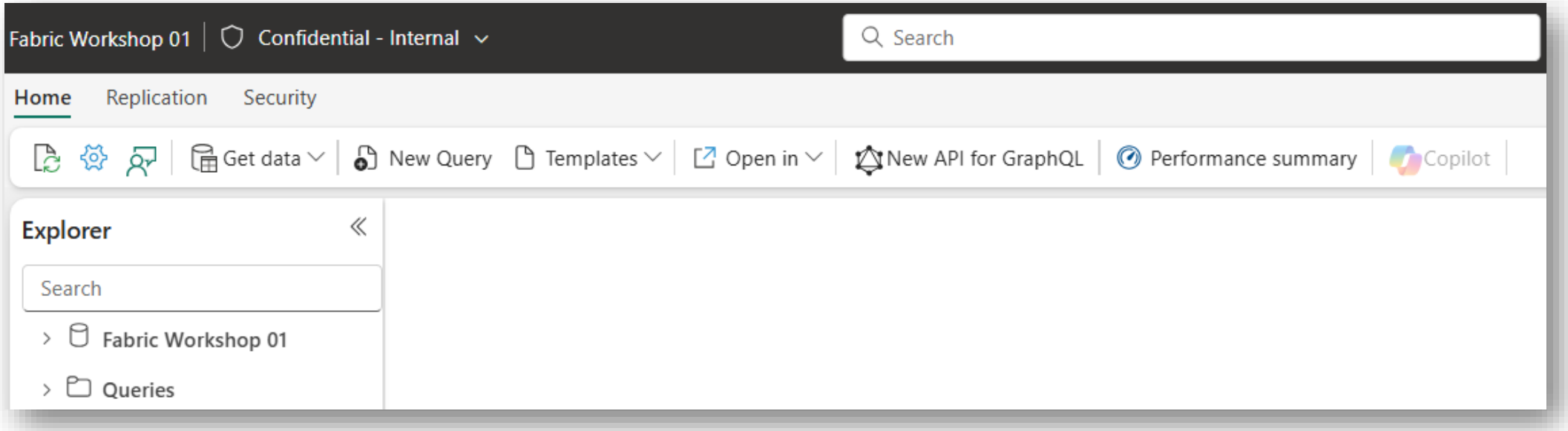




DEMO – Explore di

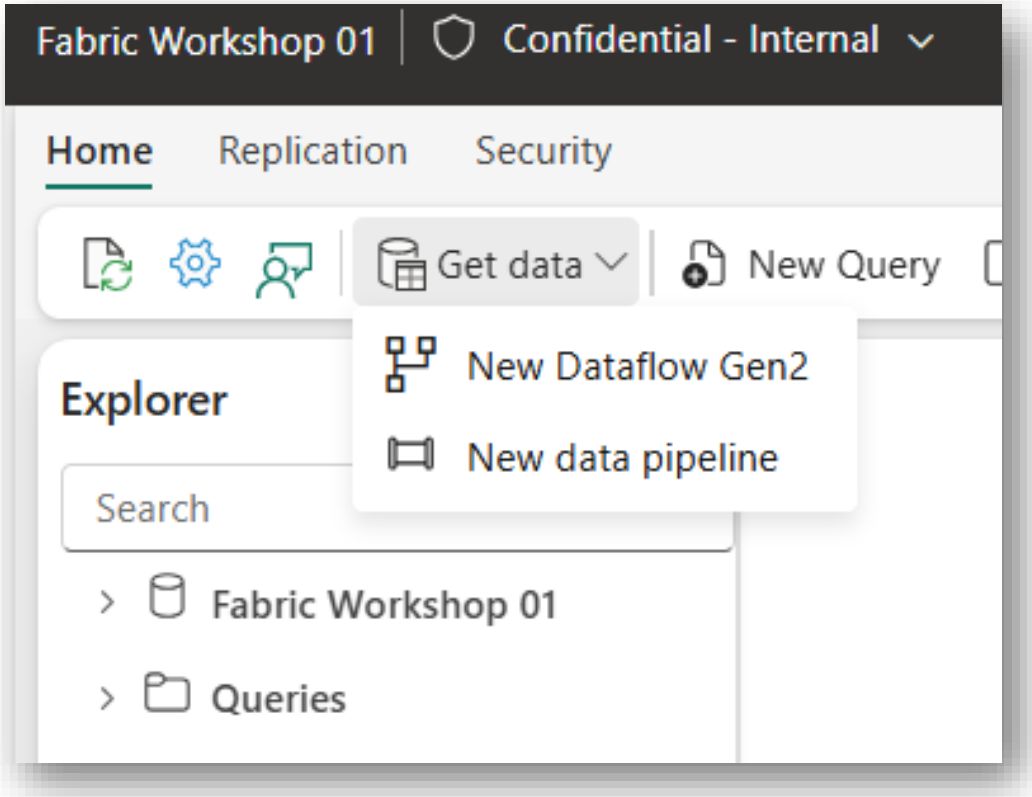
- **2. Explore different sections**





DEMO – Upload data

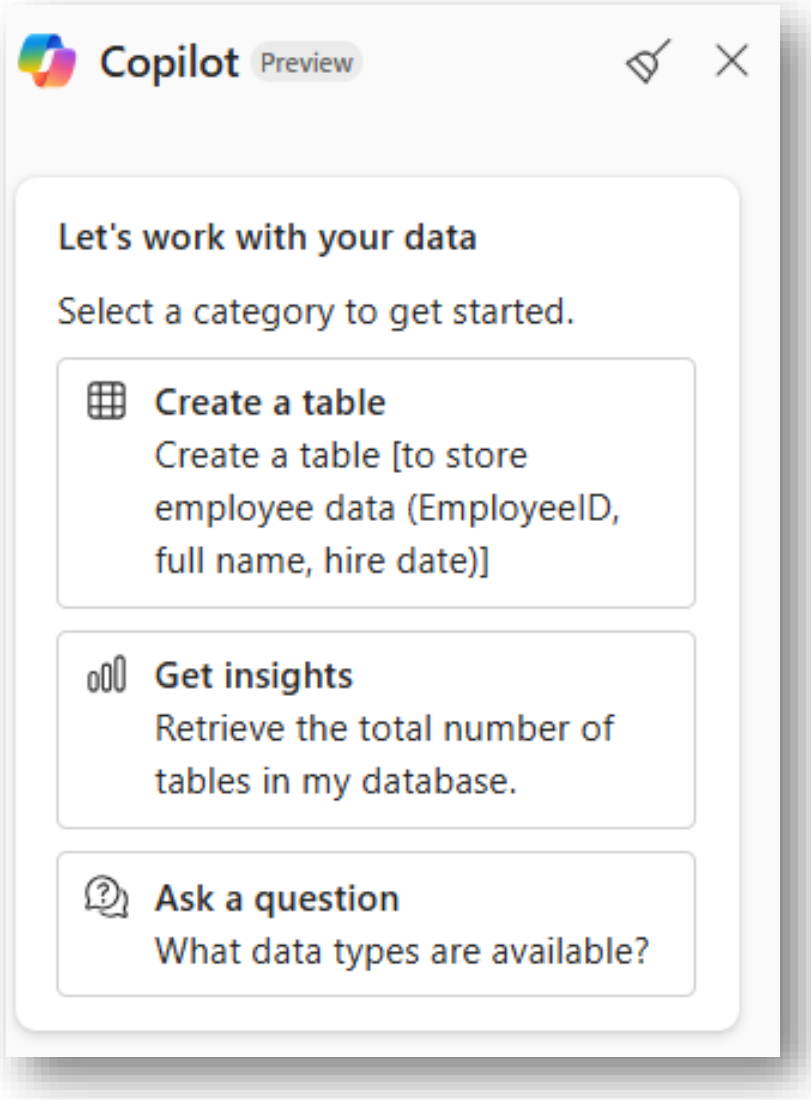
- **3. Upload data**
  - Dataflow Gen2
  - Pipeline



https://docs.google.com/spreadsheets/d/1_J5etWV3L2Or9vk5nkH2ScmhQsQx8FjY/export?format=xlsx

DEMO – Copilot

- **4. Copilot**
  - What is the product with less inventory?
  - What is the product ordered more frequently?
  - Give me the user that placed the more orders
  - Give me the products that the top customer ordered
  - Create a view to show the products with id, name, category_name and price.

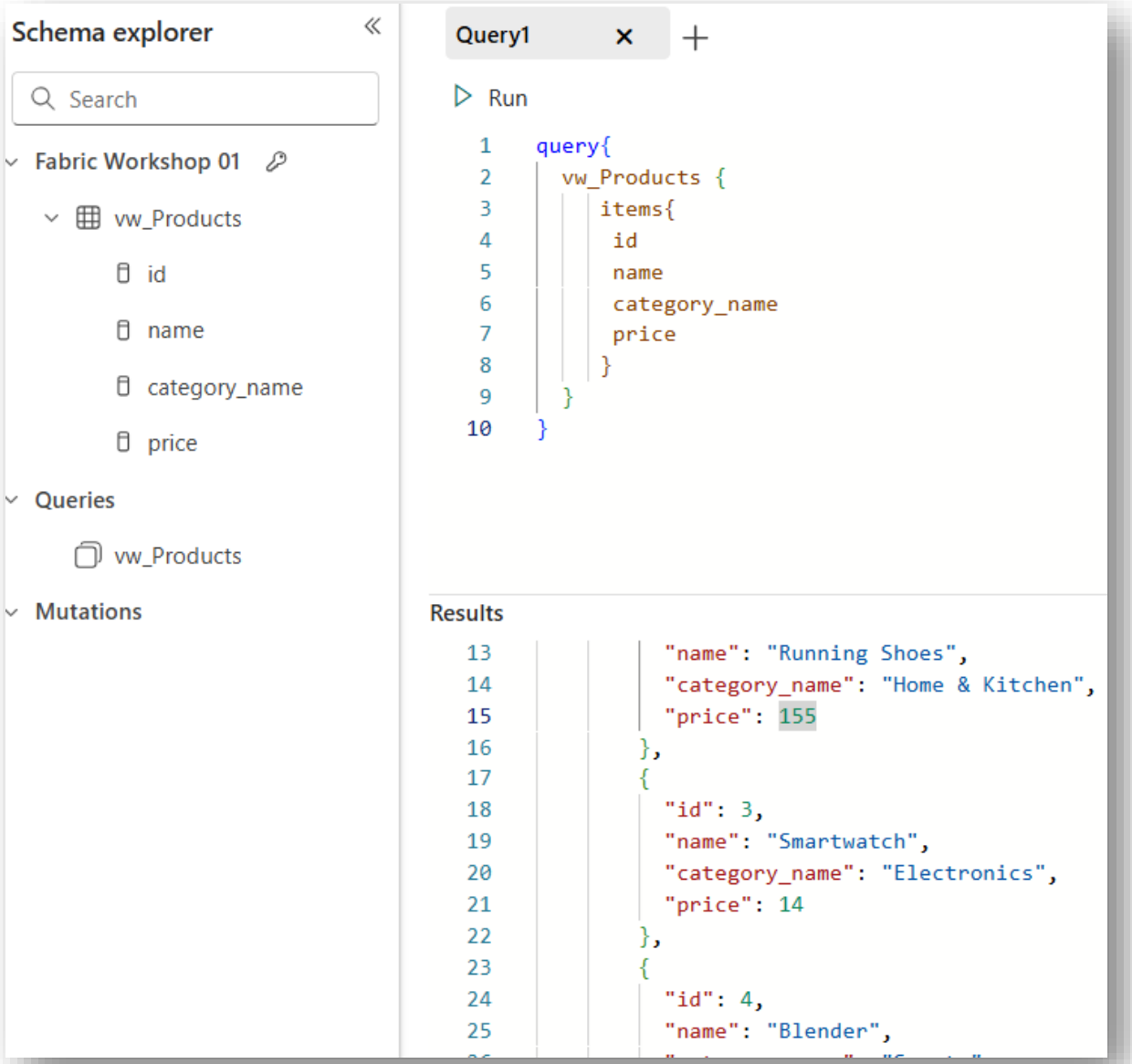


https://docs.google.com/spreadsheets/d/1_J5etWV3L2Or9vk5nkH2ScmhQsQx8FjY/export?format=xlsx

DEMO – API

- **5. API for GraphQL**

```
query{
  vw_Products {
     items{
      id
      name
      category_name
      price
     }
  }
}
```



https://docs.google.com/spreadsheets/d/1_J5etWV3L2Or9vk5nkH2ScmhQsQx8FjY/export?format=xlsx

DEMO

- **5. API for Gr**

```
query{
  vw_Prod
    item
      id
      nam
      cat
      pri
    }
  }
}
```

Generate code

Python

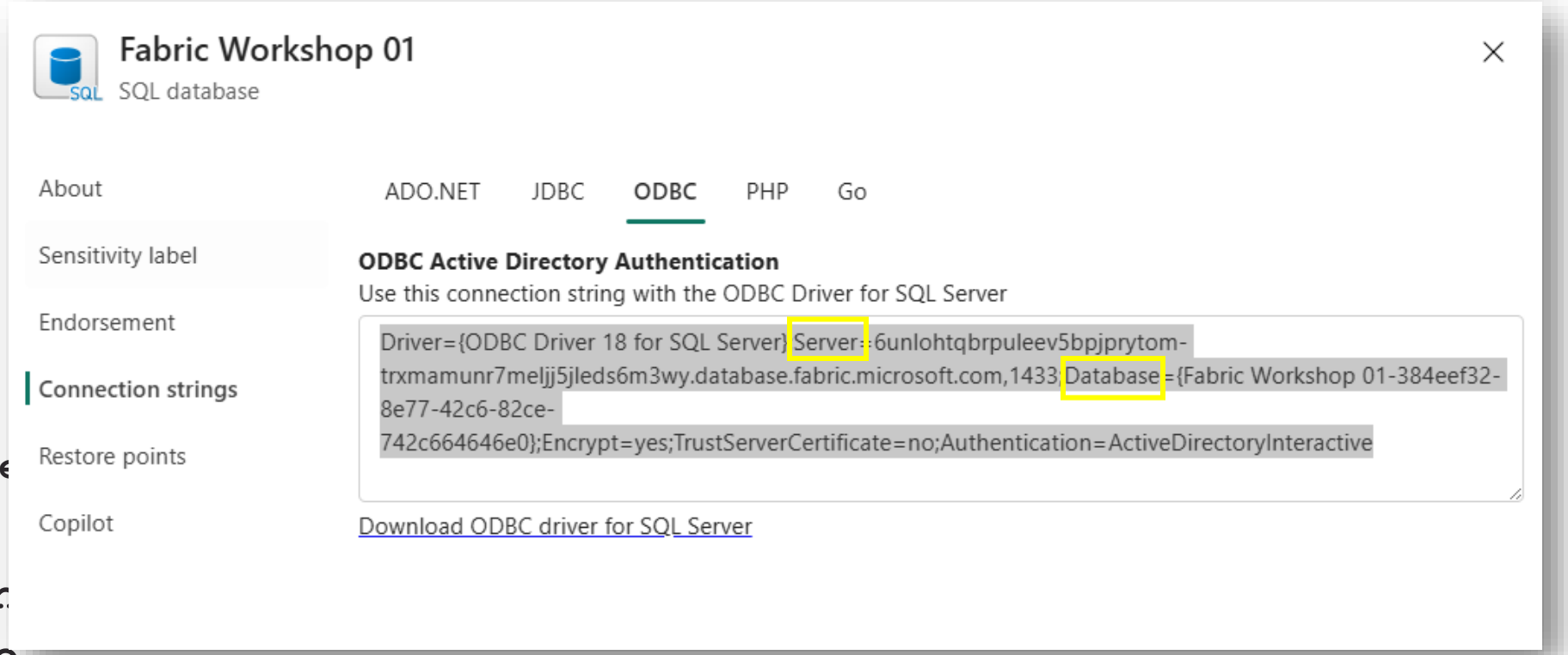
```python
from azure.identity import InteractiveBrowserCredential
import requests
import json

# Acquire a token
# DO NOT USE IN PRODUCTION.
# Below code to acquire token is for development purpose only to test the GraphQL endpoin
# For production, always register an application in a Microsoft Entra ID tenant and use t
# https://learn.microsoft.com/en-us/fabric/data-engineering/connect-apps-api-graphql#crea

app = InteractiveBrowserCredential()
scp = 'https://analysis.windows.net/powerbi/api/user_impersonation'
result = app.get_token(scp)

if not result.token:
    print('Error:', "Could not get access token")

# Prepare headers
headers = {
    'Authorization': f'Bearer {result.token}',
    'Content-Type': 'application/json'
}
```

Copy

https://docs.google.com/spreadsheets/d/1_J5etWV3L2Or9vk5nkH2ScmhQsQx8FjY/export?format=xlsx

DEMO – Connect to Apps

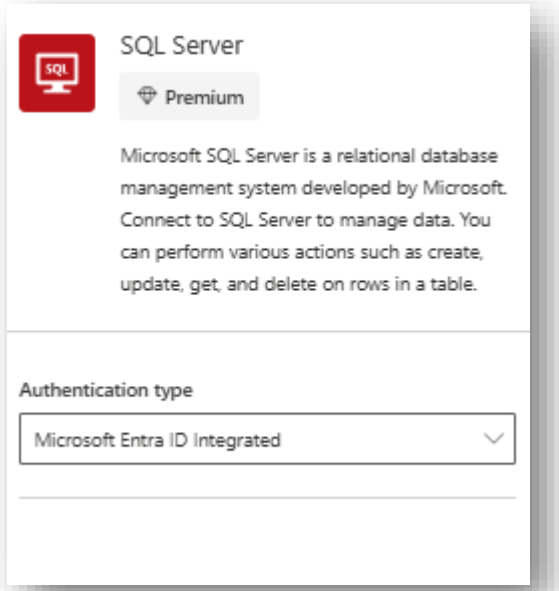
- **6. Connect**
- **Power App**
  - **Connection**
    - **Conn str:**
      - **Server**
      - **Database**
  - **Read/Write**
    - **Permissions**
    - **Use in App**

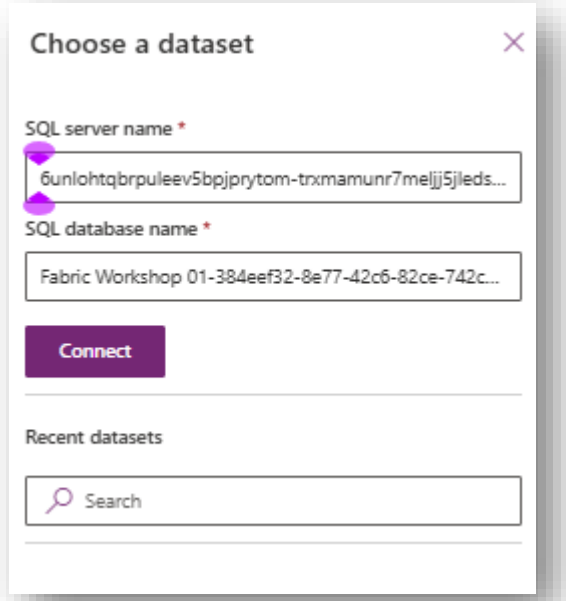


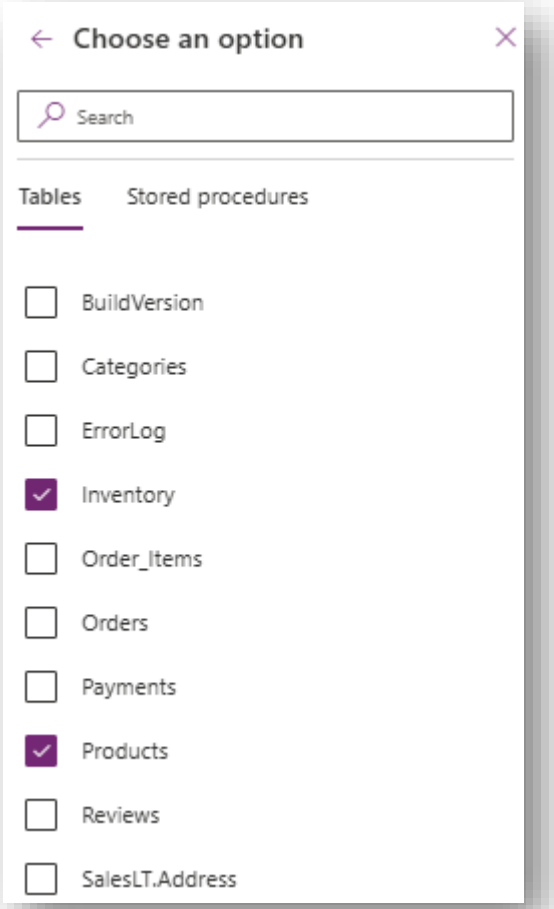
https://docs.google.com/spreadsheets/d/1_J5etWV3L2Or9vk5nkH2ScmhQsQx8FjY/export?format=xlsx

DEMO – Connect to Apps

- **6. Connect**
- **Power App**
  - **Connection**
    - **Conn str:**
      - **Server**
      - **Database**
  - **Read/Write**
    - **Permission**
    - **Use in App**





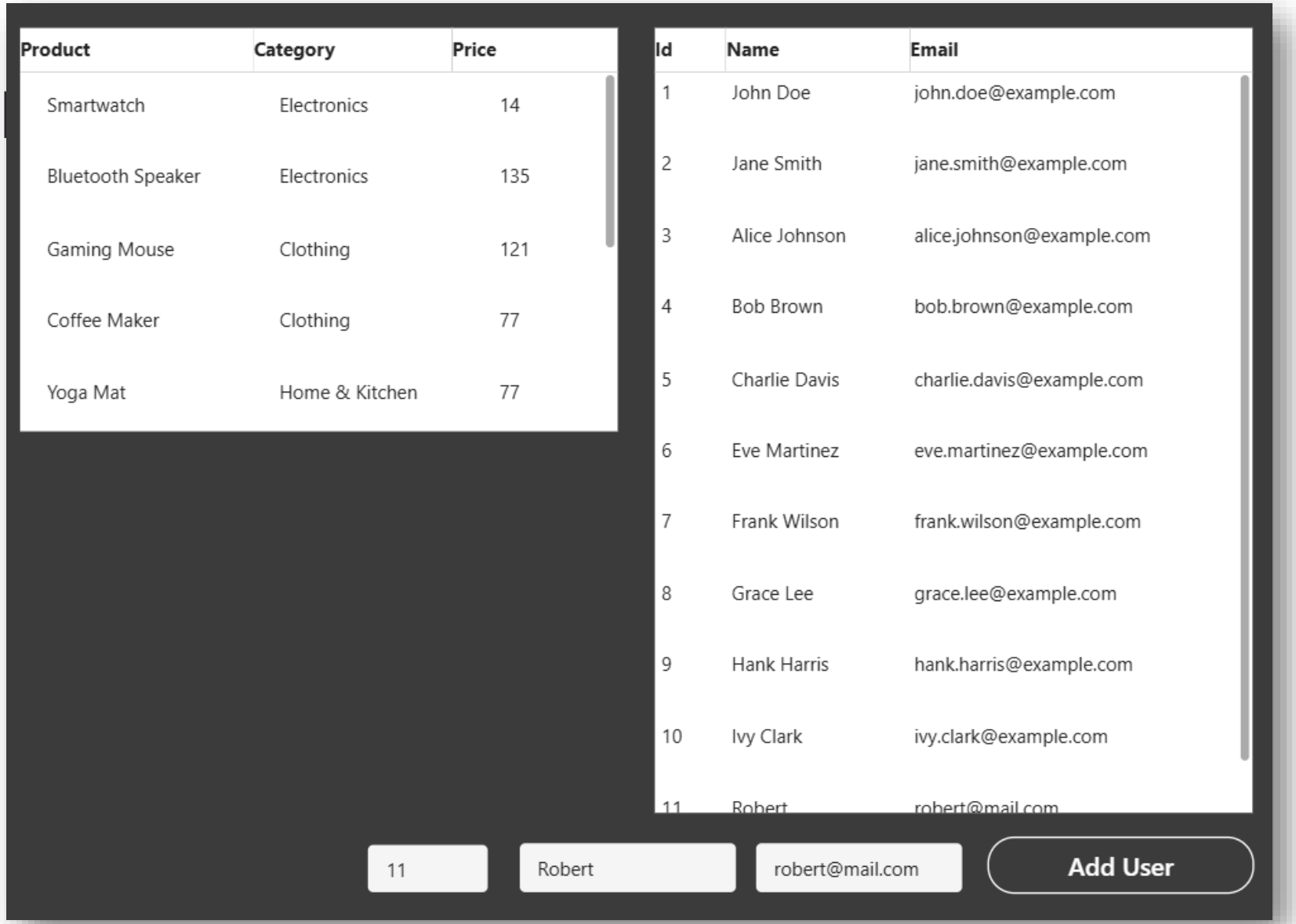



```
CREATE USER [myemail@mail.com] FROM EXTERNAL PROVIDER;
ALTER ROLE db_datareader ADD MEMBER [myemail@mail.com];   -- Read Access
ALTER ROLE db_datawriter ADD MEMBER [myemail@mail.com];   -- Insert/Update Access
```

https://docs.google.com/spreadsheets/d/1_J5etWV3L2Or9vk5nkH2ScmhQsQx8FjY/export?format=xlsx

DEMO – Con

- **6. Connect**
- **Power App**
  - **Connection**
    - **Conn str:**
      - **Server**
      - **Database**
  - **Read/Write**
    - **Permission**
    - **Use in App**



https://docs.google.com/spreadsheets/d/1_J5etWV3L2Or9vk5nkH2ScmhQsQx8FjY/export?format=xlsx

References

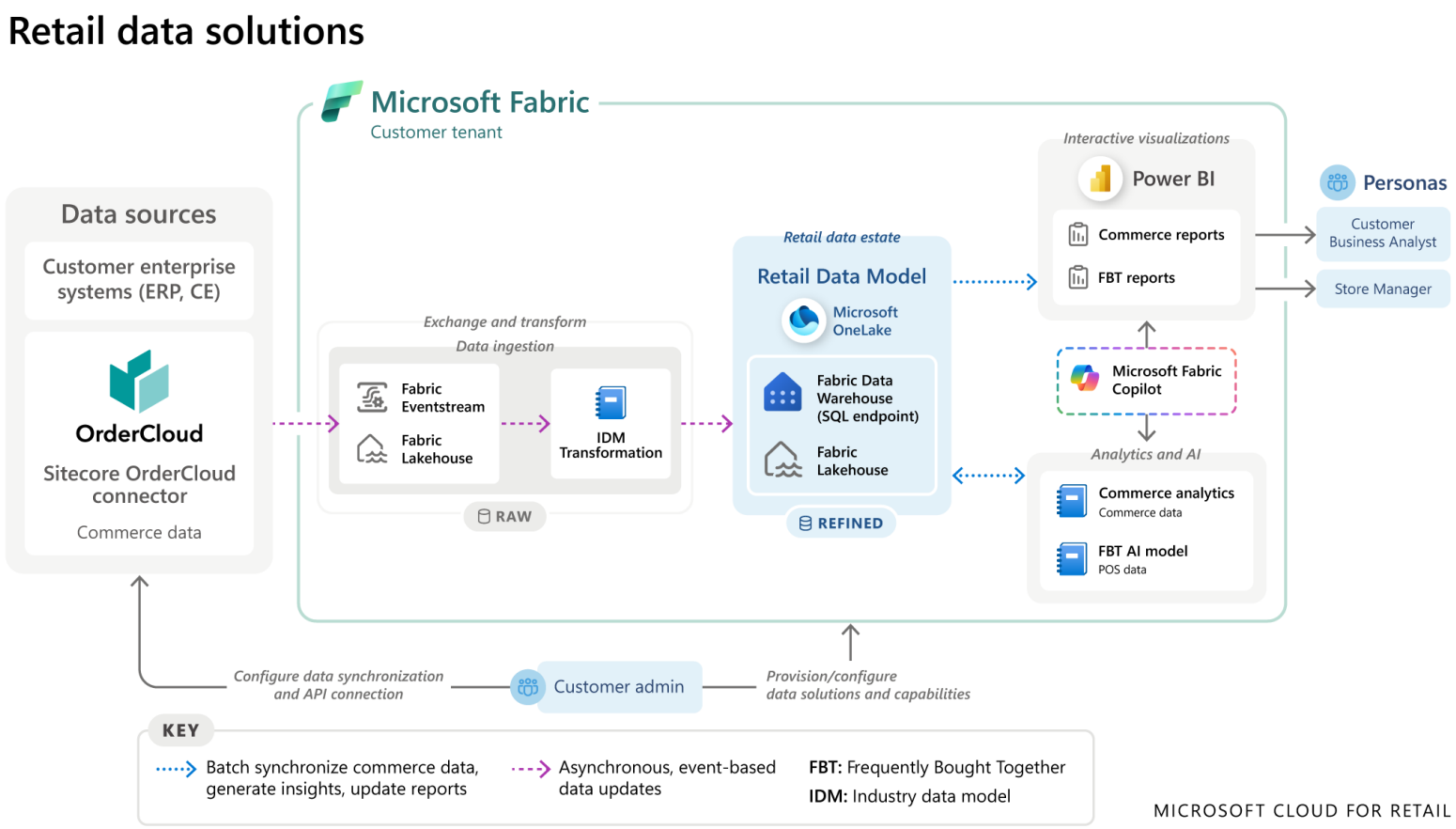
- SQL Database in Fabric
- API for GraphQL

Industry Solutions

Industry Solutions – Intro

- Industry Data Models
  - Standardization: A consistent structure for storing and organizing retail data. This simplifies data management and analysis.
  - Efficiency: Simplify data input and minimize errors by providing preset fields and formats for retail data.
  - Data integrity: Apply data validation rules and quality checks to ensure accurate reporting and analysis.
  - Consolidation: Make data easier to work with by consolidating data from different sources like POS, ecommerce, etc.
  - Analysis and reporting: Well structured data lead to high quality insights and timely decision-making.
  - Data governance: Ensure compliance, security, and privacy for data visibility.
- Pre-built AI/ML Retail Models

Retail Data Solutions - Example



Data Science with Fabric

Roberto Hernández
roberto.Hernandez@ventagium.com
https://www.linkedin.com/in/robertohj/

Data Science in Fabric

Microsoft Fabric offers Data Science experiences to empower users to complete end-to-end data science workflows for the purpose of data enrichment and business insights. You can complete a wide range of activities across the entire data science process, all the way from data exploration, preparation and cleansing to experimentation, modeling, model scoring and serving of predictive insights to BI reports.

New

Current workspace: my_workspace

Items will be saved to this workspace.

ML model

Experiment

Notebook

Environment (Preview)

Import notebook

Use a sample

Recommended

Getting Started with Data Science

Getting Started with ML Models

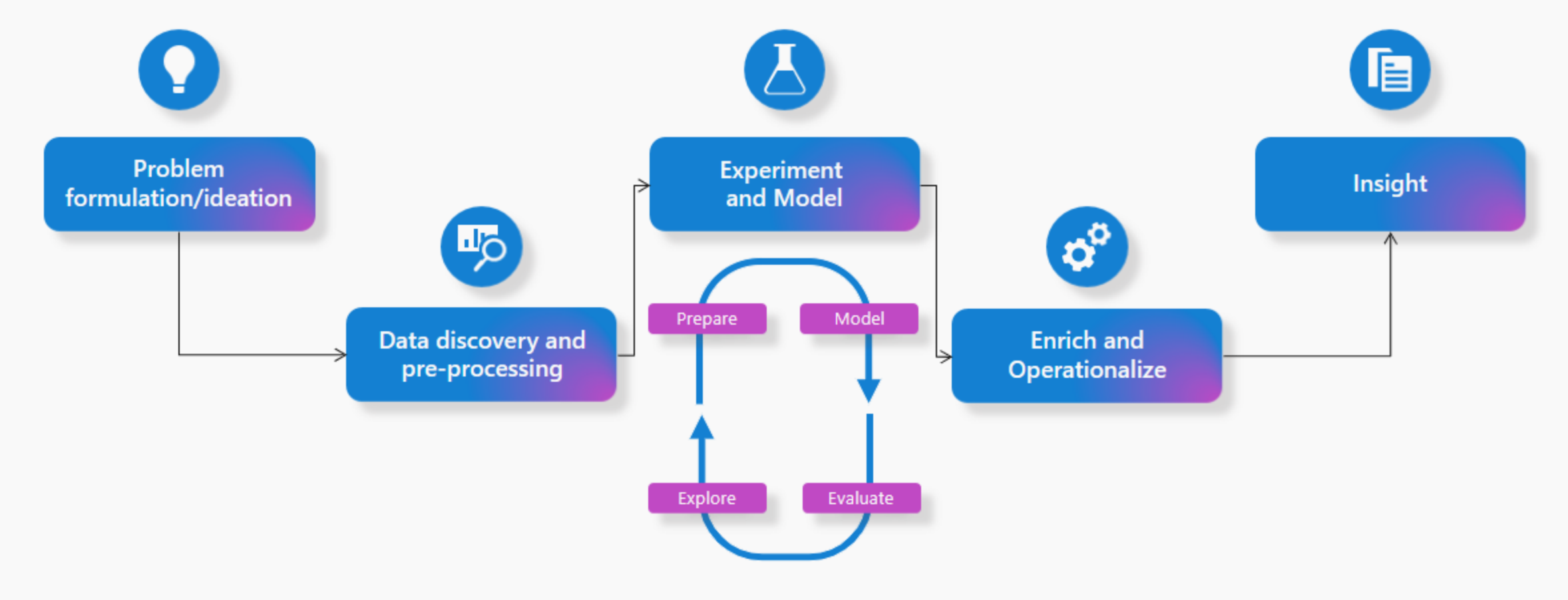
Getting Started with Data Science

Getting Started with ML Experiments

Getting Started with Data Science

Getting Started with Notebooks

The Data Science Process



Data Wrangler

- A tool in Notebooks that prepares data and generates Python code
- Useful for quick data exploration and cleansing

Machine Learning experiments

- The primary unit of organization and control for all related machine learning runsUseful for quick data exploration and cleansing
- Actions:
  - Log parameters, code versions, metrics, and output files when running their machine learning code.
  - Visualize, search for, and compare runs
  - Download run files and metadata for analysis in other tools.

References

- Data Science in Fabric
- Data Wrangler
- Machine Learning Experiments
- MLFlow

Open AI + Fabric

Overview of Integration

- APIs
- Azure OpenAI Service
- Fabric's Capabilities (Spark notebooks,Data Factory Pipelines)

Key Use Cases

- Data Summarization
- Text Classification
- Chatbot Assistance for Insights
- Automated Data Cleaning

Fabric + Open AI Key Considerations

API Rate Limits & Pricing

Data Security & Governance

Model	Input	Cached input	Output
gpt-4o ↳ gpt-4o-2024-08-06	$2.50	$1.25	$10.00
gpt-4o-audio-preview ↳ gpt-4o-audio-preview-2024-12-17	$2.50	-	$10.00
gpt-4o-realtime-preview ↳ gpt-4o-realtime-preview-2024-12-17	$5.00	$2.50	$20.00
gpt-4o-mini ↳ gpt-4o-mini-2024-07-18	$0.15	$0.075	$0.60
gpt-4o-mini-audio-preview ↳ gpt-4o-mini-audio-preview-2024-12-17	$0.15	-	$0.60
gpt-4o-mini-realtime-preview ↳ gpt-4o-mini-realtime-preview-2024-12-17	$0.60	$0.30	$2.40
o1 ↳ o1-2024-12-17	$15.00	$7.50	$60.00
o3-mini ↳ o3-mini-2025-01-31	$1.10	$0.55	$4.40
o1-mini ↳ o1-mini-2024-09-12	$1.10	$0.55	$4.40

Case Study: Home Delivery with Open AI

- Context:
  - Thousands of reviews with comments.
  - 5 Carriers
  - 100 Hubs
- Problem:
  - Providing Actionable Insights from Reviews

Fabric + Open AI Demo