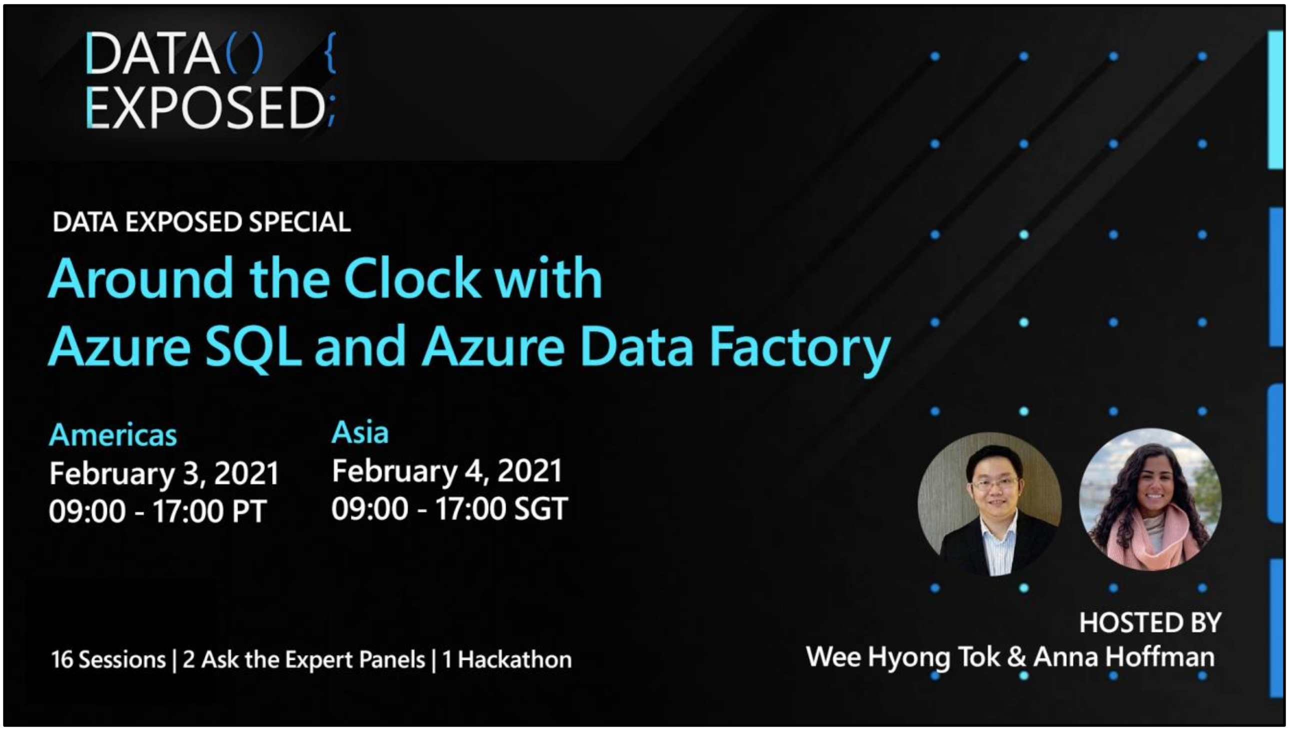


DATA() {
EXPOSED;

DATA EXPOSED SPECIAL

# Around the Clock with Azure SQL and Azure Data Factory

**Americas**
February 3, 2021
09:00 - 17:00 PT

**Asia**
February 4, 2021
09:00 - 17:00 SGT

16 Sessions | 2 Ask the Expert Panels | 1 Hackathon

HOSTED BY
Wee Hyong Tok & Anna Hoffman

© Microsoft Corporation

DATA() {
EXPOSED;

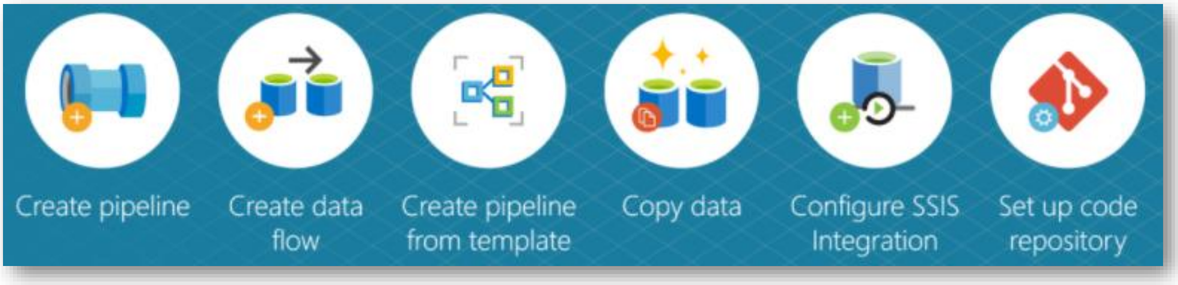
# Designing a Data Architecture with Azure Data Factory

Buck Woody - Microsoft

More about the presenter here: https://www.linkedin.com/in/buckwoody/

© Microsoft Corporation

## Centering on the Pipeline



©Microsoft Corporation
Azure

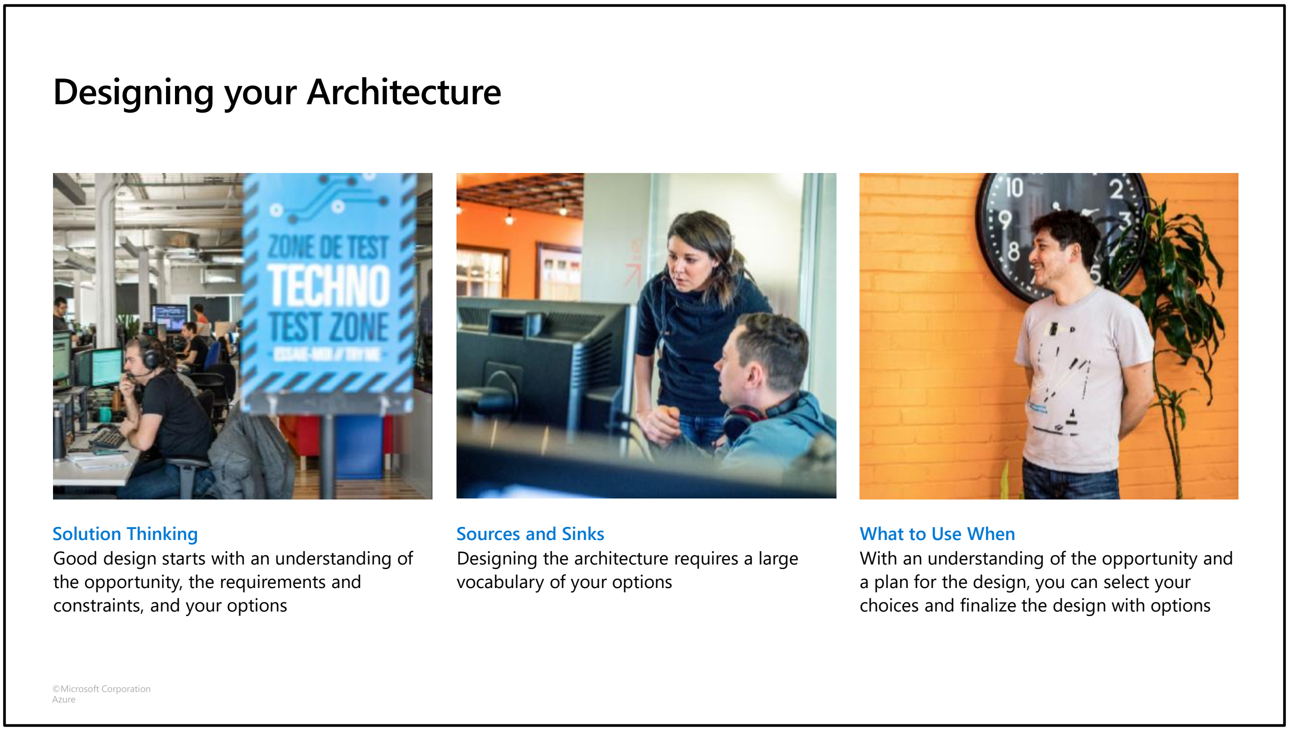
Creating an architecture starts with the data use, and next to the data flow. You want to identify, automate, manage and monitor the workflow of the data from source to application.

Overview: https://docs.microsoft.com/en-us/azure/data-factory/introduction

Primary Concepts:

1. Pipelines
2. Activities
3. Datasets
4. Linked services
5. Data Flows
6. Integration Runtimes

© Microsoft Corporation



# Designing your Architecture

**Solution Thinking**
Good design starts with an understanding of the opportunity, the requirements and constraints, and your options

**Sources and Sinks**
Designing the architecture requires a large vocabulary of your options

**What to Use When**
With an understanding of the opportunity and a plan for the design, you can select your choices and finalize the design with options

©Microsoft Corporation
Azure

As you design your architecture, you'll want to focus on three parts:

1. Solution thinking
2. Sources and Sinks
3. And what to use for each part of the solution

This session covers each of these in more depth

Great reference on considerations here: https://www.dbta.com/Editorial/Think-About-It/8-Steps-to-Building-a-Modern-Data-Architecture-101417.aspx

Data Lake Research and Credits:
https://www.guru99.com/data-lake-architecture.html
https://dzone.com/storage/temp/4214871-zaloni-data-lake-architecture.png

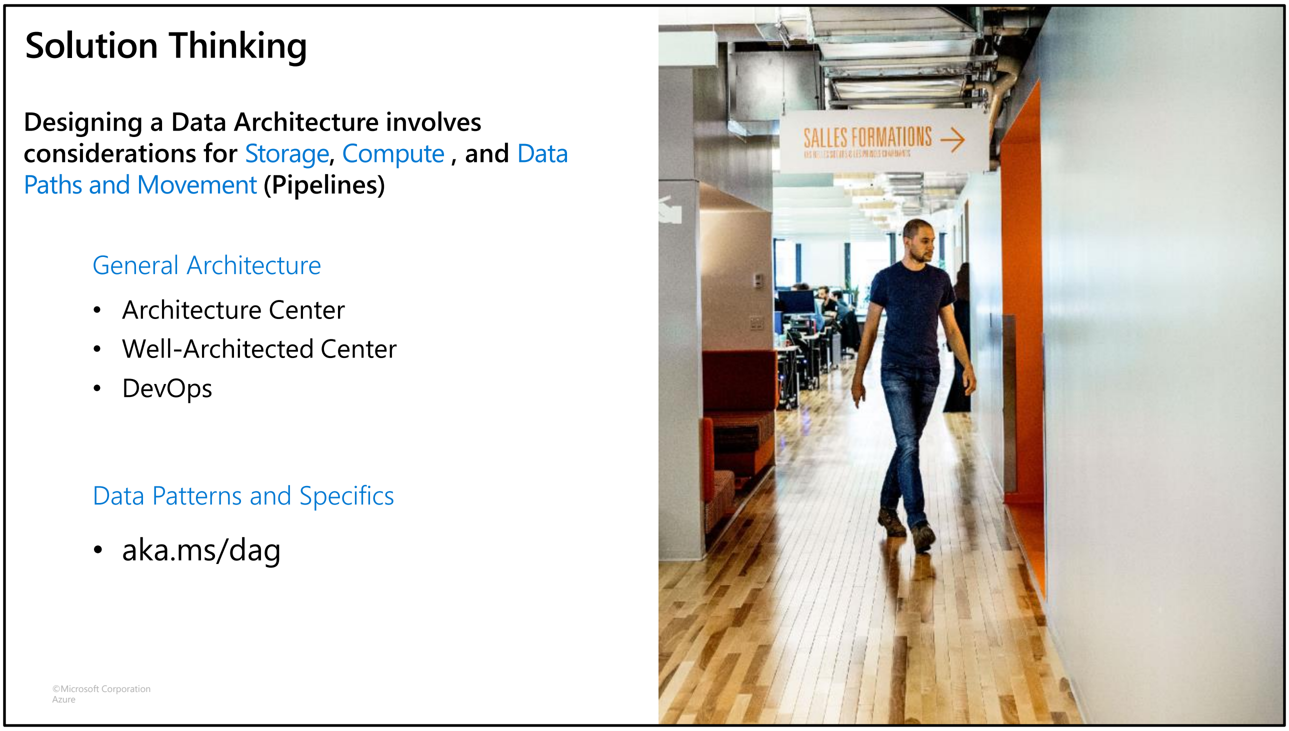
https://www.zaloni.com/resources-all-content/
https://www.zaloni.com/themencode-pdf-viewer-sc/?tnc_pvfw=ZmlsZT1odHRwczovL3d3dy56YWxvbmkuY29tL3dwLWNvbnRlbnQvdXBsb2Fkcy8yMDIwLzA0L1VuZGVyc3RhbmRpbmctTWV0YWRhdGEtZUJvb2sucGRmJnNldHRpbmdzPTExMTExMTExMTExMTExJmxhbmc9ZW4tVVM=#page=&zoom=&pagemode=
https://www.zaloni.com/resources/briefs-papers/papers/architecting-data-lakes/
https://www.zaloni.com/themencode-pdf-viewer-sc/?tnc_pvfw=ZmlsZT1odHRwczovL3d3dy56YWxvbmkuY29tL3dwLWNvbnRlbnQvdXBsb2Fkcy8yMDIwLzAzL0FyY2hpdGVjdGluZy1EYXRhLUxha2VzLTJlLnBkZiZzZXR0aW5ncz0xMTExMTExMTExMTExMSZsYW5nPWVuLVVT#page=&zoom=&pagemode=
https://www.zaloni.com/themencode-pdf-viewer-sc/?tnc_pvfw=ZmlsZT1odHRwczovL3d3dy56YWxvbmkuY29tL3dwLWNvbnRlbnQvdXBsb2Fkcy8yMDIwLzA0L0RhdGEtTGFrZS1NYXR1cml0eS1Nb2RlbC1lQm9vay0yLjQuMTkucGRmJnNldHRpbmdzPTExMTExMTExMTExMT

ExJmxhbmc9ZW4tVVM=#page=&zoom=&pagemode=

https://www.microsoft.com/en-us/itshowcase/powering-digital-transformation-at-microsoft-with-modern-data-foundations

https://en.wikipedia.org/wiki/Data_lake

https://www.qubole.com/blog/data-warehouses-vs-data-lakes/



## Solution Thinking

**Designing a Data Architecture involves considerations for Storage, Compute , and Data Paths and Movement (Pipelines)**

General Architecture

- Architecture Center
- Well-Architected Center
- DevOps

Data Patterns and Specifics

- aka.ms/dag

©Microsoft Corporation
Azure

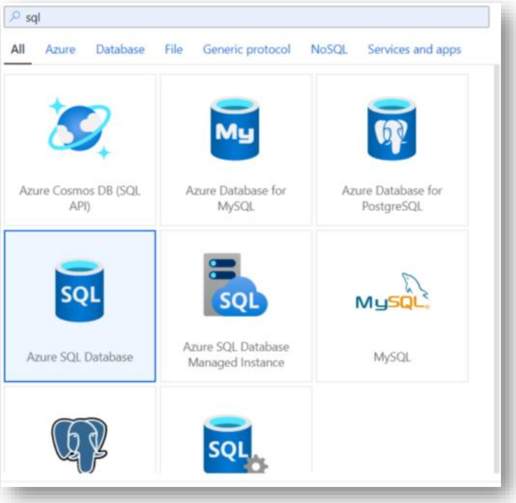
SALLES FORMATIONS →

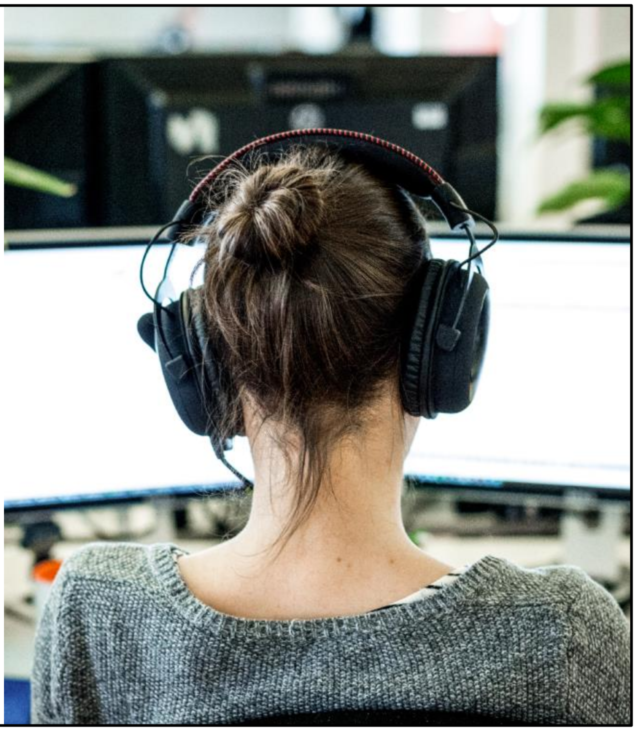
Start Here - Azure Architecture Center - Azure Architecture Center | Microsoft Docs
Move here - Microsoft Azure Well-Architected Framework - Microsoft Azure Well-Architected Framework introduction | Microsoft Docs
End here - Azure Data Architecture Guide - Azure Architecture Center | Microsoft Docs
Review the Cloud Adoption Framework even if you are on-prem or hybrid - Microsoft Cloud Adoption Framework for Azure - Cloud Adoption Framework | Microsoft Docs And msft.it/6010p3Pbs

Security is a huge part of your design, and should permiate throughout the process: https://docs.microsoft.com/en-us/azure/data-factory/data-movement-security-considerations
Note that you can now include Data Documentation as part of ADF: https://docs.microsoft.com/en-us/azure/data-factory/how-to-discover-explore-purview-data

Microsoft Azure

# Knowing your options – Sources and Sinks



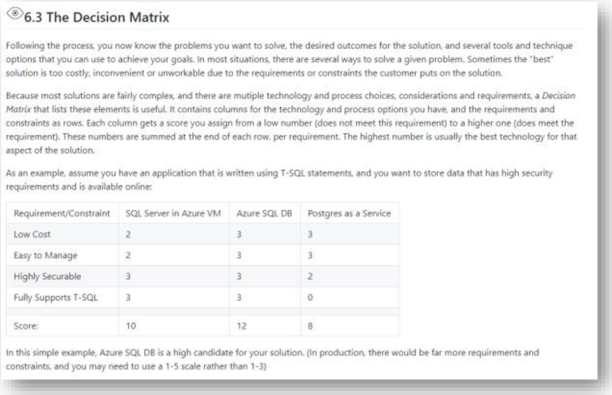


Involves two components: Linked Services and Datasets - [https://docs.microsoft.com/en-us/azure/data-factory/concepts-linked-services](https://docs.microsoft.com/en-us/azure/data-factory/concepts-linked-services)

© Microsoft Corporation

Microsoft Azure

# What to Use When – pairing up the options with the requirements



### 6.3 The Decision Matrix

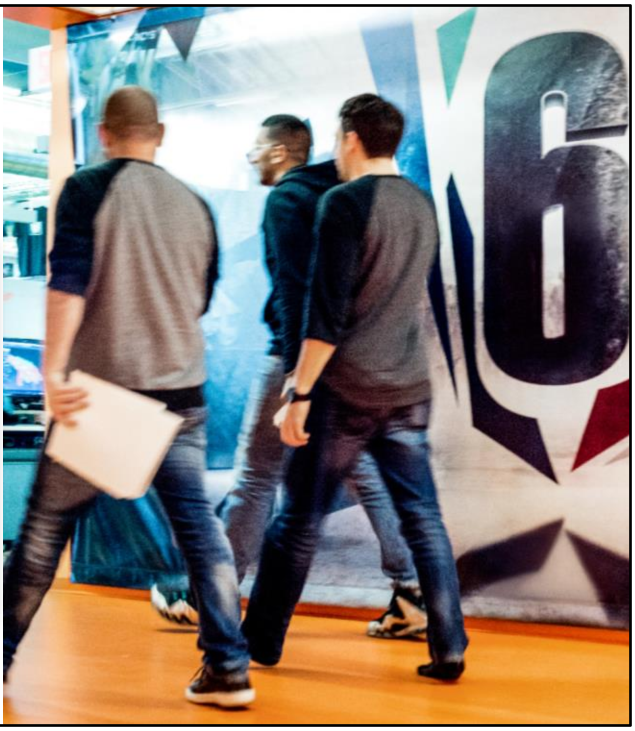
Following the process, you now know the problems you want to solve, the desired outcomes for the solution, and several tools and technique options that you can use to achieve your goals. In most situations, there are several ways to solve a given problem. Sometimes the "best" solution is too costly, inconvenient or unworkable due to the requirements or constraints the customer puts on the solution.

Because most solutions are fairly complex, and there are mutiple technology and process choices, considerations and requirements, a *Decision Matrix* that lists these elements is useful. It contains columns for the technology and process options you have, and the requirements and constraints as rows. Each column gets a score you assign from a low number (does not meet this requirement) to a higher one (does meet the requirement). These numbers are summed at the end of each row, per requirement. The highest number is usually the best technology for that aspect of the solution.

As an example, assume you have an application that is written using T-SQL statements, and you want to store data that has high security requirements and is available online:

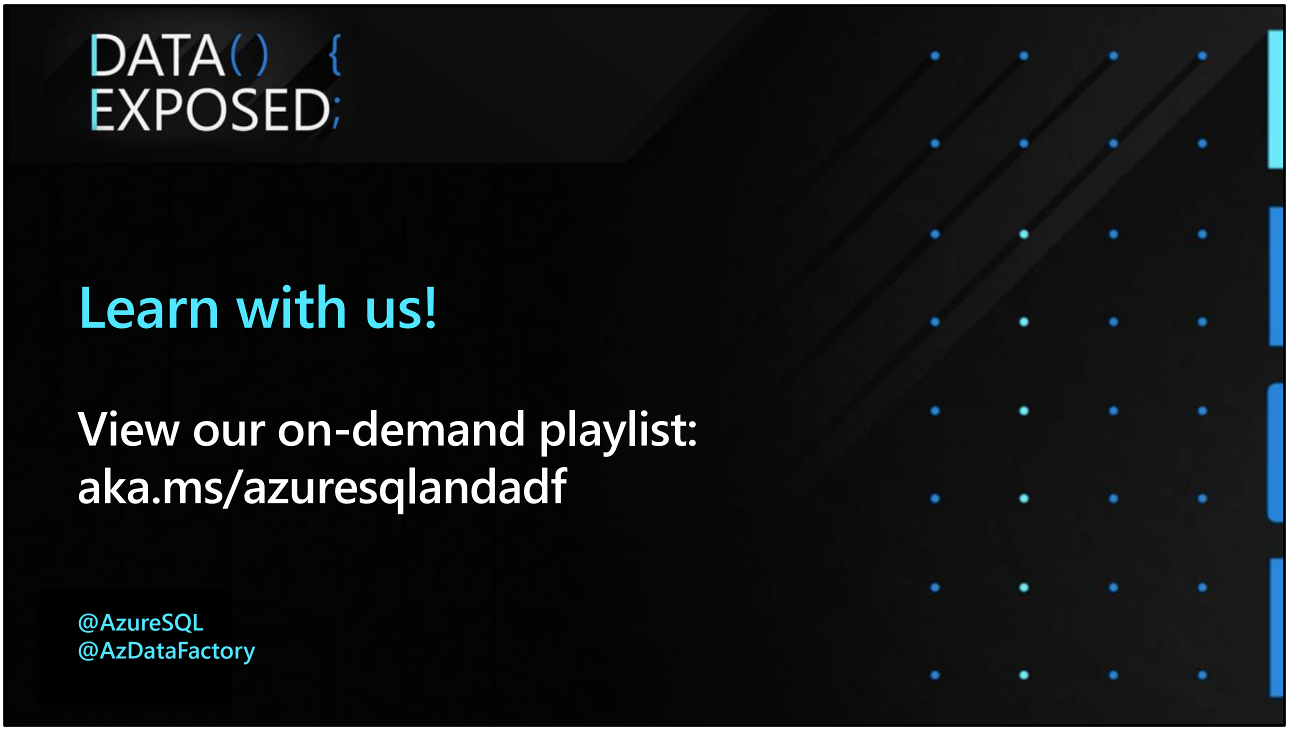
| Requirement/Constraint | SQL Server in Azure VM | Azure SQL DB | Postgres as a Service |
|---|---|---|---|
| Low Cost | 2 | 3 | 3 |
| Easy to Manage | 2 | 3 | 3 |
| Highly Securable | 3 | 3 | 2 |
| Fully Supports T-SQL | 3 | 3 | 0 |
| | | | |
| Score: | 10 | 12 | 8 |

In this simple example, Azure SQL DB is a high candidate for your solution. (In production, there would be far more requirements and constraints, and you may need to use a 1-5 scale rather than 1-3)



https://github.com/microsoft/sqlworkshops-sqlg2c/blob/master/sqlgroundtocloud/06-WhatToUseWhen.md

© Microsoft Corporation



DATA( ) {
EXPOSED;

# Learn with us!

View our on-demand playlist:
aka.ms/azuresqlandadf

@AzureSQL
@AzDataFactory

© Microsoft Corporation

Microsoft Azure

© Copyright Microsoft Corporation. All rights reserved.