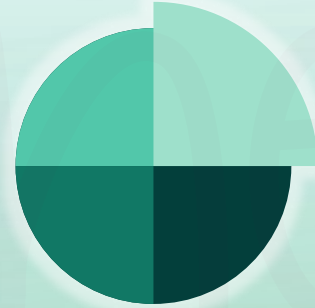




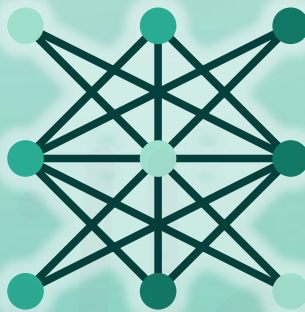
**Data Warehousing**



**Big Data and Analytics**



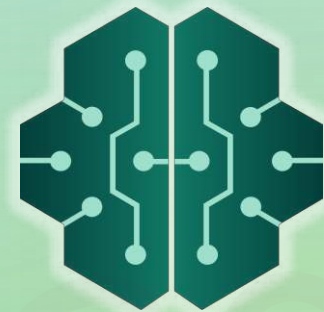
**Business Intelligence**



**Artificial Intelligence**



**Data Science**



**Machine Learning**



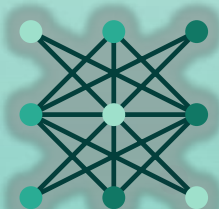
**Data Warehousing**



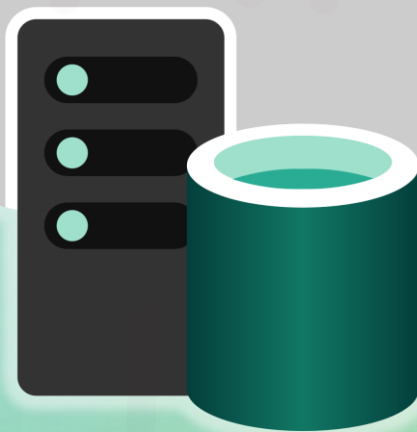
**Big Data and Analytics**



**Business Intelligence**



**Artificial Intelligence**



**Machine Learning**



**Data Science**

What has happened?

# What?

What will happen?

When did it happen?

# When?

When will it happen?

# Why?

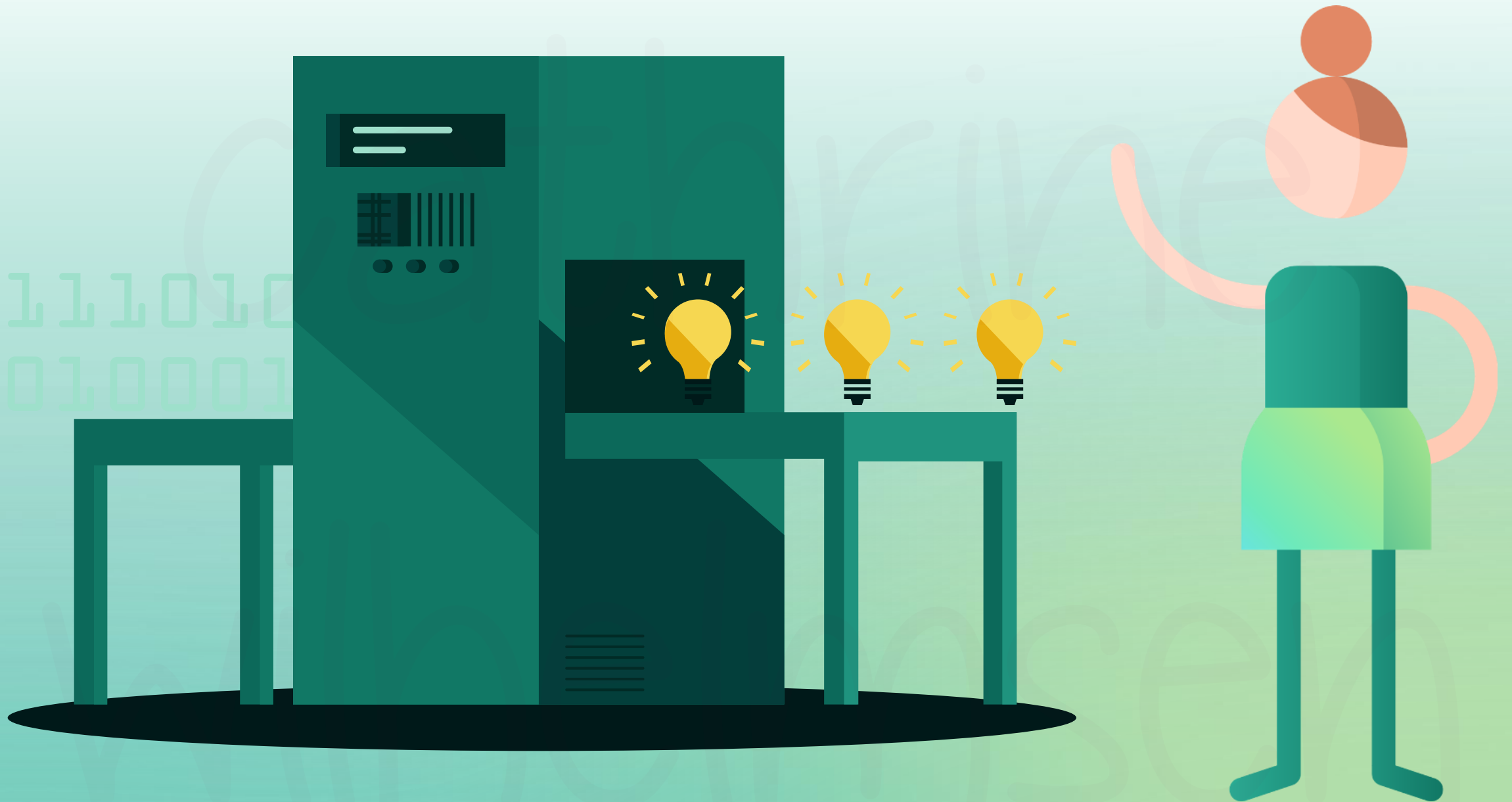
Why did it happen?





**Collect**  
**Store**  
**Transform**  
**Integrate**  
**Prepare**







# Data Integration using Data Factory in Microsoft Fabric

Cathrine Wilhelmsen • Microsoft Fabric Week • February 19th, 2024



# Cathrine Wilhelmsen

*She / Her*

Solutions Architect, Evidi



I love data and coding, as well as teaching and sharing knowledge 🧐

✉ hi@cathrinew.net

🏠 cathrinew.net

💬 @cathrinew

★ Microsoft Data Platform MVP

❄ Organizing Fabric February

🔧 Renovating a house



# Microsoft Fabric



# What is Microsoft Fabric?

## End-to-end Analytics:

- Software-as-a-Service
- Lake-centric architecture
- Role-tailored tools





**Data  
Factory**



**Data  
Engineering**



**Data  
Warehouse**



**Data  
Science**



**Real-Time  
Analytics**



**Power  
BI**



**Data  
Activator**



**OneLake storing one copy of all your data in an open format**

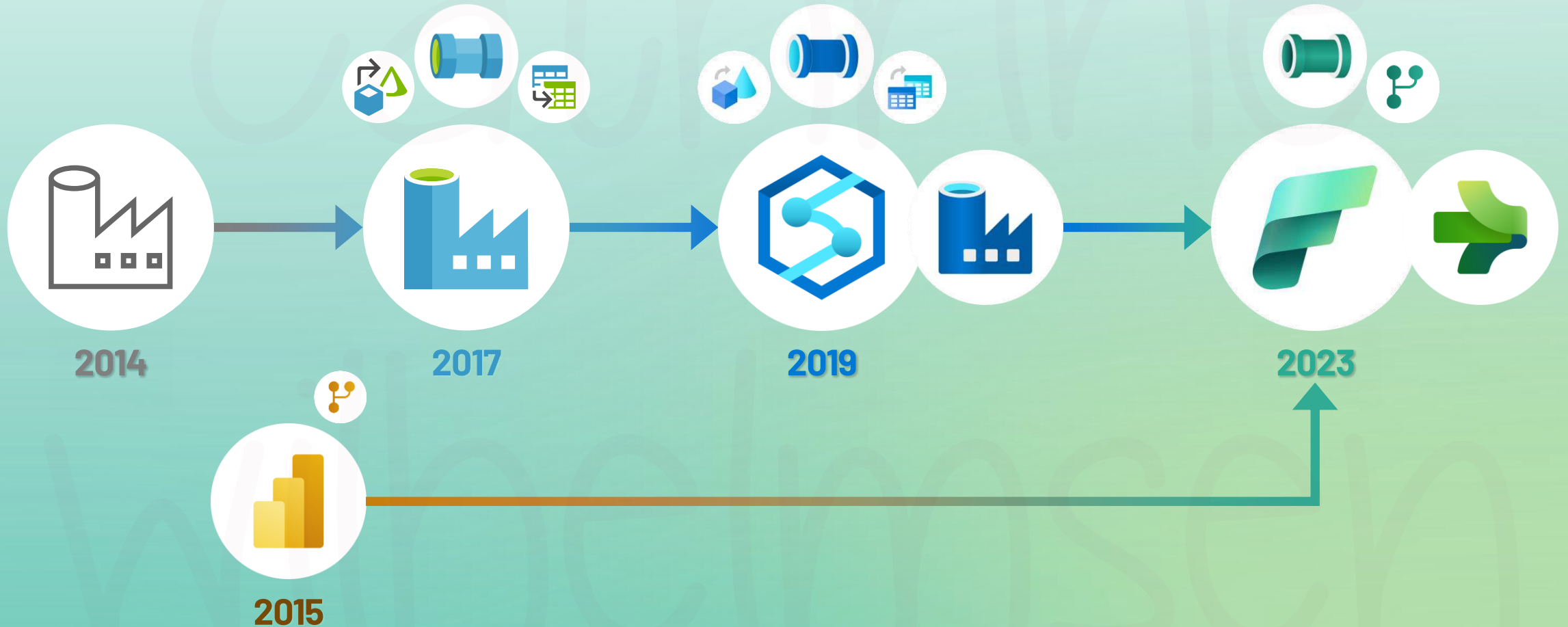


**Built-in security, governance and compliance**



# Data Factory

# The evolution of Data Factory



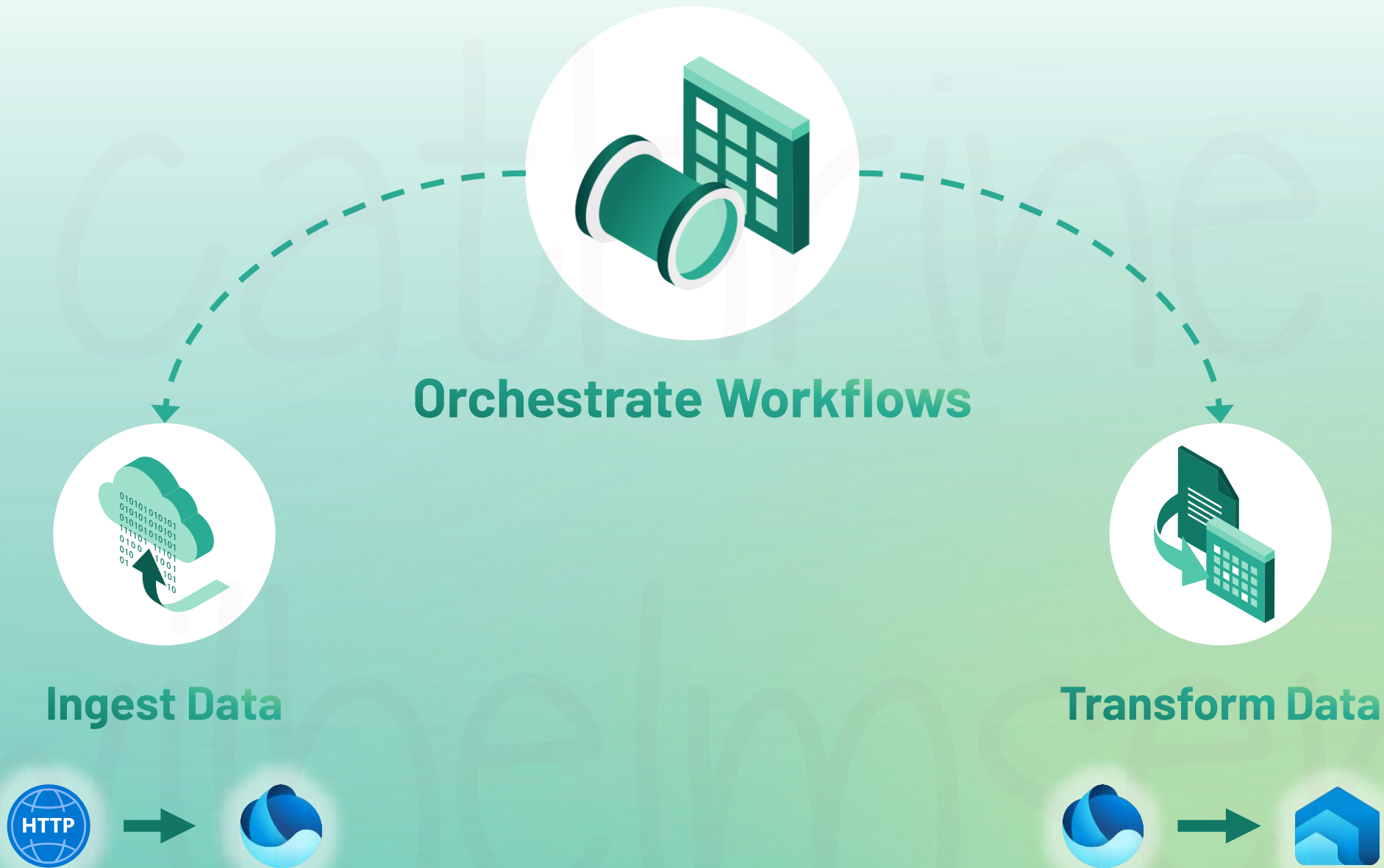


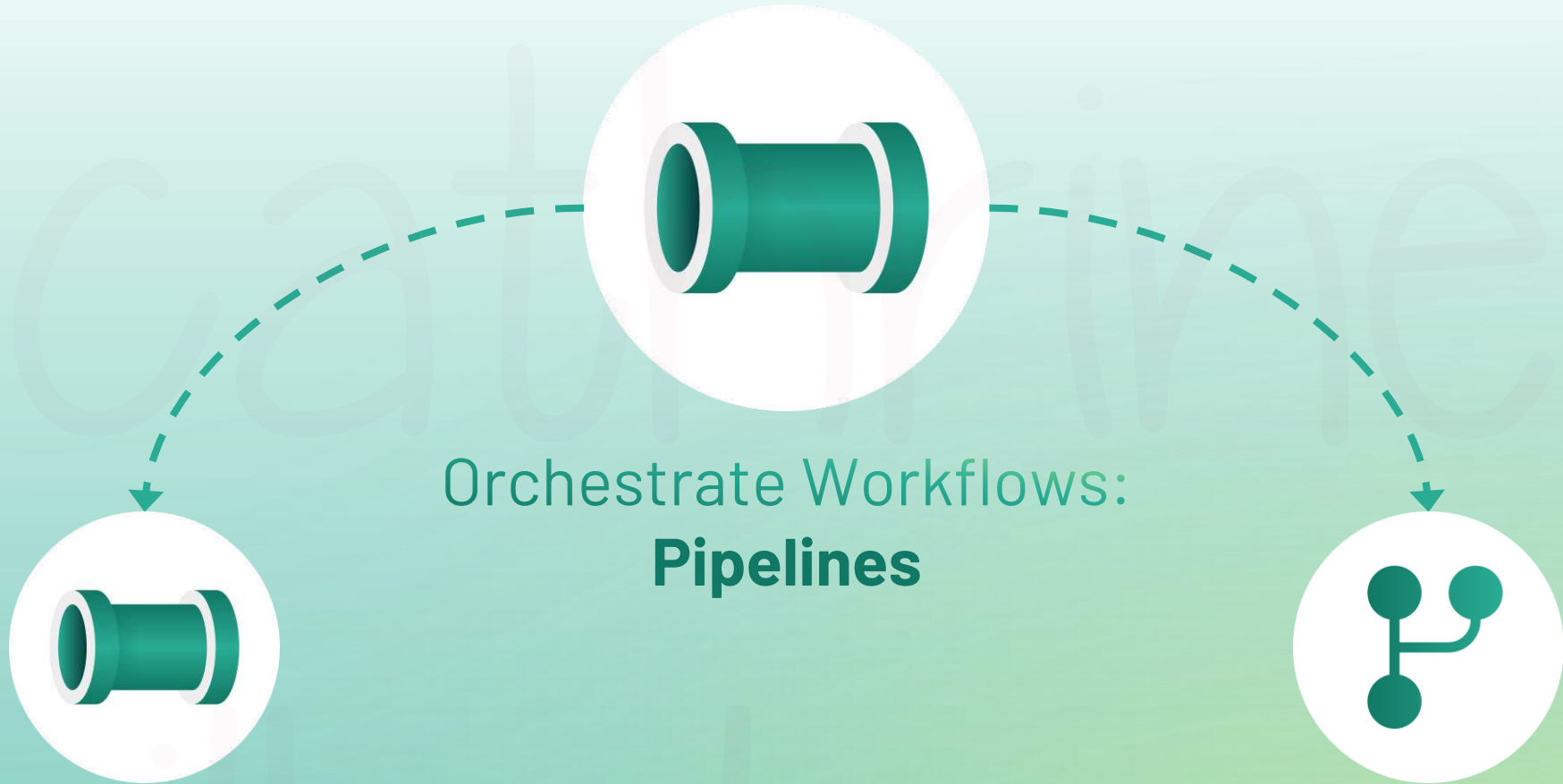
**Ingest Data**



**Transform Data**







Ingest Data:  
**Pipelines**

Orchestrate Workflows:  
**Pipelines**

Transform Data:  
**Dataflows**

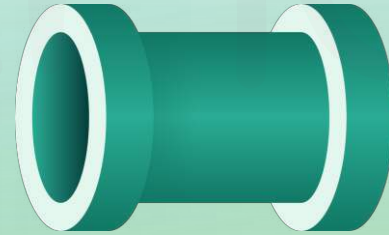


# Pipelines



# What are Pipelines?

Pipelines are what  
you execute or run



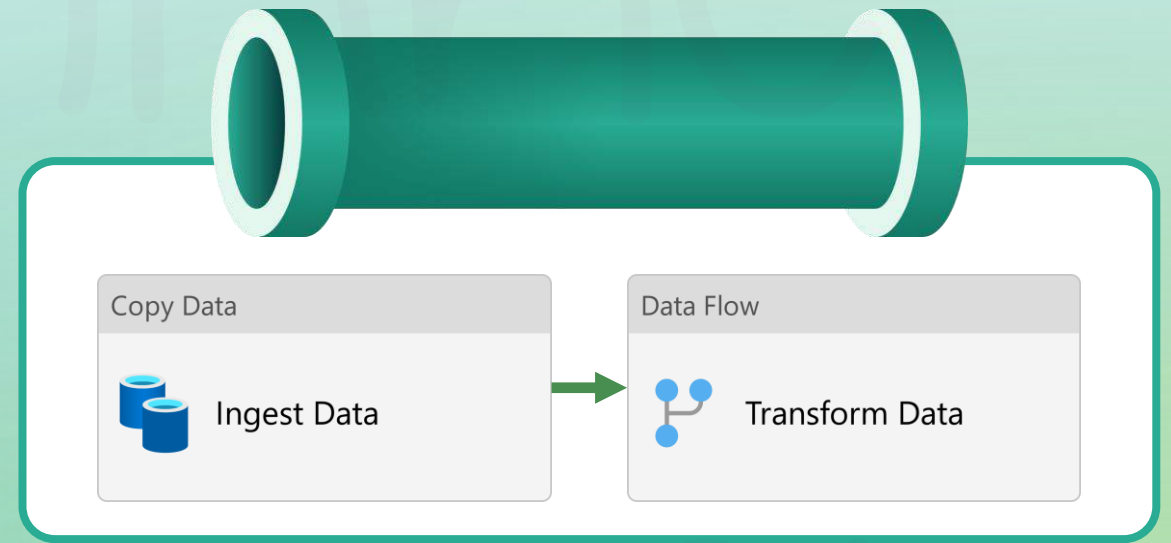
They define workflows:  
what to do in which order



# What are Activities?

Activities are individual steps inside pipelines

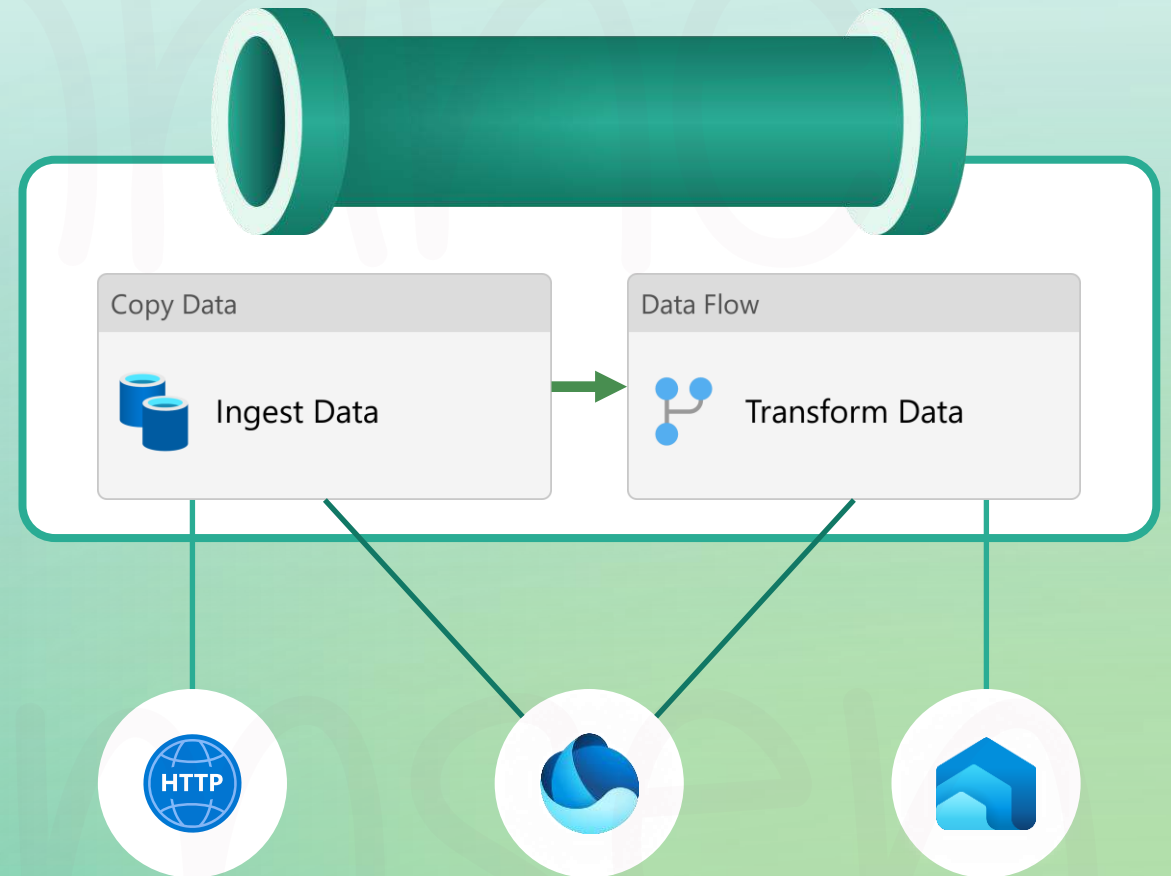
They can be chained or run in parallel



# What are Connections?

Connections specify how to connect and authorize

Connect to workspace or external source



# Copy Data Activity

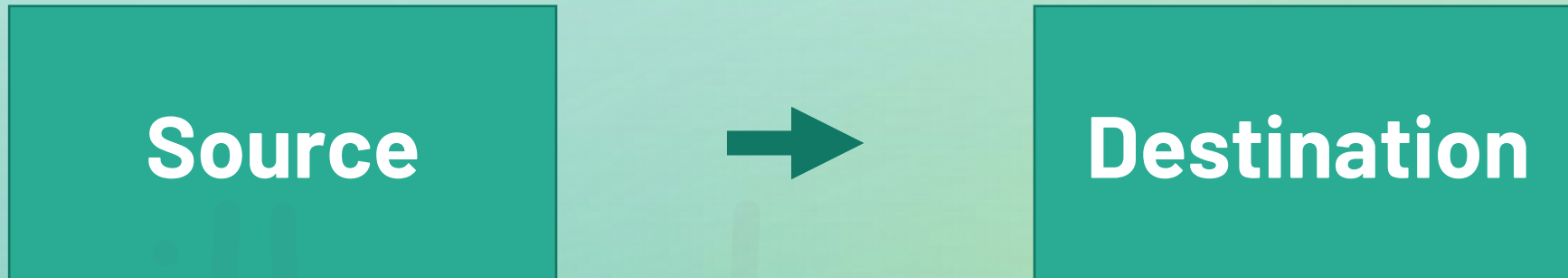
Powerful built-in capabilities

Copy Data



Copy Data

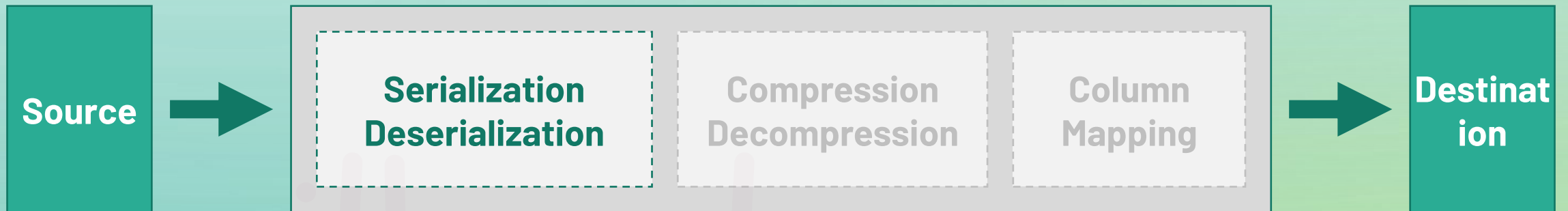
# Copy Data Activity: Binary Files



# Copy Data Activity: Complex Data



# Copy Data Activity: Complex Data



Convert file formats

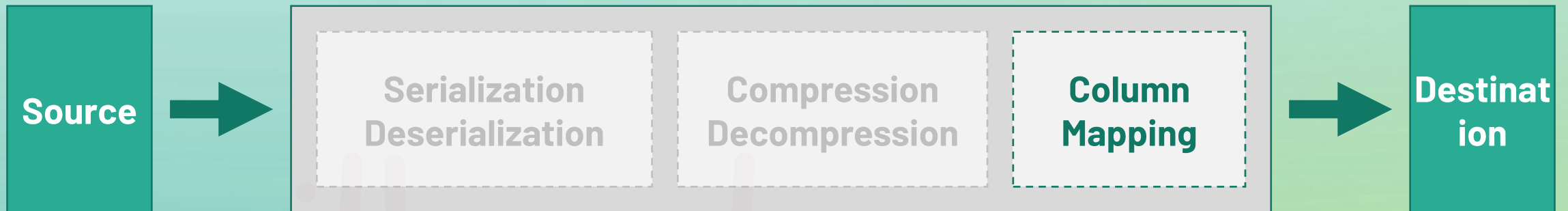
# Copy Data Activity: Complex Data



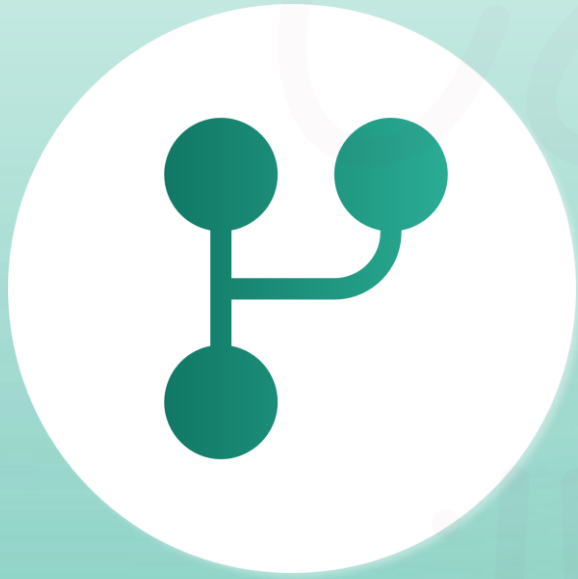
Zip or unzip files



# Copy Data Activity: Complex Data



Map columns implicitly or explicitly



# Dataflows

# What are Dataflows?

Power Query!

Data transformation at scale

Visual editor, low-code experience

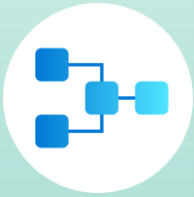
# Let's demo!





# Orchestration

# What does orchestration mean?



## Defining Workflows

Which activities to run in which order?



## Configuring Alerts and Error Handling

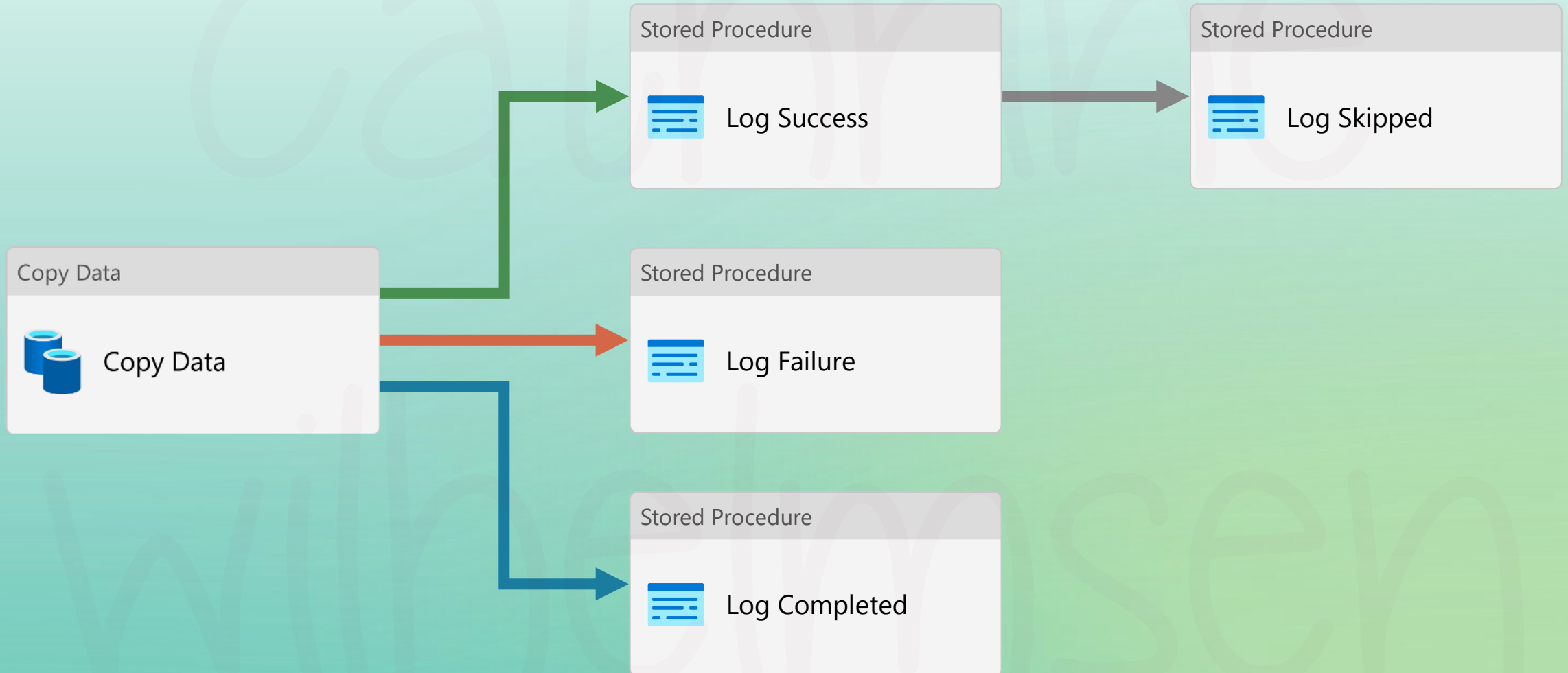
How to handle unexpected results and failures?



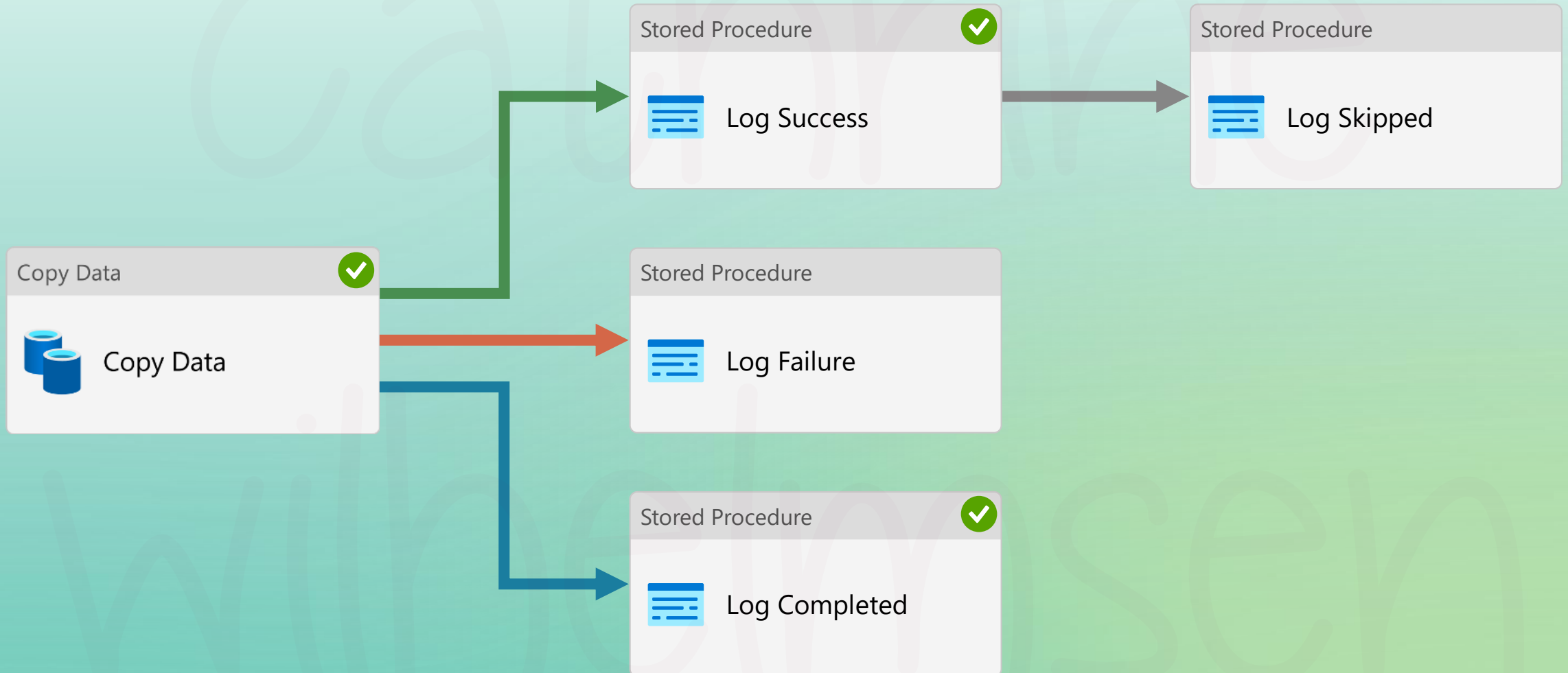
## Adding Schedules

When to execute pipelines?

# Activity Dependencies

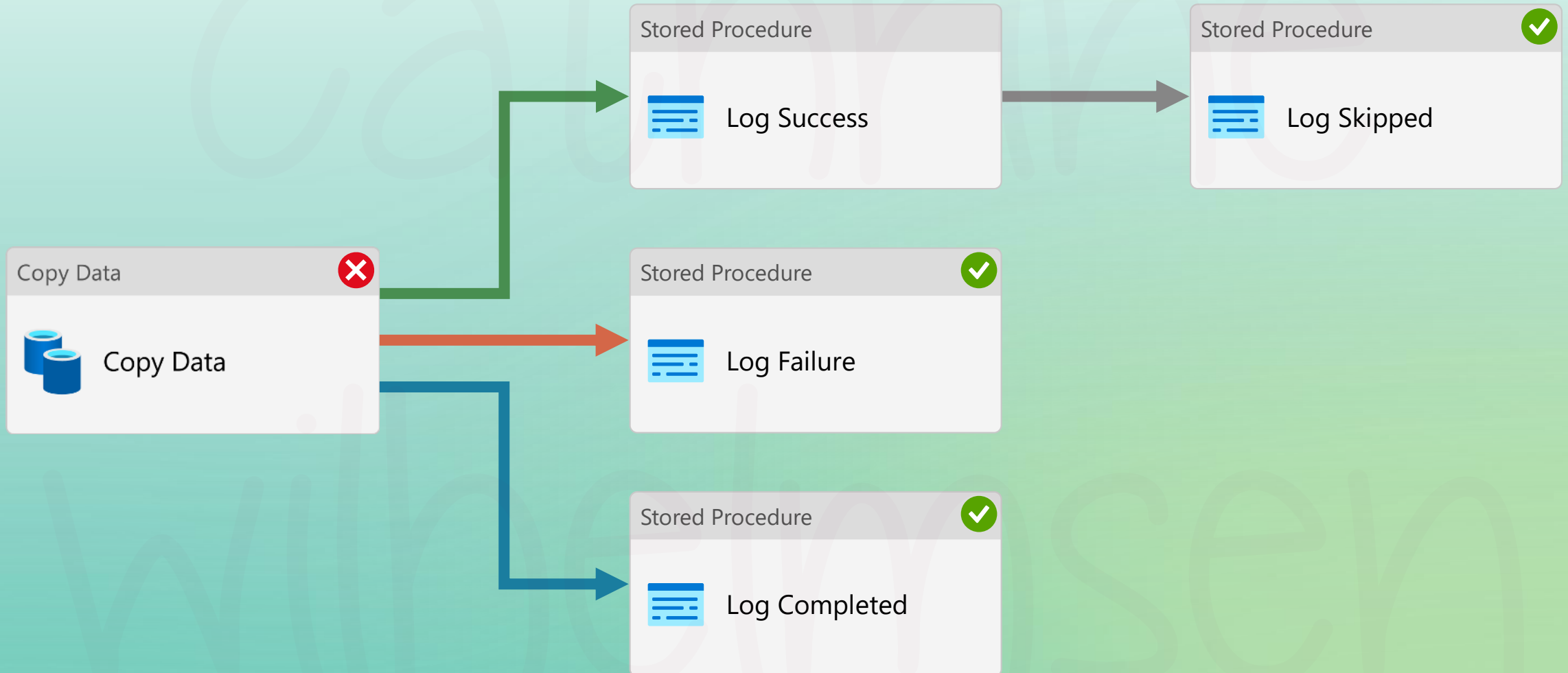


# Activity Dependencies

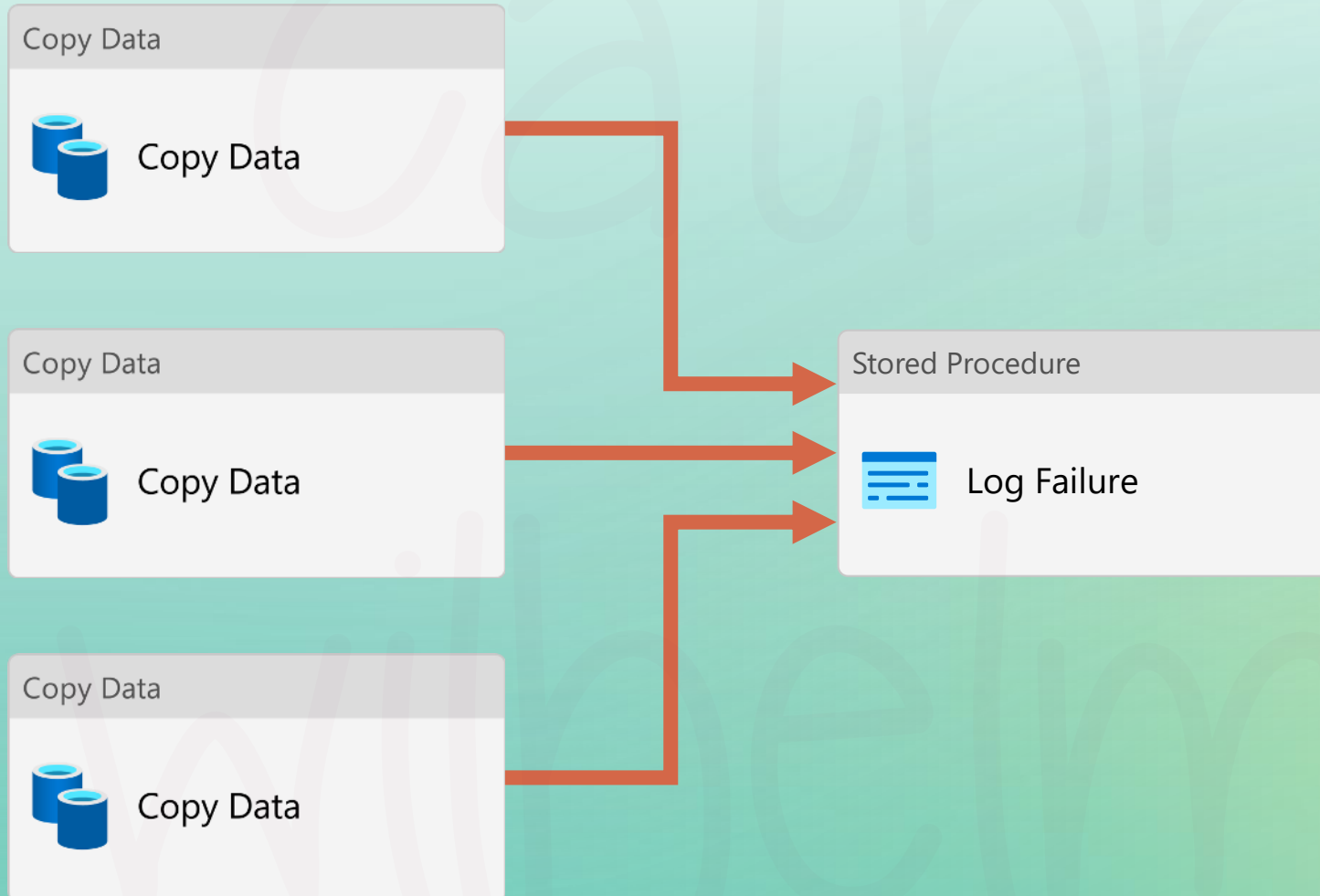




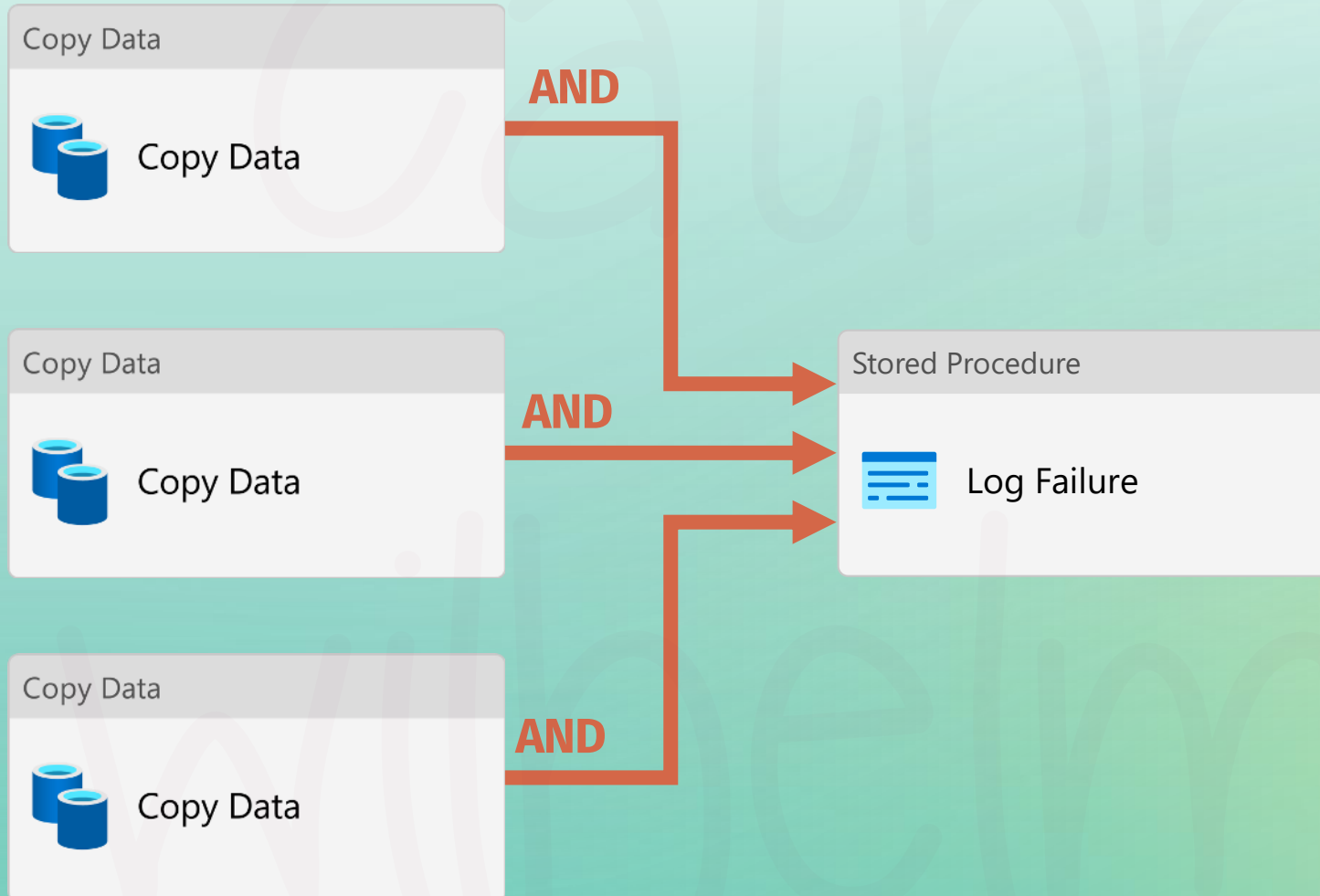
# Activity Dependencies



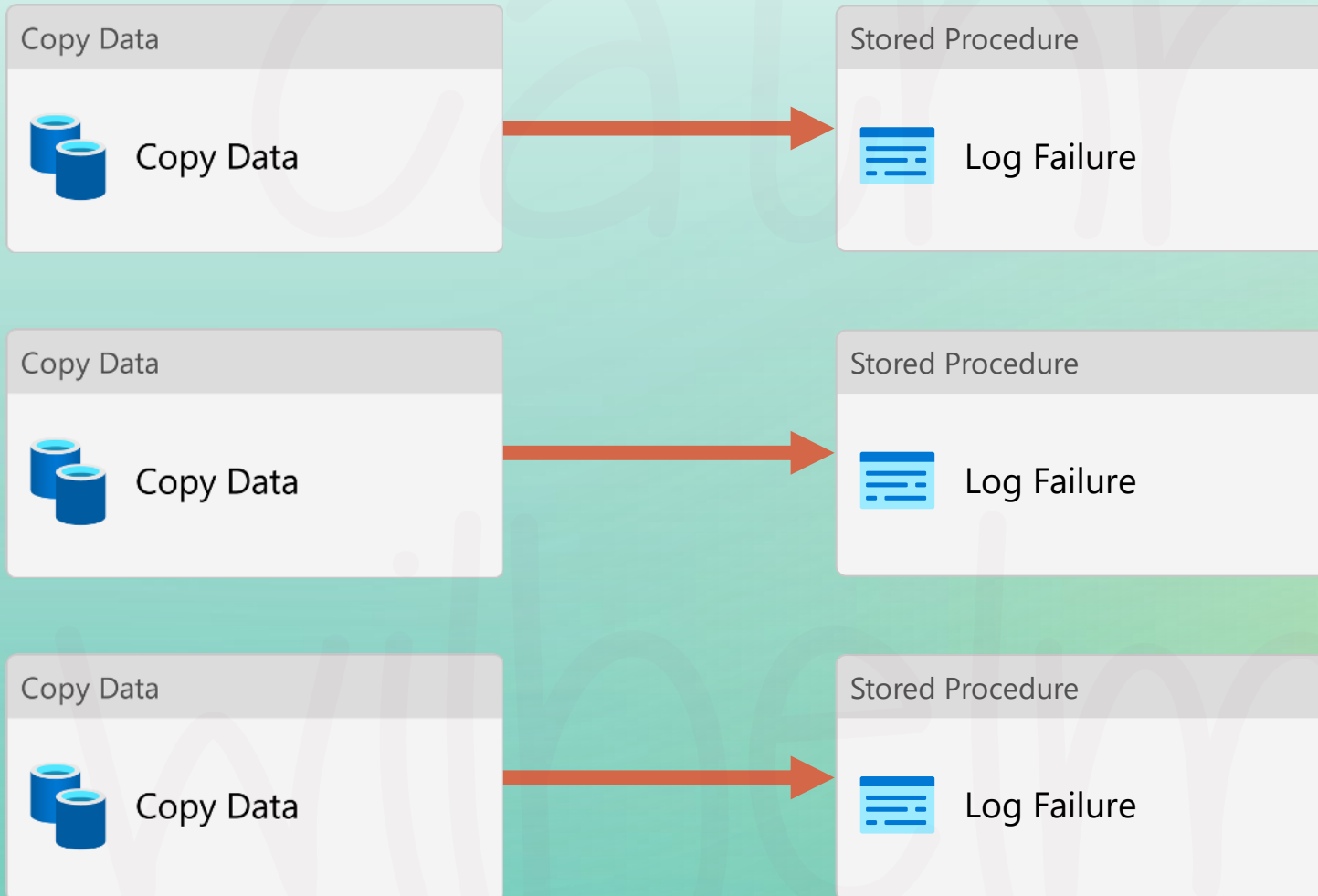
# Activity Dependencies: Logical... ?



# Activity Dependencies: Logical AND



# Activity Dependencies: Logical OR



# Questions?



# Thank you!



hi@cathrine.wilhelmsen.net



cathrine.wilhelmsen.net



@cathrine.wilhelmsen