



# Azure Data Factory For Beginners

DARIUS LIKTORIUS

Please Silence  
Your Devices



# Contact & Follow Me

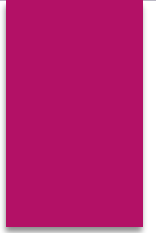
**Presentation Landing Page & Resources:**

[Liktorius.com/go/ADFBeginners](https://liktorius.com/go/ADFBeginners)

- ▶ Twitter: [@DLiktorius](https://twitter.com/DLiktorius)
- ▶ Blog: [Liktorius.com](https://liktorius.com)
- ▶ Follow me on LinkedIn:  
[linkedin.com/in/DariusLiktorius](https://linkedin.com/in/DariusLiktorius)



# Introductions



# My Background

- ▶ Cloud Architect with Cognizant Softvision
- ▶ Founder of NetCorp – since 1997
- ▶ Using and Certified in SQL Server since 6.5
- ▶ SQL: Architecture, Development, Infrastructure & DBA
- ▶ Certified on both AWS and Azure  
MCSE, MCSD, MCDBA & MCT – since 1996
- ▶ Hobby: Professional Photography

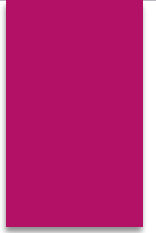
# About You

- ▶ DBA's?
- ▶ DB or BI Developers?
- ▶ Architects?
- ▶ Who's used SQL in the Azure?

# Our Agenda

- ▶ Intro to Azure Data Factory (ADF)
- ▶ ADF Components
- ▶ Demo: Creating a Data Factory
- ▶ Demo: Creating and Debugging a Pipeline
- ▶ Demo: Triggers & Monitoring
- ▶ ADF Pricing
- ▶ Q&A (?)

# Intro to Azure Data Factory (ADF)

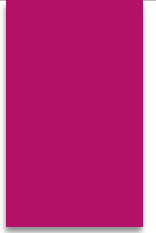




# What is Azure Data Factory?

- ▶ Microsoft Azure Platform as a Service (PaaS) offering
  - ▶ Runs in the Cloud but Hybrid with on-prem features
- ▶ Data Integration Service (remember SSIS?)
  - ▶ ETL (Extract Transform and Load)
  - ▶ ELT (Extract Load and Transform)
- ▶ Create pipelines to Copy and Transform your data

# Azure Data Factory Components



# Components of Azure Data Factory

- ▶ Data Factory
- ▶ Pipelines
- ▶ Activities
- ▶ Data Flows
- ▶ Datasets
- ▶ Linked Services
- ▶ Integration Runtimes
- ▶ Triggers
- ▶ Templates

# ADF Components – Data Factory

- ▶ Environmental Boundary
- ▶ Container to hold all other ADF components
- ▶ Deploy as many as you want/need (Dev, QA, Prod)
- ▶ Security (IAM) Boundary
- ▶ Integrates with Git and Azure DevOps
- ▶ Deployable as a single ARM template

# ADF Components – Pipelines

- ▶ Definition of executable ETL/ELT workflow
- ▶ Similar to an SSIS package
- ▶ Canvas designer interface
- ▶ Collection of units of work (Activities)
- ▶ Create new or from a Template

# ADF Components – Activities

- ▶ Single unit of work / task
- ▶ Chained in sequence or Parallelized
- ▶ Multiple categories:
  - ▶ Move & transform
  - ▶ Iteration & conditionals
  - ▶ Azure Function, Databricks, Machine Learning
  - ▶ *More...*

# ADF Components – Data Flows

- ▶ Recently introduced
- ▶ Separate type of activities with their own canvas
- ▶ Mapping Data Flows – Transform Data
- ▶ Wrangling Data Flows (Preview) – Prepare Data
- ▶ Executed from a Pipeline
- ▶ Run on Databricks and Spark environments

# ADF Components – Datasets

- ▶ Data reference objects
- ▶ Define format of input/output data
- ▶ Represent single table/view, file or folder



# ADF Components – Linked Services

- ▶ Data source / service connection object
- ▶ Ninety (90) supported services – *and growing*
  - ▶ From SQL Server to Cosmos DB to MySQL to Oracle
  - ▶ From Amazon S3 to SAP to SFTP to QuickBooks
- ▶ Maintains connection and credential properties

# ADF Components – Integration Runtimes

- ▶ Reference infrastructure to run Activities on
- ▶ Three types:
  - ▶ Azure
  - ▶ Self-Hosted
  - ▶ Azure-SSIS

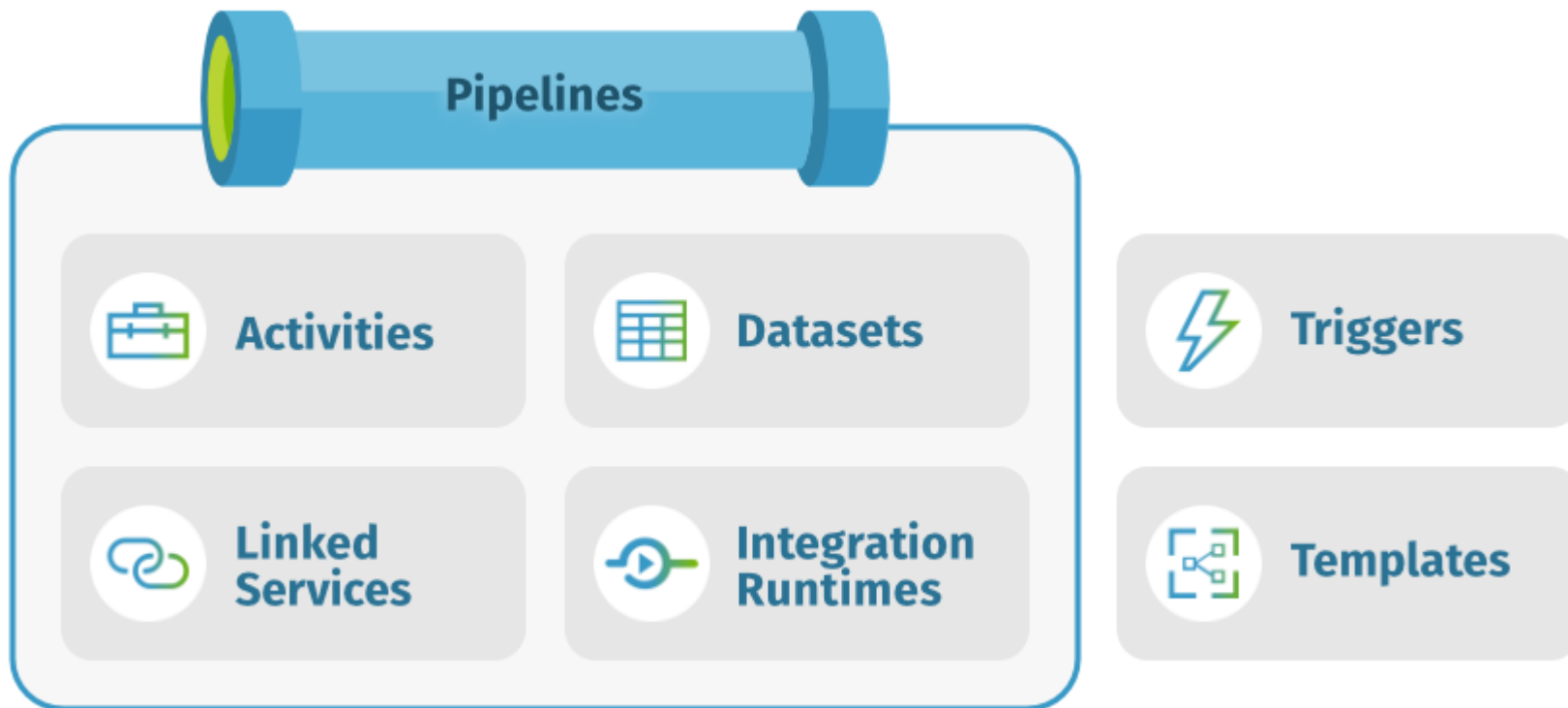
# ADF Components – Triggers

- ▶ Define when to execute your Pipelines
- ▶ Use a recurring schedule, interval or event based

# ADF Components – Templates

- ▶ Created from your own Pipeline
- ▶ Samples available from Microsoft

# ADF Component Summary



*Credit: cathrinewilhelmsen.net*



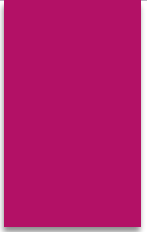
DEMO

Creating a  
Data Factory



DEMO

Creating &  
Debugging  
Pipelines



DEMO

Triggers &  
Monitoring



# Azure Data Factory Pricing



# Azure Data Factory - Pricing

- ▶ Prices vary based on Azure vs Self-hosted Runtime
- ▶ Orchestration charges (per 1,000 runs)
- ▶ Execution charges (per hour, by activity type)
- ▶ Data Flows (per vCore-hour)
- ▶ Data Factory Operations
  - ▶ Read/Write: per 50k modified/referenced entities
  - ▶ Monitoring: per 50k run records retrieved



Q&A

*Don't be shy!*

# Contact & Follow Me

**Presentation Landing Page & Resources:**

**[Liktorius.com/go/ADFBeginners](https://liktorius.com/go/ADFBeginners)**

- ▶ Twitter: **[@DLiktorius](https://twitter.com/DLiktorius)**
- ▶ Blog: **[Liktorius.com](https://liktorius.com)**
- ▶ Follow me on LinkedIn:  
**[linkedin.com/in/DariusLiktorius](https://linkedin.com/in/DariusLiktorius)**

