Exploring Data Integration Solutions with Azure Data Factory

Naveen Kumar M

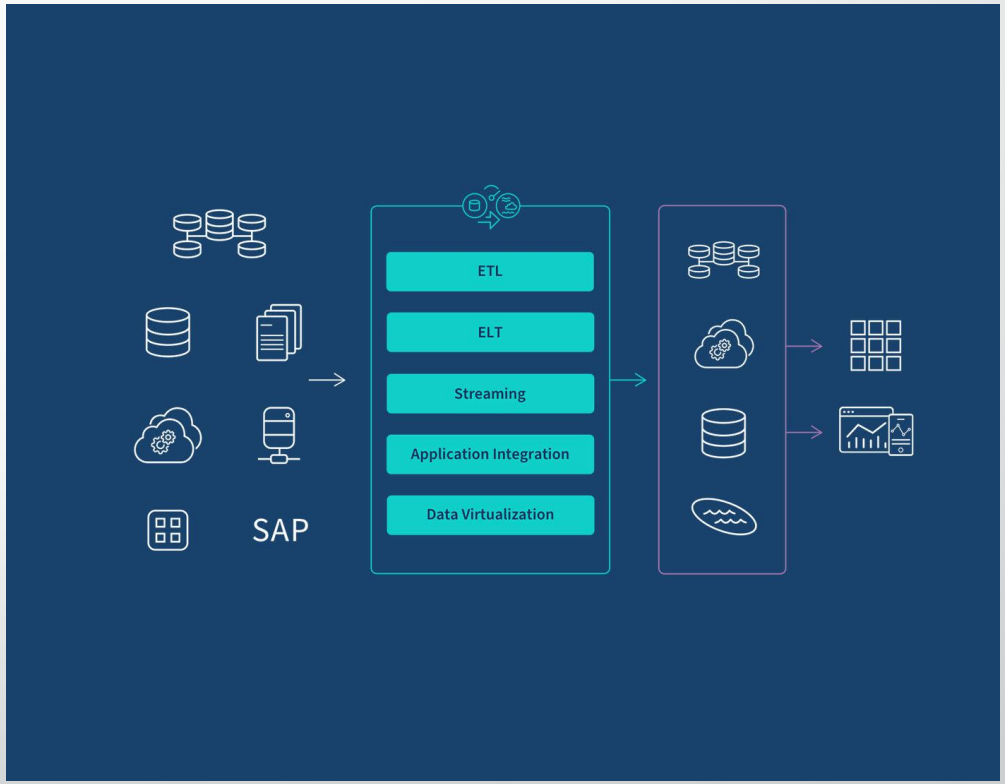
Data Engineer | C# Corner MVP | Blogger

Agenda

- Understanding Data Integration
- Introduction to Azure Data Factory
- Components of Azure Data Factory
- Demo
  - Data Movement and Transformation in Azure Data Factory
  - Monitoring and Management
- Real-world Use Cases
- Q&A

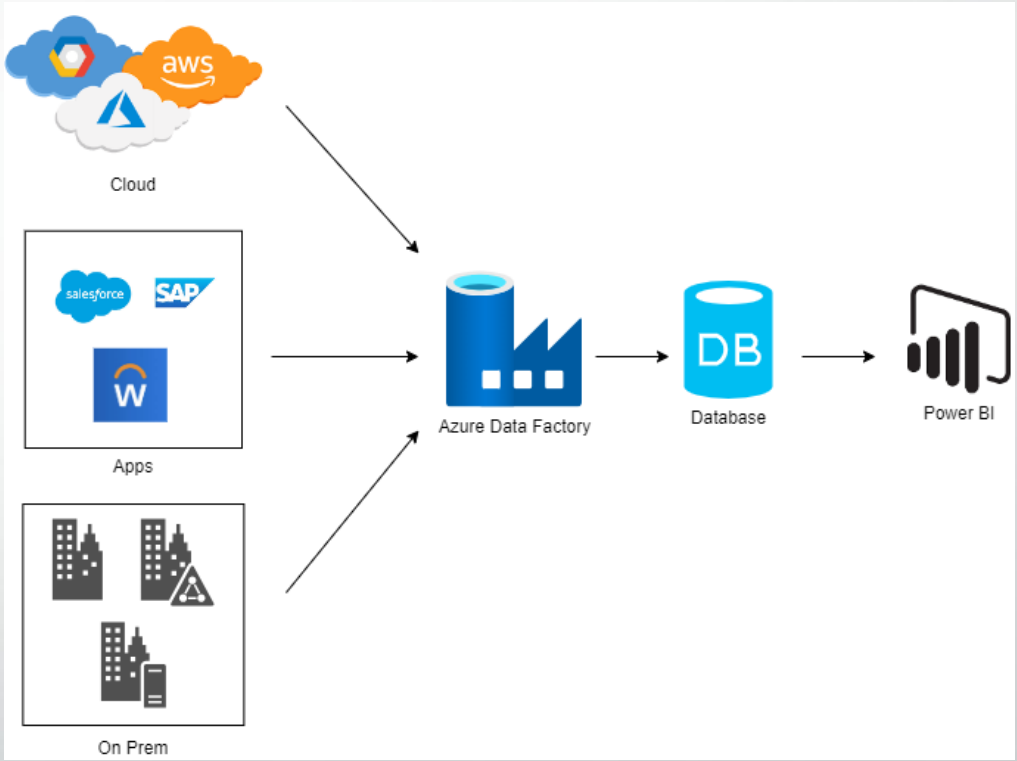
Understanding Data Integration

- Process of combining data from different sources into a unified view.
- Transforming it into a consistent format that can be analyzed and used effectively.
- Making it accessible and valuable for various business purposes.



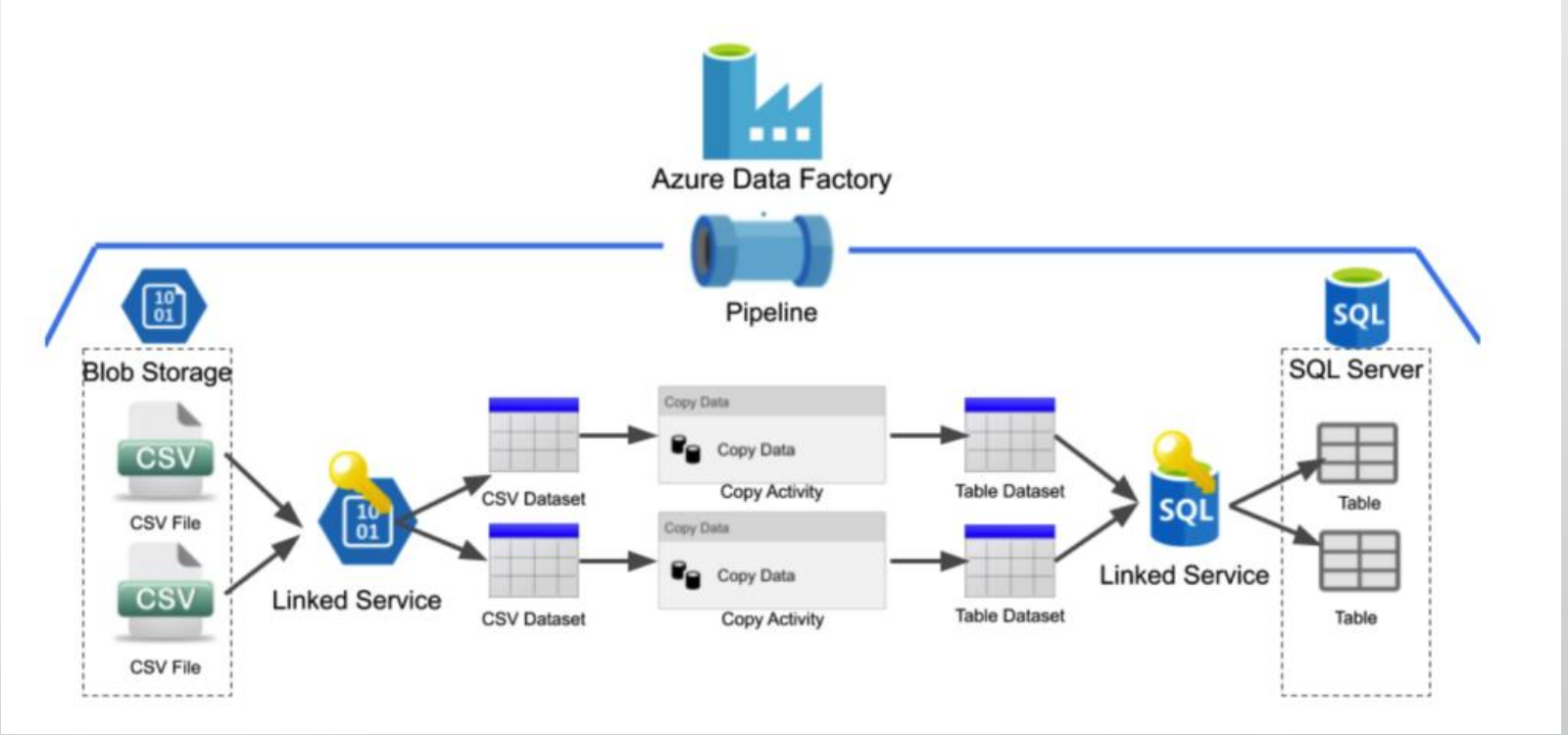
Introduction to Azure Data Factory

- Cloud bases ETL as a Service.
- Allows to create, schedule, and manage data-driven workflows.
- Enables seamless movement of data between various data stores, both on-premises and in the cloud.

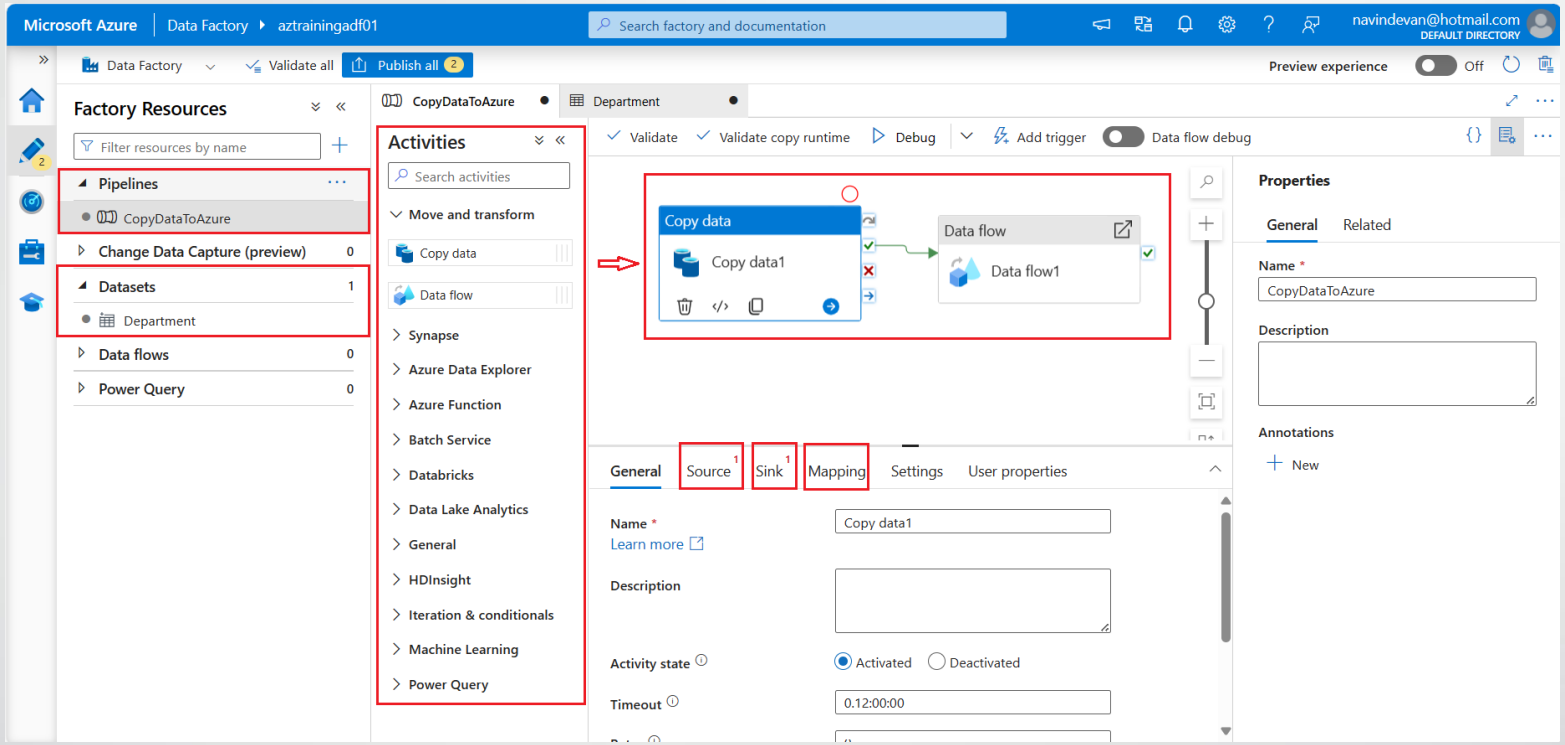


Components of Azure Data Factory

- Pipelines
- Activities
- Linked Services
- Datasets
- Triggers
- Integration Runtime



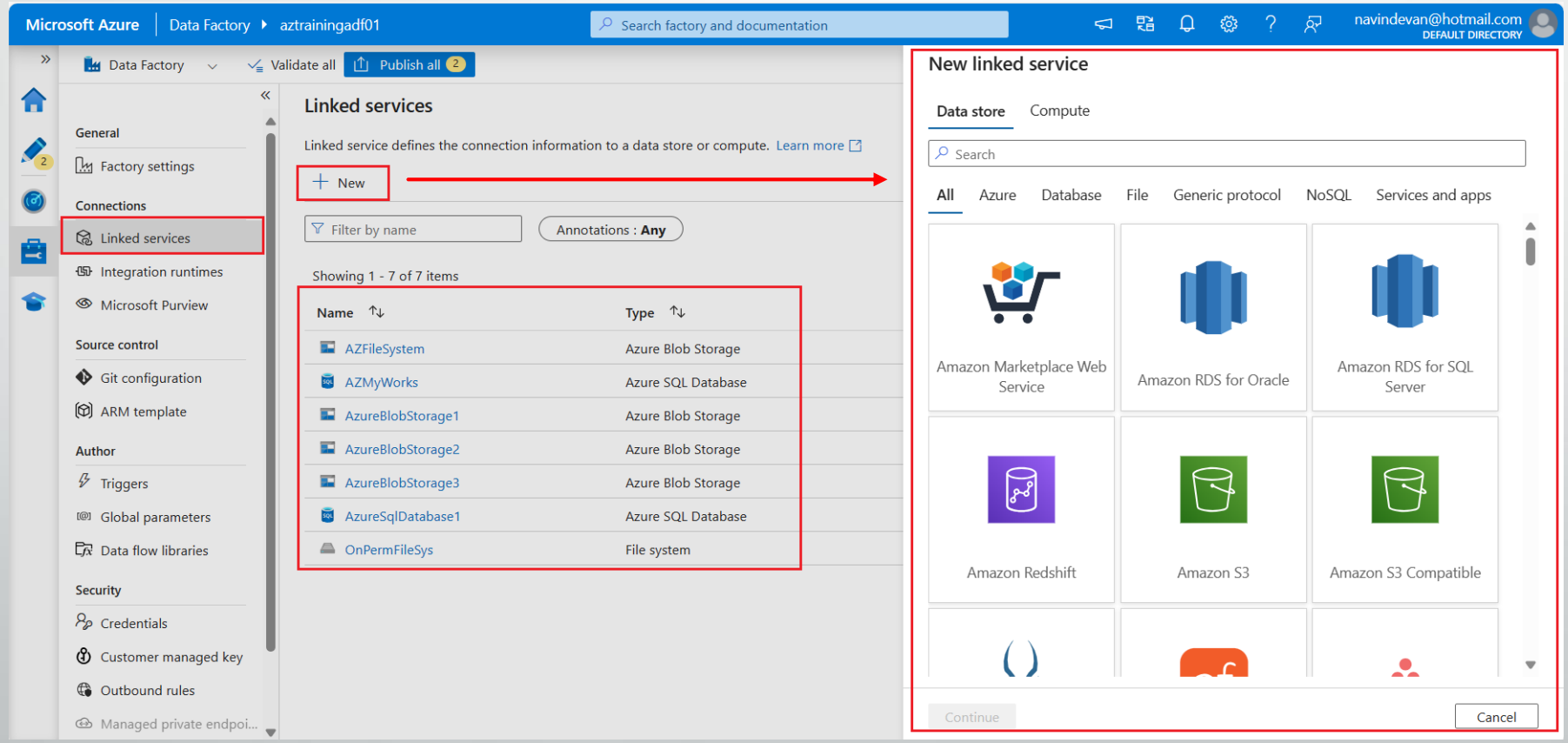
Components of Azure Data Factory Cont..



- **Pipelines** – It's the logical groups of activities that perform a specific job.
- **Activities** – An activity is just like a logical operation or the action that we perform on our data.
- **Datasets** – Datasets represent designated perspectives of data, serving as pointers or references to the specific data required for your tasks, whether as inputs or outputs.

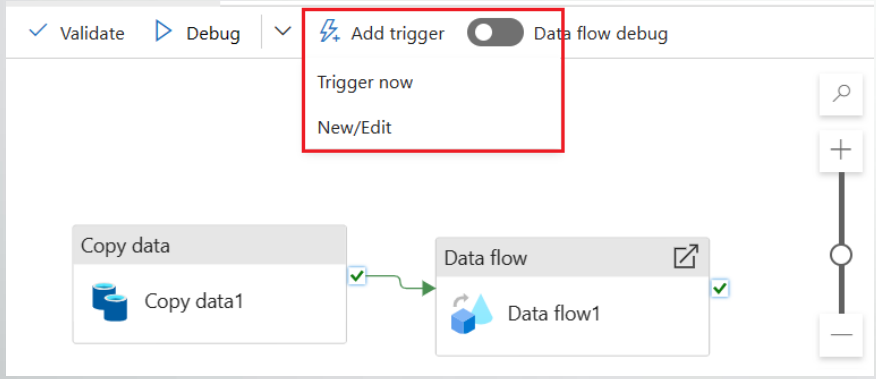
Components of Azure Data Factory Cont..

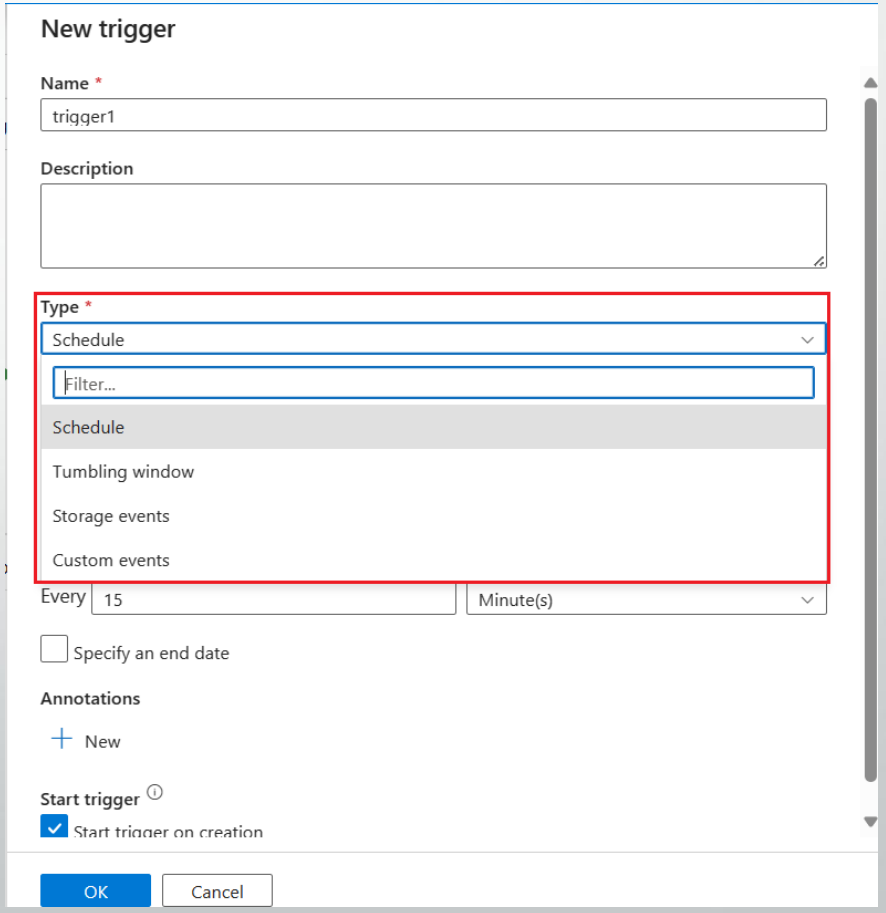
- **Linked Services** – Linked services are a very important component to link your data store to the ADF.



Components of Azure Data Factory Cont..

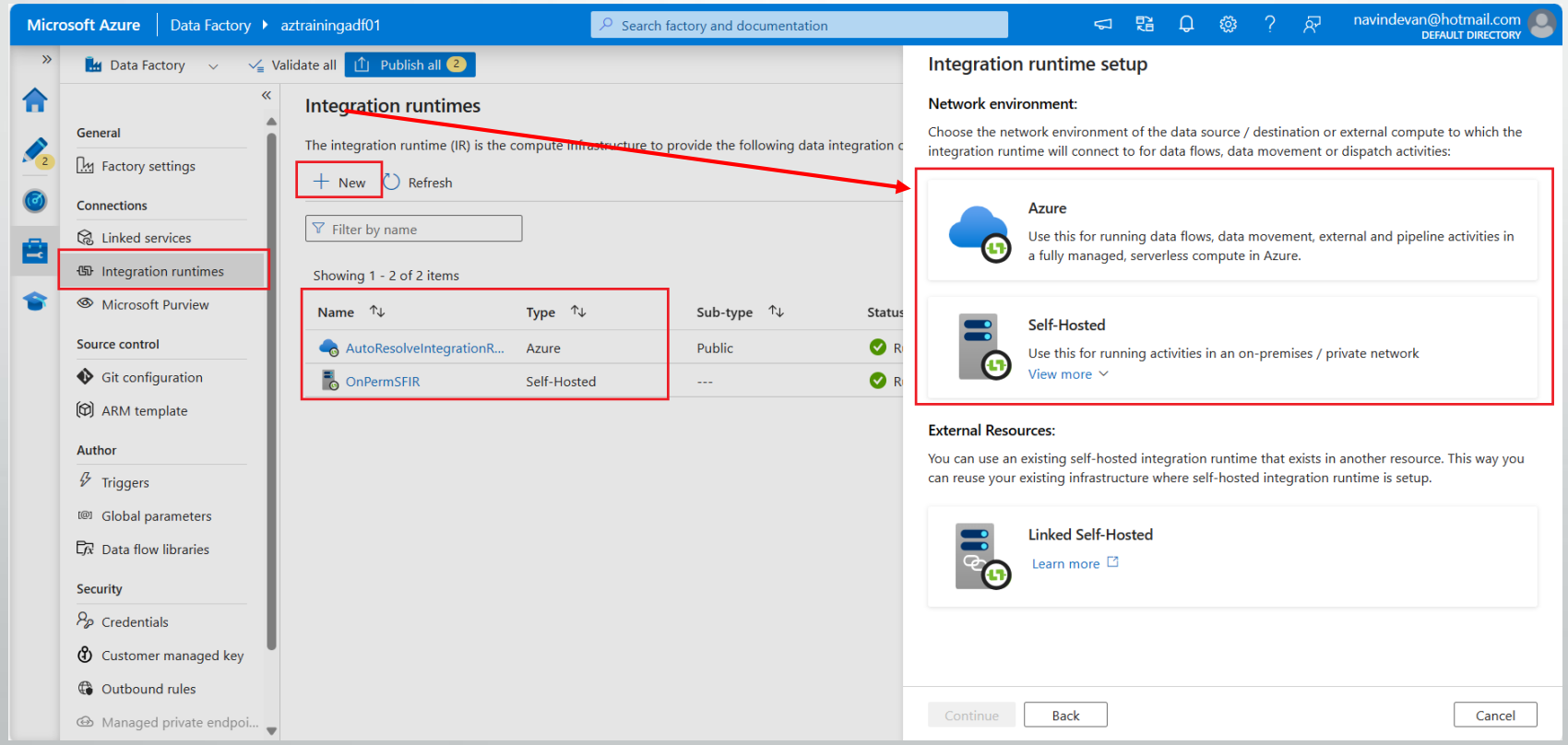
- **Triggers** – The triggers within ADF offer an alternative method for initiating a pipeline run.





Components of Azure Data Factory Cont..

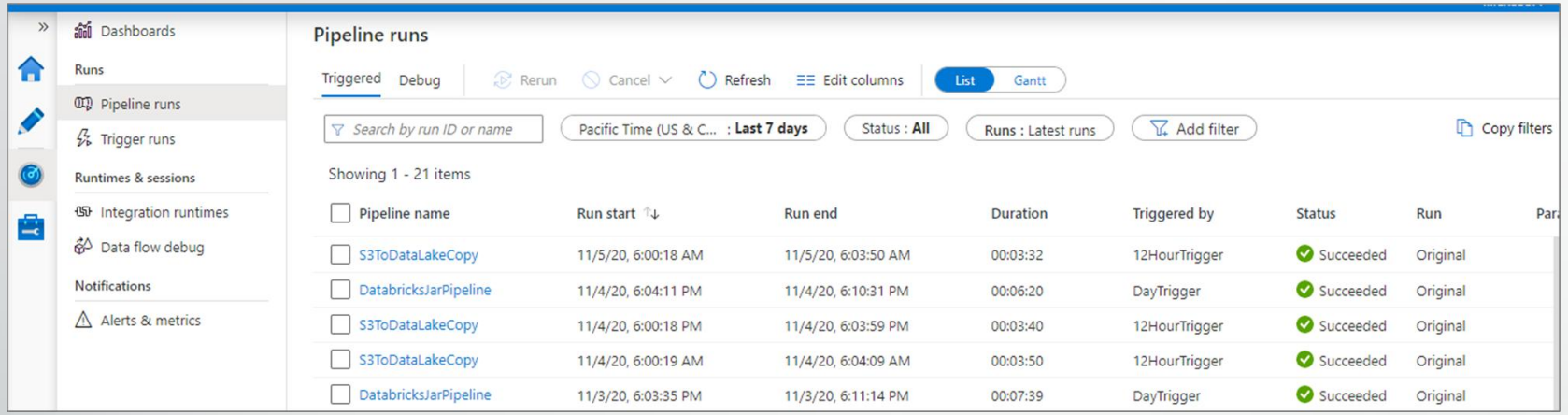
- **Integration Runtime** - Integration Runtime in Azure Data Factory is a managed compute infrastructure used for data integration and data transfer scenarios.
  - Azure Integration Runtime
  - Self Hosted

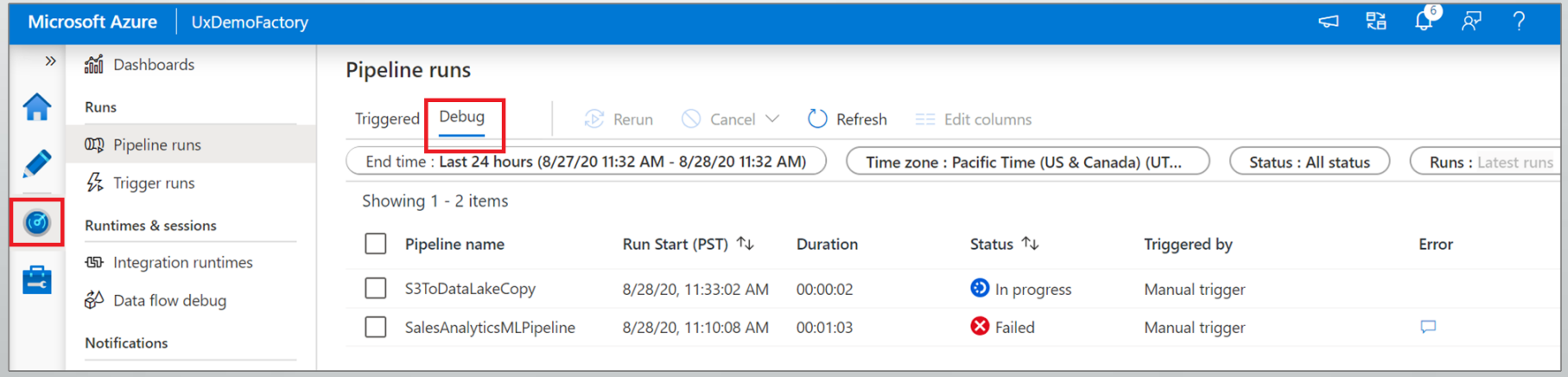


Data Movement and Transformation in ADF

- Get data integration with more than 90 built-in connectors
- Copy Activity
- Data Flow Activity
- HDInsight Activities

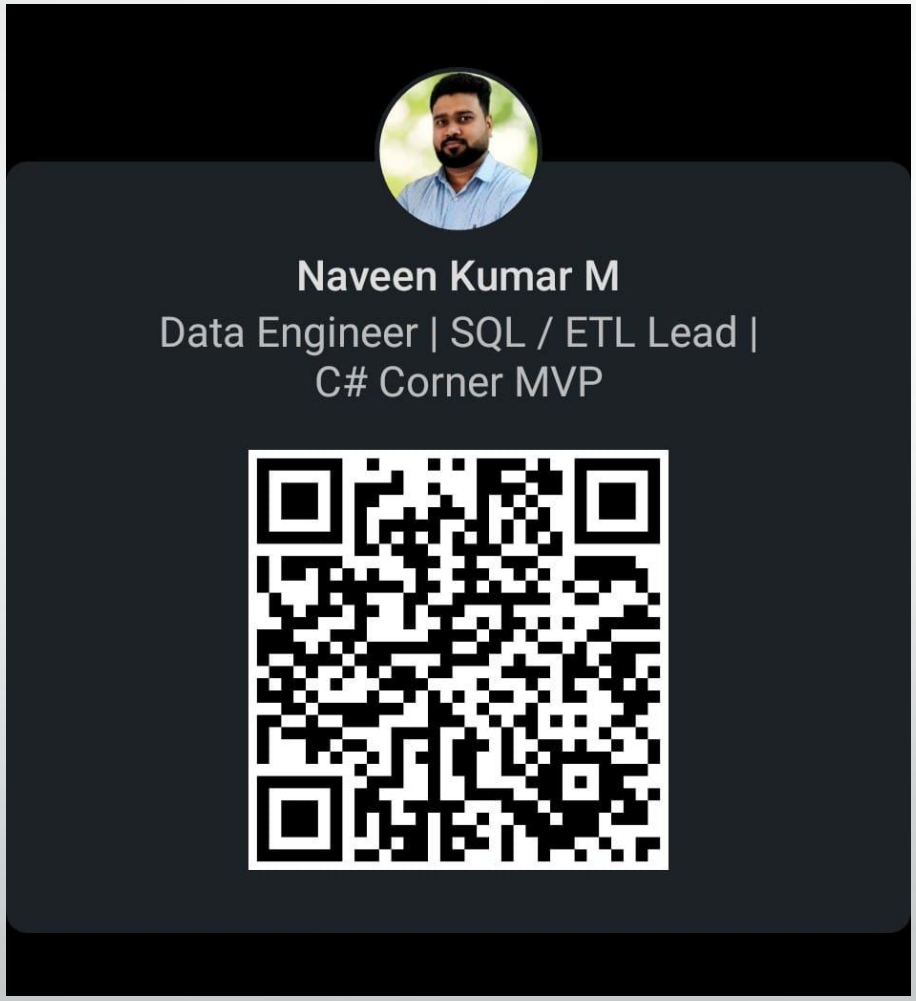
Monitoring and Management





Real-world Use Cases

- Utilized to facilitate data migration for advanced analytics initiatives.
- Offers the capability to streamline the ETL process, transitioning from SQL Server Integration Services to additional data sources.
- Provides a solution for extracting data from client servers or online sources and transferring it to an Azure Data Lake through pipeline creation to manage data flow.
- Serves as a versatile tool for executing various data integration tasks and is recognized as one of the leading ETL solutions.
- Enables the integration of data from diverse ERP systems into Azure Synapse for reporting purposes, with support for analysis and reporting via Power BI.
- The ADF tools boast seamless integration, enabling rapid development of ETL, big data, data warehousing, and machine learning solutions, while maintaining the flexibility to evolve and adapt to changing or enhanced requirements.

?
Q
&
A

Let's connect on LinkedIn