



Don't  
repeat  
yourself!

## Metadata-Driven Approach in ADF / Synapse

15.02.2023

Microsoft  
Partner

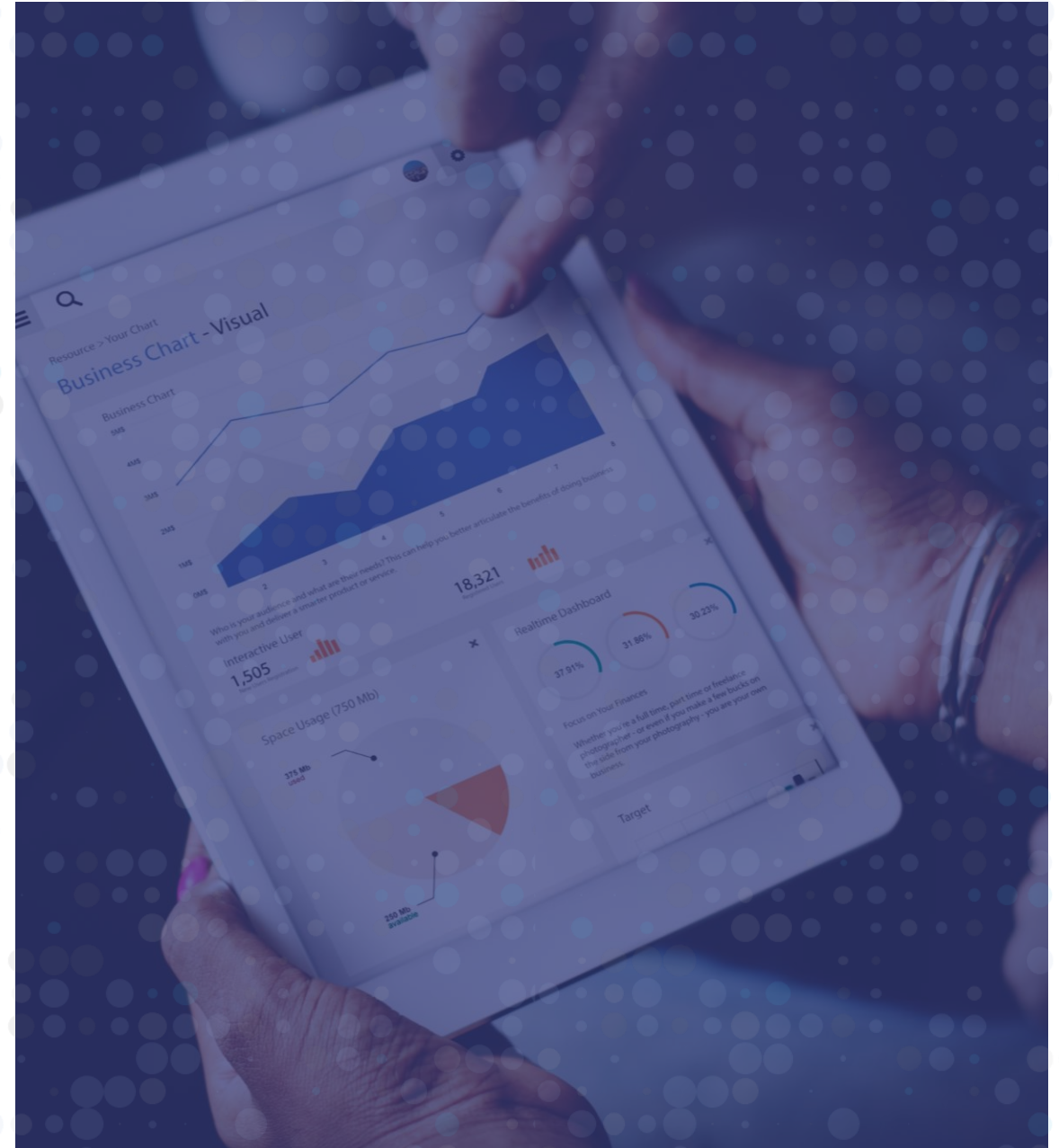


Gold Data Analytics  
Gold Data Platform  
Gold Datacenter  
Gold Cloud Platform  
Gold Cloud Productivity  
Gold Application Development



# Agenda

- Introduction
- Metadata-Driven Pipelines
- Demo-Case „Adventure Works“
- Conclusion



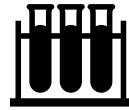


1.

# Introduction

# Who the hell is talking to me?

Robin Hübner  
Data & Analytics Consultant



- Education
  - Business & Engineering
  - Business Administration



- Work Experience
  - E-Commerce
  - Data & Analytics

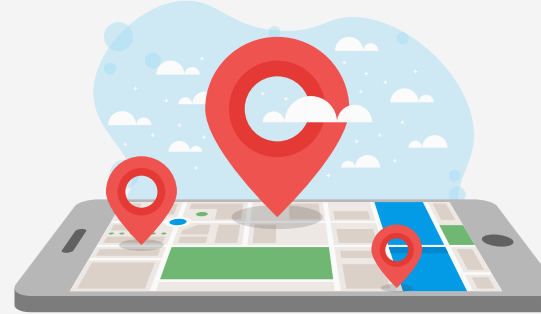


- Favorite Tools
  - Synapse / ADF
  - Power BI
  - Azure SQL

# Who is this great employee OBUNGI?



Experts for:  
Cloud based Data,  
Analytics & AI solutions



Location:  
In Hamburg, Germany,  
City Center



Team:  
More than 40 software  
developer und BI  
consultants



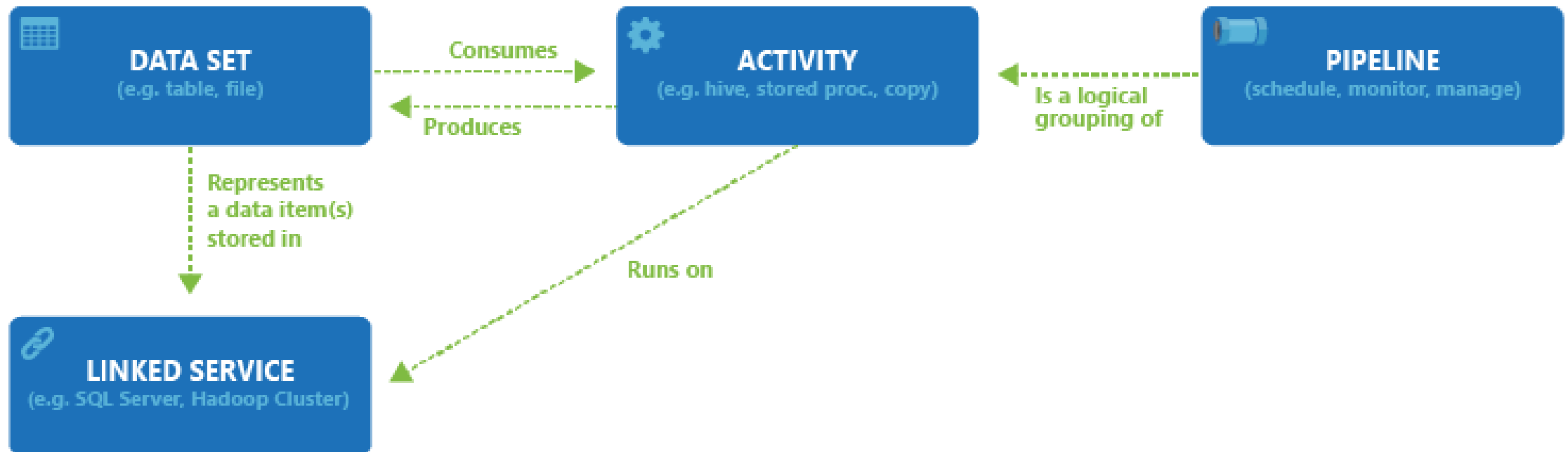
Deep Dive in:

- Banking & insurances
- Energy supplier
- wholesale & consumer goods
- Transportation & logistics

# 2.

## **Metadata-Driven Pipelines**

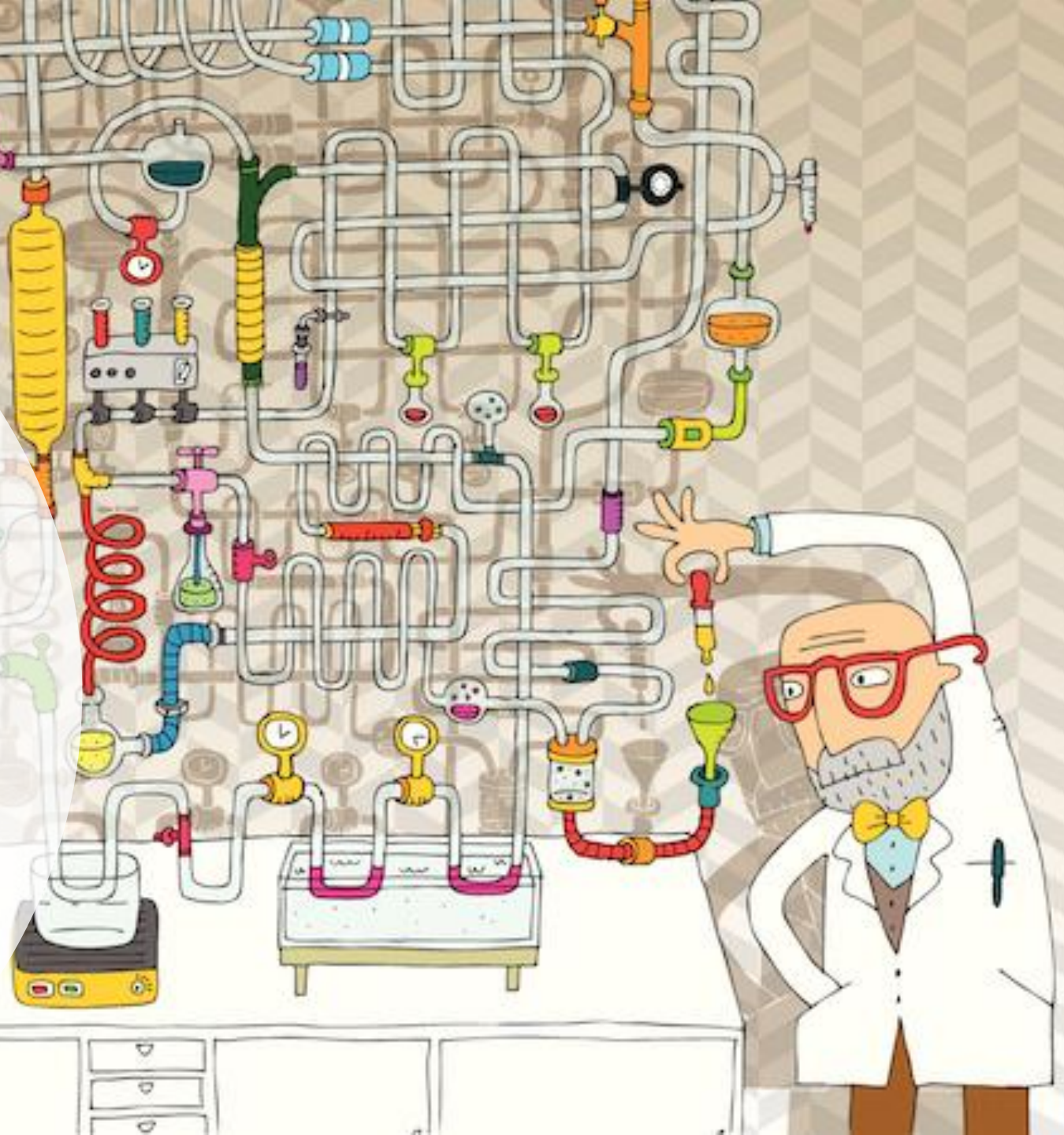
# Azure Data Factory Structure





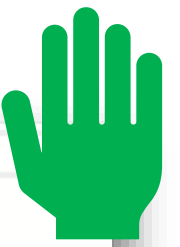
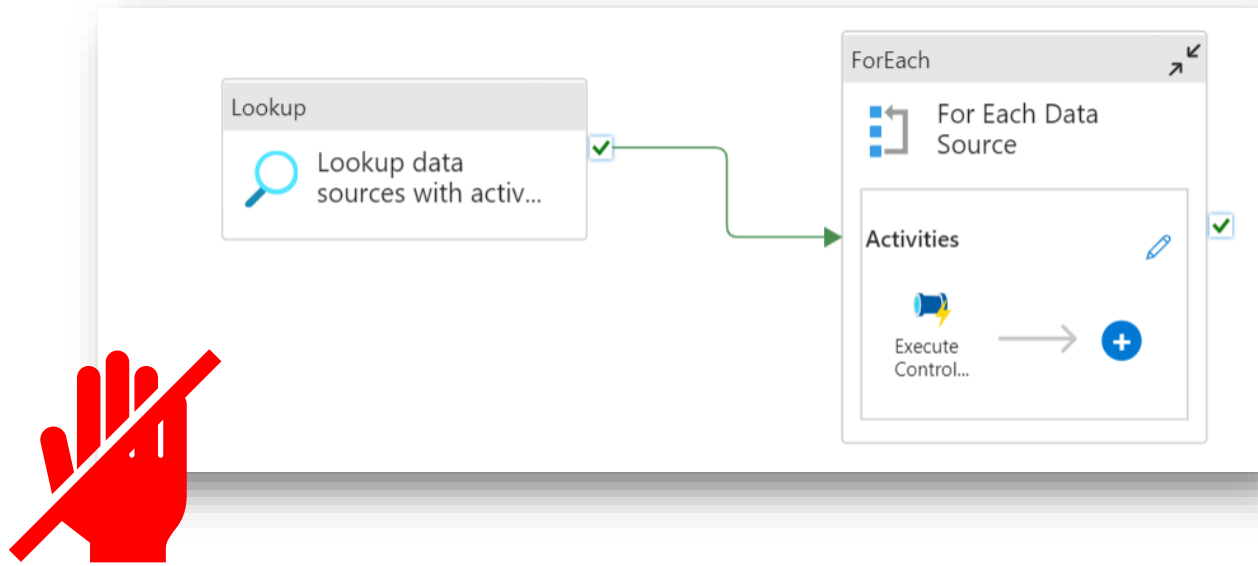
# PROBLEM

- Within a project multiple data sources (Linked Services) must be connected to Azure Data Factory / Synapse
- Each of the data sources has multiple tables (Datasets)
- Data integration and transformation can become on the one hand an annoying and repetitive task and on the other hand an unmanageable mess





# Metadata-Driven Pipelines (Concept)

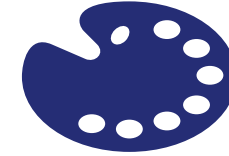


```
SELECT TOP (1000) [Id]
, [DataSourceName]
, [LinkedServiceName]
, [SourceSchemaName]
, [SourceTableName]
, [SinkFileDirectoryName]
, [SourceQueryDefinition]
, [LoadBatchGroup]
, [IsActive]
FROM [ctrl].[SourceToBronze]
```

# Individual Solution vs. ADF Integrated Solution



- No conceptual work is required
- SQL code for metadata-database is automatically generated (table and data)
- Pipeline and Datasets in ADF are created and configured automatically



- A concept must be elaborated
- SQL code for metadata-database must be created manually
- Pipelines and Datasets in ADF must be created and configured manually

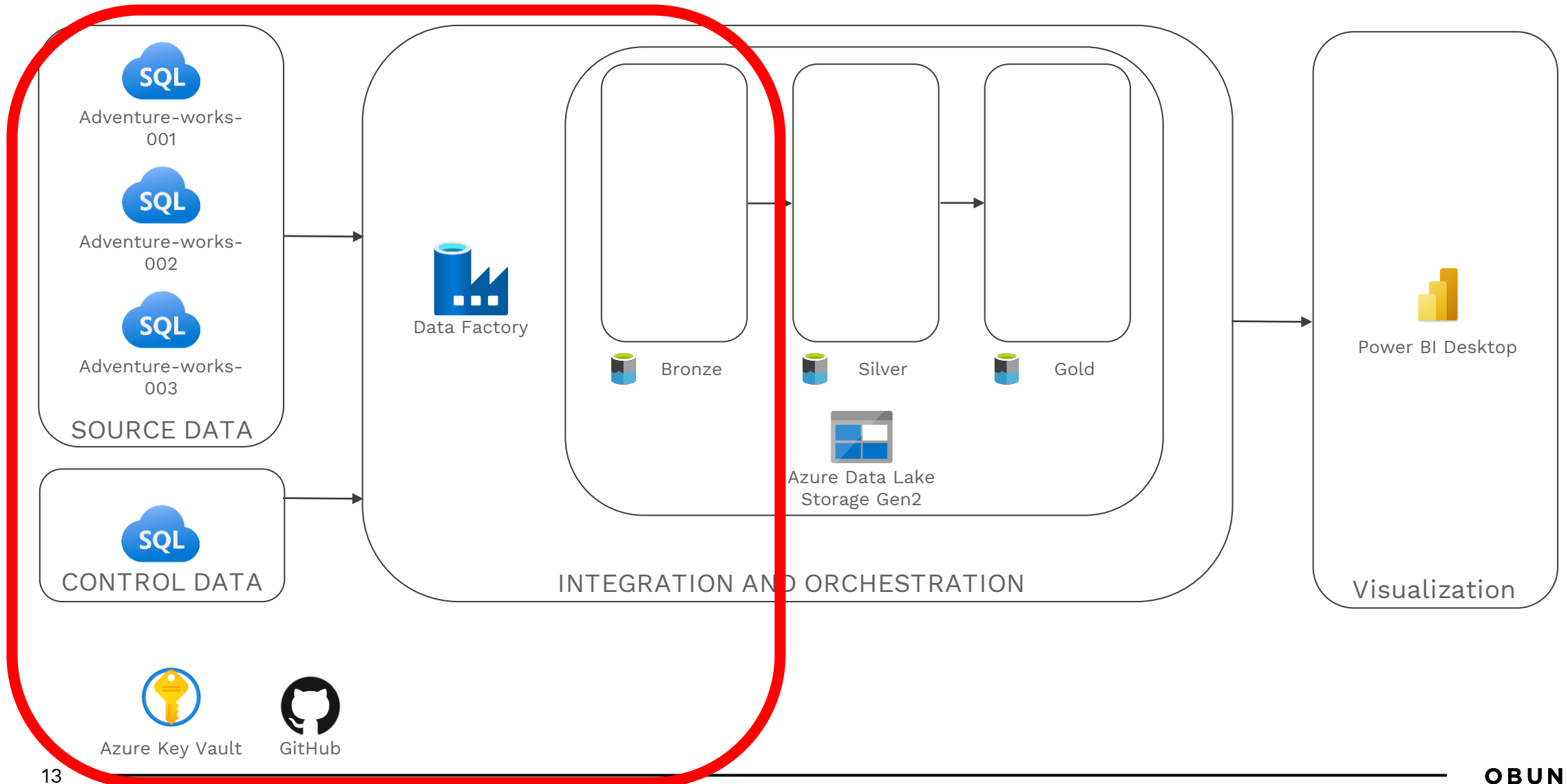
# 3.

**Demo-Case**

**“Adventure Works”**



# Architecture of the Demo Case



# 4.

## Conclusion

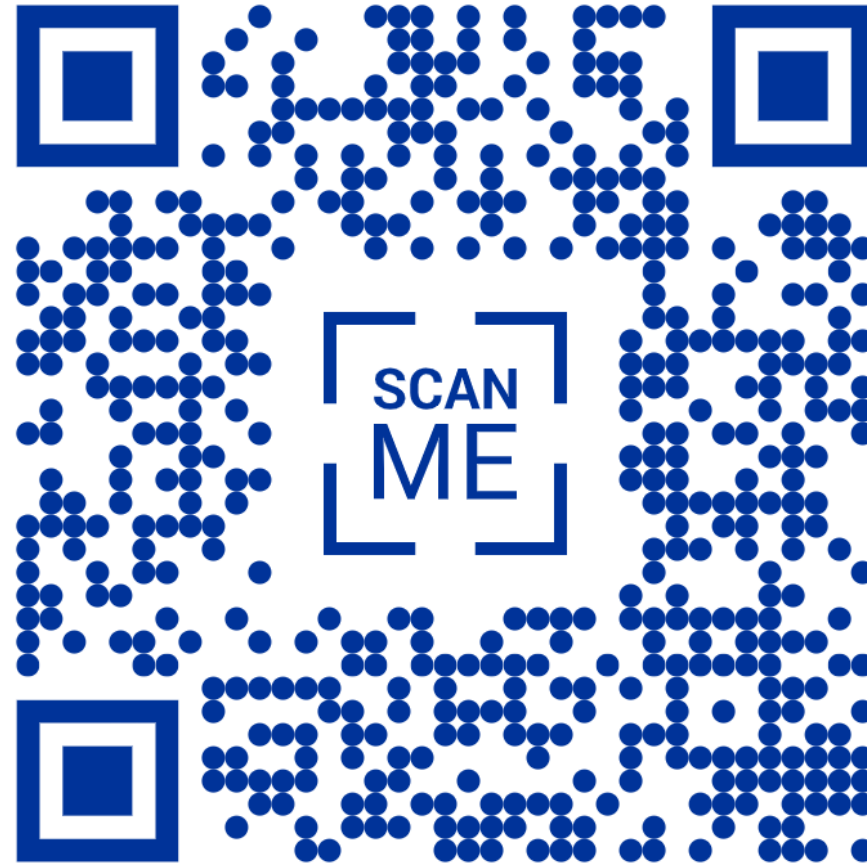
# Conclusion

- Metadata-driven pipelines are a must have in an ADF analytics solution
- Build-in metadata-driven pipelines are a good starting point
- Self-created and defined concepts are a powerful tool to build up large and scalable solutions in an easy manageable way





# Let's keep in touch & thank you for the attention



[www.linkedin.com/in/robinhuebner](https://www.linkedin.com/in/robinhuebner)