



From Source data to OneLake



Martin Catherall

Data Analytics Consultant

 @MartyCatherall

www.MartinCatherall.com

- Data Architect

- SQL Server

- Since 2002

- Microsoft Azure

- Since 2015

- Microsoft Fabric

- Since 2022

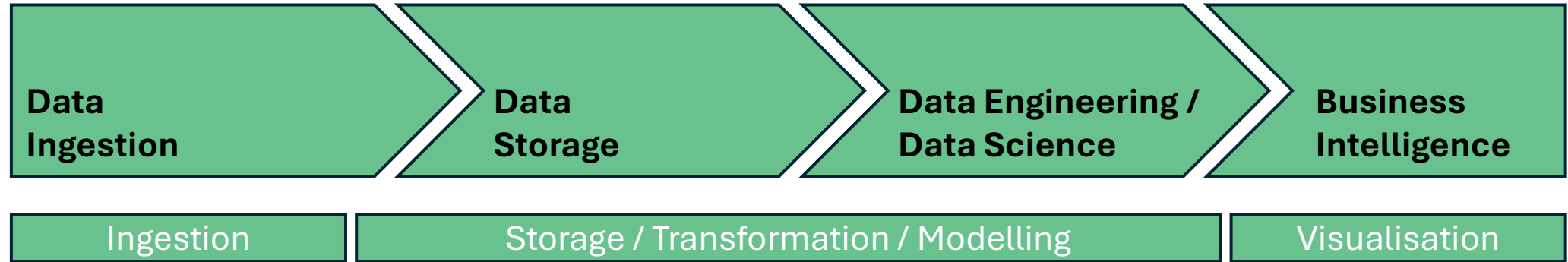


Agenda and content

- Brief Intro
- Data in One-Lake
- A number of ways to get data into Fabric
- Summary

The Data Journey.

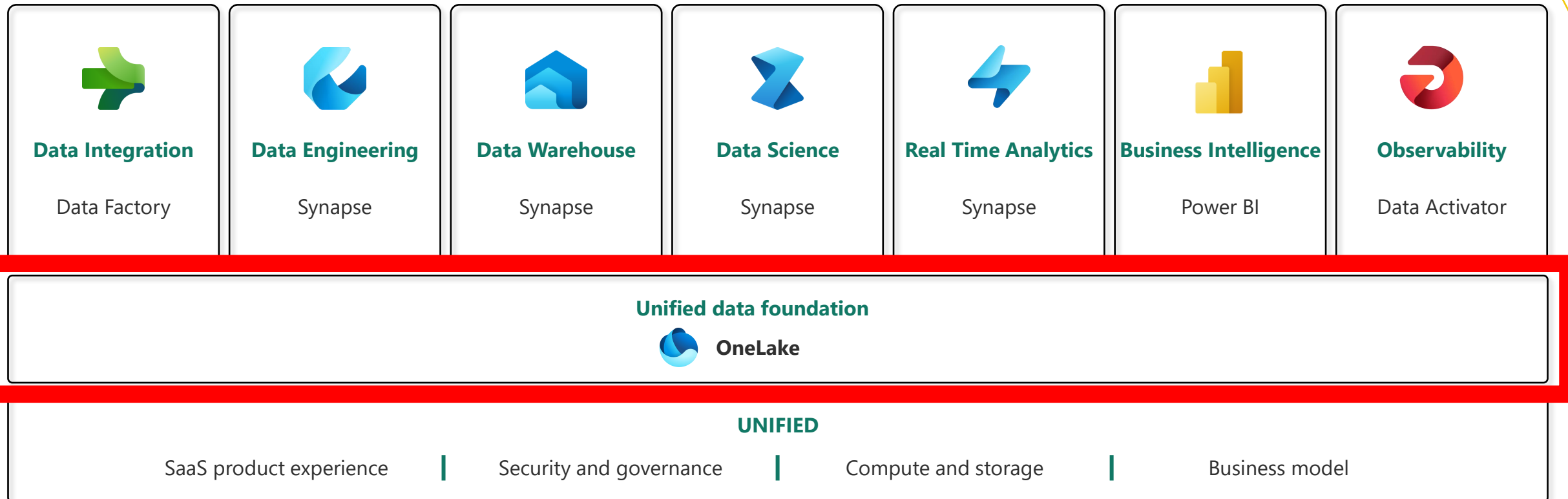
From Source to Visualisation.



- Excel (Power Pivot)
- Power BI
- Azure (Synapse)
- Microsoft Fabric

Microsoft Fabric does it all—in a unified solution

An end-to-end analytics platform that brings together all the data and analytics tools that organizations need to go from the data lake to the business user



One Lake - Acquiring Data

- | | | | |
|---|--------------------|----|---------------|
| 1 | Manually | 9 | SQL on Fabric |
| 2 | Data Integration | 10 | Warehouse |
| | - Notebooks | | COPY |
| | | | OPENROWSET |
| | | | T-SQL (CTAS) |
| 3 | Data Engineering | 11 | Mount ADF |
| | - (Data) Pipelines | | |
| | - Dataflow Gen 2 | | |
| 4 | Mirroring | | |
| 5 | Shortcuts | | |
| 6 | Semantic Model | | |
| 7 | Programmatically | | |
| 8 | Real Time | | |

One Lake - Acquiring Data

We'll need a place to store the data

Lakehouse

Warehouse

EventHouse

SQL

So, where do we put it? What do we need?

One Lake - Acquiring Data

Workspace and capacity

Workspace settings

General

License info

Azure connections

System storage

Git integration

OneLake

Workspace identity

Network security

Monitoring

Power BI

Delegated Settings

Data Engineering/Science

Data Factory

License info

Choose a license for this workspace.

License Configuration

Current license

Fabric capacity

License capacity

acapacity

SKU: F2, Region: Australia East

Semantic model storage format

Small semantic model storage format

Edit

Connection link

Use this link to connect third-party software to the workspace. Copy the link and add it to your third-party software.

powerbi://api.powerbi.com/v1.0/myorg/DataPlat01

Home >

acapacity

Fabric Capacity

Search

Pause

Refresh

Move

Delete

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Resource visualizer

Scale

Essentials

Resource group (move) : Capacities

Status : Active

Location : Australia East

Subscription (move) : Microsoft Azure Sponsorship (MVP)

Subscription ID : 2635e25c-2d0f-4b09-ad52-1a62452a94e9

Tags (edit) : Add tags

Resource name : acapacity

SKU : F2

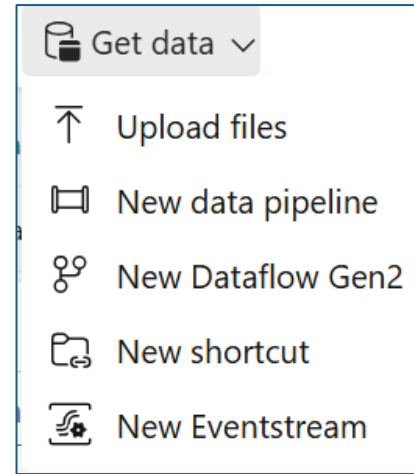
One Lake - Acquiring Data

1 Manually

Get data

OneLake File Explorer

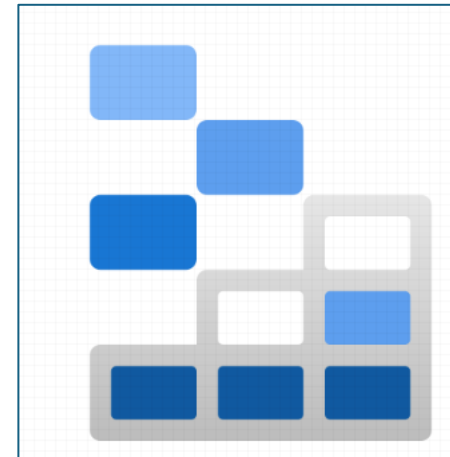
Azure Storage Explorer



OneLake File Explorer

The OneLake file explorer application seamlessly integrates OneLake with Windows File Explorer

[Download OneLake app](#) 



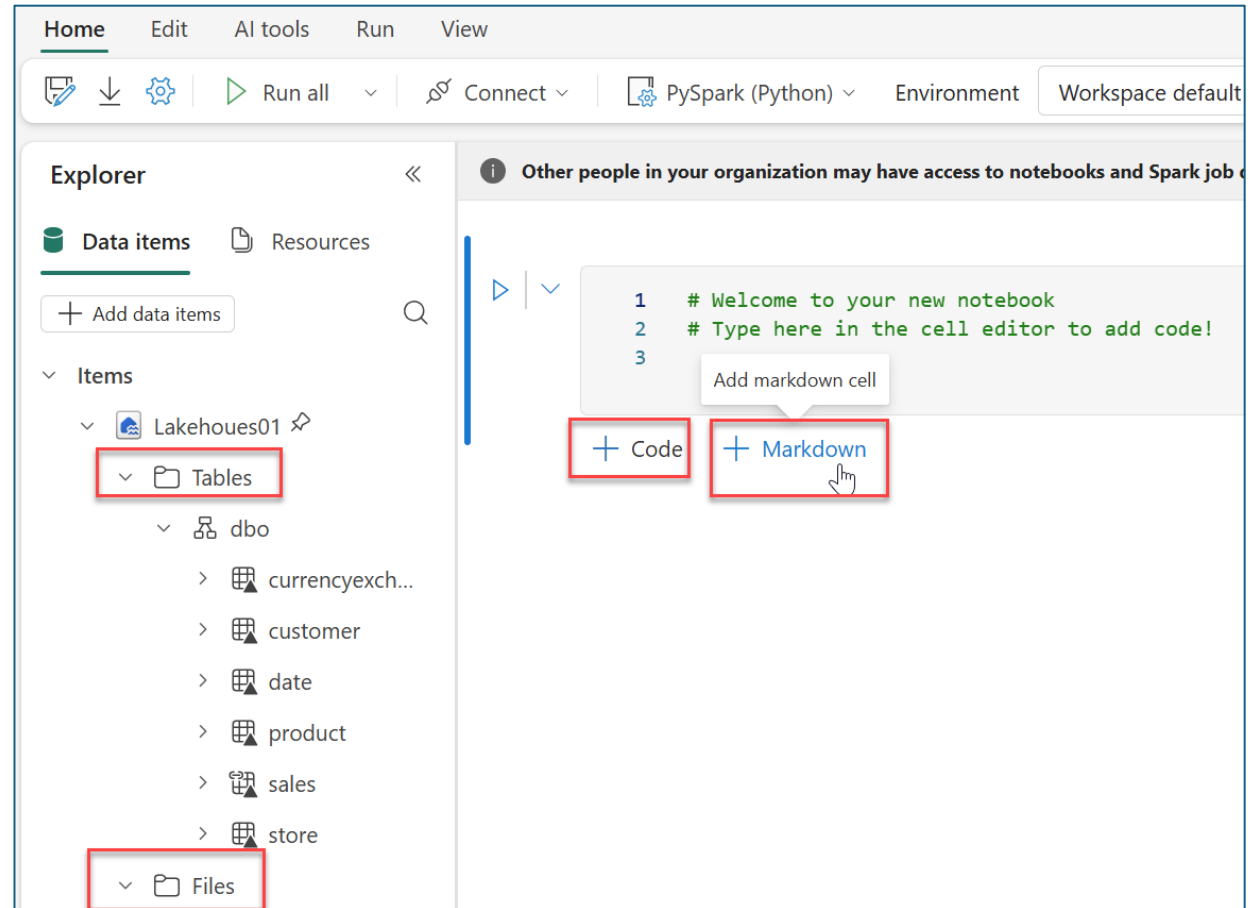
One Lake - Acquiring Data

2 Data Engineering (Notebooks)

Get Data

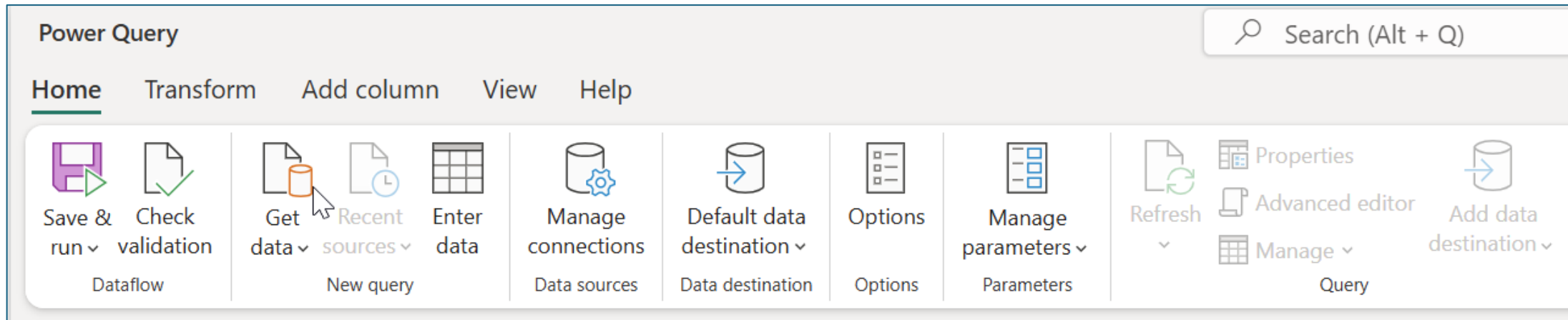
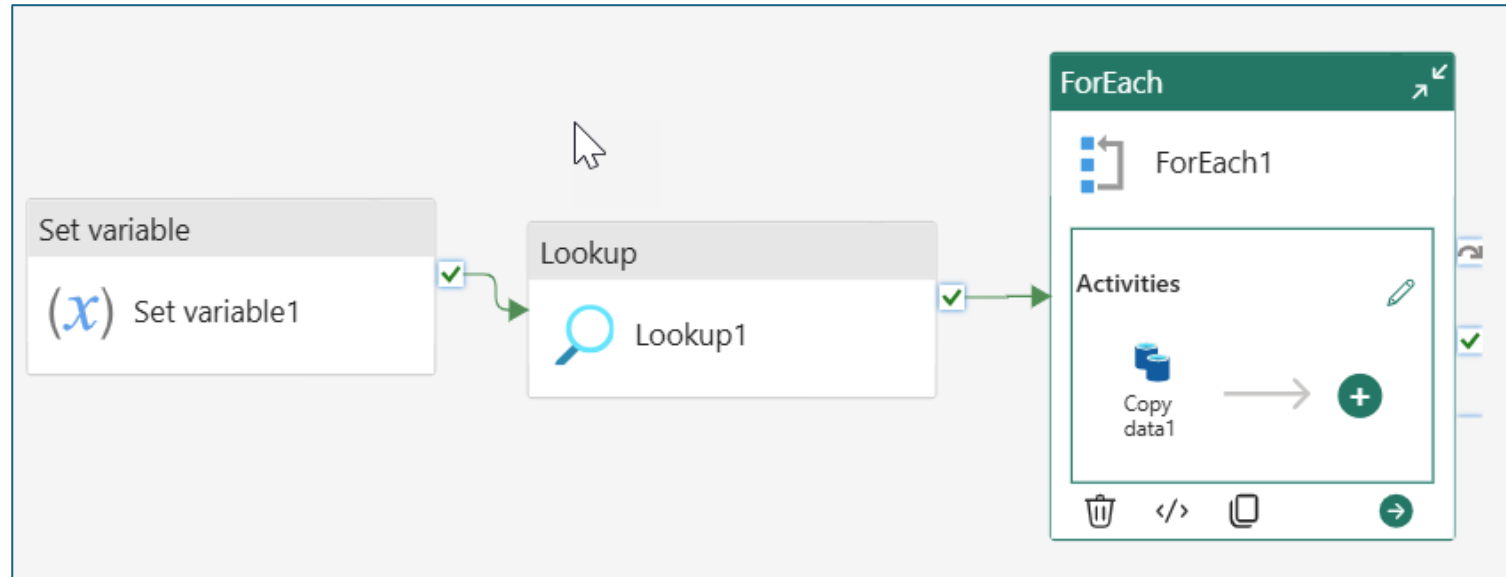
Prepare data

Analyse and Train Data



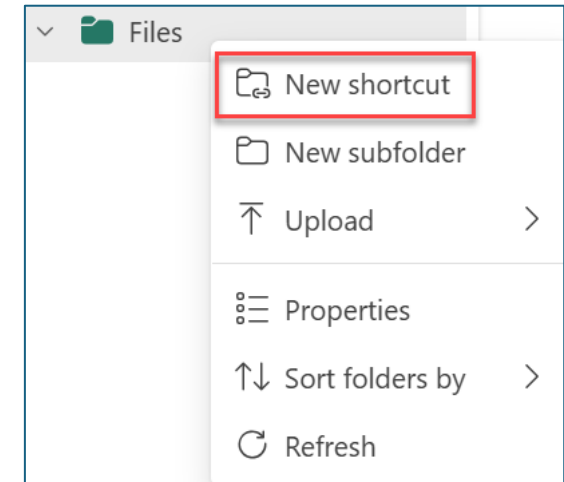
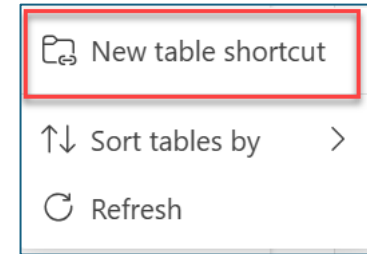
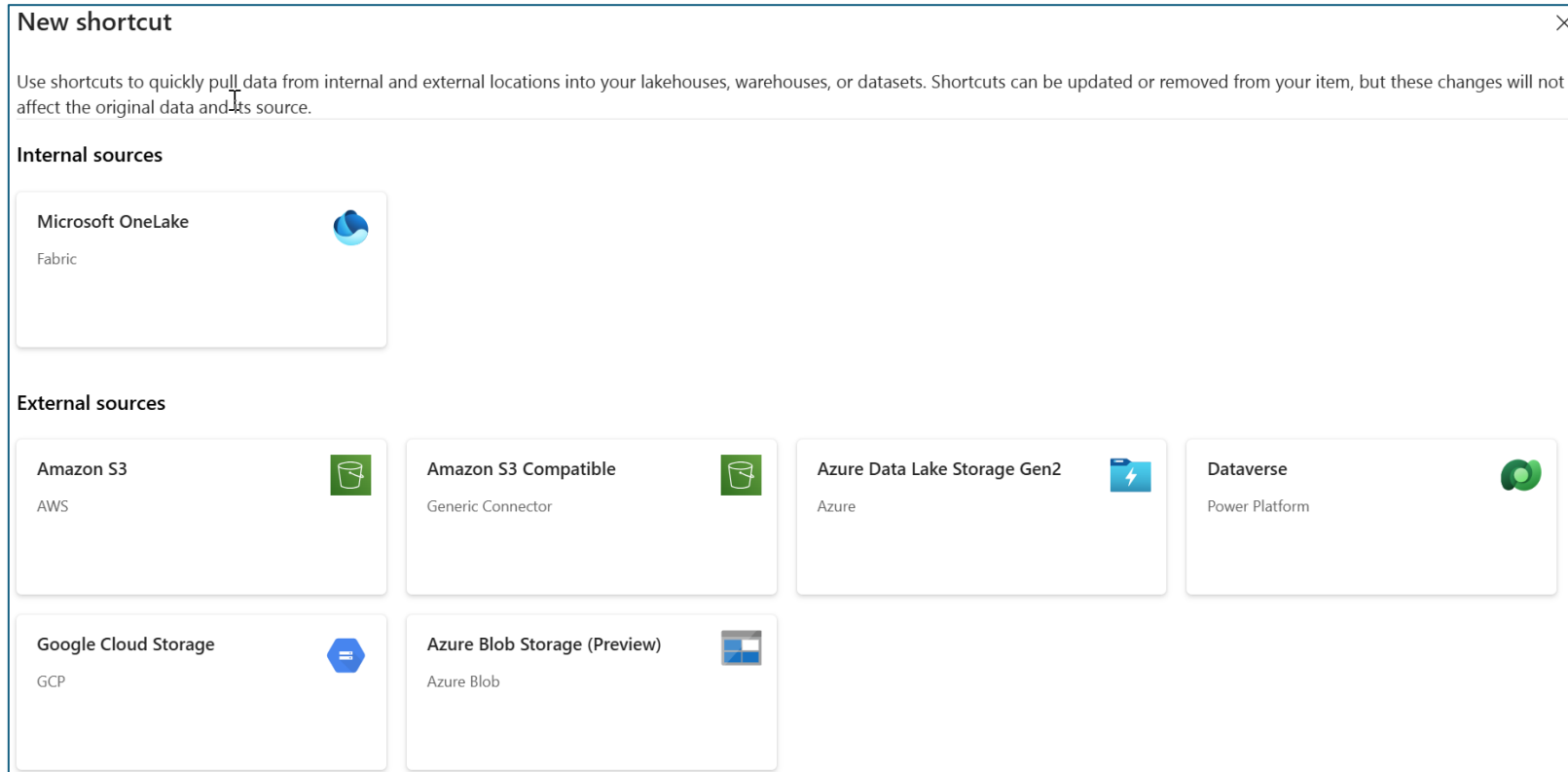
One Lake - Acquiring Data

3 Data Integration



One Lake - Acquiring Data

4 Shortcuts



One Lake - Acquiring Data

5 Mirroring

Get data

Ingest batch and real-time data into a single location within your Fabric workspace.

Mirrored Azure Cosmos DB (pre...



Easily replicate data from an existing source into an analytics-friendly format.



Mirrored Azure Database for Po...



Easily replicate data from an existing source into an analytics-friendly format.



Mirrored Azure Databricks catalog



Explore Unity Catalog Tables



Mirrored Azure SQL Database



Easily replicate data from an existing source into an analytics-friendly format.



Mirrored Azure SQL Managed In...



Easily replicate data from an existing source into an analytics-friendly format.



Mirrored database



Easily replicate data from an existing source into an analytics-friendly format.



Mirrored Snowflake



Easily replicate data from an existing source into an analytics-friendly format.



Mirrored SQL Server (preview)

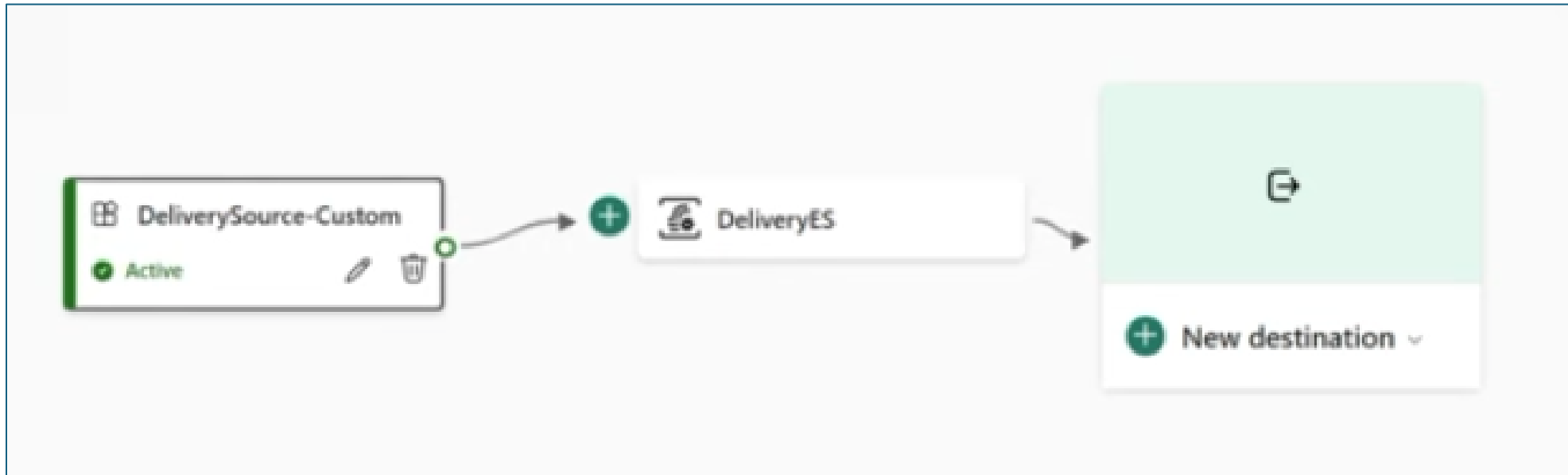


Easily replicate data from an existing source into an analytics-friendly format.



One Lake - Acquiring Data

6 Real Time



One Lake - Acquiring Data

7 Semantic Model

Upload an existing Semantic Model

Connect to it – example using a notebook

Acquire data

[Read data from semantic models and write data that semantic models can consume using Spark - Microsoft Fabric | Microsoft Learn](#)

One Lake - Acquiring Data

8 **Programmatically**

ReST API's

Various API's

Fabric CLI

One Lake - Acquiring Data

9 **SQL on Fabric**

Azure SQL Database inside Fabric

2 End Points

Database – Read & Write

Database Warehouse – Read Only

One Lake - Acquiring Data

10 Warehouse - TSQL

COPY

```
COPY INTO [database].[schema].[table]
FROM 'https://...'
WITH (
    FILE_TYPE = 'CSV',
    FIRSTROW = 2
);
```

OPENROWSET

```
SELECT *
FROM OPENROWSET(BULK
    'https://...',
    FORMAT='CSV',
    HEADER_ROW=True,
    ROW_TERMINATOR='\n',
    FIELD_TERMINATOR=',') AS data;
```

CTAS

```
CREATE TABLE [database].[schema].[table]
AS
SELECT DATEPART(YEAR, updated) AS [year],
    DATEPART(MONTH, updated) AS [month],
    DATEPART(DAY, updated) AS [dayofmonth],
    *
FROM dbo.[table];
```

One Lake - Acquiring Data

- | | | | |
|---|--------------------|----|---------------|
| 1 | Manually | 9 | SQL on Fabric |
| 2 | Data Integration | 10 | Warehouse |
| | - Notebooks | | COPY |
| 3 | Data Engineering | | OPENROWSET |
| | - (Data) Pipelines | | T-SQL (CTAS) |
| | - Dataflow Gen 2 | 11 | Mount ADF |
| 4 | Mirroring | | |
| 5 | Shortcuts | | |
| 6 | Semantic Model | | |
| 7 | Programmatically | | |
| 8 | Real Time | | |

Thank You