



# Module 1: Data Factory Fundamentals

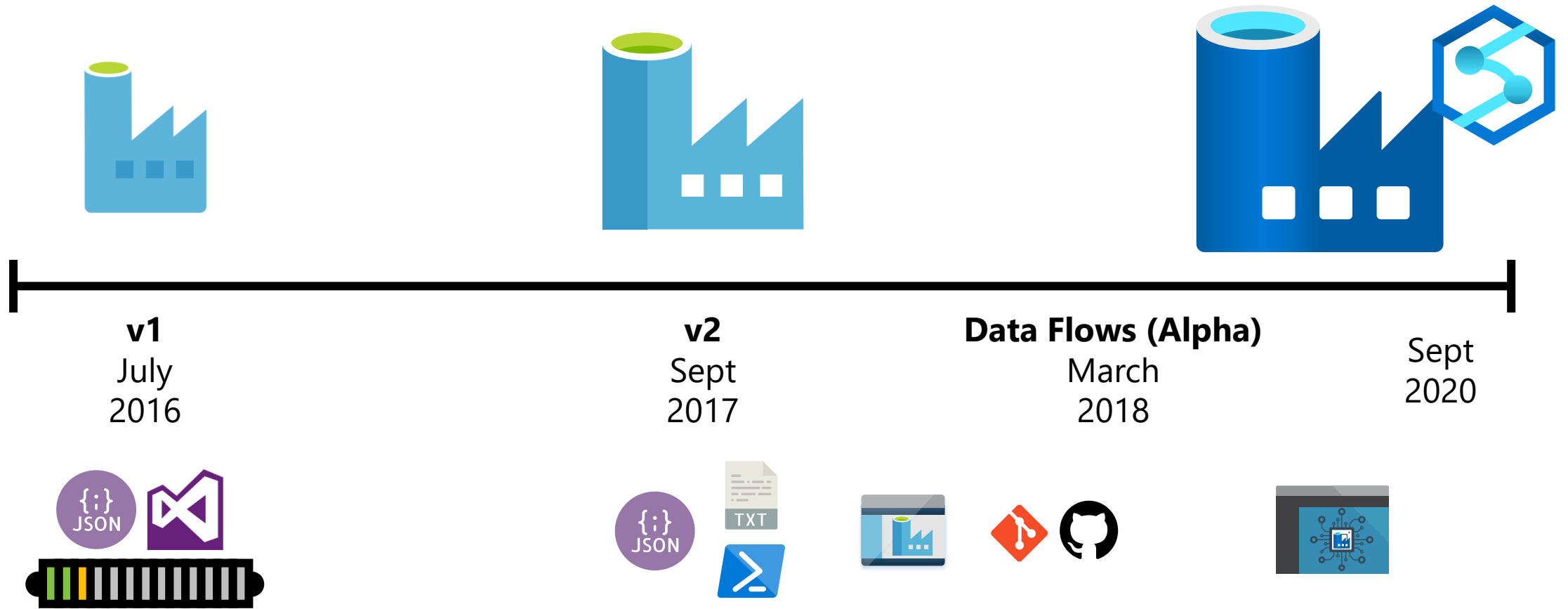
- 🔊 What is it and why use it?
- 🔊 Resource Components
- 🔊 Common Activities
- 🔊 Execution Dependencies

# Azure Data Factory –

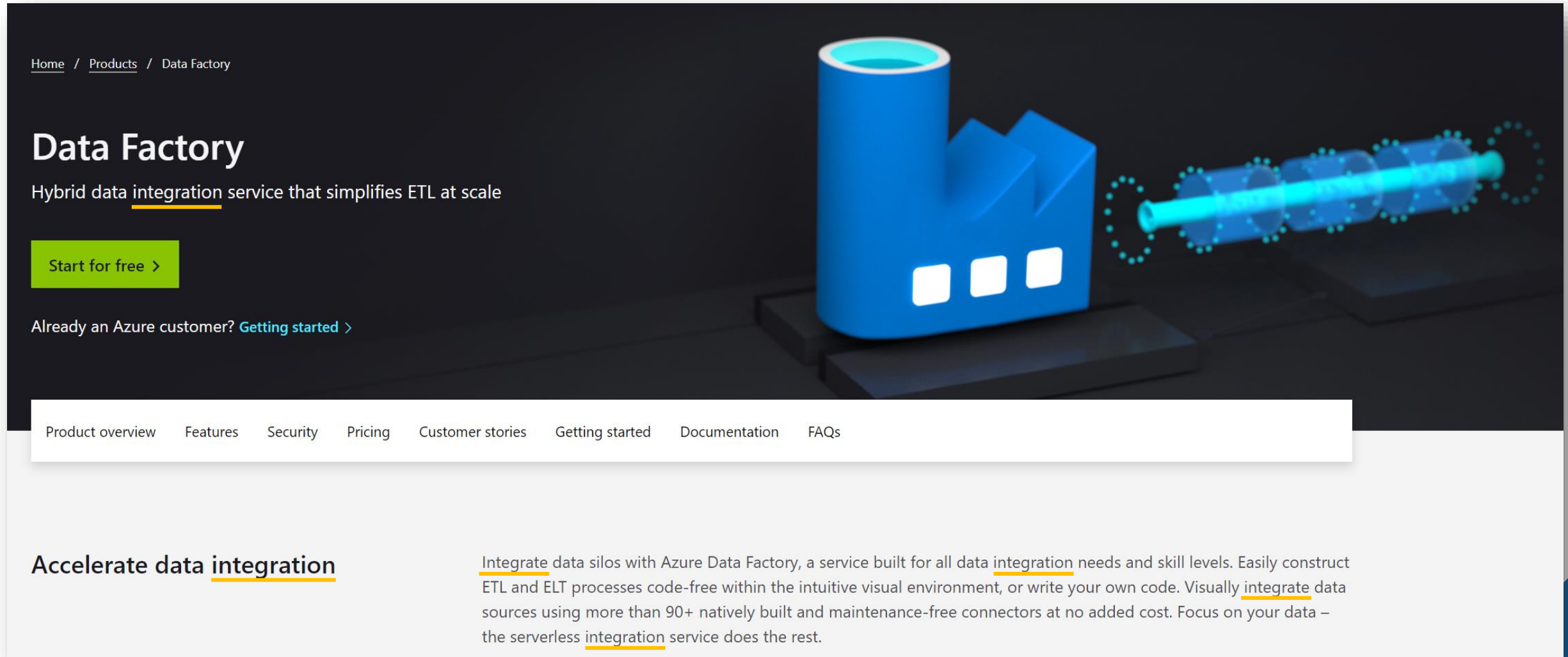
What is it?  
Why use it?



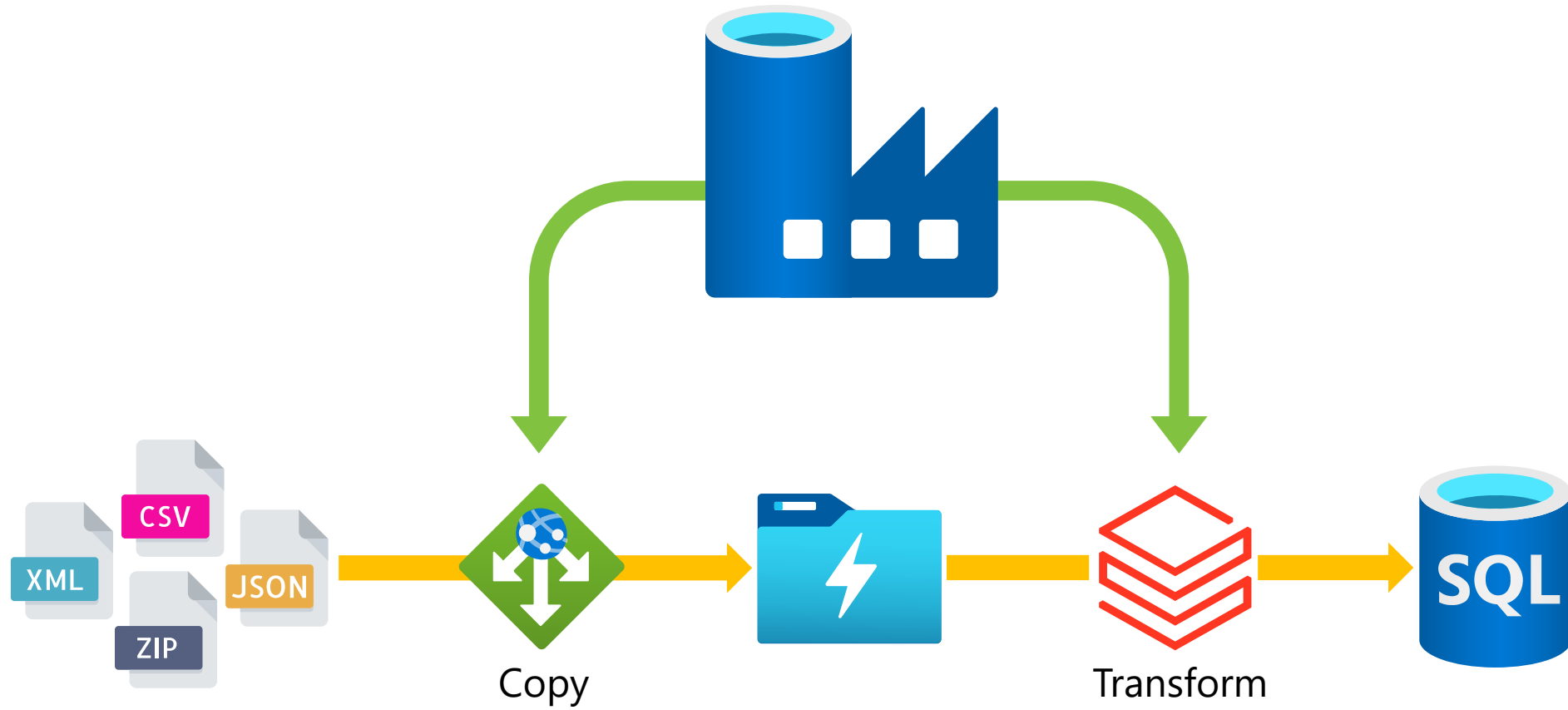
# A Quick History Lesson



# What is Azure Data Factory (ADF)?

The image is a screenshot of the Azure Data Factory product page. At the top, there is a navigation breadcrumb: 'Home / Products / Data Factory'. Below this, the heading 'Data Factory' is displayed in a large, bold font. Underneath the heading, a sub-headline reads 'Hybrid data integration service that simplifies ETL at scale'. A prominent green button with the text 'Start for free >' is located to the left. Below the button, a link states 'Already an Azure customer? [Getting started >](#)'. The right side of the page features a 3D illustration of a blue industrial factory building with a tall chimney, emitting a glowing blue beam of light that passes through a series of translucent blue cylinders, each surrounded by a dotted blue circle, representing a data pipeline. A horizontal navigation bar is positioned below the main content, containing links for 'Product overview', 'Features', 'Security', 'Pricing', 'Customer stories', 'Getting started', 'Documentation', and 'FAQs'. At the bottom, a section titled 'Accelerate data integration' is followed by a descriptive paragraph about the service's capabilities in integrating data silos and simplifying ETL/ELT processes.

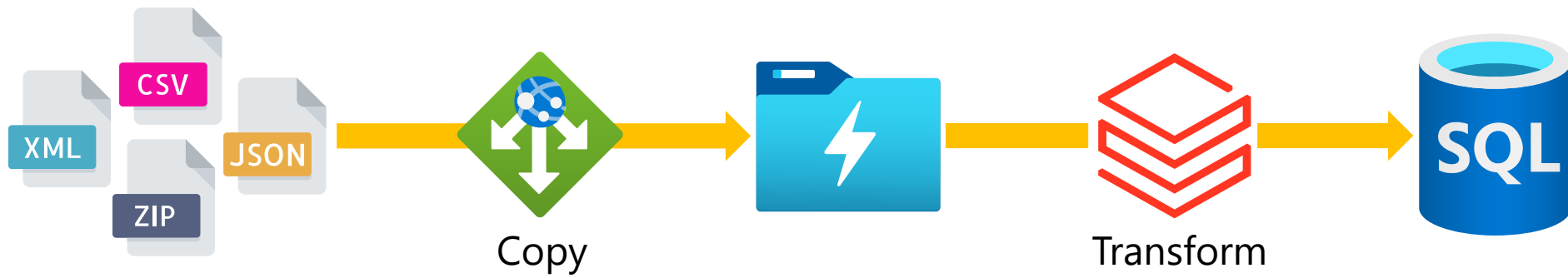
# What is Azure Data Factory (ADF)?



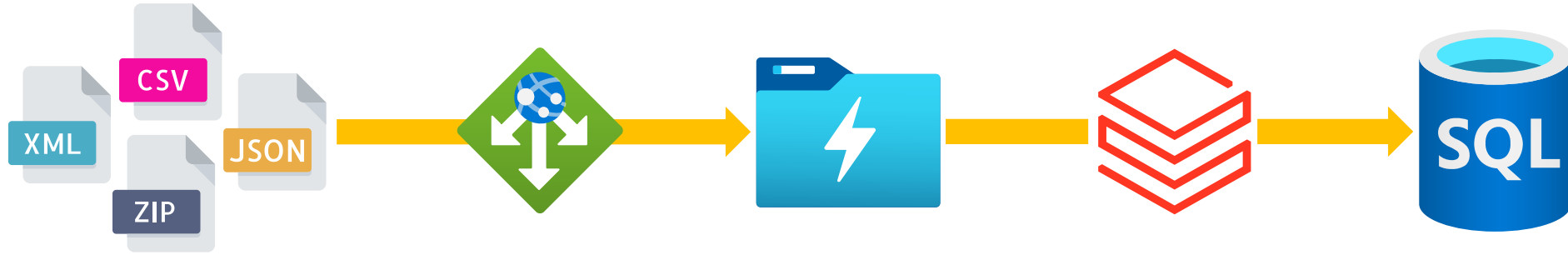
# Resource Components

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

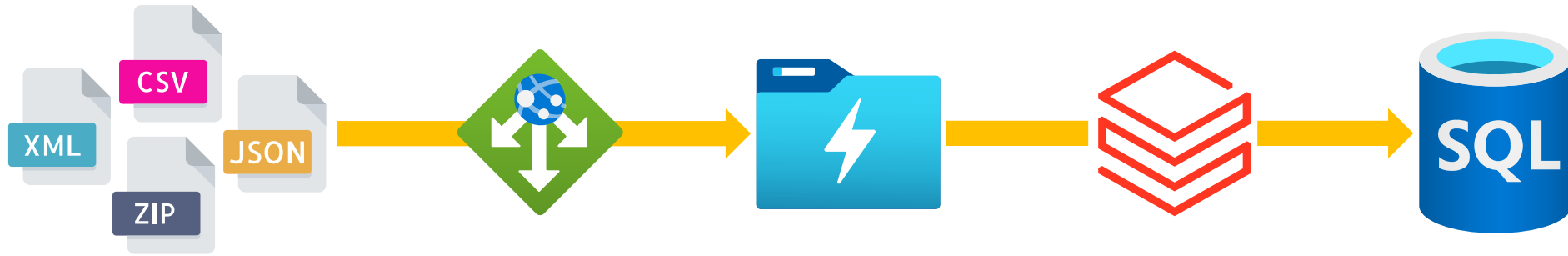


# Data Factory Components



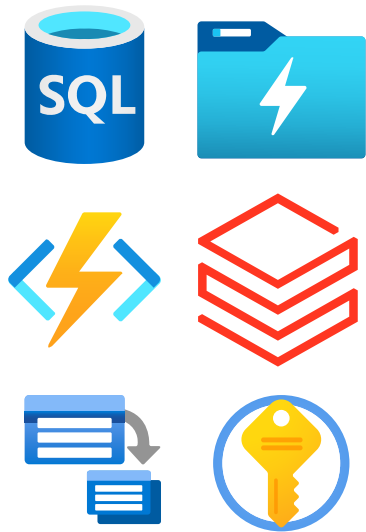


# Data Factory Components



1

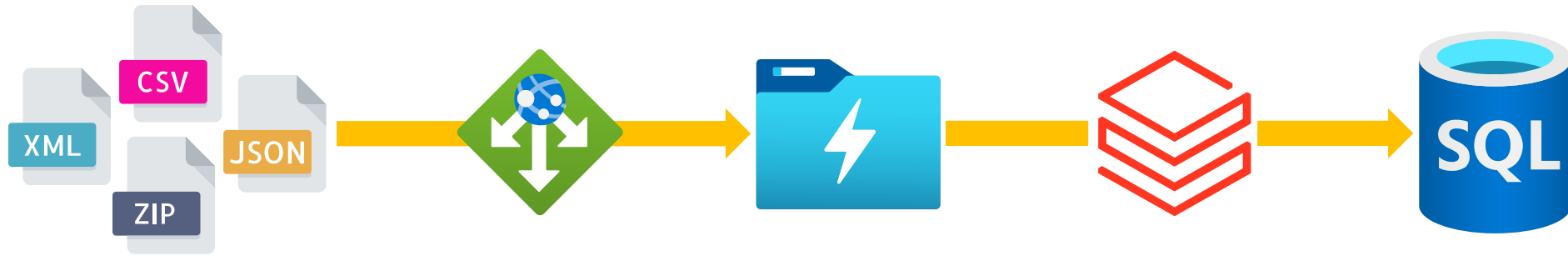
## Linked Services – What to interact with and how?



SQLDBLinkedService

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

# Data Factory Components



1

## Linked Services

2

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

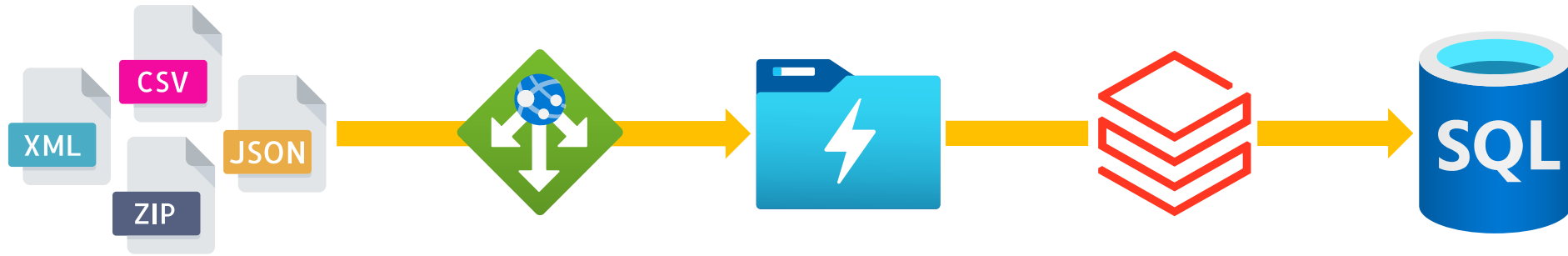


[dbo].[SalesOrders]



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

# Data Factory Components



## 1 Linked Services

## 2 Datasets

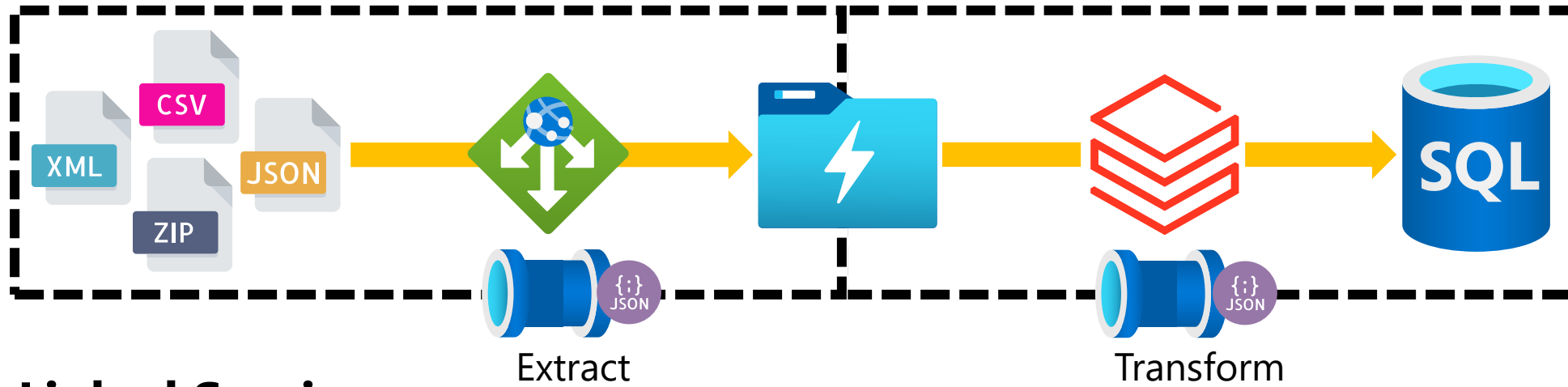
3 **Activities** – What do we want to happen when we invoke a Linked Service? With what conditions?

{:}  
JSON

### Databricks Notebook Activity

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

# Data Factory Components

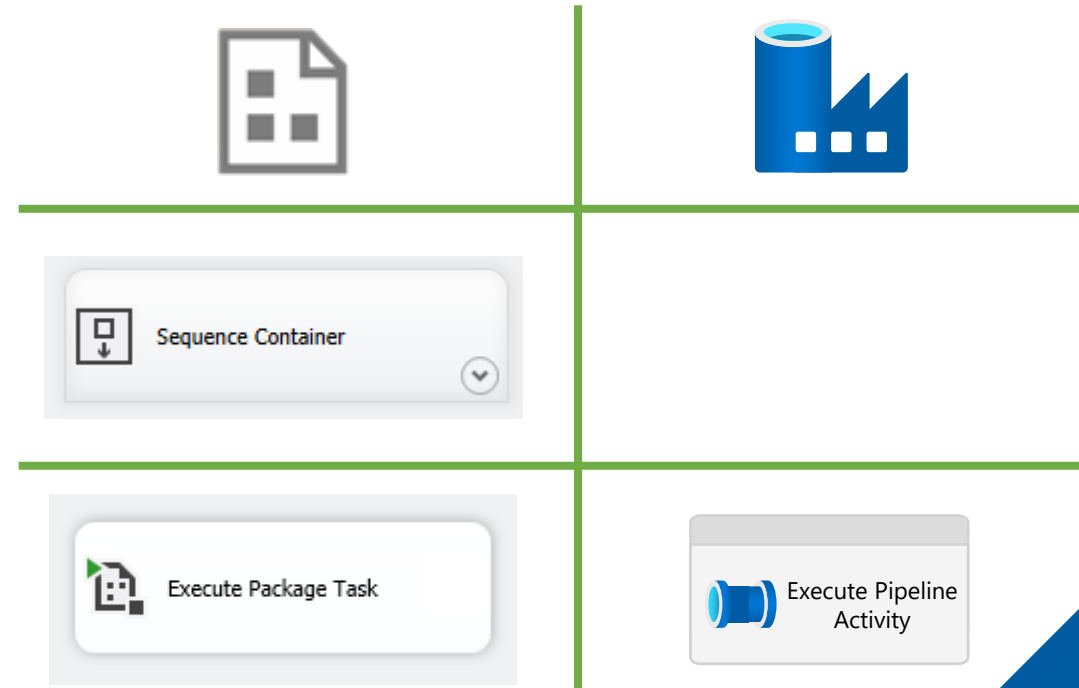


1 **Linked Services**

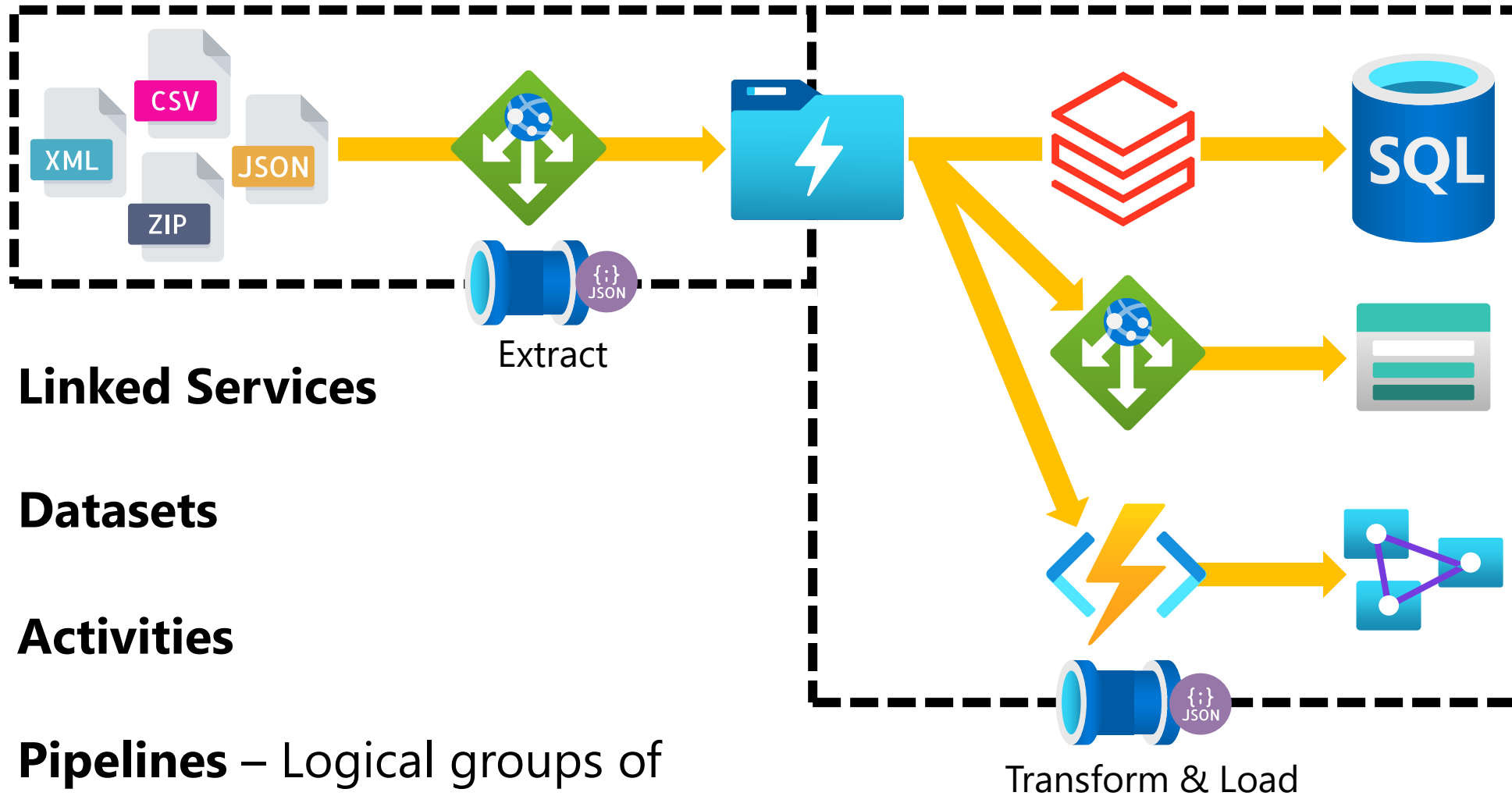
2 **Datasets**

3 **Activities**

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



# Data Factory Components



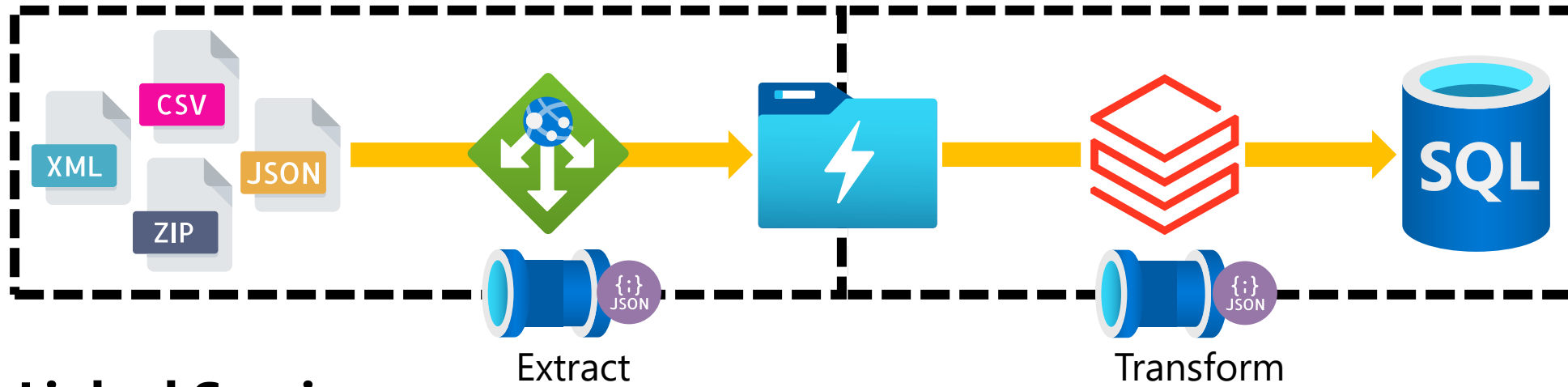
**1 Linked Services**

**2 Datasets**

**3 Activities**

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

# Data Factory Components



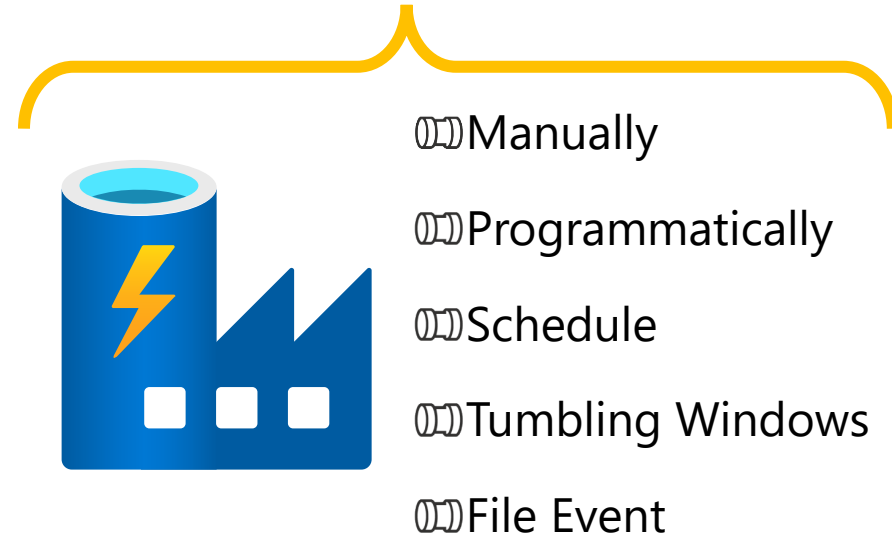
1 **Linked Services**

2 **Datasets**

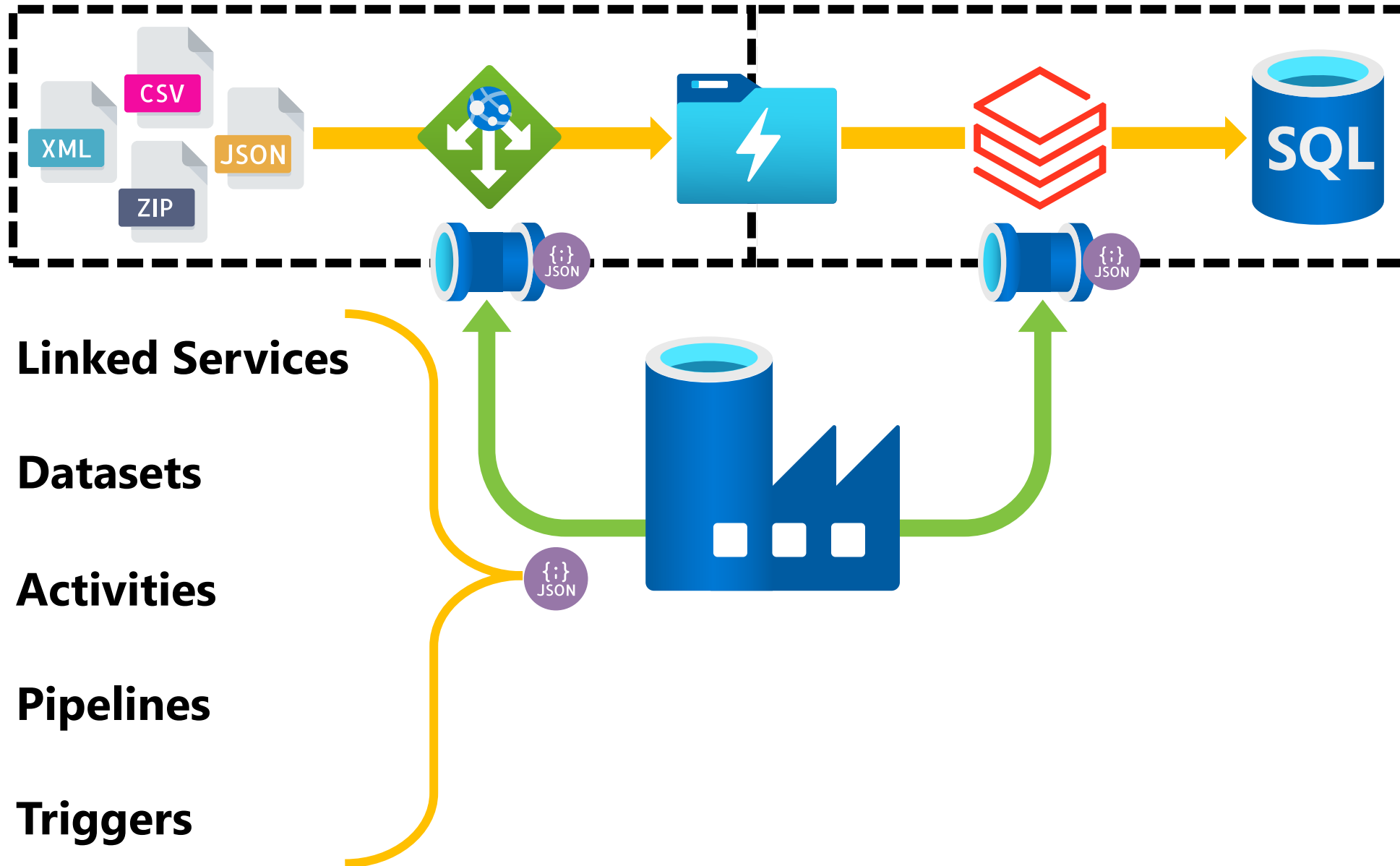
3 **Activities**

4 **Pipelines**

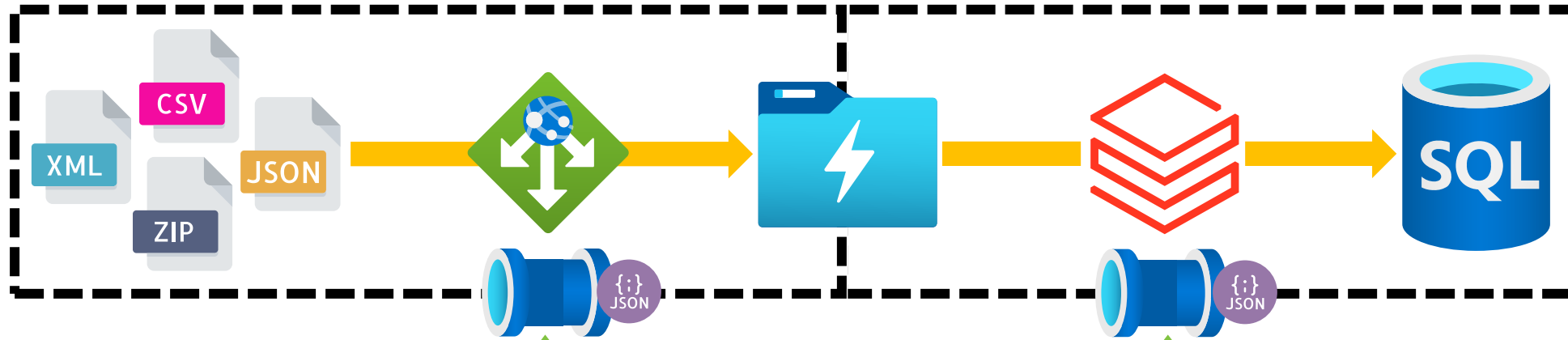
5 **Triggers** – Telling our when pipelines to run.



# Data Factory Components



# Data Factory Control Flow Components



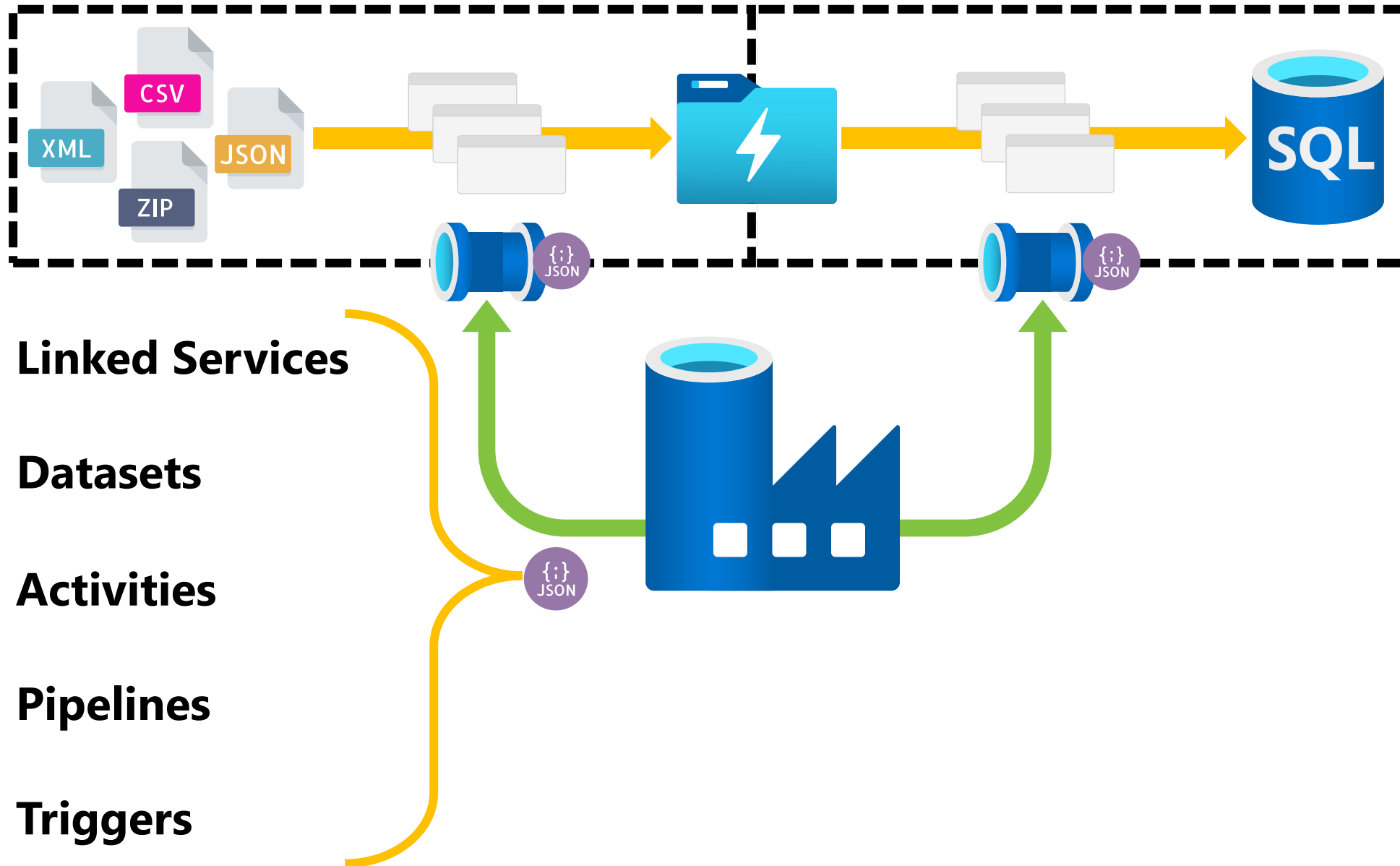
- 1 Linked Services
- 2 Datasets
- 3 Activities
- 4 Pipelines
- 5 Triggers



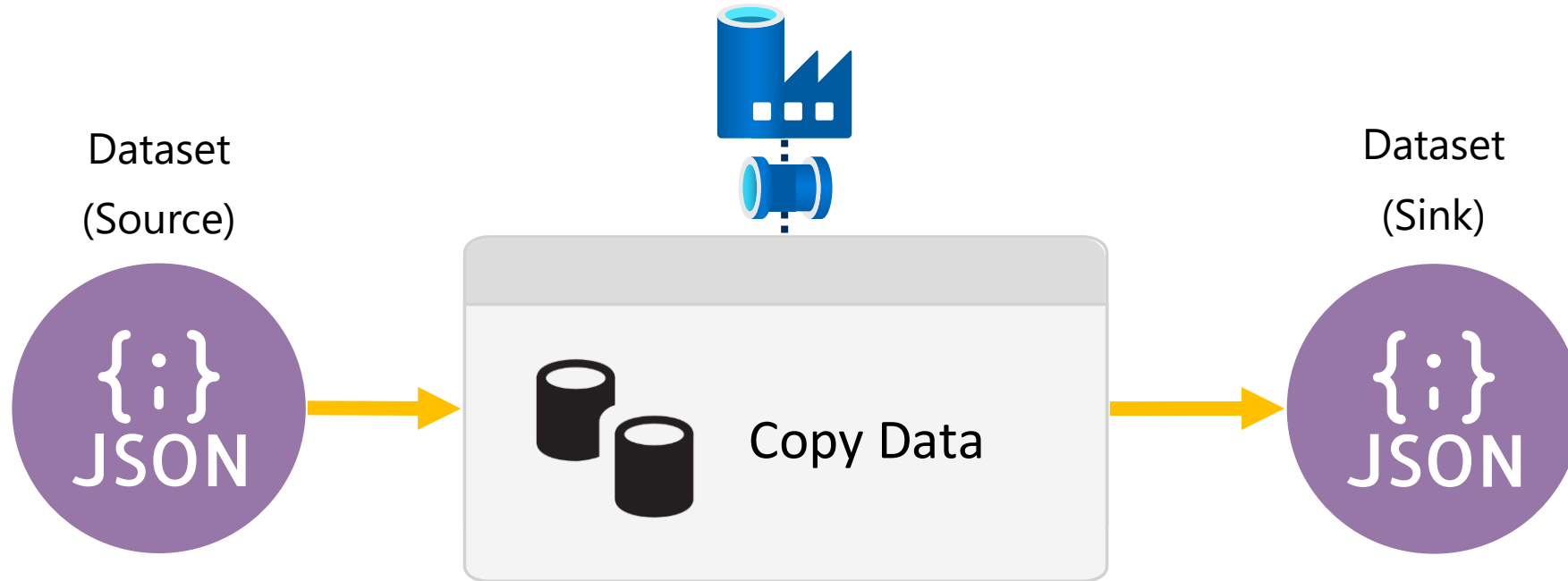
Common Activities

*Paul's Favourites*

# Data Factory Common Activities



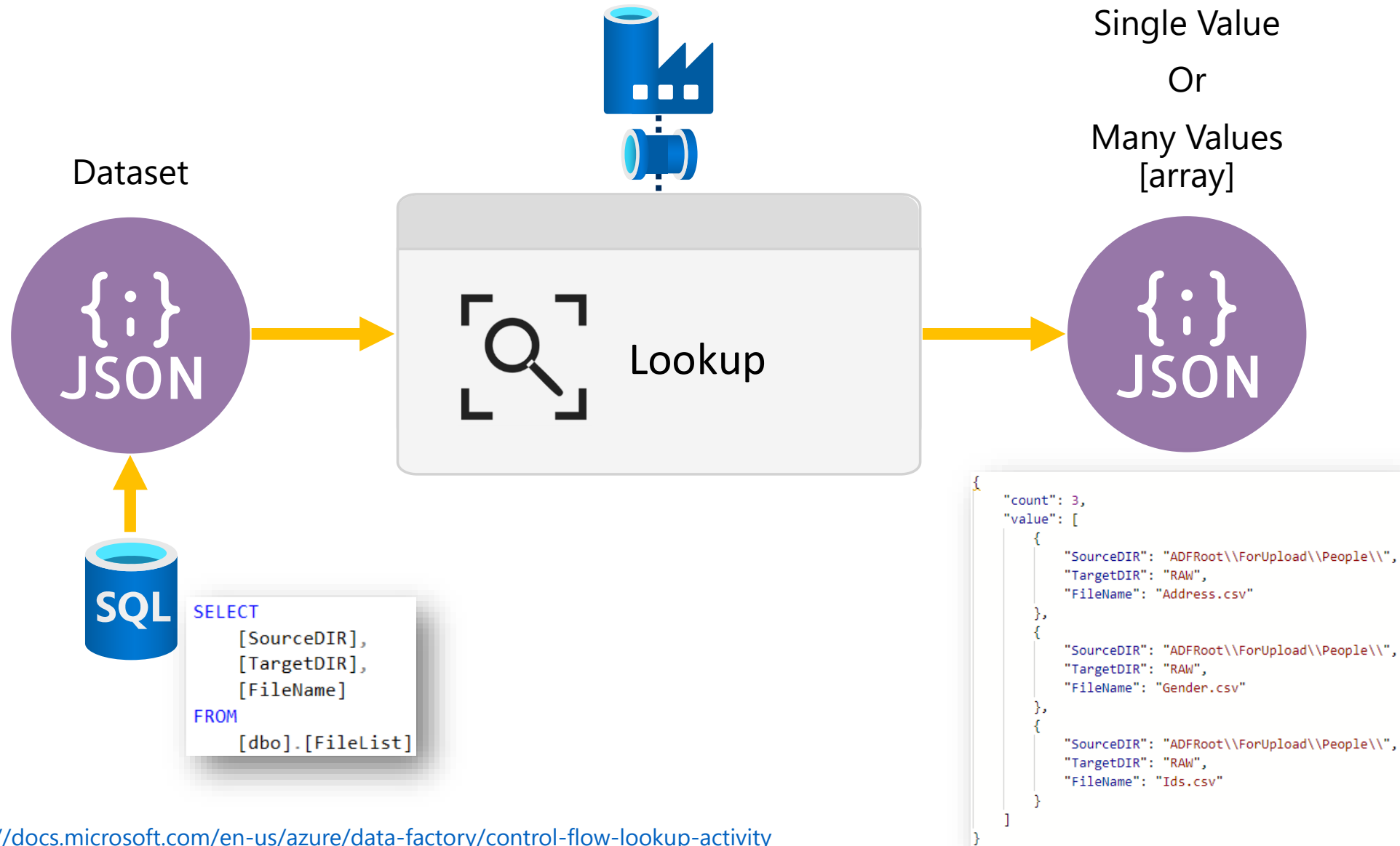
# Copy



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

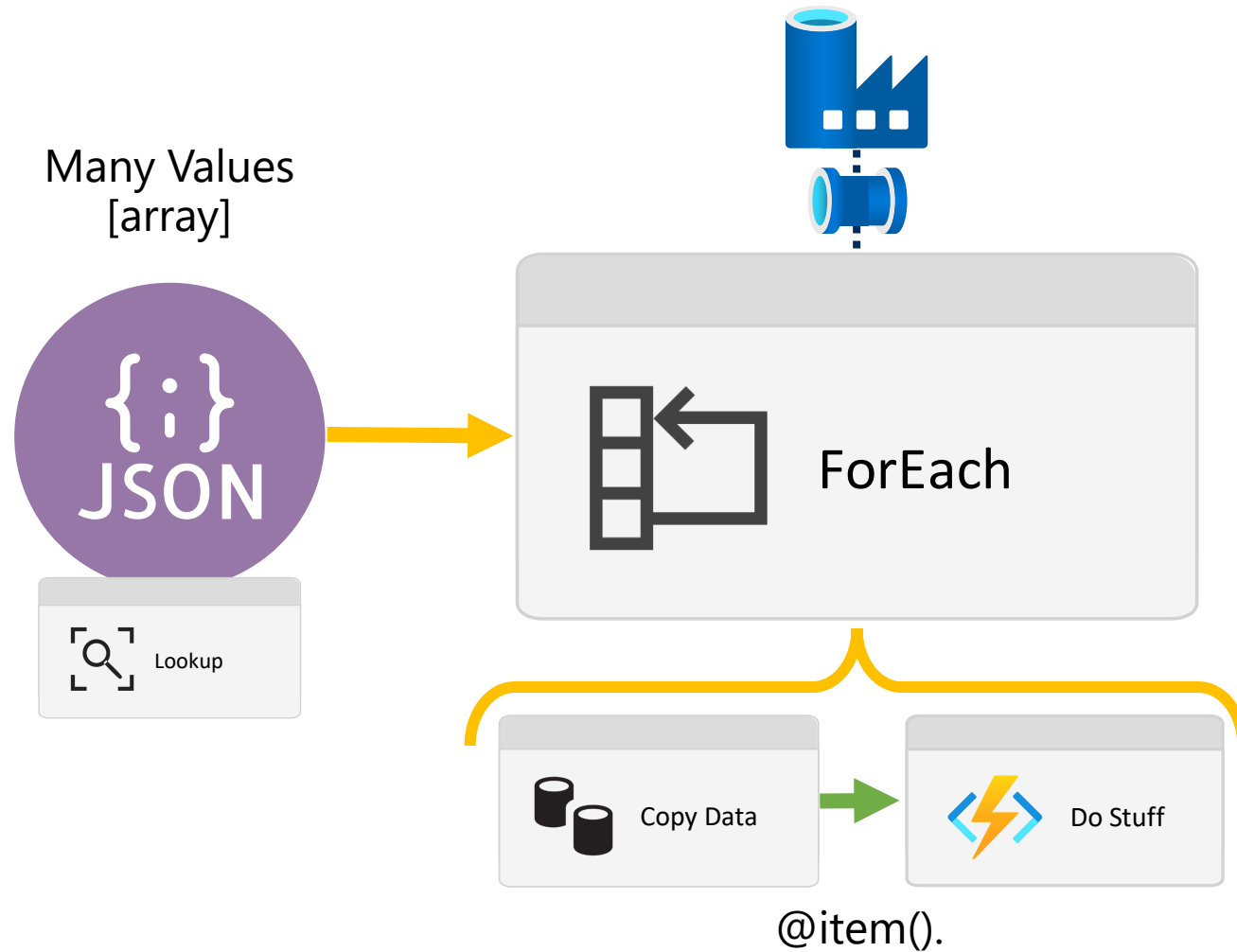
# Lookup

Get value to support other control flow activities

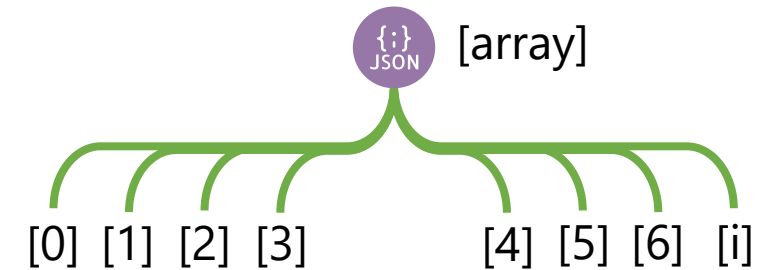
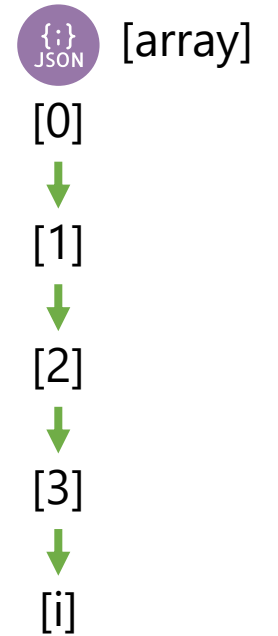


# ForEach

## Scaling Out Control Flow Activities



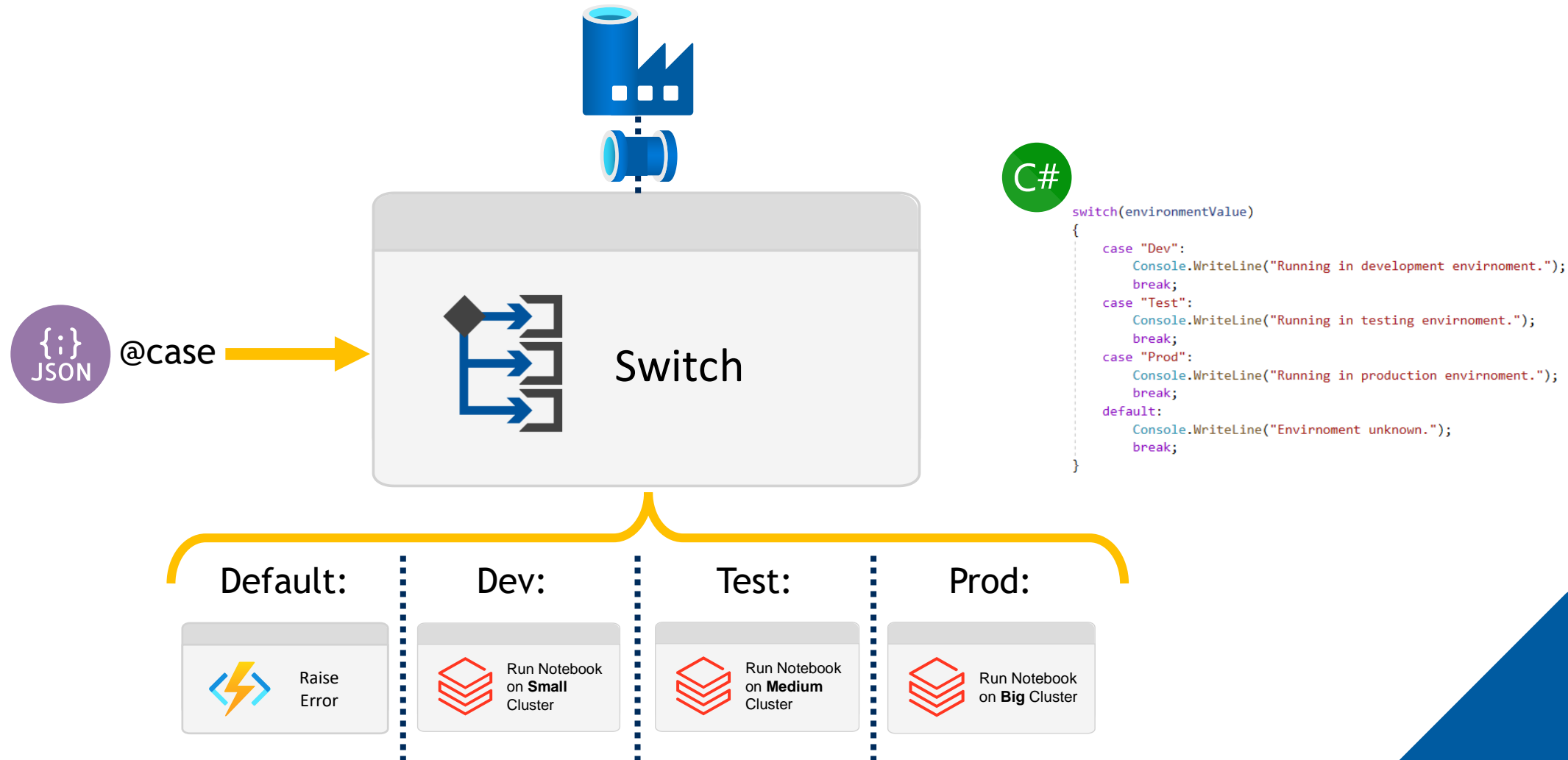
IsSequential:  
true



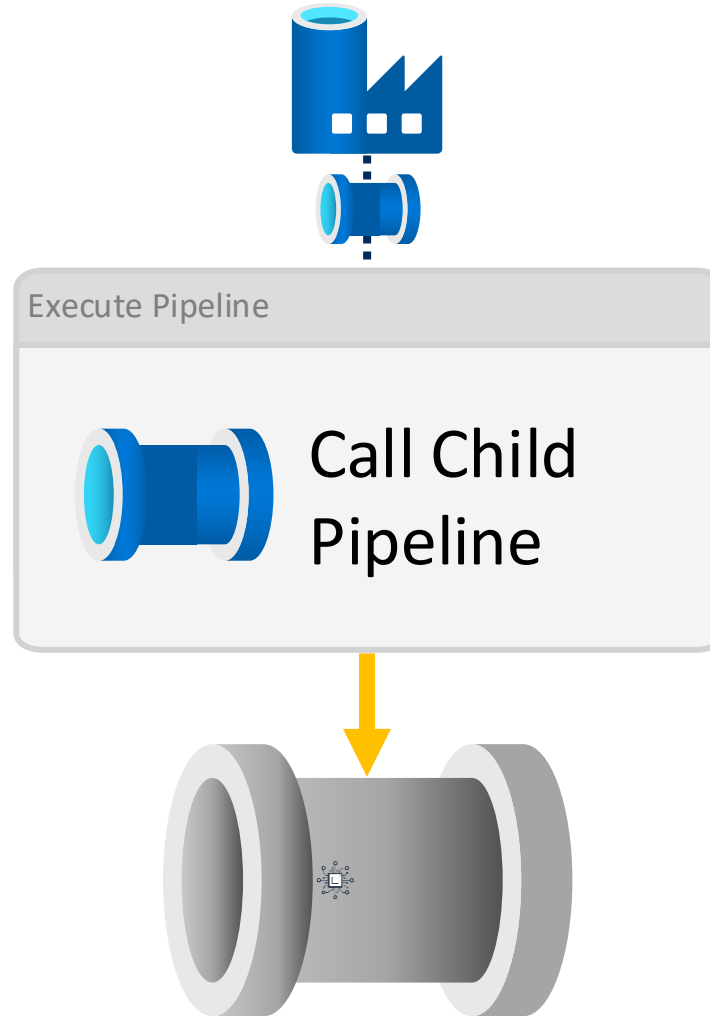
Batch Count Default: 20

Batch Count Max: 50

# Switch



# Execute Pipeline



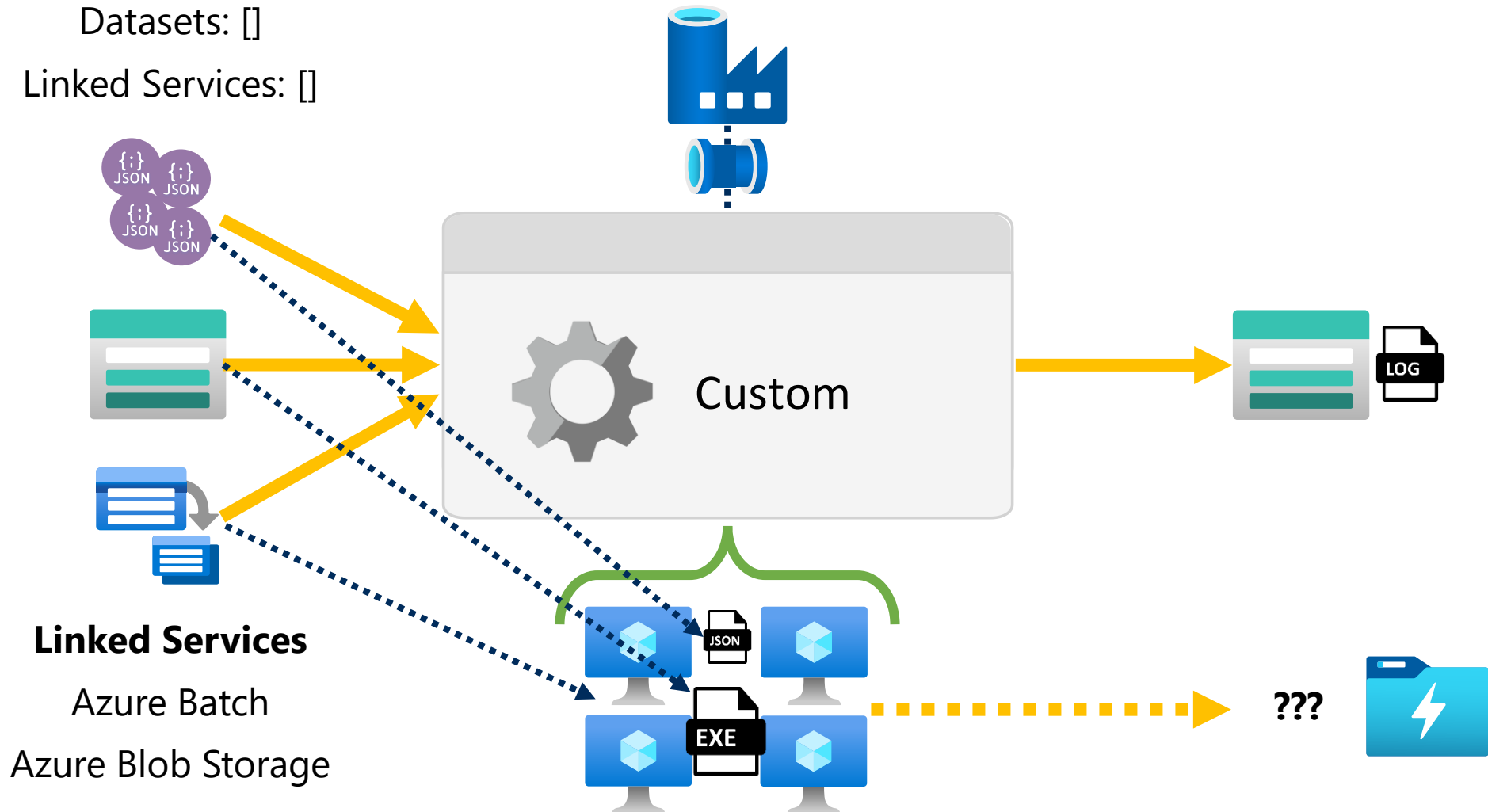
# Custom

Extend Data Factory with Custom Code

## References Objects

Datasets: []

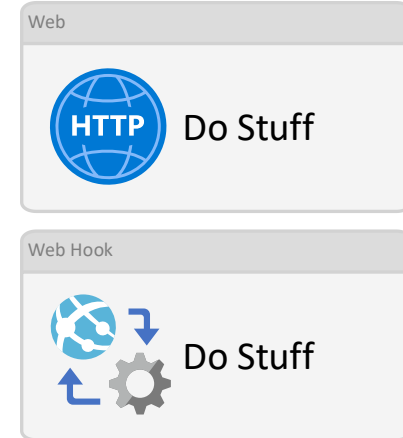
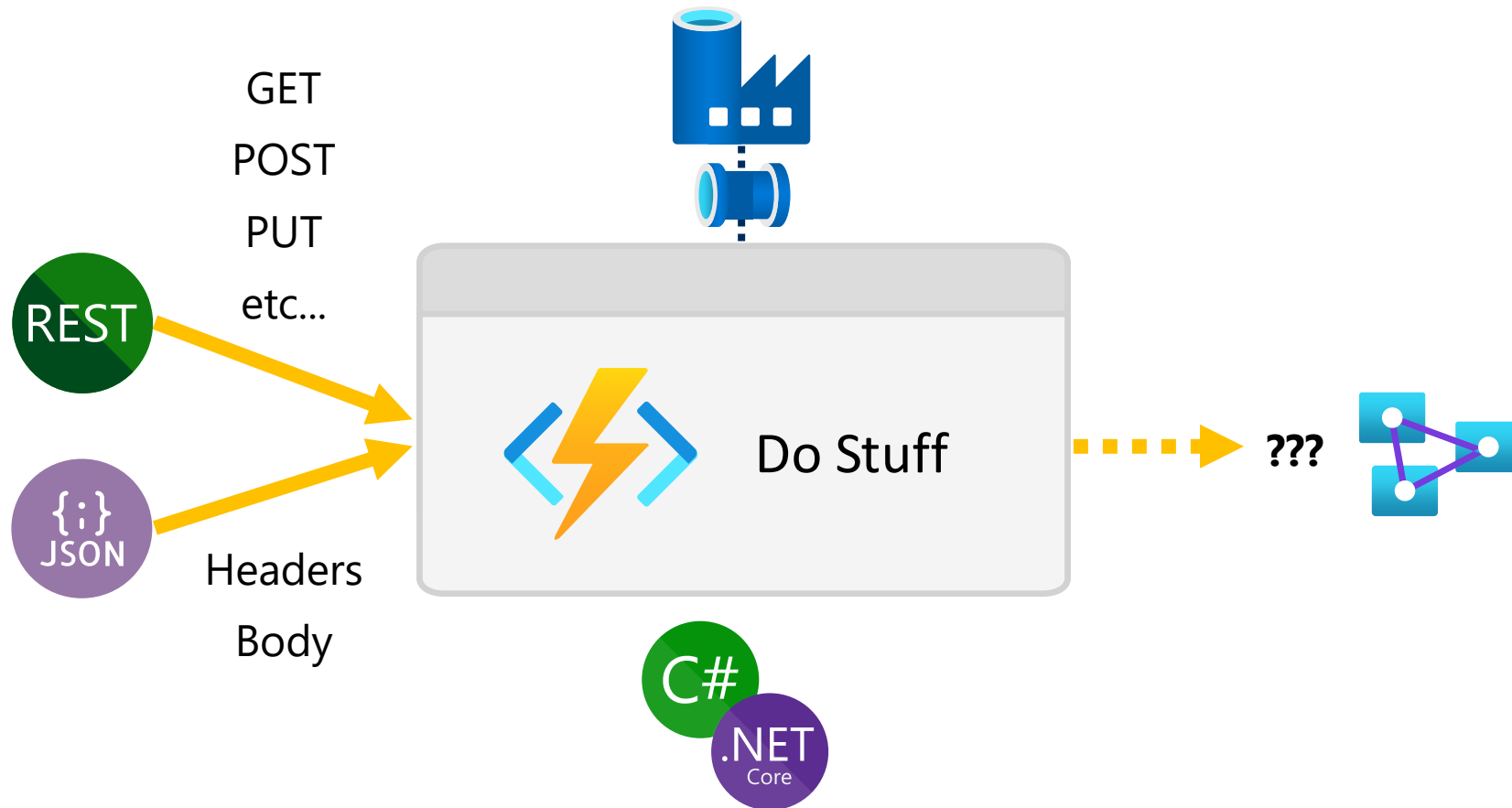
Linked Services: []





# Azure Function

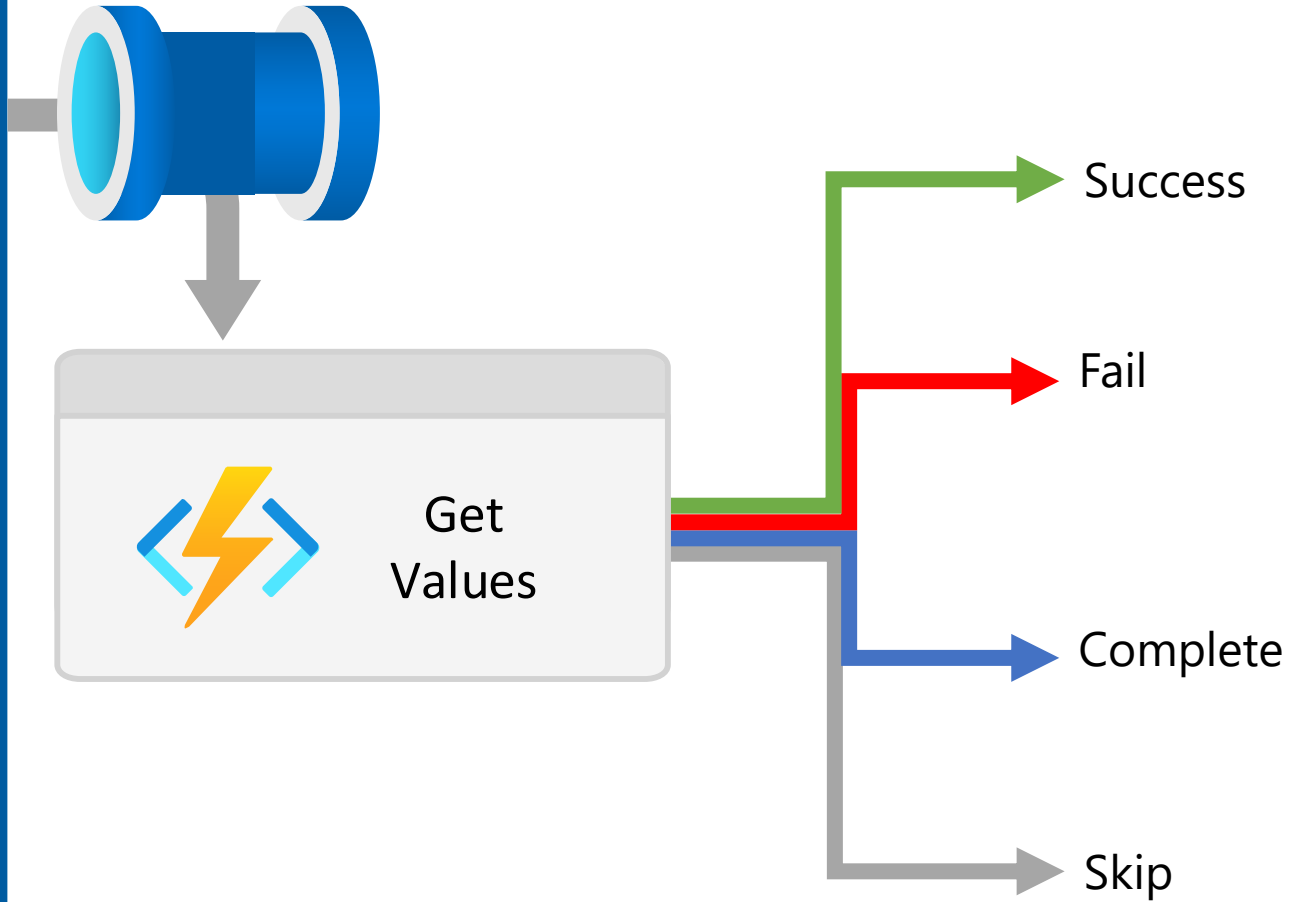
Extend Data Factory with Rest Calls



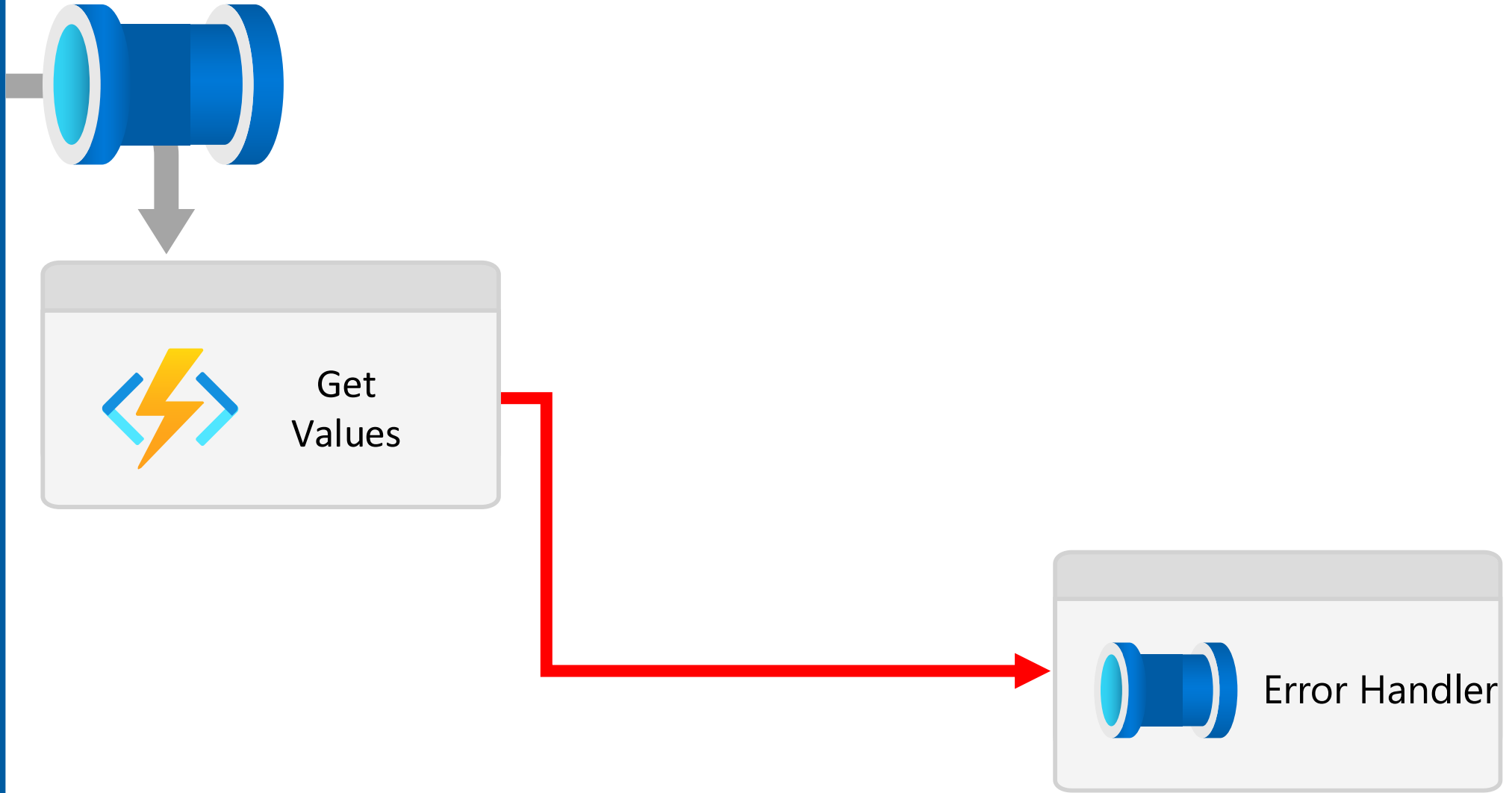
# Execution Dependencies

A large, solid blue triangle is positioned in the bottom right corner of the slide, pointing towards the top right. It serves as a decorative background element.

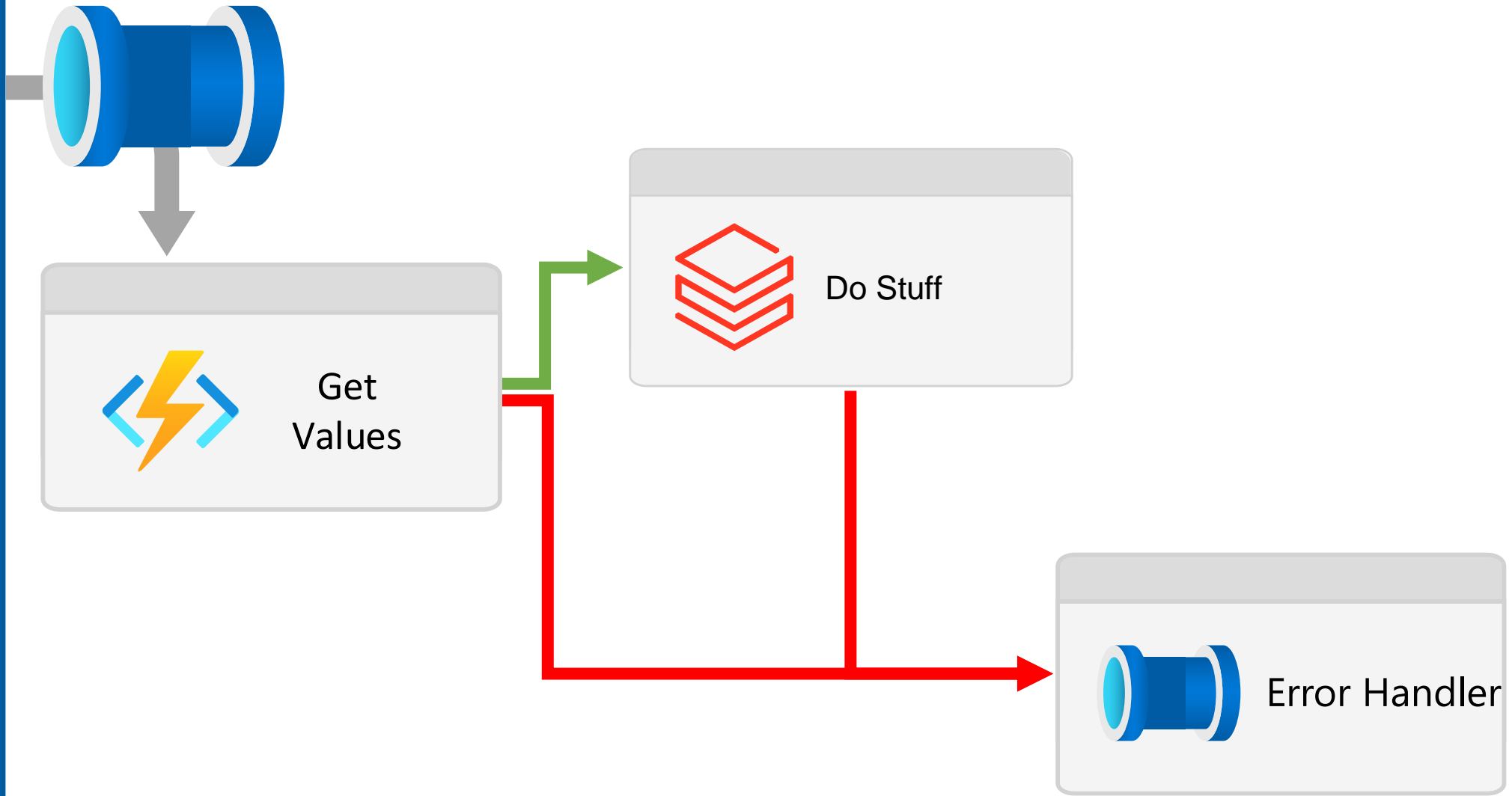
# Execution Dependency Options



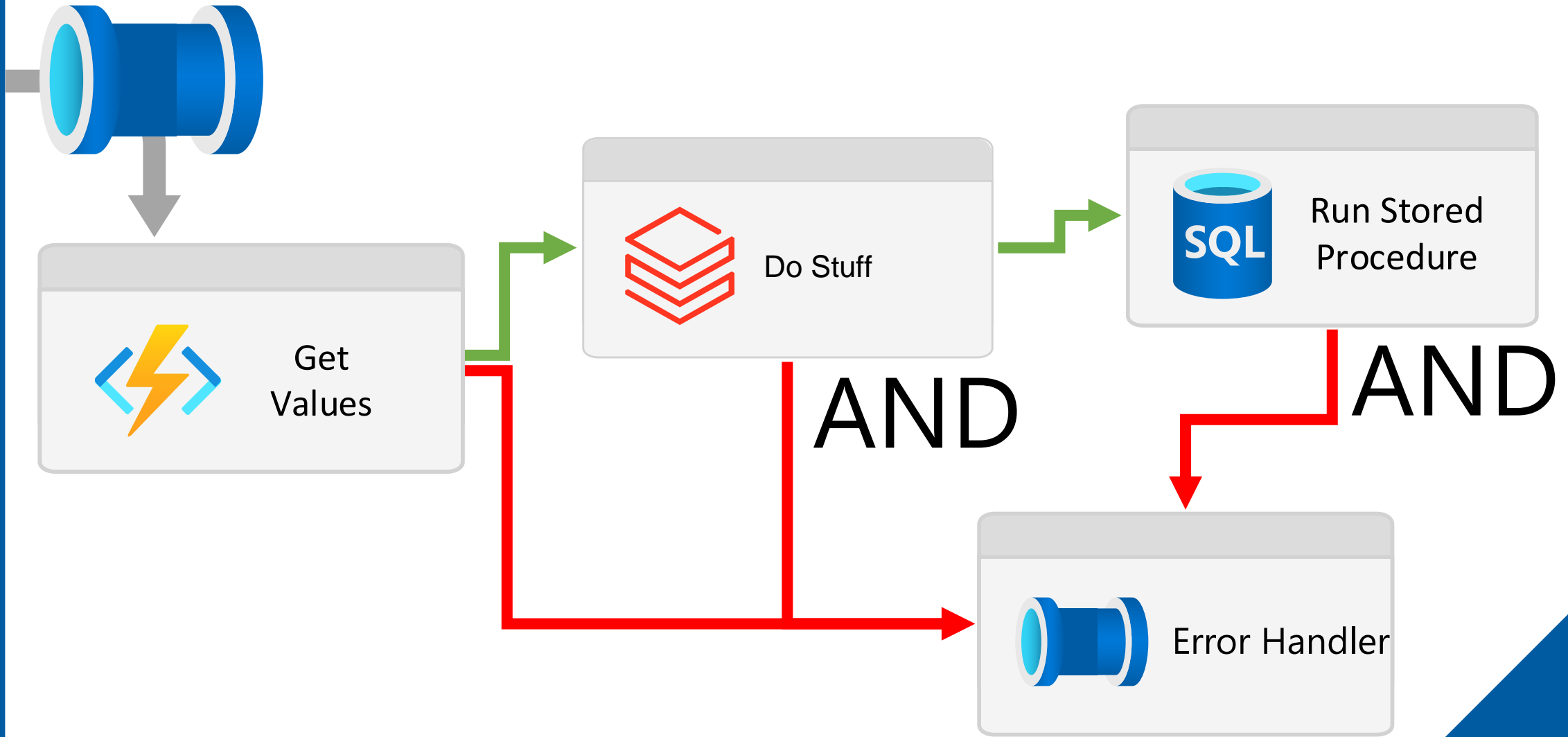
# Execution On Failure



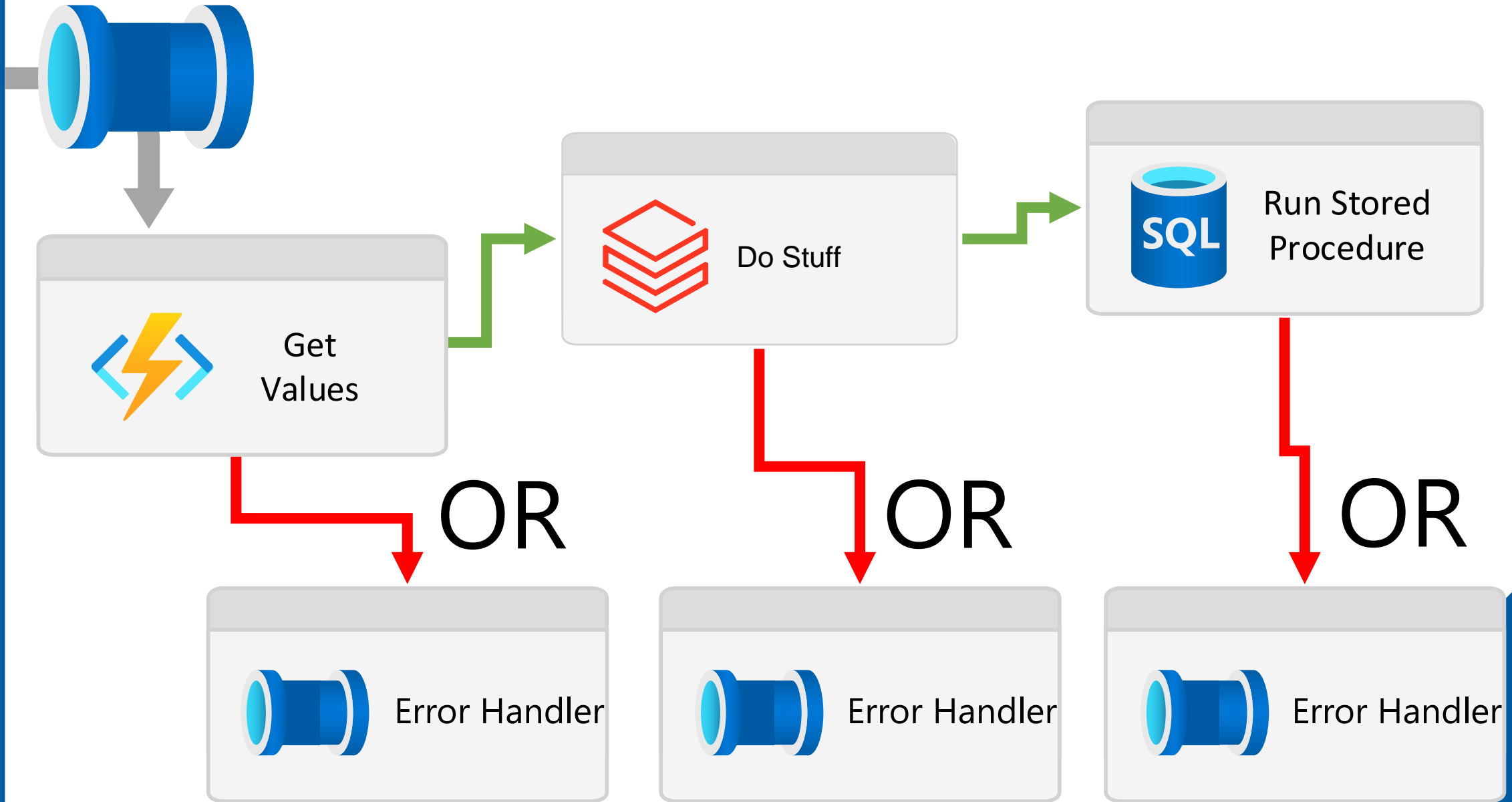
# Execution On Failure or On Success



# 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 ✓