

Putting Fabric to the Test

Building an Enterprise Data Fabric



- **Jacob Rønnow Jensen**
- Head of Data Platform, AP Pension
- Enterprise Voice Contributor
- Microsoft Denmark Data & AI Executive Board

Data Platform, AP Pension

As simple as possible – but not simpler



Data Platform

As little data movement as possible

- but not less

As few technologies as possible

- but not fewer

As cheap as possible

- but not cheaper

Product Management
Data & Platform Architect
Data Engineer
Data Analyst
Business Analyst



Microsoft Fabric

The unified data platform for the era of AI



Data
Factory



Synapse Data
Engineering



Synapse Data
Science



Synapse Data
Warehousing



Synapse Real
Time Analytics



Power BI



Data
Activator



AI



OneLake



Purview



Unified architecture



Unified experience



Unified governance



Unified business model

OneLake – OneDrive for your data

Home

Reporting

SQL analytics endpoint

New SQL query

New visual query

New report

New measure

This lakehouse automatically adds new objects to its default Power BI semantic model. Don't want to sync? Go to settings to turn this off. [Learn more](#)

Manage default Power BI semantic model

Warehouses

ado_lakehouse

Schemas

dbo

Tables

fabconfest

Data preview

Showing 1000 rows

Search

	ID	Title	Revision	Parent	createdBy	
1	228010	Poc på lognignsframework	1	228009	Sæbo Davtians	SDA
2	228010	Poc på lognignsframework til fremvisning	2	228009	Sæbo Davtians	SDA
3	228010	Poc på lognignsframework til fremvisning	3	228009	Sæbo Davtians	SDA
4	228010	Poc på lognignsframework til fremvisning	4	228009	Sæbo Davtians	SDA
5	228032	Implementering af logningsframework version 0.1	1	228009	Sæbo Davtians	SDA
6	228032	Implementering af logningsframework version 0.1	2	228009	Sæbo Davtians	SDA
7	228036	Skriv slutbruger dokumentation	1	228008	Sæbo Davtians	SDA

Mirroring – Azure SQL Database

The screenshot displays the Azure SQL Database interface. On the left, the SQL query editor shows a script to create a table, declare a variable, and insert data. The script is as follows:

```
CREATE TABLE #TEST( A INT, B INT)
DECLARE @I INT=0
WHILE @I < 1000000
BEGIN
    INSERT INTO #TEST VALUES(@I,@I*10)
    SET @I=@I+1
END
insert into mirrortest1
SELECT * FROM #TEST
where A > 99999
```

Below the query editor, the Messages pane shows the execution status: "(900000 rows affected)" and "Completion time: 2024-01-17T17:48:53.5367493+01:00".

On the right, the SQL query 1 window is open, showing the query:

```
1 SELECT count(*)
2 FROM [T32Life].[dbo].[dbo_mirrortest1]
```

The Results tab is selected, displaying a table with one row and one column:

	123 untitled1
1	1000000

Mirroring - CosmosDB

The screenshot displays the AP Pension CosmosMirror web application. The interface includes a top navigation bar with the AP Pension logo, a search bar, and user profile controls. A dropdown menu for 'Mirrored Cosmos DB' is open, showing options for 'Mirrored Cosmos DB' (Configure and monitor mirrored database) and 'SQL analytics endpoint' (Query data using SQL).

The main content area is divided into two panels. The left panel, titled 'NOSQL API', shows a sidebar with 'DATA' and a list of collections: 'Commitment' and 'Voucher'. The 'Voucher' collection is selected, displaying a table of documents. The table has columns 'id' and '/ID'. The first document is highlighted with a blue background.


id	/ID
2343efc3-5da7-4ccb-92c6-aa3...	
1889f34c-027e-4837-8560-72...	
433de3b7-3d75-4c87-9d31-3...	
0333b861-0399-4802-aa08-b2...	
df0aa66c-4668-48a3-b7c6-b9...	
627906c0-aaca-4f76-b2c0-818...	
c4c8f26c-8140-4d1a-a9cb-06...	
c89fcb9f-3441-45b4-838e-d4...	
cf3d7fca-5e61-4156-b907-5e3...	
254f380d-dd8e-46ae-a8d9-97...	
6a784e2e-837f-4bcf-b1a6-755...	
61f154e4-f7b3-4ec9-8214-0e5...	
401d3b39-5e47-48c4-bd43-ce...	






A 'Load more' button is located at the bottom of the table.

The right panel displays a detailed JSON document for the selected voucher. The document structure is as follows:





```
1 {
2   "_type": "MoneyInVoucher",
3   "_baseType": "Voucher",
4   "_id": "2343efc3-5da7-4ccb-92c6-aa36896126e7",
5   "amount": 700000,
6   "commitmentOwner": {
7     "_type": "PersonRef",
8     "externalRef": "1305651234"
9   },
10  "companyAgreementId": "04001",
11  "companyIdentifier": "26259495",
12  "contributionOptionality": false,
13  "creationStamp": "2023-11-07T16:19:27.5166525",
14  "employerContributionPercentage": 0.0815,
15  "grossSalary": 85000,
16  "labourAgreement": "UL341",
17  "paymentDate": "2022-11-01",
18  "salaryFromDate": "2022-10-01",
19  "salaryToDate": "2022-10-31",
20  "stamp": "2023-11-07T16:19:27.5166525",
21  "tenantId": "AP",
22  "totalContributionPercentage": 0.1215,
23  "userId": "c054a925-4e4c-4e04-9c87-25e7bc30afc9",
24  "id": "2343efc3-5da7-4ccb-92c6-aa36896126e7",
25  "_rid": "Ku8NAMFBZwoCAAAAAAAAAA==",
26  "_self": "dbs/Ku8NAA==/colls/Ku8NAMFBZwo=/docs/Ku8NAMFBZwoCAAAAAAAAAA==/",
27  "_etag": "\"22048246-0000-0100-0000-65c26bef0000\"",
28  "_attachments": "attachments/",
29  "_ts": 1707240431
30 }
```

Leverage data with the SQL Endpoint


CosmosMirror

8






Home
Reporting

New SQL query
New visual query
New report
New measure

+

Warehouses

▼

CosmosMirror

▼

Schemas

▼

dbo

▼

Tables

▼

[CosmosMirror].[dbo].[life_Voucher]

▼

Views

▼

life_Voucher

▼

life_Commit...

▼

life_Commit...

▼

life_Commit...

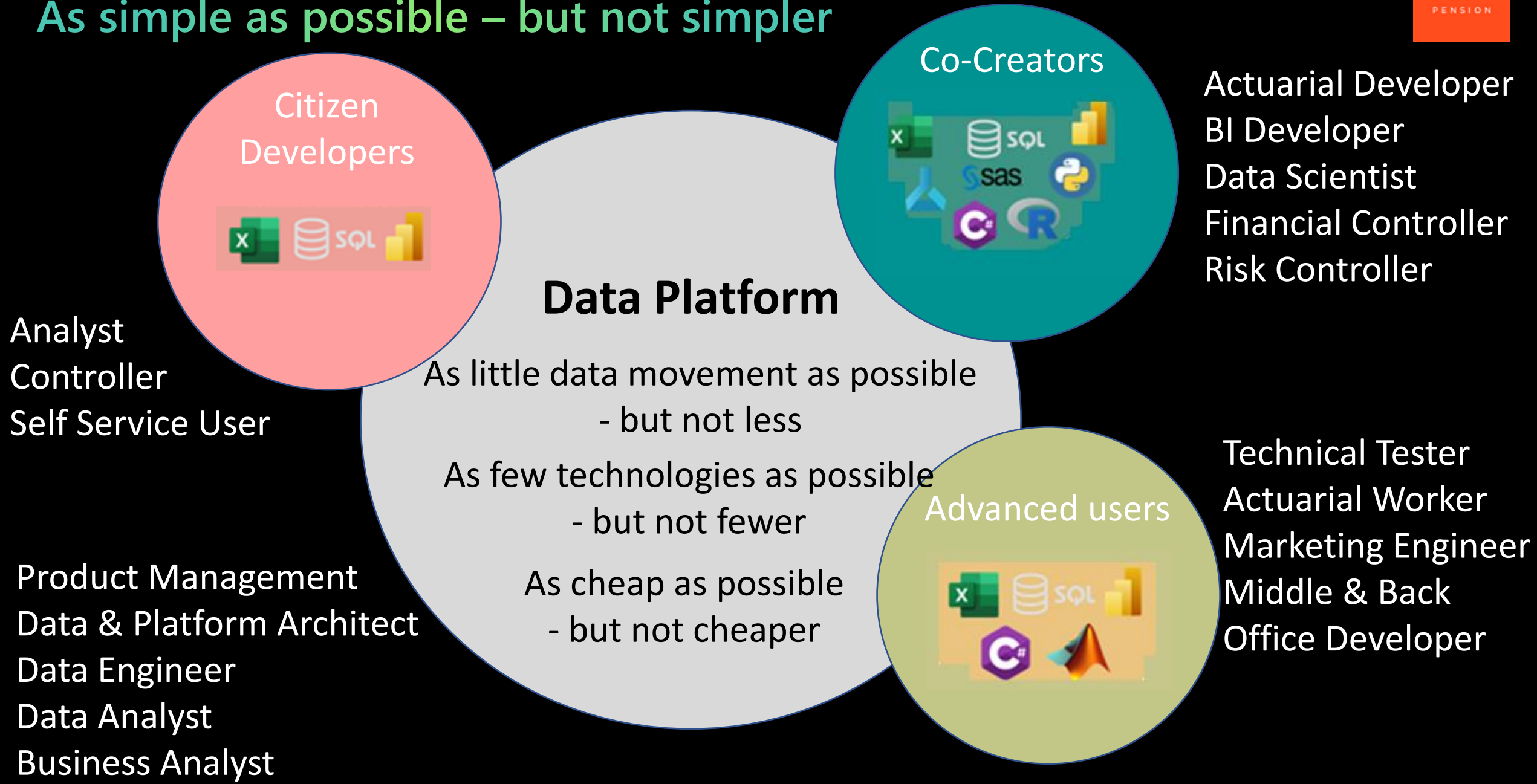
Showing 13 rows

	ABC _rid	ABC _type	ABC _baseType	ABC _id	12L amount	ABC commitmentOwner	ABC companyAgreementId
1	Ku8NAMfBZwoCAAAAAAAAAA==	MoneyInVoucher	Voucher	2343efc3-5da7-4ccb-92c6-aa36896126e7	700000	{"_type":"PersonRef","externalRef":"1305651234"}	04001
2	Ku8NAMfBZwoDAAAAAAAAA==	MoneyInVoucher	Voucher	1889f34c-027e-4837-8560-724e1073a0a0	465000	{"_type":"PersonRef","externalRef":"3103571234"}	04001
3	Ku8NAMfBZwoEAAAAAAAAA==	MoneyInVoucher	Voucher	433de3b7-3d75-4c87-9d31-37747e6272e8	400000	{"_type":"PersonRef","externalRef":"0106561234"}	04001
4	Ku8NAMfBZwoFAAAAAAAAAA==	MoneyInVoucher	Voucher	0333b861-0399-4802-aa08-b22c1cd0e70b	900000	{"_type":"PersonRef","externalRef":"2302611234"}	04002
5	Ku8NAMfBZwoGAAAAAAAAA==	MoneyInVoucher	Voucher	df0aa66c-4668-48a3-b7c6-b929b9da577a	1100000	{"_type":"PersonRef","externalRef":"2201611234"}	04002
6	Ku8NAMfBZwoHAAAAAAAAA==	MoneyInVoucher	Voucher	627906c0-aaca-4f76-b2c0-818f415cc8cc	1000000	{"_type":"PersonRef","externalRef":"2302701234"}	04002
7	Ku8NAMfBZwoIAAAAAAAAAA==	BillingVoucher	Voucher	c4c8f26c-8140-4d1a-a9cb-067a40c32963	null	{"_type":"PersonRef","externalRef":"0106561234"}	null
8	Ku8NAMfBZwoJAAAAAAAAA==	BillingVoucher	Voucher	c89fcb9f-3441-45b4-838e-d42670015392	null	{"_type":"PersonRef","externalRef":"2201611234"}	null
9	Ku8NAMfBZwoKAAAAAAAAA==	BillingVoucher	Voucher	cf3d7fca-5e61-4156-b907-5e3d1694d8f5	null	{"_type":"PersonRef","externalRef":"3103571234"}	null
10	Ku8NAMfBZwoLAAAAAAAAA==	BillingVoucher	Voucher	254f380d-dd8e-46ae-a8d9-97ab23c855ee	null	{"_type":"PersonRef","externalRef":"3103571234"}	null
11	Ku8NAMfBZwoMAAAAAAAAAA==	BillingVoucher	Voucher	6a784e2e-837f-4bcf-b1a6-7553991a388b	null	{"_type":"PersonRef","externalRef":"2302701234"}	null
12	Ku8NAMfBZwoNAAAAAAAAA==	BillingVoucher	Voucher	61f154e4-f7b3-4ec9-8214-0e578ae7a1a5	null	{"_type":"PersonRef","externalRef":"1305651234"}	null
13	Ku8NAMfBZwoOAAAAAAAAA==	BillingVoucher	Voucher	401d3b39-5e47-48c4-bd43-ceb868343ad7	null	{"_type":"PersonRef","externalRef":"2302611234"}	null

SQL analytics endpoint

Data Platform, AP Pension

As simple as possible – but not simpler



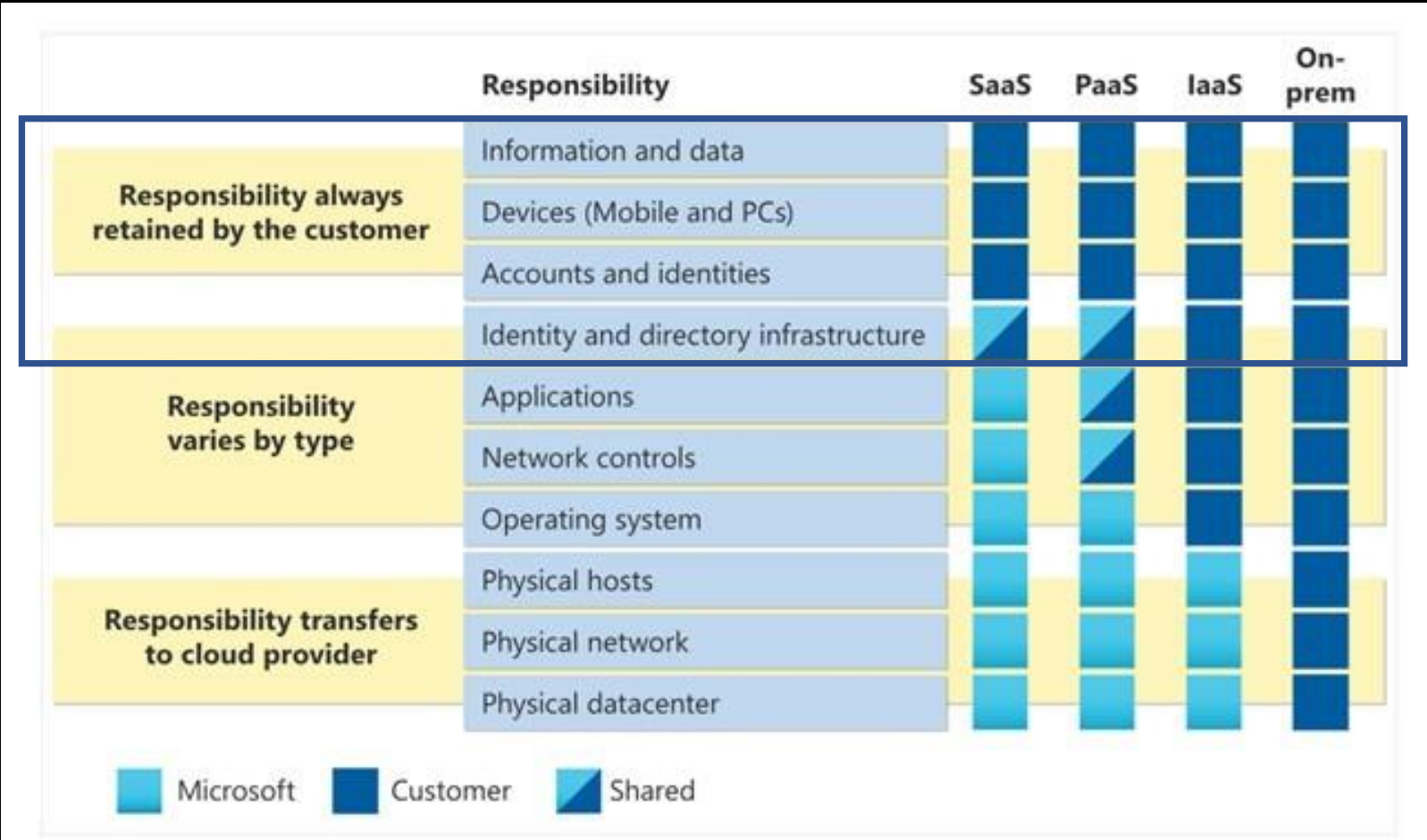
Enterprise Data Fabric – the PoC

“A data fabric delivers a unified, integrated, and intelligent end-to-end data platform. It automates all data management functions — including ingestion, transformation, orchestration, governance, security, preparation, quality, and curation.”

Noel Yuhanna

The Forrester Wave™: Enterprise Data Fabric, Q2 2022

Shared responsibility with SaaS



The risk of non-compliance



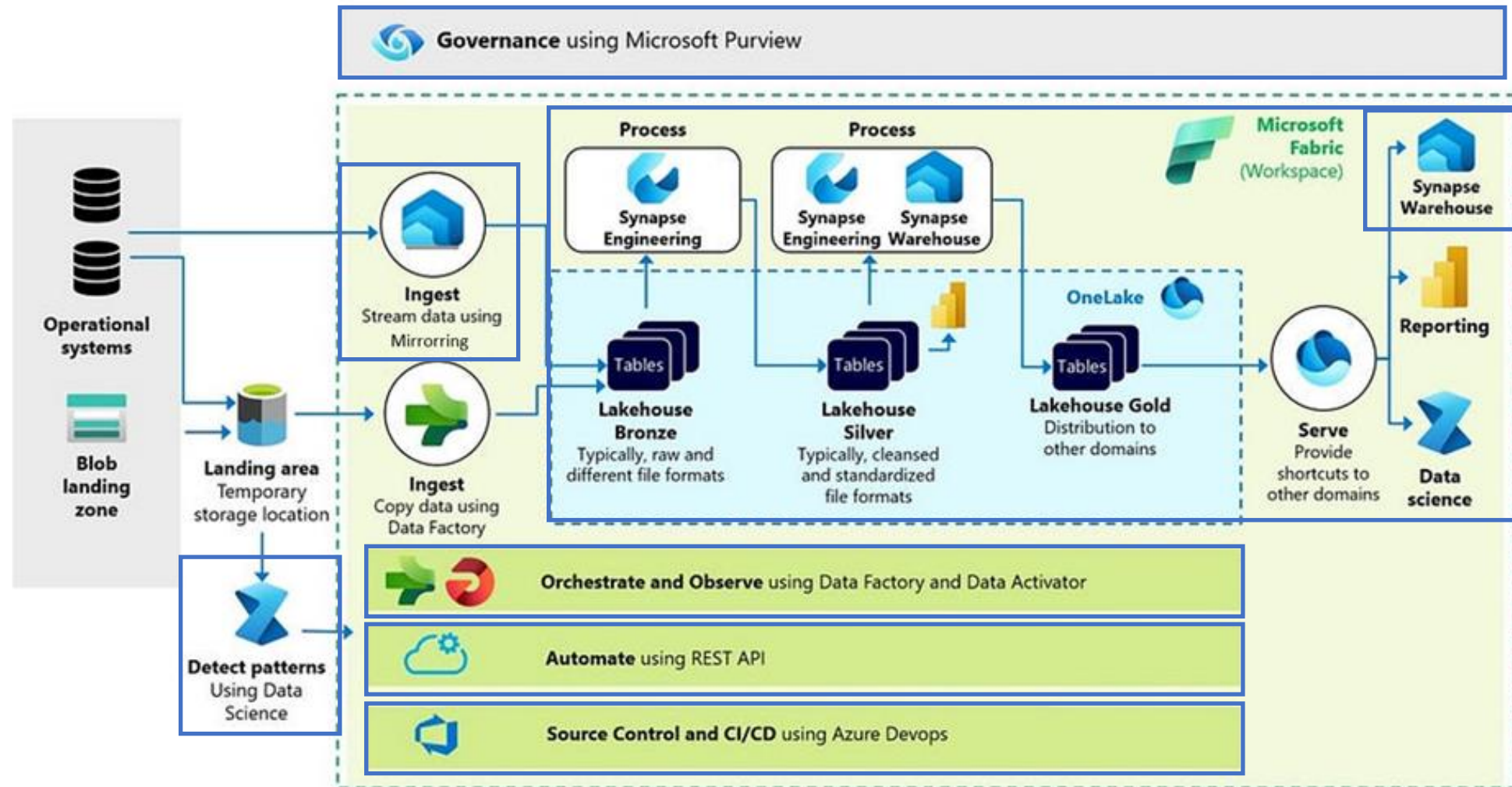
The EDPS has therefore decided to order the Commission, effective on **9 December 2024**, to **suspend all data flows** resulting from its use of Microsoft 365 **to Microsoft** and to its affiliates and sub-processors **located in countries outside the EU/EEA** not covered by an **adequacy decision**. The EDPS has also decided to order the Commission to **bring the processing operations** resulting from its use of Microsoft 365 **into compliance** with Regulation (EU) 2018/1725. The Commission must demonstrate compliance with both orders by 9 December 2024.

The PoC



Is it possible to build a high performance, metadata driven and template-based framework in Fabric to enable/handle

- all our known ingestion types**
- the encryption of PII-data at load time**
- slowly changing dimensions and bi-temporal timelines**
- orchestration of asynchronous updates**
- One Enterprise Data Platform**
- controlled collaboration and co-creation**

Enterprise Data Fabric




Microsoft Fabric Release plan

[Learn](#) /  

Microsoft Fabric release plan documentation

The Microsoft Fabric release plan documentation announces the latest updates and timelines to customers as features are prepared for future releases.

Get started


 [OVERVIEW](#)

[Admin and governance](#)

[OneLake](#)

[Fabric shared experiences](#)

Synapse

 [OVERVIEW](#)


[Synapse Data Warehouse](#)

[Synapse Data Engineering](#)

[Synapse Data Science](#)


[Synapse Real-Time Analytics](#)

Data Factory

 [OVERVIEW](#)

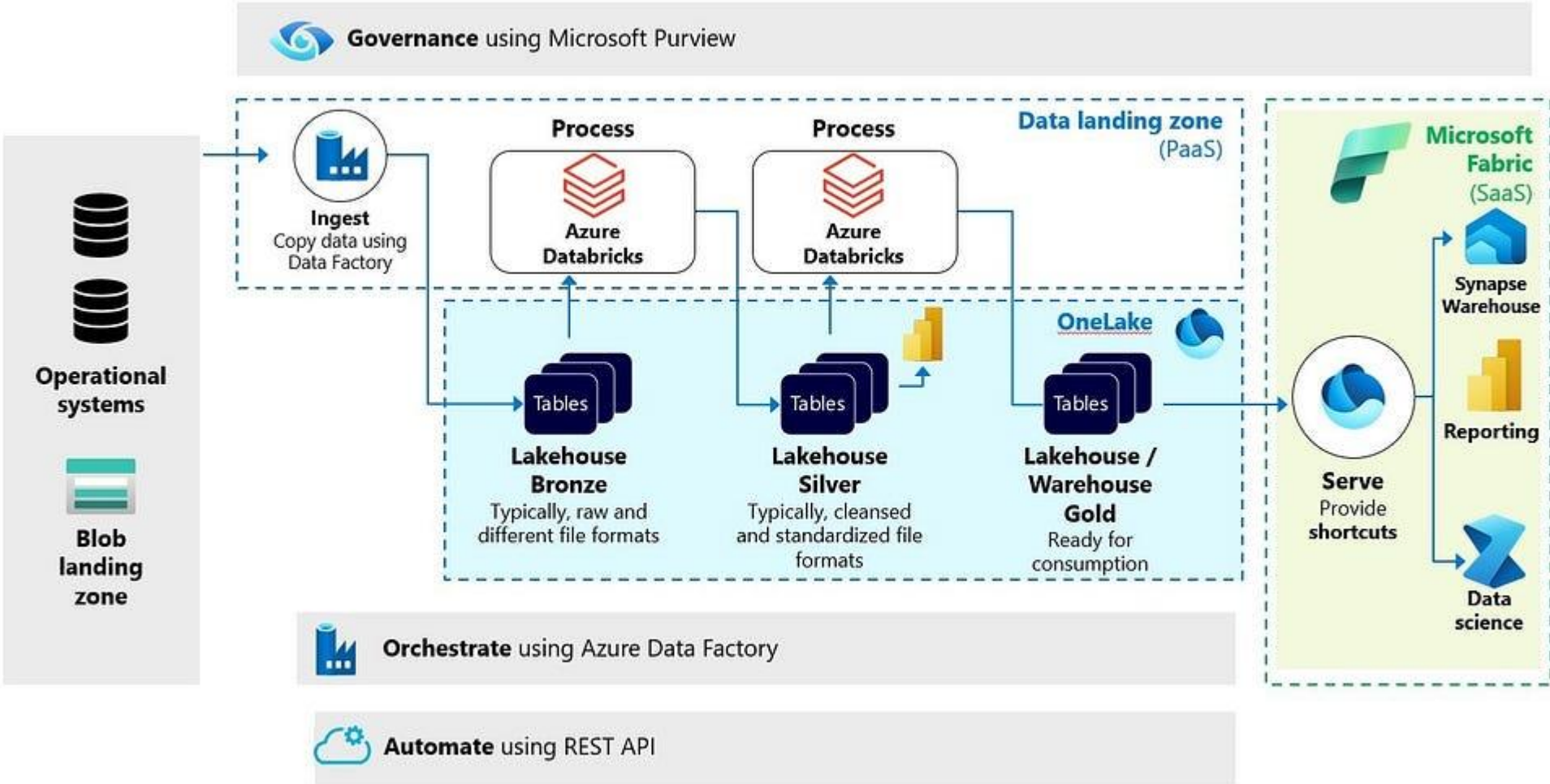
[Data Factory](#)

Power BI

 [OVERVIEW](#)

[Power BI](#)

Enterprise Data Fabric, Databricks/OneLake



Achievements

- ✓ **First-hand experience with Fabric.**
- ✓ **Best practices that support Databricks and Fabric.**
- ✓ **Encryption of PII-data at time of ingestion**
- ✓ **Performance measures of Databricks vs. Lakehouse vs. DataWarehouse.**
- ✓ **List of blocking issues in Fabric.**
- ✓ **A coherent design that supports our MVP Enterprise Data Fabric in 2024 and our roadmap for 2025.**

Performance

Fact load: 57 million rows, 200 columns

Optimization on source	Databricks 40 Cores	Fabric Lakehouse (F64)	Fabric Warehouse (F64)
V-Order	7 min 24 sec	3 min 34 sec	4 min 15 sec
V-Order + Z-order	7 min 50 sec	3 min 25 sec	4 min 15 sec

	V-Order	No Optimizations	Z-Order	V-Order + Z-order
Databricks		12,9 GB	10,4 GB	6,6 GB
Fabric	18 GB			

Achievements



Encryption of PII-data at time of ingestion

Explorer

lakehouse_archive_puzzel

Tables

dbo_Agent_Event_Sour...

dbo_call_events

dbo_enqreg_header

Unidentified

Files

dbo_call_events

Showing 1000 rows

APDATA_So...	ABC CprNummer	123 IntegrationK...	123 caller_on_ho...	ABC add_originat...	ABC leg_type	123
3/18/2024 10:39:...	gAAAAAABl-Bsr_K...	1109914				
3/18/2024 10:39:...	gAAAAAABl-Bvhvf...	1096578				
3/18/2024 10:39:...	gAAAAAABl-Bvhsk...	1096578	0		DirectCall	0
3/18/2024 10:39:...	gAAAAAABl-Bvhm...	1096578				
3/18/2024 10:39:...	gAAAAAABl-BvhK...	1096578				
3/18/2024 10:39:...	gAAAAAABl-Bvhm...	1096578				
3/18/2024 10:39:...	gAAAAAABl-BvhN...	1096578				
3/18/2024 10:39:...	gAAAAAABl-BvhN...	1096578				
3/18/2024 10:39:...	gAAAAAABl-BvhH...	1096578	0		Agent	0
3/18/2024 10:39:...	gAAAAAABl-BvhK...	1096578				
3/18/2024 10:39:...	gAAAAAABl-BvhD...	1096578				
3/18/2024 10:39:...	gAAAAAABl-Bvhw...	1096578				
3/18/2024 10:39:...	gAAAAAABl-BvhR...	1096578	0		DirectCall	0
3/18/2024 10:39:...	gAAAAAABl-Bvhja...	1096578				
3/18/2024 10:39:...	gAAAAAABl-Bvhu...	1096578				
3/18/2024 10:39:...	gAAAAAABl-BvhlT...	1096578				

Succeeded (18 sec 107 ms)

Columns 51 Rows 1,000

Achievements



Metadata driven ingestion framework – from Databricks to OneLake

puzzel_stream_append users/sda/puzzel Python Last edit was yesterday New cell UI: OFF

Workspace

- base
- translations
- validations
- load_base_portman_transformation...
- portman_stream_append_non_UC
- puzzel_stream_append

Parameters / setup

```

1  ## Set metadata and parameters
2  table_prefix = "puzzel"
3  # Paths
4  base_location = f'/mnt/onelake_base_{table_prefix}/'
5  base_location_azure = f'{config.azure_storage_mount_location}_base'
6
7  # Configs
8  incremental_load = True
9  source_system = "Puzzel"
10
11 # Get metadata for source
12 source_object_metadata = metadata.source_object_metadata(source=source_system)
13 print(source_object_metadata)
14

```

Achievements

 Template based dimensional framework – transferrable to Fabric

Parameters

Cmd 4

```
1  ## Set metadata and parameters
2
3  # Parameters
4  table_name = "enq_topic"
5  primary_key_column = ['enq_topic_key']
6  full_load = False
7
8  # Paths
9  curated_location = f'/mnt/onelake_curated_{config.dim_prefix}/'
10
11 # Naming
12 dimension_name = f"{table_name}"
13 dimension_table = f"{config.hive_catalog}.curated_{config.dim_prefix}.{dimension_name}"
```

Cmd 5

Cmd 9

Load

Cmd 10

```
1  ## Load dimension table using apply_changes pattern.
2  write.upsert_type_1(
3      destination_table=dimension_table,
4      df=df,
5      primary_key_columns= primary_key_column
6  )
```

Achievements

 Access to data from PowerBI, Fabric, SAS, SSMS with little to no additional set-up

Power BI Dataplatform Dev

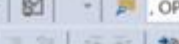
Dataplatform Dev

+ New ▾ ↑ Upload ▾ ⚙ Create deployment pipeline 📦 Create app

Name	Actions
Lakehouse_curated_dim	<ul style="list-style-type: none"> Open Open with > Explore this data (preview) Analyze in Excel New report Favorite Manage permissions Copy SQL connection string Settings View lineage View details
Lakehouse_curated_fact	




















SQLQuery4.sql - [je@appension.dk (171)]* - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

Execute ✓  OPTIMIZE_FOR_SEQUENTIAL_K

Lakehouse_curated_dim

Object Explorer

Connect ▾                   

Performance

Joining the fact with a couple of dimensions before grouping (output: 91 rows)

Query Compute Engine	Compute Size	Duration
Databricks SQL cell in Notebook	40 Core Cluster	32,5 sec (2,2 sec cached)
Databricks SQL cell in Notebook	40 Core Cluster (Photon)	19,7 sec (1,2 sec cached)
Fabric Lakehouse Spark SQL in Notebook	40 Core Cluster	50,2 sec (2,5 sec cached)
Fabric Lakehouse SQL Endpoint, Web UI	32 V-cores (0,5*64)	4,5 sec (2,5 sec cached)
Fabric Lakehouse SQL Endpoint, SSMS	32 V-cores (0,5*64)	7,5 sec (0,5 sec cached)
Fabric Warehouse SQL Endpoint, WEB UI	32 V-cores (0,5*64)	13,6 sec (2,5 sec cached)

Achievements



Purview integration

Microsoft Purview

PREVIEW

Search

Ny Microsoft Purview-portal

1

Jacob Rønnow Jen...

JJ

Startside

Datakatalog

Oversigt

Gennemse

Ordliste over erhvervsudtryk

Datadeling

Insights for dataejendom

Løsninger

Relaterede løsninger

Datatilknytning

Data Catalog

Browse, search, and discover data assets across your organization.

No sources

422+ assets

No glossary terms

Search catalog

Explore your data

Microsoft Azure

View Azure subscriptions and their contents.

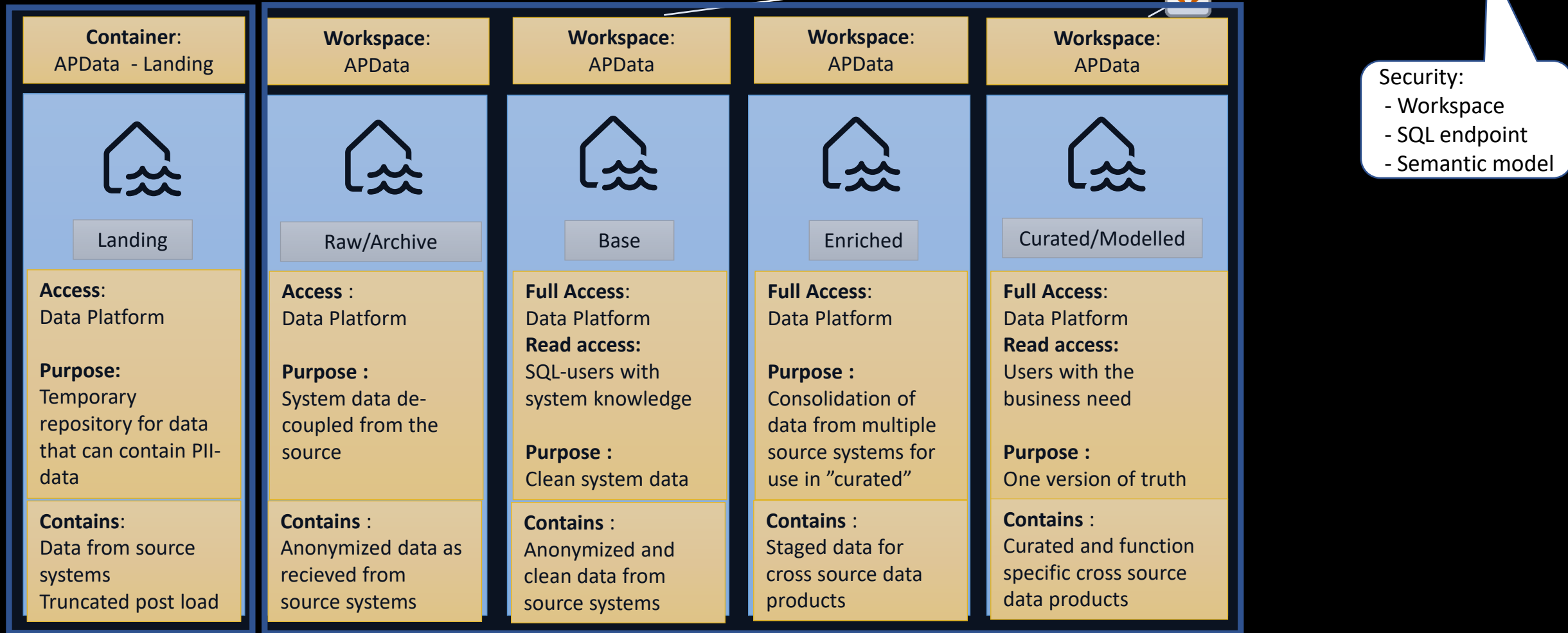
Microsoft Fabric

View all workspaces that you have access to.

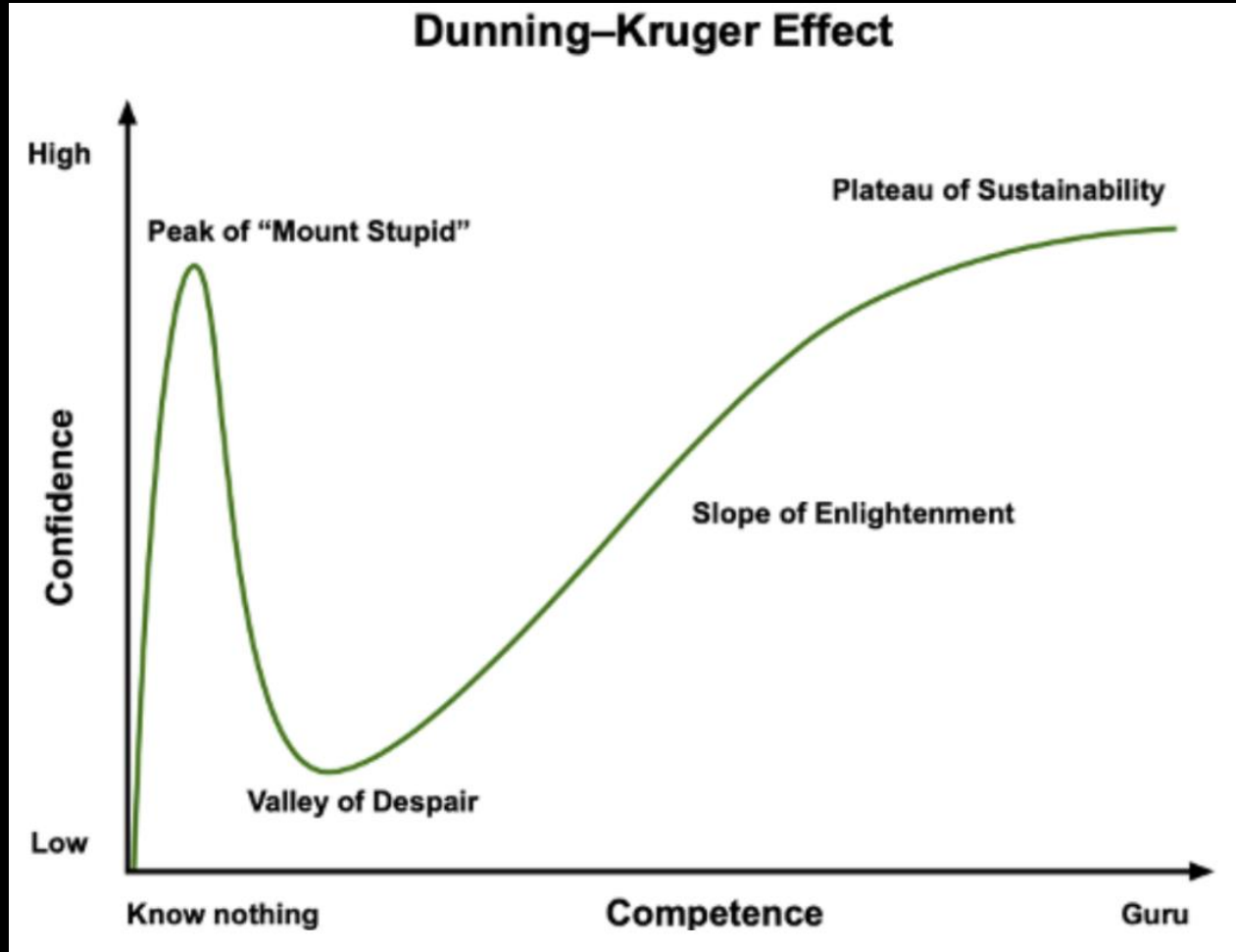
Quick access

From PoC to MVP

Focus on user experience and security



Starting a Fabric Journey?



Stick your toes in the water

Make a list of your organization's must-haves

Be aware of the roadmap

SaaS is not PaaS

Be curious (DP-600 is a good starting point, but the resources are endless.


Use your network and have an expert on retainer

Make a plan for adoption

Sharing is caring

Microsoft Microsoft Security Azure Dynamics 365 Microsoft 365 Microsoft Teams Windows 365 All Microsoft

Customer Stories Search



AP Pension revolutionizes its data governance and analytics with Microsoft Fabric

May 23, 2024

Print

Learn More

Microsoft Fabric
Mirroring in Fabric

AP Pension had decades' worth of business and investment data and needed to establish a consolidated approach to analytical data that fits its principles and digital strategy. The company is redefining its insights with Microsoft Fabric to make actuarial teams, advanced users, and citizen developers benefit from the enhanced capabilities and flexibility of the platform. Co-creating and prototyping products in a secure, collaborative environment facilitates innovation, accelerating the development of solutions tailored to the firm's needs.

Microsoft Customer Story-AP Pension revolutionizes its data governance and analytics with Microsoft Fabric



Posts Comments Images Articles Reactions

Responsibility SaaS

Information and data
Devices (Mobile and PCs)
Accounts and identities

Shared

Being responsible for data in OneLake

"OneLake is the OneDrive for your data!". As a user, I understand what it...

by Jacob Rønnow Jensen • 5 min read

Introducing Mirroring in Fabric

At the first annual Microsoft Fabric Community Conference, Microsoft...

by Jacob Rønnow Jensen • 3 min read

Learning from the future as it emerges

A structured and democratized approach to analytical data is one of...

by Jacob Rønnow Jensen • 2 min read

The 2023 Ignite Book of News

With more than 100 announcements, Microsoft Ignite is pretty intense this...

by Jacob Rønnow Jensen • 1 min read

Putting Fabric to the test

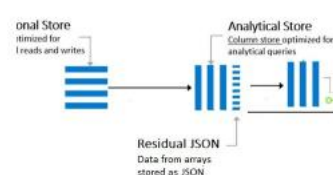
Having worked most of my professional career with data...

by Jacob Rønnow Jensen • 3 min read

Working with JSON and nested arrays in Microsoft Fabric

As Data Engineers, we get more and more semi-structured data as input f...

by Jacob Rønnow Jensen • 4 min read



Activity | Jacob Rønnow Jensen | LinkedIn

Questions?



Let's continue the dialogue

Jacob Rønnow Jensen
jje@appension.dk
www.appension.dk

Data Engineer til Data Platform i AP Pension - AP Pension

OM AP PENSION

Ledige stillinger

Ledige stillinger

Data Engineer til Data Platform i AP Pension

Vil du arbejde med de nyeste teknologier og være med til at gøre en forskel i AP Pension? Så har du chancen nu, for vi søger en erfaren og dygtig Data Engineer til vores Data Platform afdeling.

Om jobbet

Du bliver en del af en afdeling, som har ansvaret for at op- og udbygge en moderne og intuitiv dataplatform til understøttelse af den digitale strategi og en effektiv dagligdag i AP Pensions analyseenheder.

Vi er midt i en re-implementering af vores analytiske dataplatform, hvor du vil indgå i et team med forskelligartede opgaver - herunder:

- vedligehold af de gamle sql-baserede warehouses (on-prem og i azure)
- migrering til en Lakehouse arkitektur med Databricks og Microsoft Fabric
- indhentning og modellering af data fra nye datakilder – og automatisering af denne proces
- sammen med IT at optimere vores nuværende og fremtidige Azure-forbrug
- Drift og DevOps

Din profil

Vi forestiller os, at du er nysgerrig og en holdspiller af natur, og at du herudover har erfaring med data warehousing, Microsoft-platformen og Azure, herunder

- Databricks og SQL-databaser
- Azure Data Factory
- Azure DevOps
- tSQL og pySpark

Har du erfaring med forskellige former for datamodellering og en længerevarende relevant uddannelse inden for IT, økonomi eller naturvidenskab, vil det også være en fordel.

Vi tilbyder

- En udfordrende og spændende hverdag, hvor der er masser at tage fat i, og hvor du får mulighed for at arbejde med de nyeste værktøjer på Microsoft-platformen.
- Et motiverende arbejdsmiljø med gode udviklingsmuligheder.
- En arbejdsplads med fokus på trivsel og arbejdsglæde og en uformel omgangstone.
- Attraktive ansættelsesvilkår, såsom pensionsordning, sundhedsforsikring og tandforsikring
- Vi ligger tæt på offentlig transport. Et stenkast fra Nordhavn St.

Vi ser frem til at høre fra dig

Ønsker du yderligere oplysninger, er du velkommen til at kontakte:

Jacob Rønnow Jensen
Afdelingschef
Mobil: 2912 4284