



Module 1: Data Factory Fundamentals

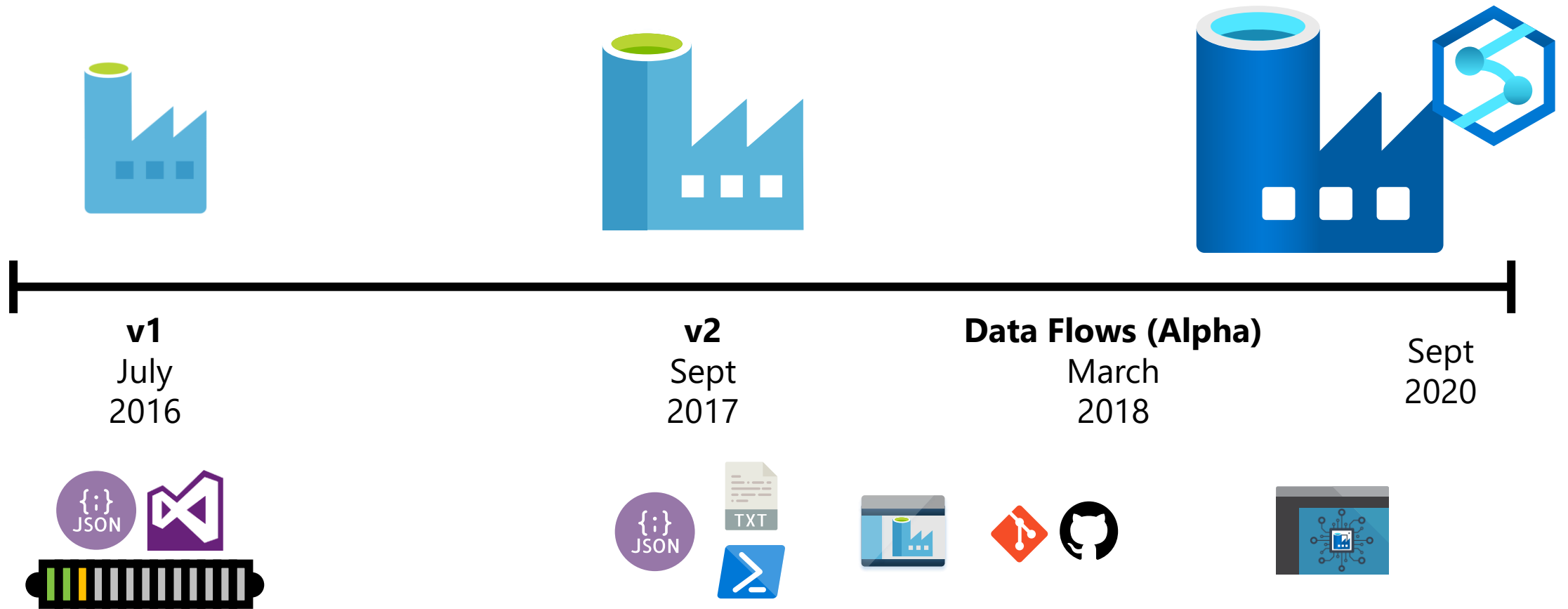
- 🔊 What is it and why use it?
- 🔊 Resource Components
- 🔊 Common Activities
- 🔊 Execution Dependencies

Azure Data Factory –

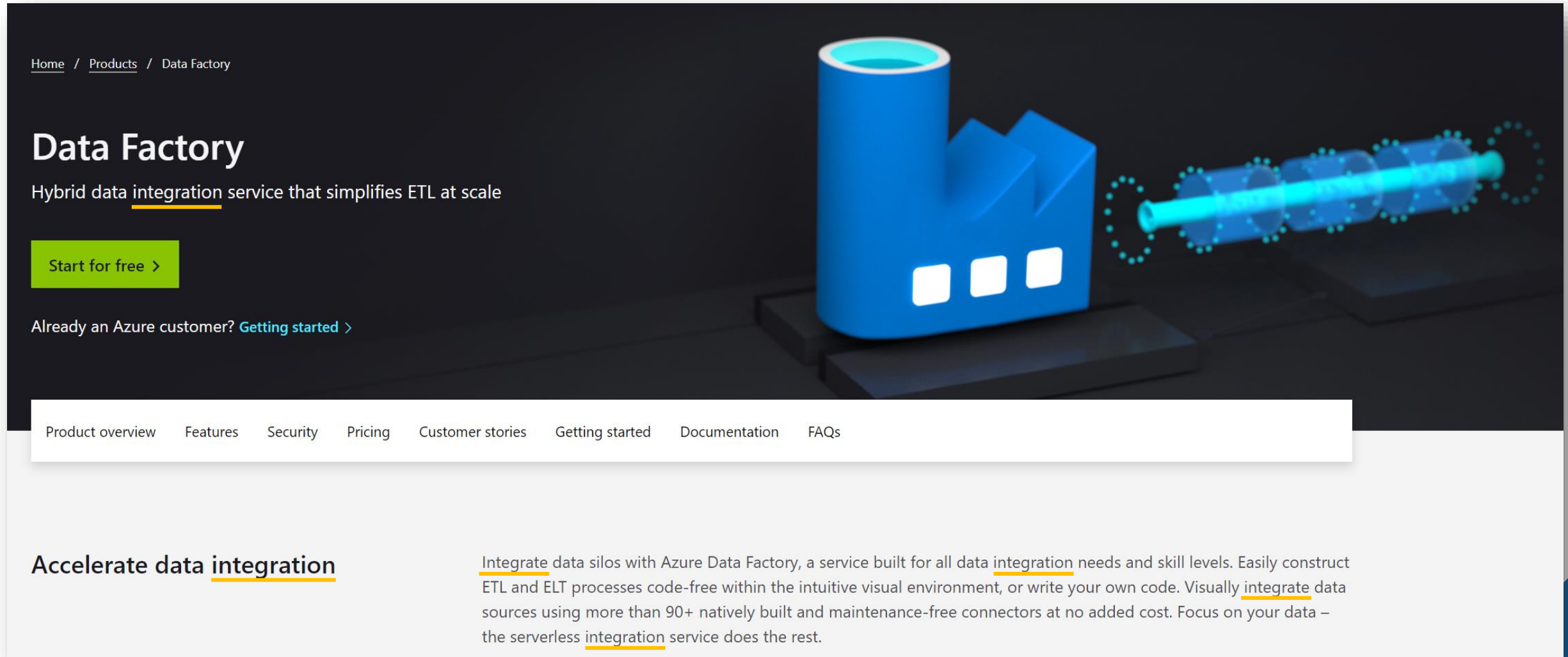
What is it?
Why use it?



A Quick History Lesson



What is Azure Data Factory (ADF)?

The image is a screenshot of the Azure Data Factory (ADF) landing page. The background is dark with a blue 3D illustration of a factory building on the left and a glowing blue data pipeline with circular nodes extending to the right. The page layout includes a top navigation bar, a main header section with a title and description, a green 'Start for free' button, a link for existing customers, a horizontal menu of product links, and a main content area with a heading and descriptive text.

[Home](#) / [Products](#) / [Data Factory](#)

Data Factory

Hybrid data integration service that simplifies ETL at scale

[Start for free >](#)

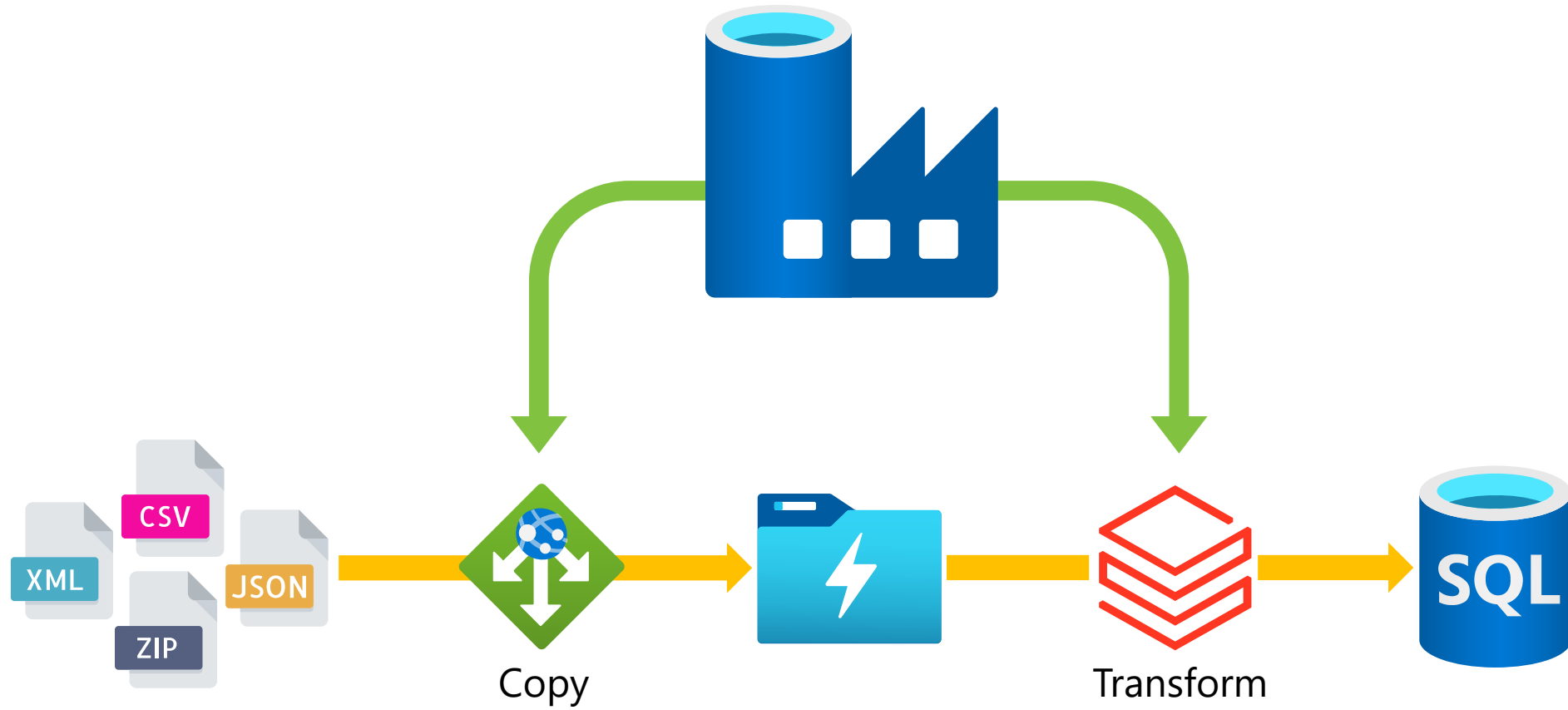
Already an Azure customer? [Getting started >](#)

[Product overview](#) [Features](#) [Security](#) [Pricing](#) [Customer stories](#) [Getting started](#) [Documentation](#) [FAQs](#)

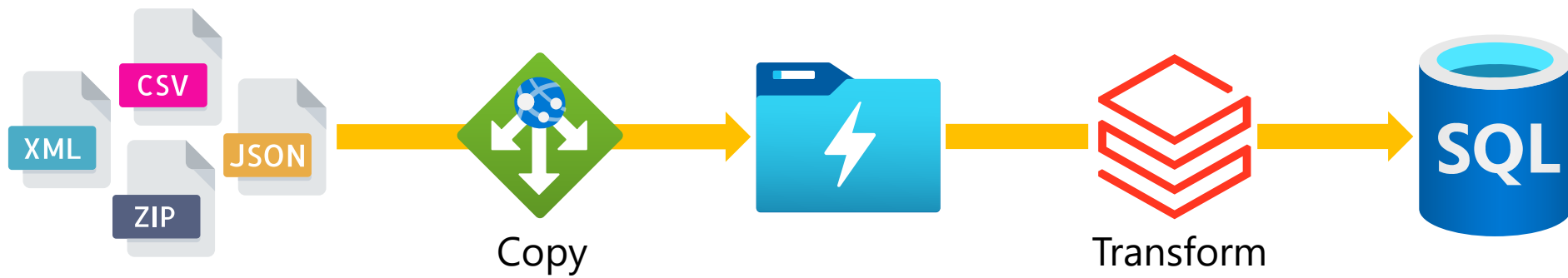
Accelerate data integration

Integrate data silos with Azure Data Factory, a service built for all data integration needs and skill levels. Easily construct ETL and ELT processes code-free within the intuitive visual environment, or write your own code. Visually integrate data sources using more than 90+ natively built and maintenance-free connectors at no added cost. Focus on your data – the serverless integration service does the rest.

What is Azure Data Factory (ADF)?



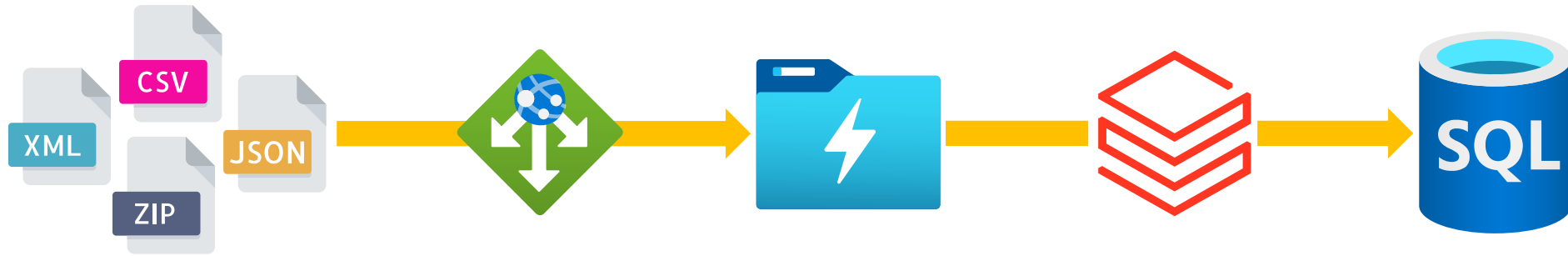
What is Azure Data Factory (ADF)?



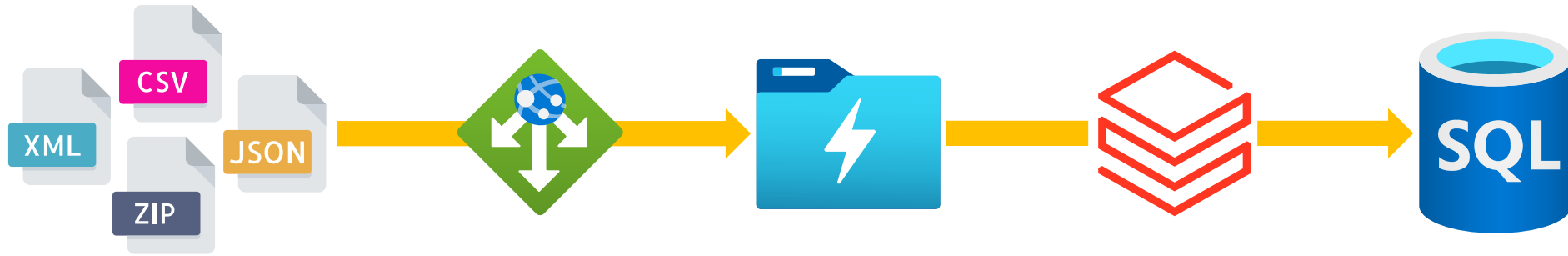
Resource Components

A large, solid blue shape that starts as a diagonal line from the bottom left and extends towards the top right, filling the right side of the slide.

Data Factory Components

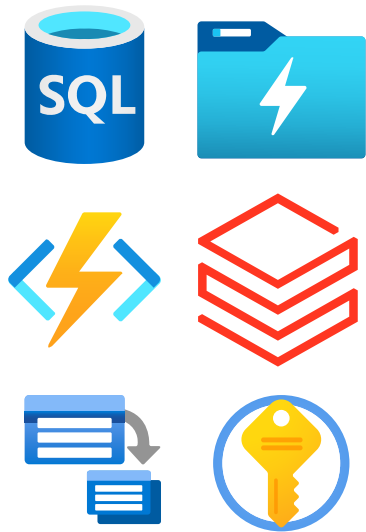


Data Factory Components



1

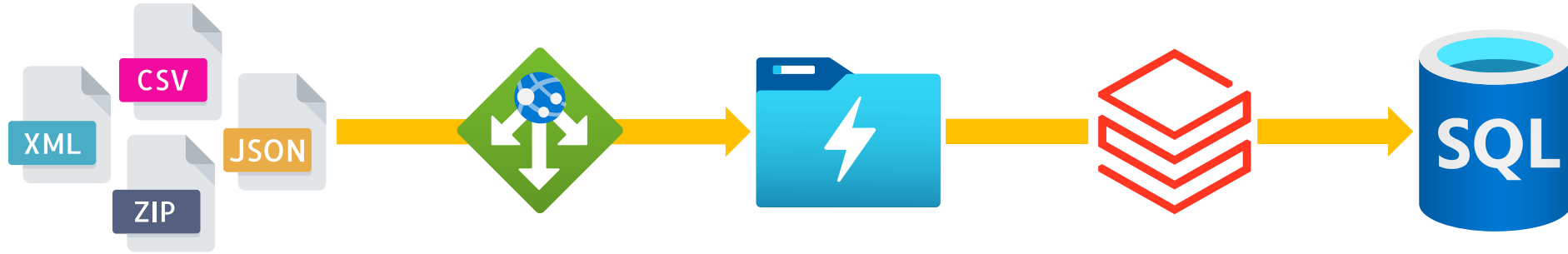
Linked Services – What to interact with and how?



SQLDBLinkedService

ConnectionString: *Server=MyServer;Database=myDataBase*
UserName: *"MrPaulAndrew"*
Password: *******

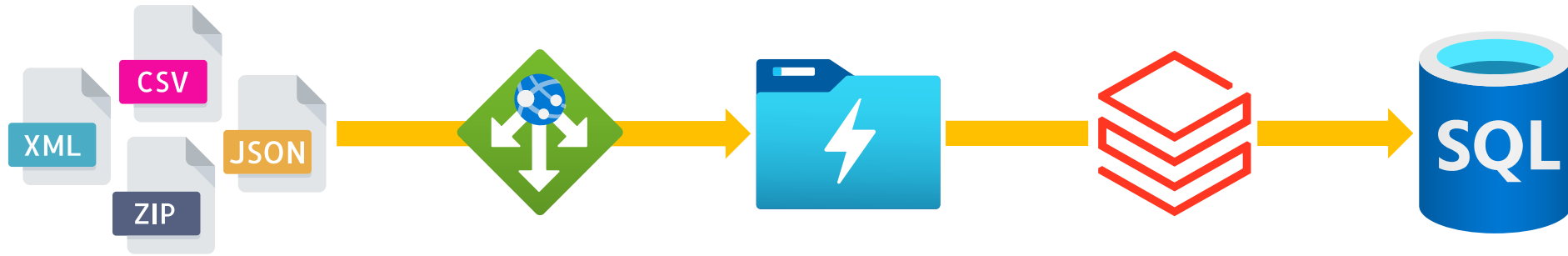
Data Factory Components



1

Linked Services

Data Factory Components



1 Linked Services

2 **Datasets** – Where is my data? What format? What file path/table do I need?

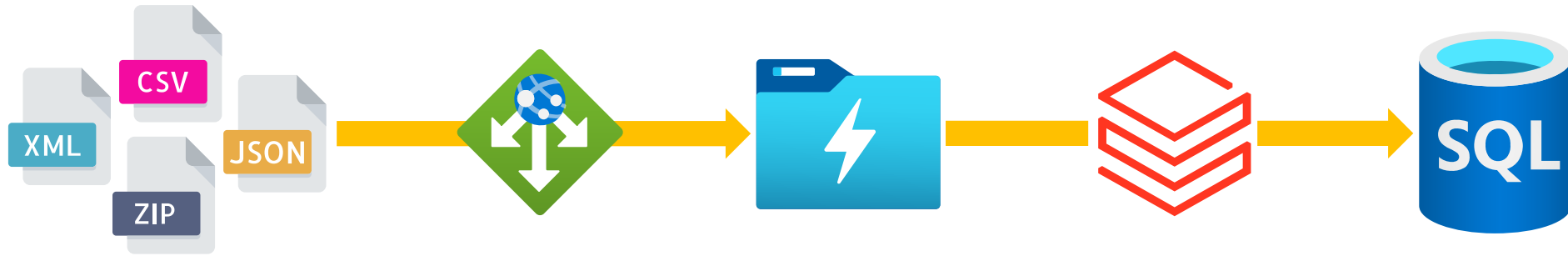


[dbo].[SalesOrders]



/RAW/Orders/2018/01/01/SalesOrders.csv

Data Factory Components



1 Linked Services

2 Datasets

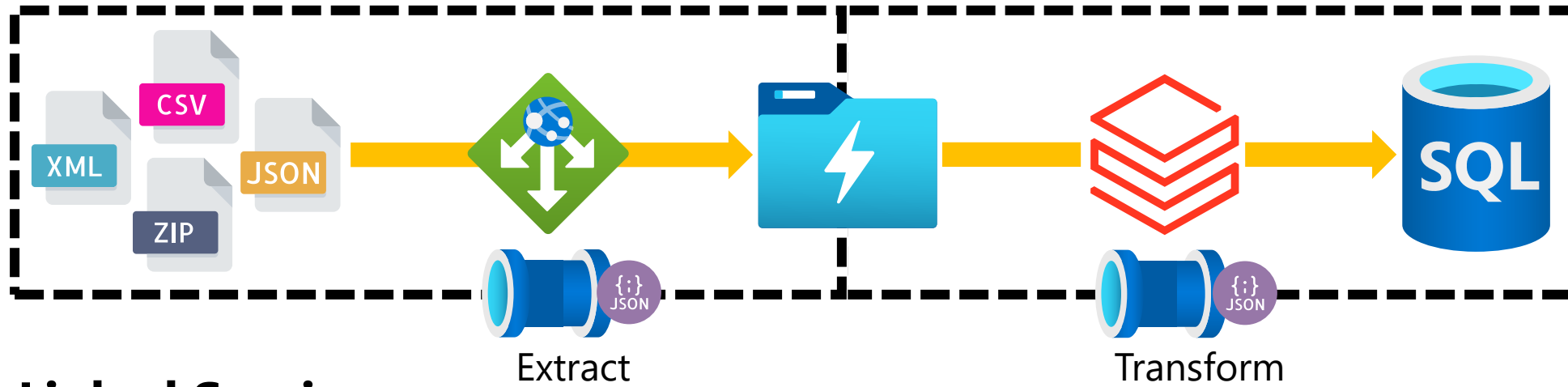
3 **Activities** – What do we want to happen when we invoke a Linked Service? With what conditions?



Databricks Notebook Activity

```
notebookPath: /Playground/Playing
baseParameters: Testing
libraries[jar]: dbfs:/lib1.jar
linkedServiceName: BricksOfData01
```

Data Factory Components

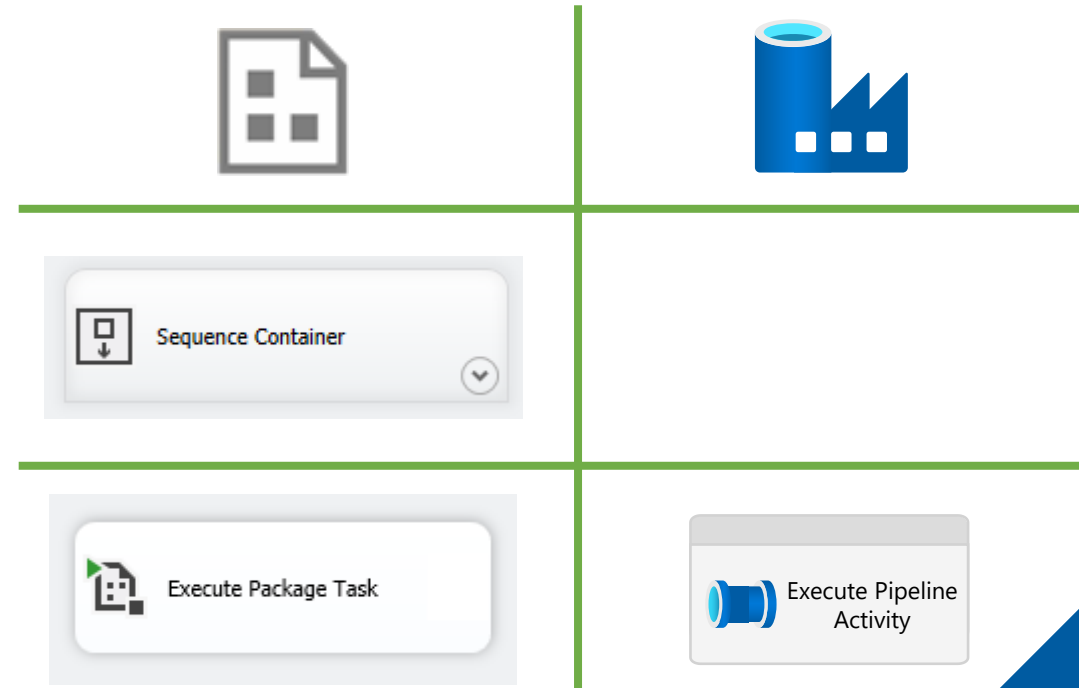


1 **Linked Services**

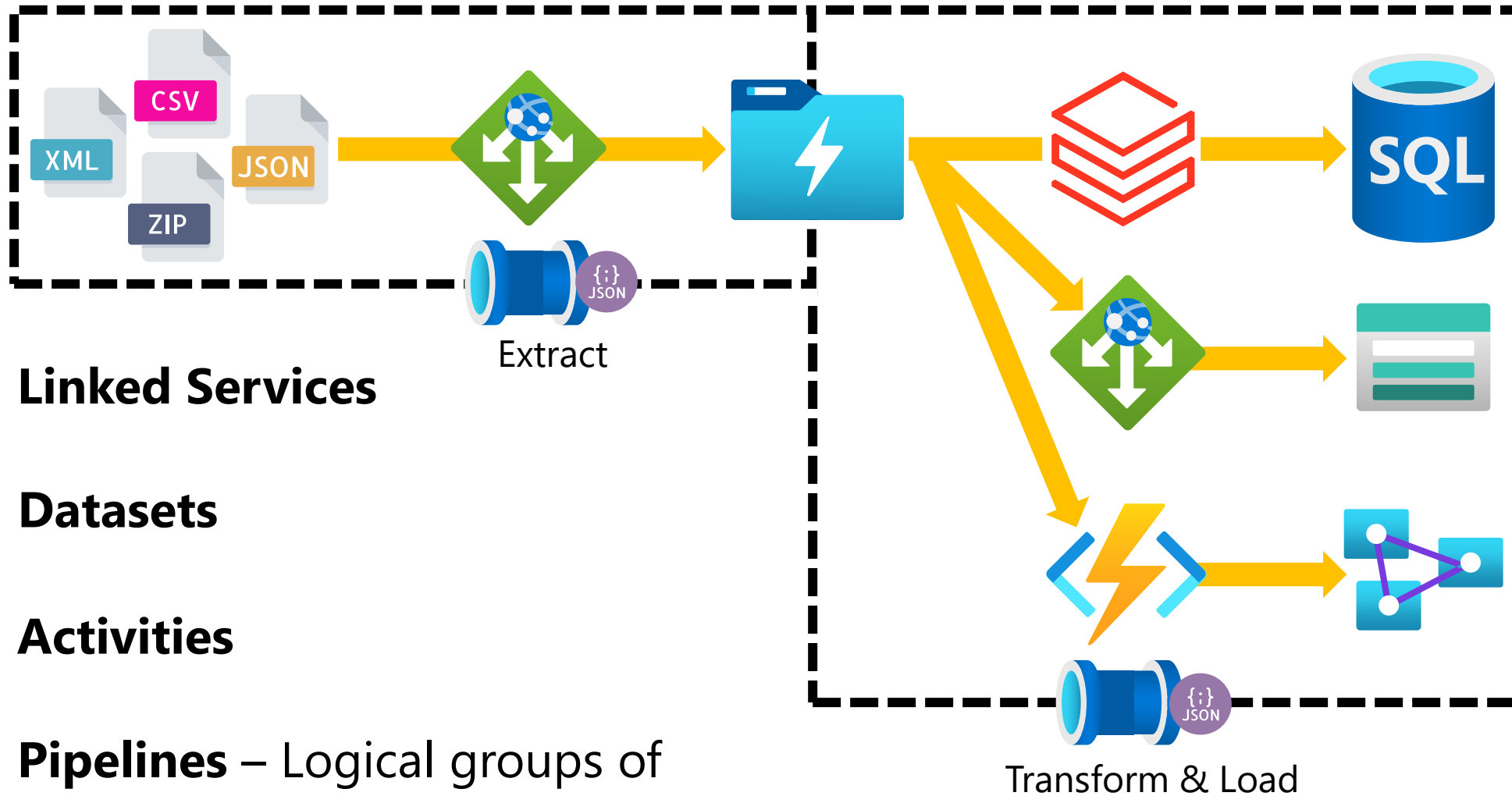
2 **Datasets**

3 **Activities**

4 **Pipelines** – Logical groups of work that can be executed.



Data Factory Components



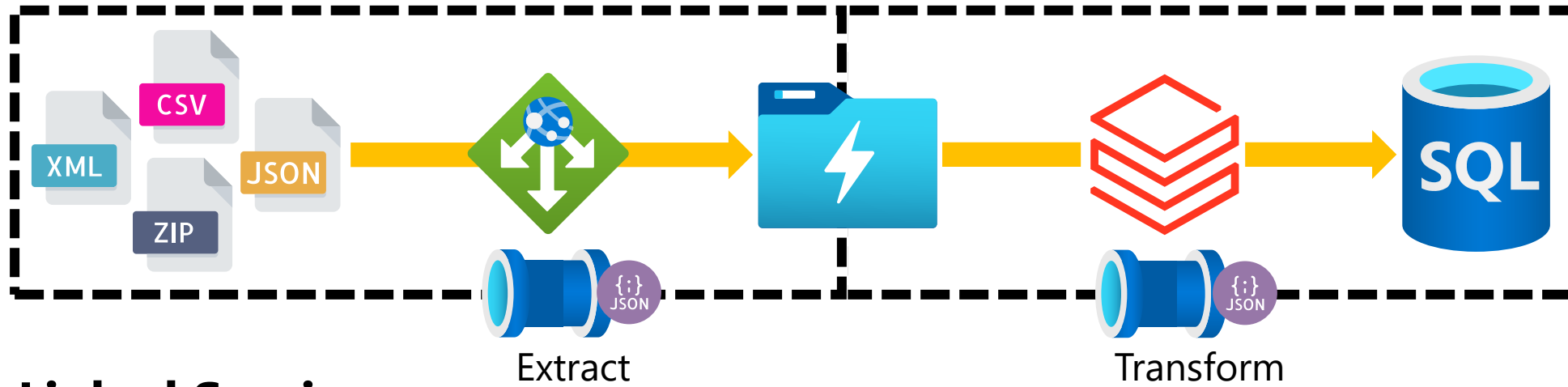
1 Linked Services

2 Datasets

3 Activities

4 Pipelines – Logical groups of work that can be executed.

Data Factory Components



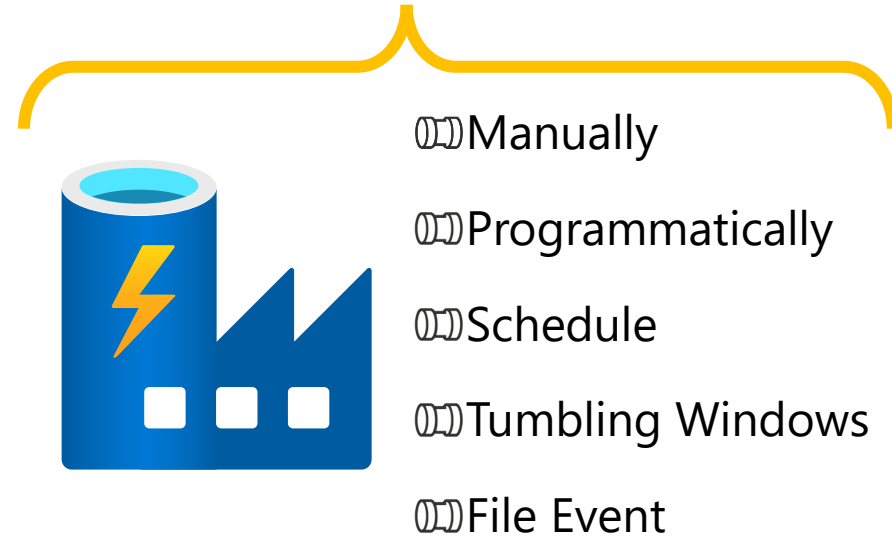
1 **Linked Services**

2 **Datasets**

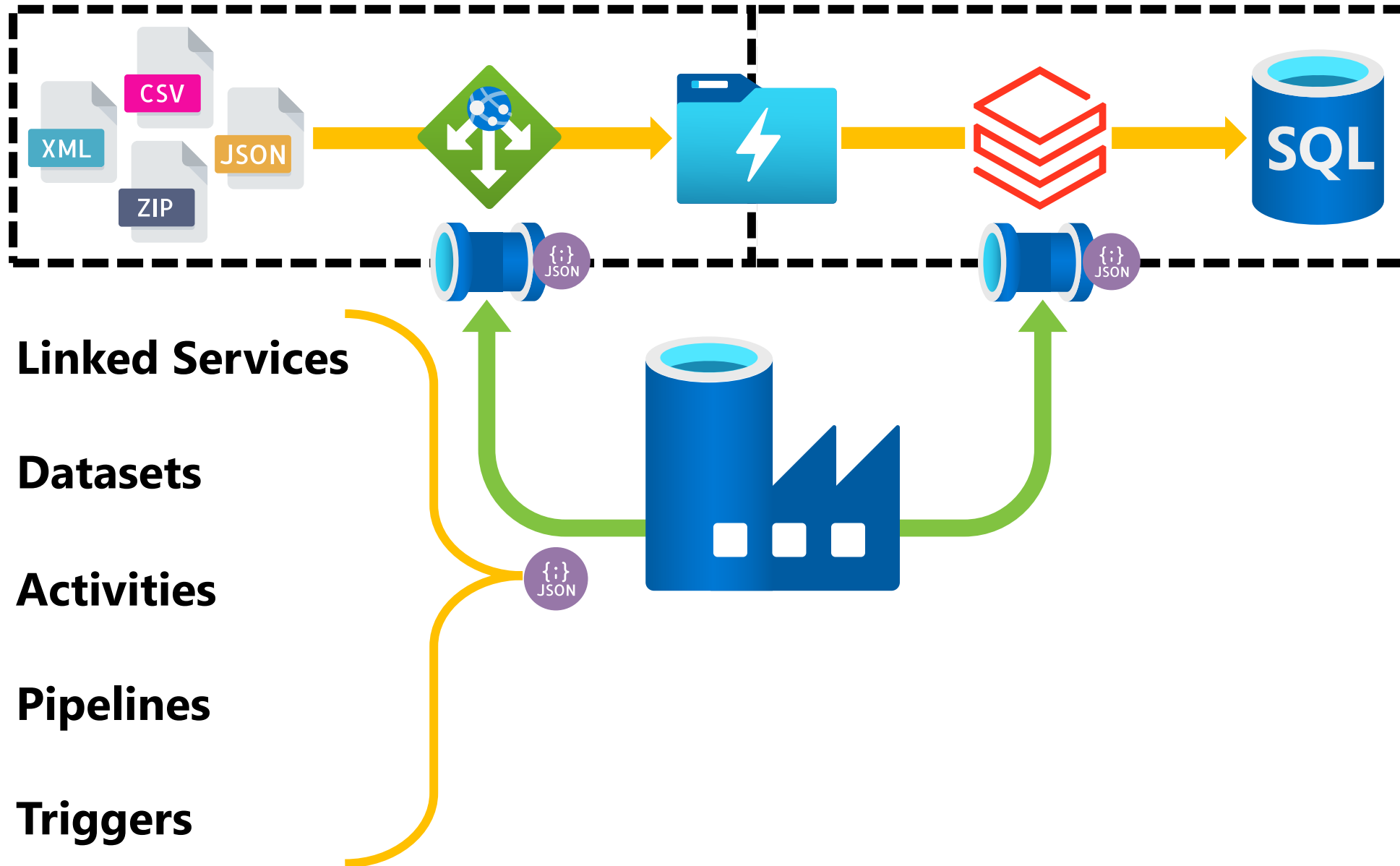
3 **Activities**

4 **Pipelines**

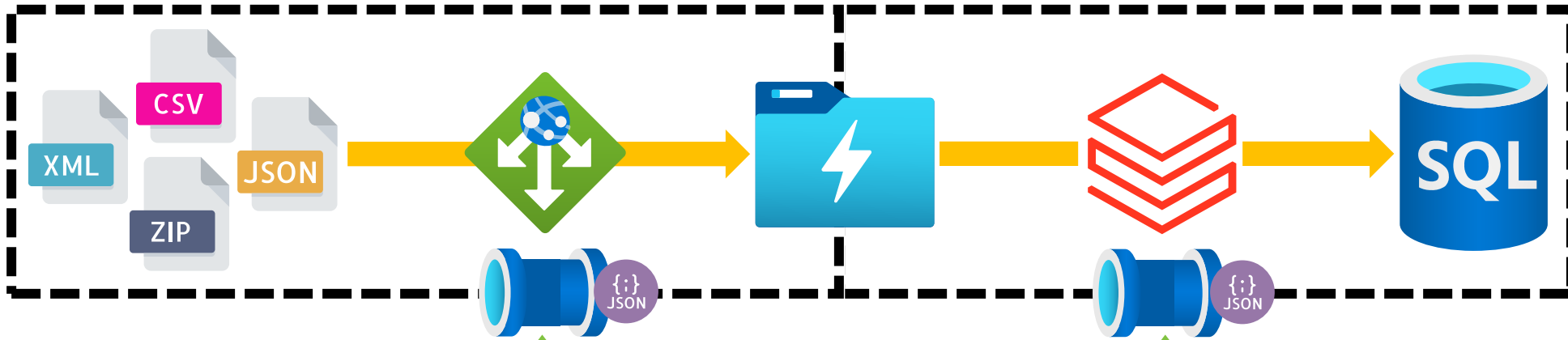
5 **Triggers** – Telling our when pipelines to run.



Data Factory Components



Data Factory Control Flow Components

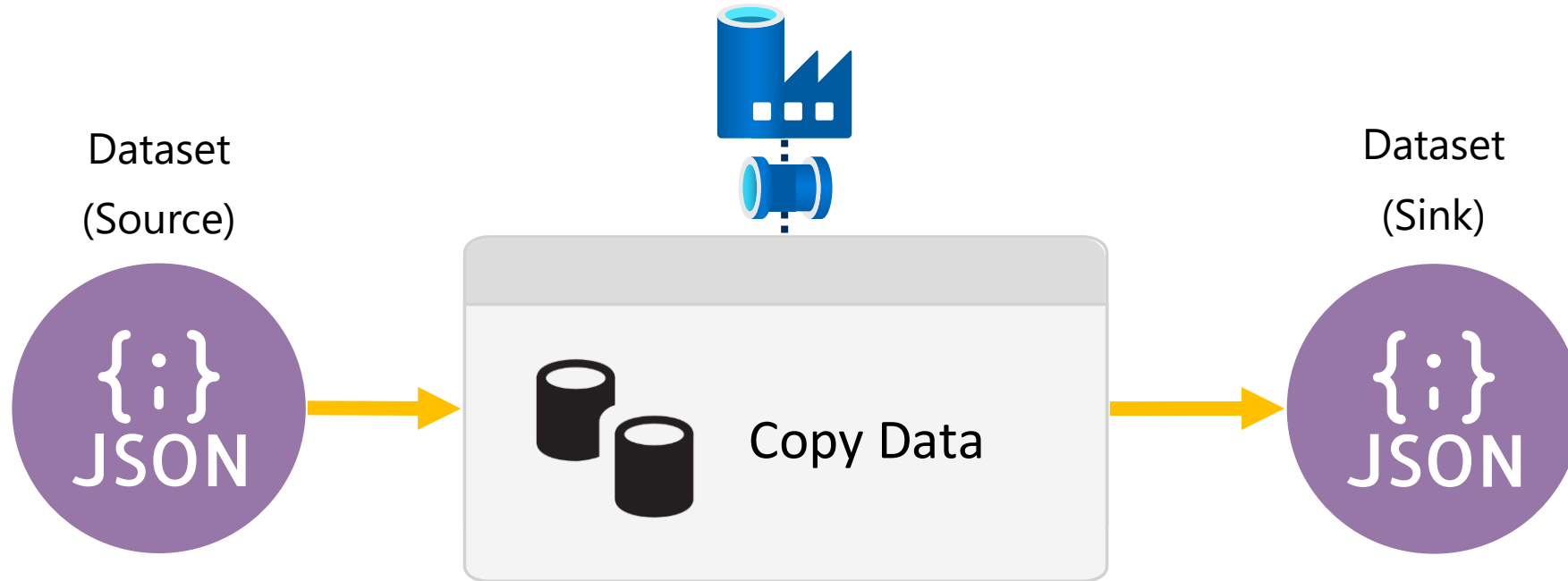


- 1 Linked Services
- 2 Datasets
- 3 Activities
- 4 Pipelines
- 5 Triggers

Common Activities

A solid blue triangular shape pointing towards the top right corner of the slide, occupying the right half of the image.

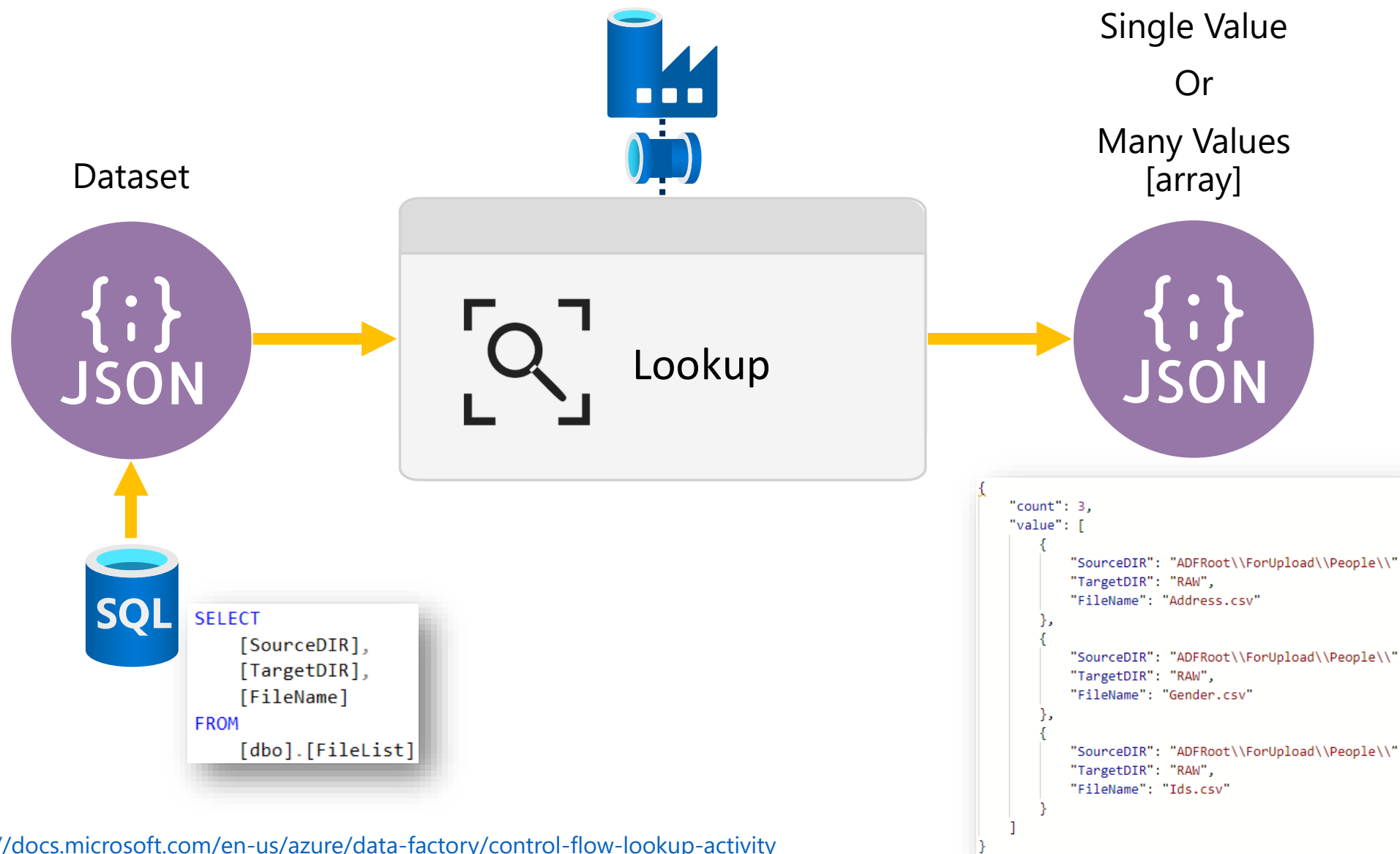
Copy



- ☞ Auto Scaling
- ☞ Transactional Restarts
- ☞ Handle Zip Compression
- ☞ Attribute Mapping and Schema Drift
- ☞ Handle Failed Rows
- ☞ Add Custom Attributes
- ☞ Parse Excel & JSON Files

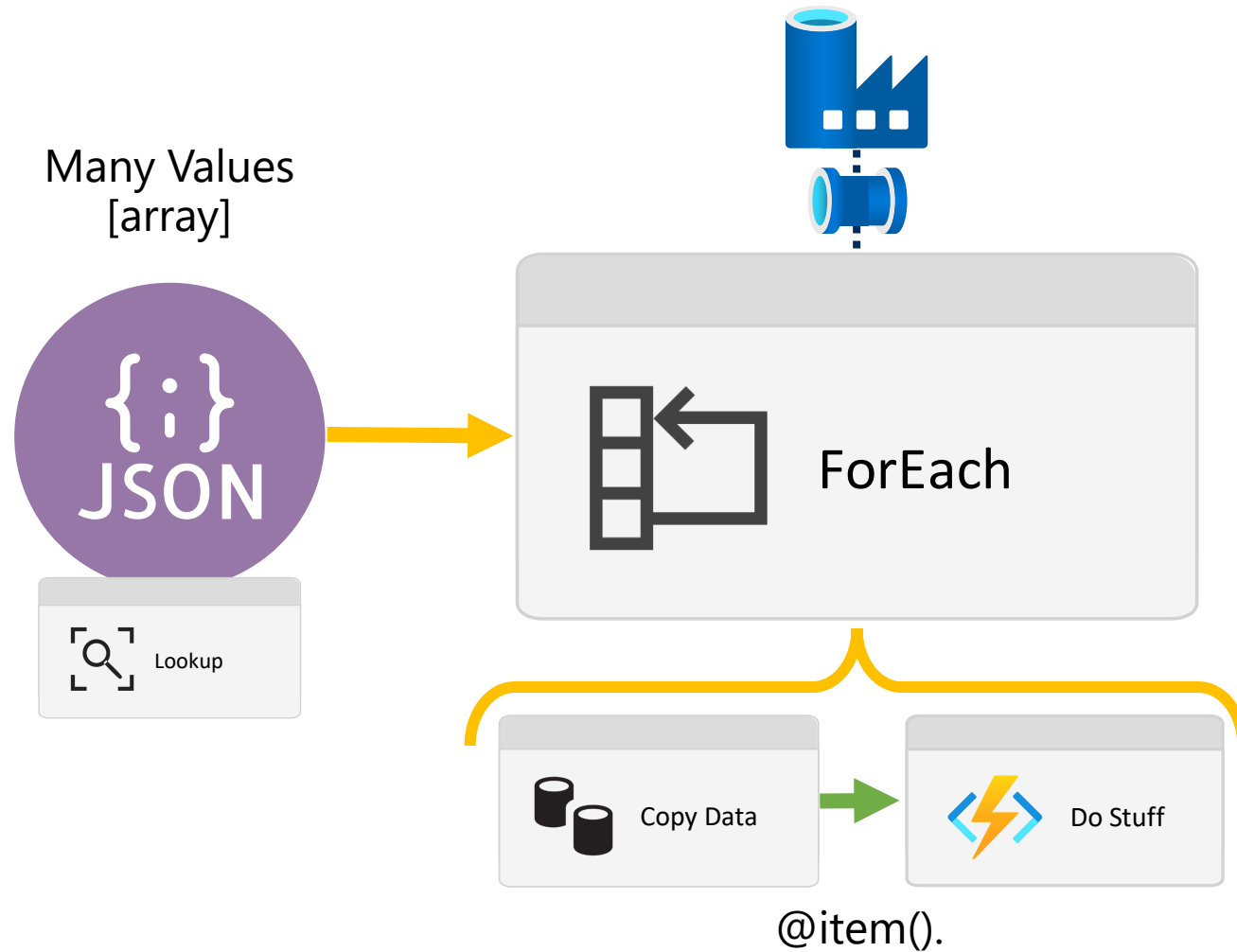
Lookup

Get Metadata to Support Other Control Flow Activities

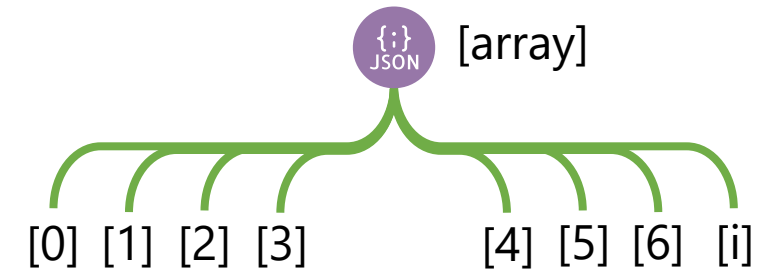
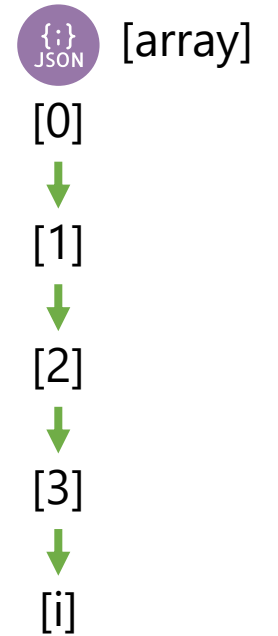


ForEach

Scaling Out Control Flow Activities



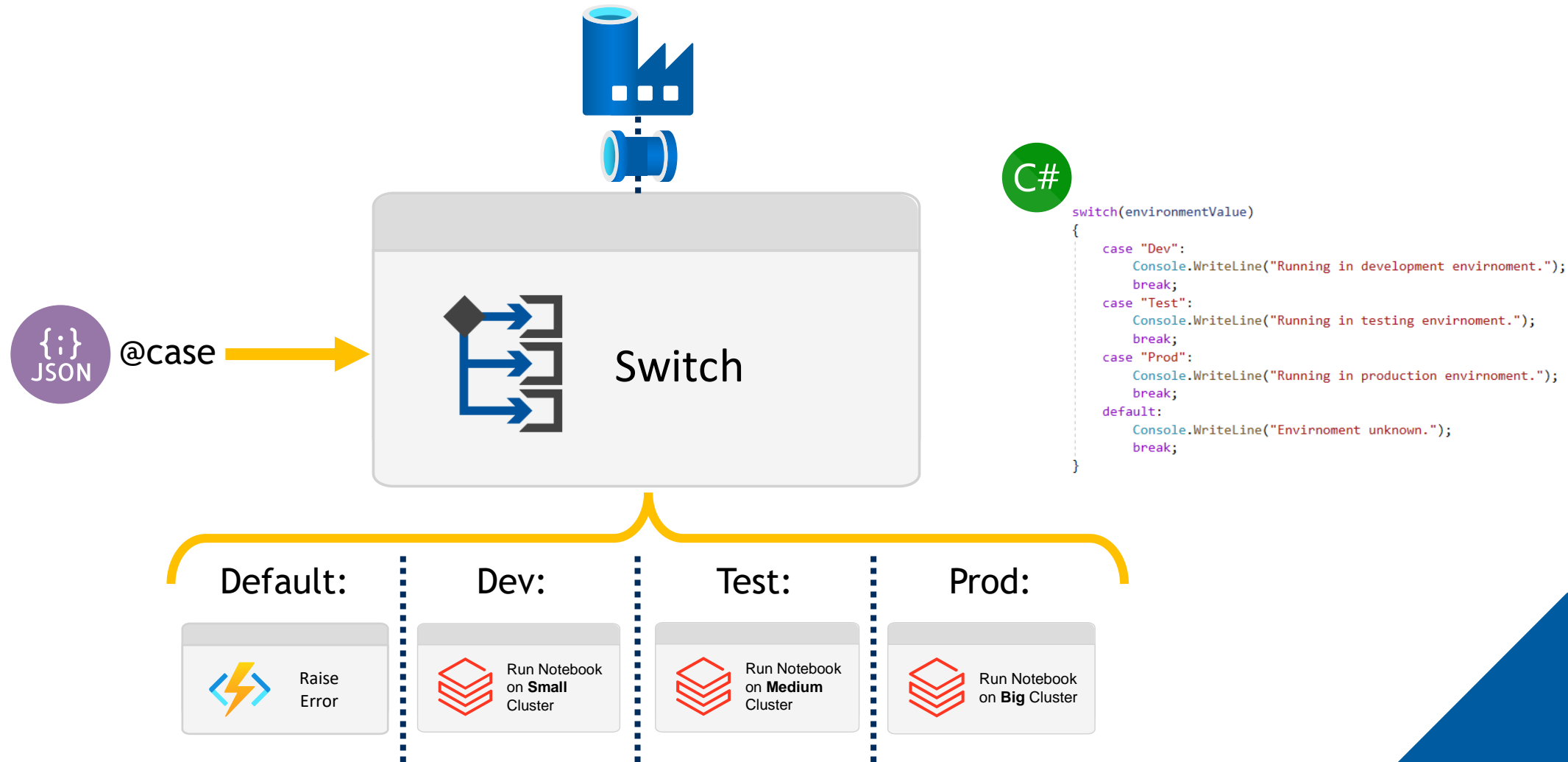
IsSequential:
true



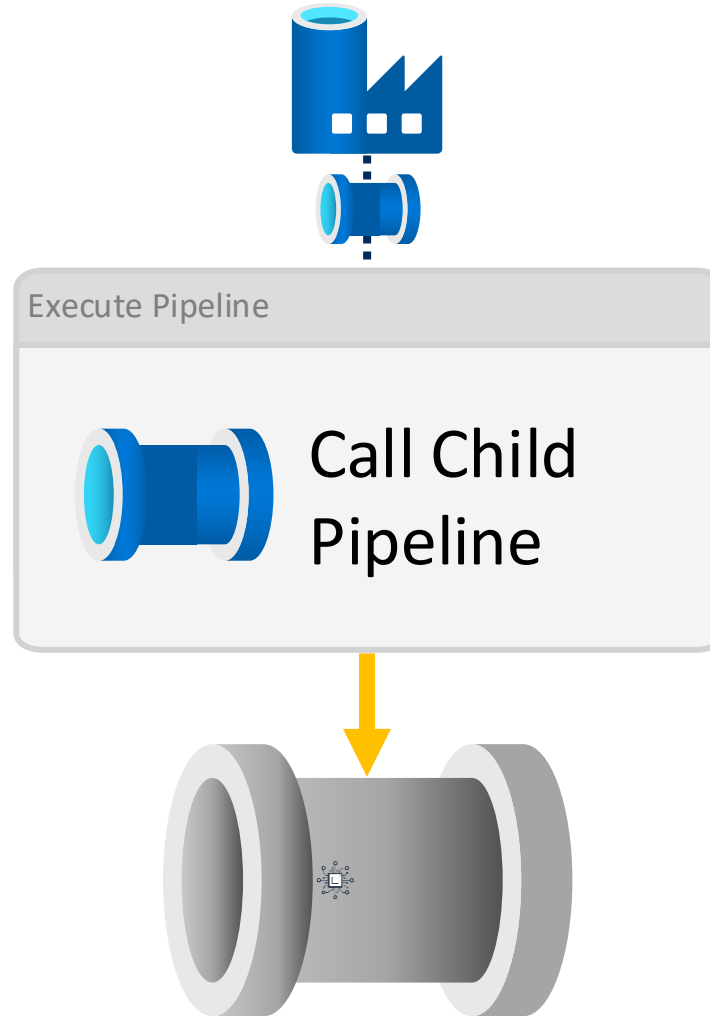
Batch Count Default: 20

Batch Count Max: 50

Switch



Execute Pipeline



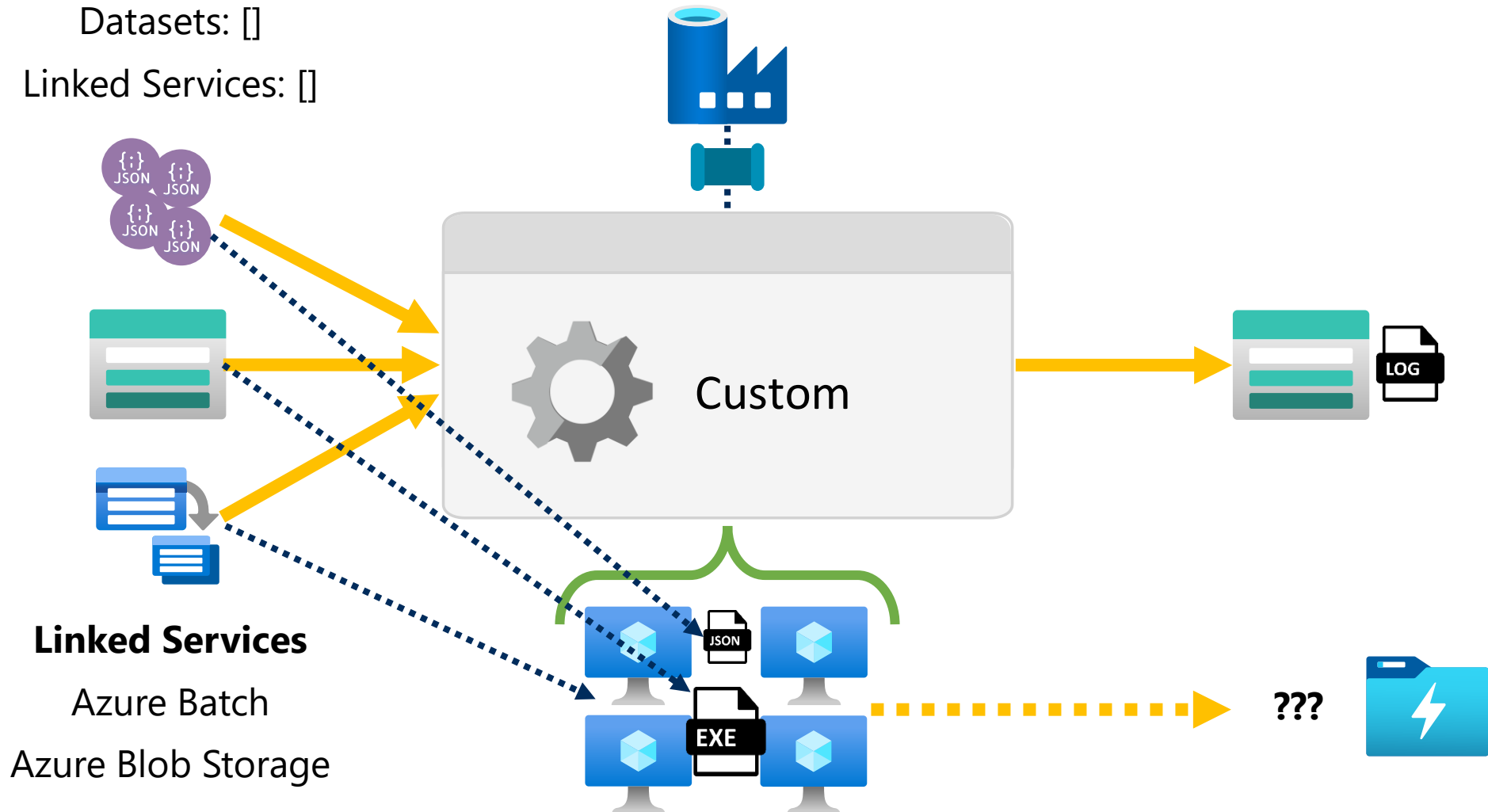
Custom

Extend Data Factory with Custom Code

References Objects

Datasets: []

Linked Services: []



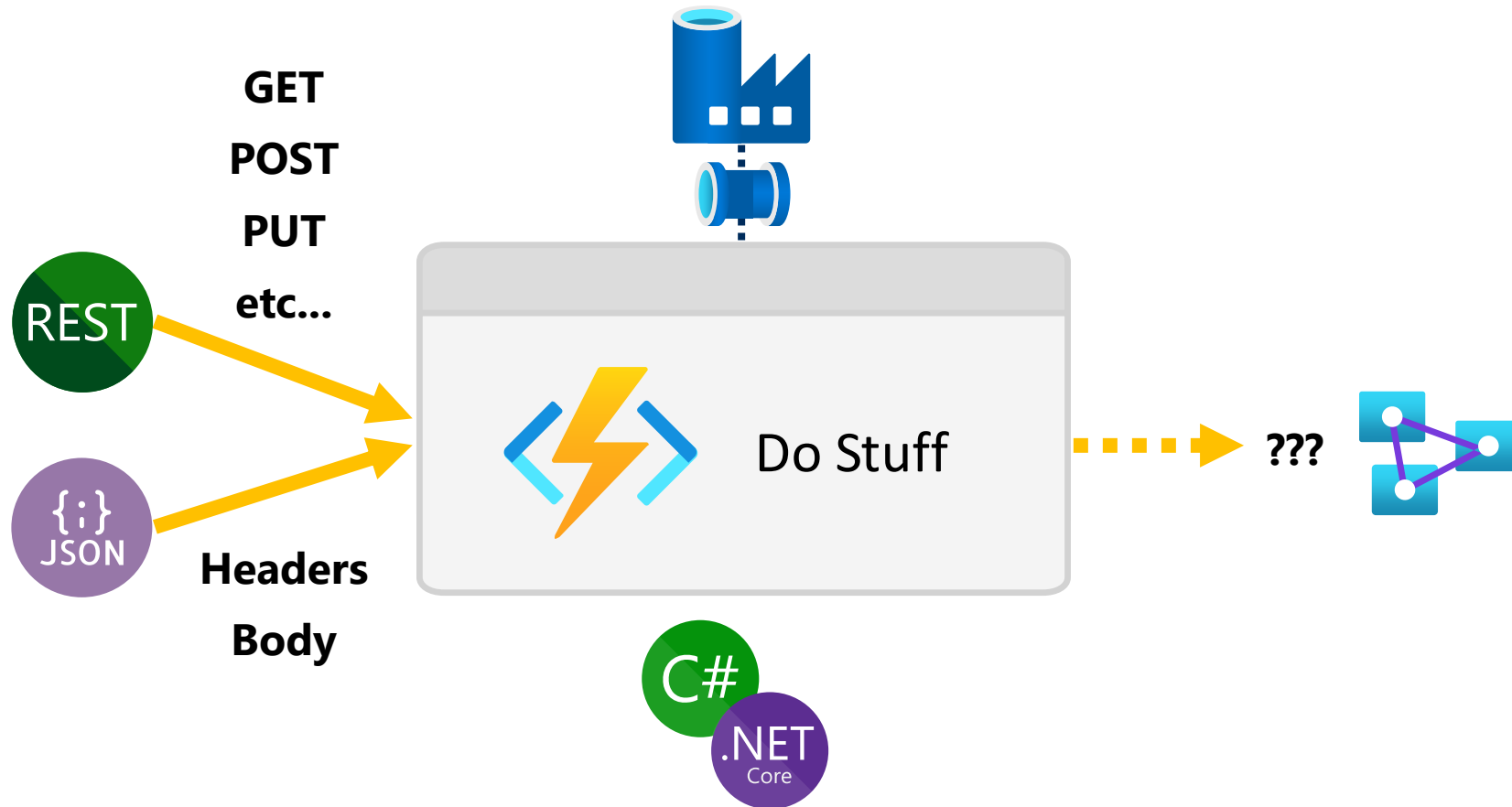
Linked Services

Azure Batch

Azure Blob Storage

Azure Function

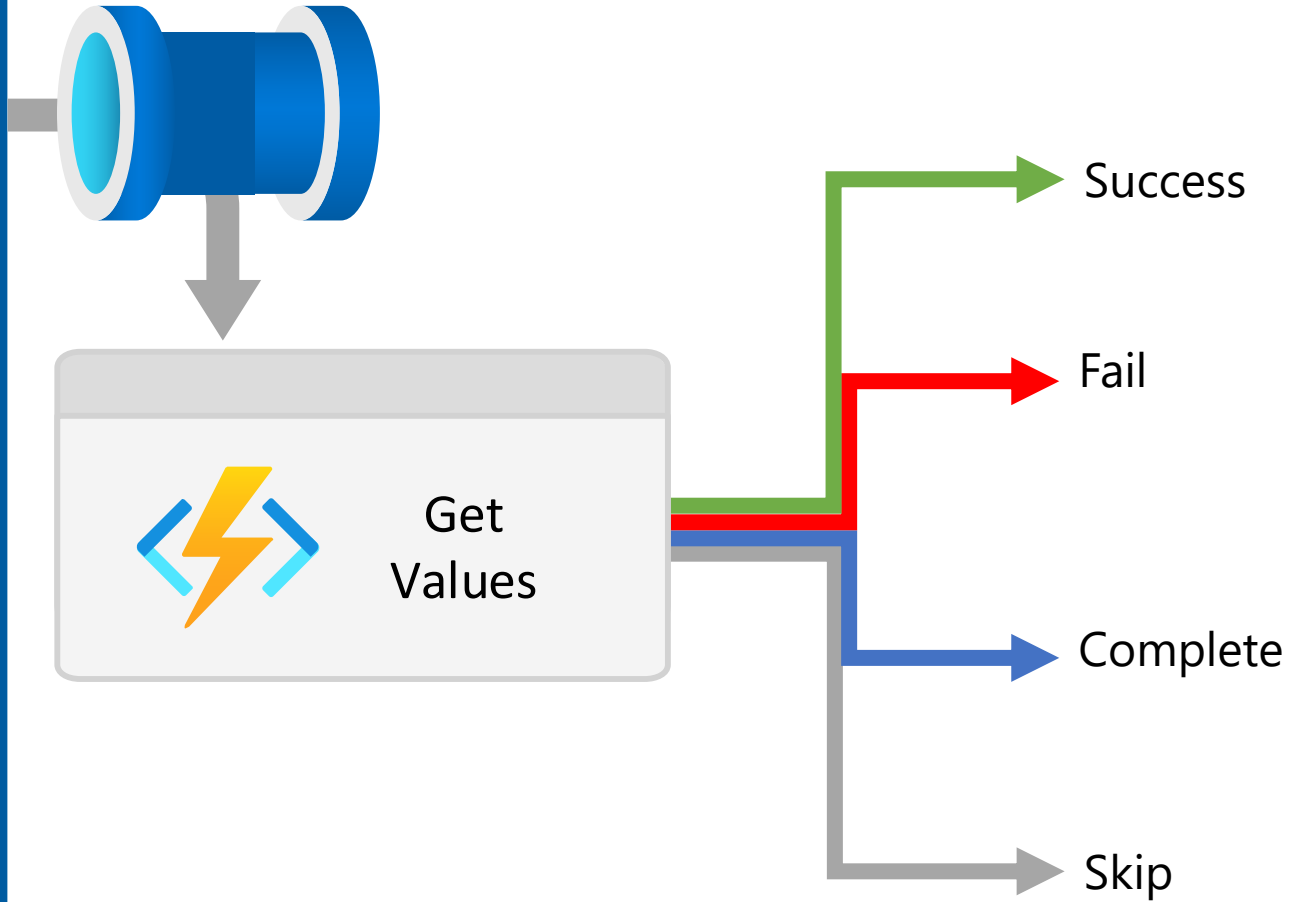
Extend Data Factory with Rest Calls



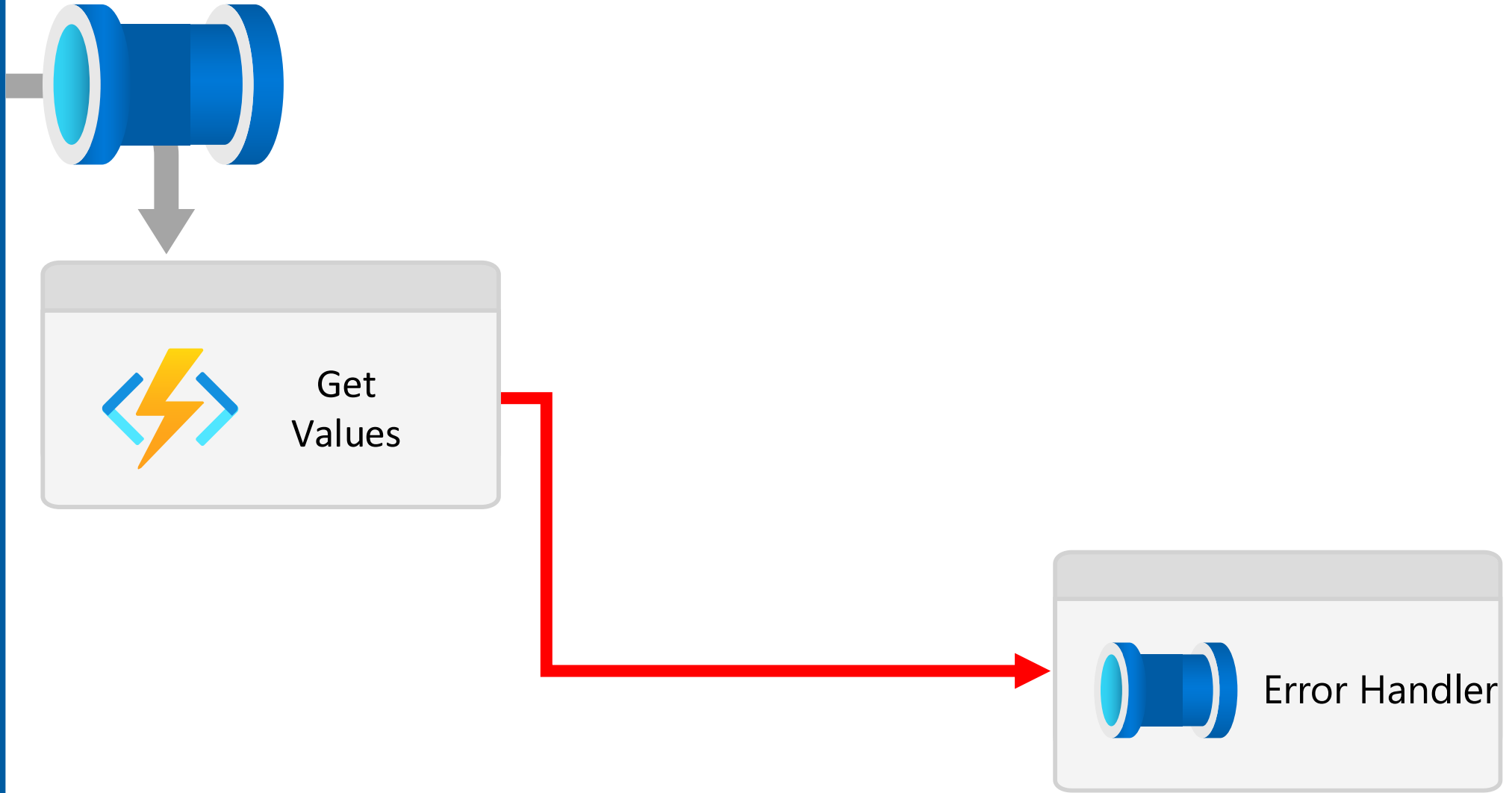
Execution Dependencies

A large blue triangle is positioned on the right side of the slide, pointing towards the top right corner.

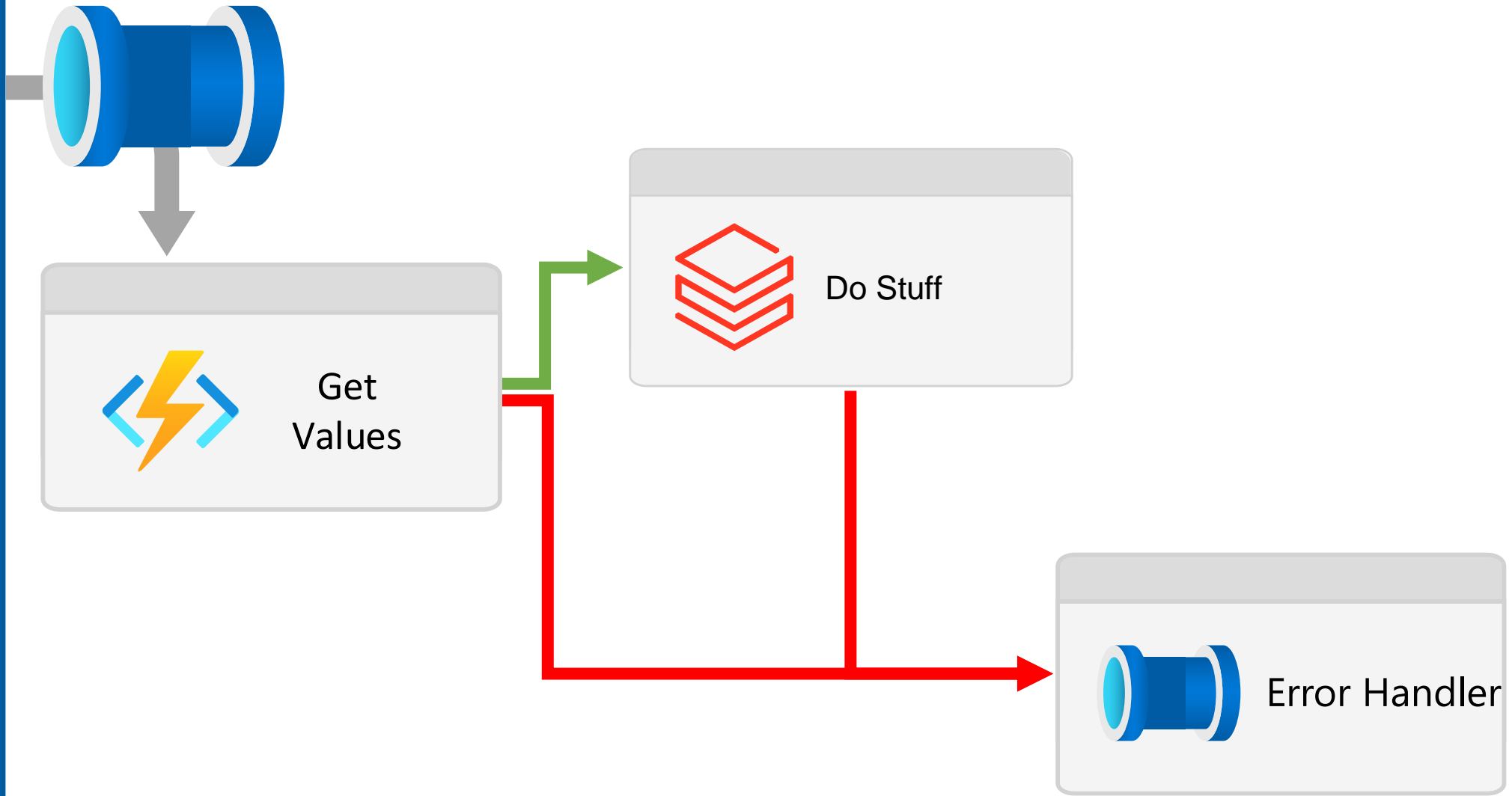
Execution Dependency Options



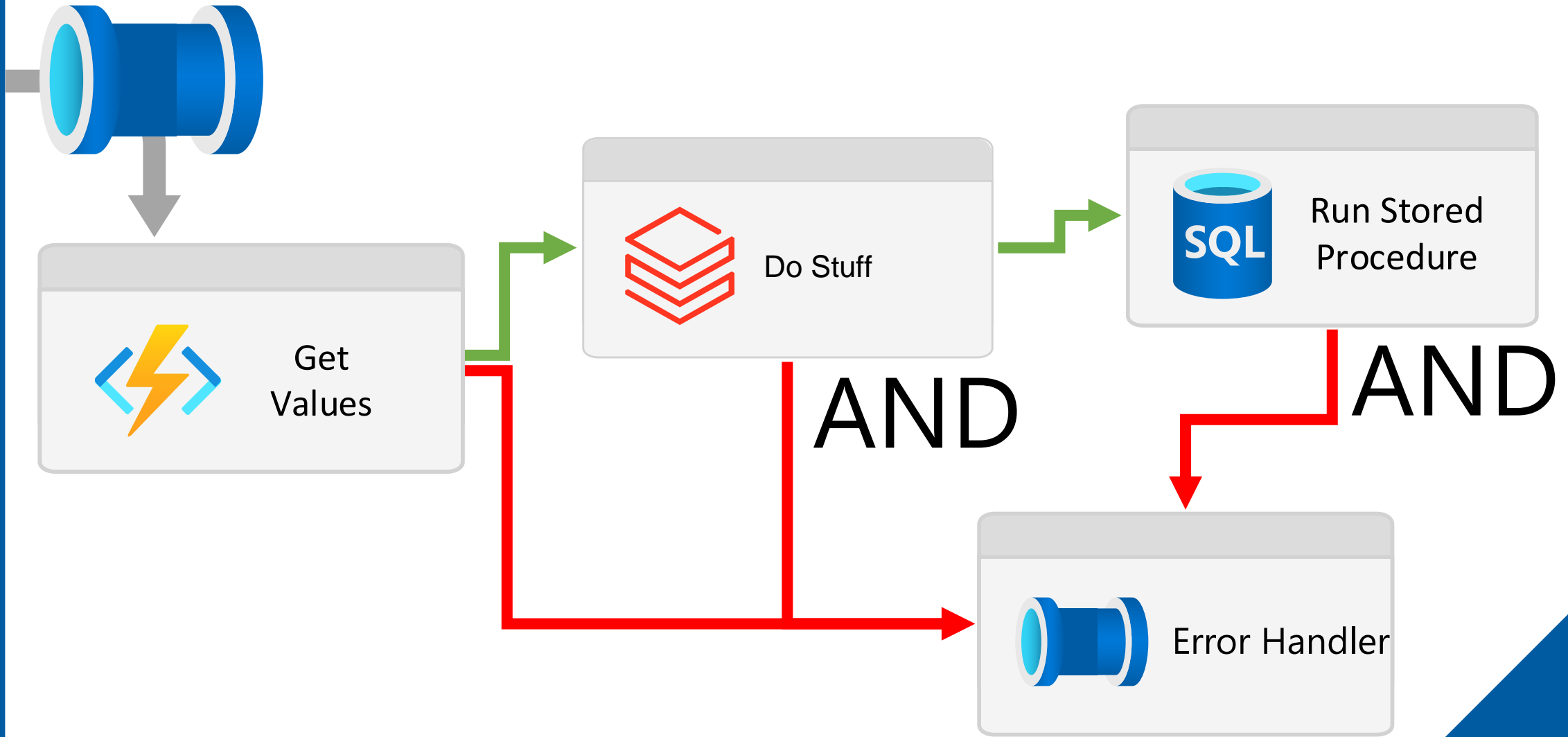
Execution On Failure



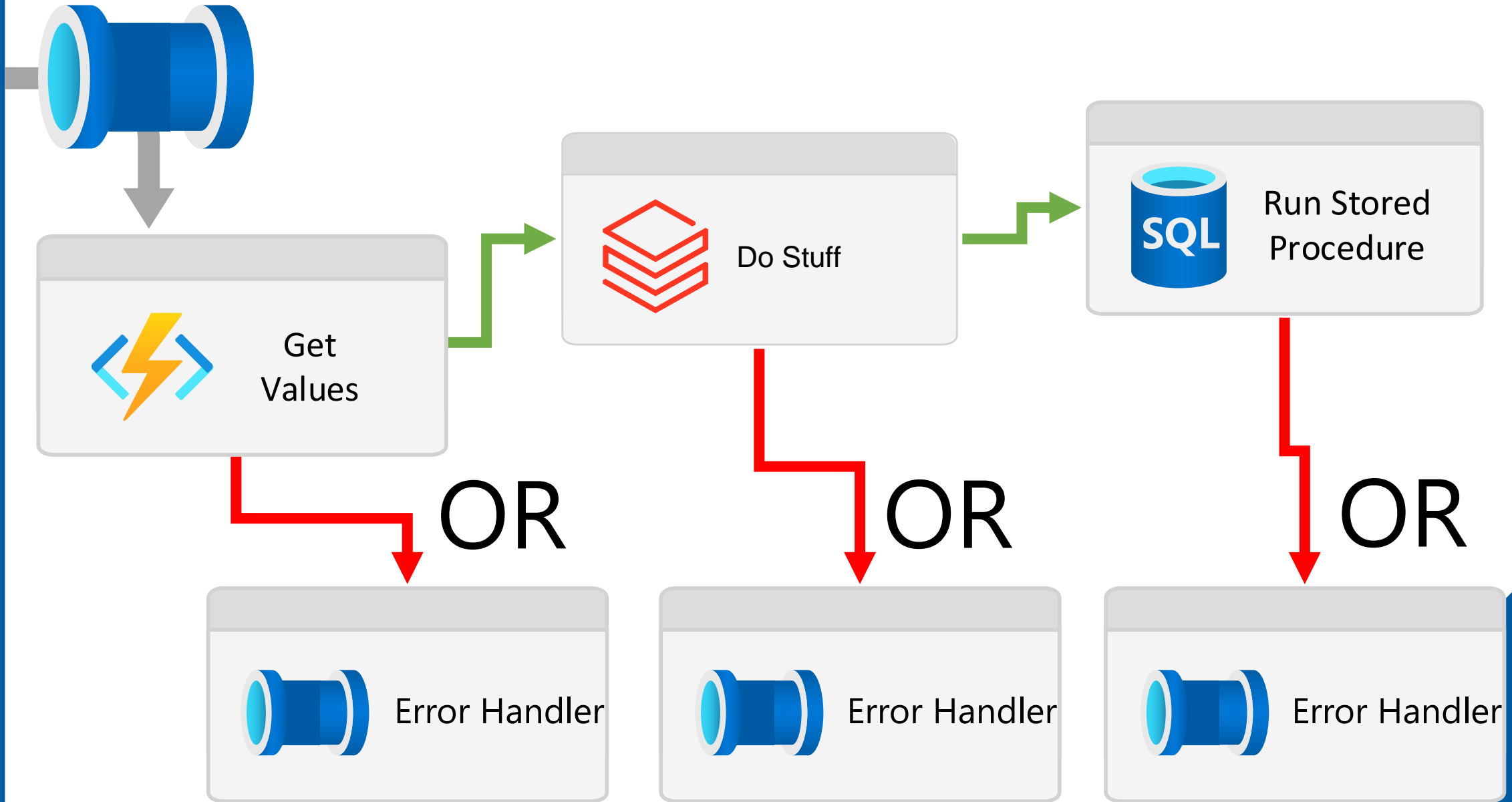
Execution On Failure or On Success



Execution On ???



Execution On Failure or On Success



Module 1:

Data Factory Fundamentals

- 🔊 What is it and why use it? ✓
- 🔊 Resource Components ✓
- 🔊 Common Activities ✓
- 🔊 Execution Dependencies ✓