

Benni De Jagere



Power BI Dataflows?
Why you need to use them!

Sponsors



business.
people.
technology.



Many thanks to our sponsors, without whom
such an event would not be possible.



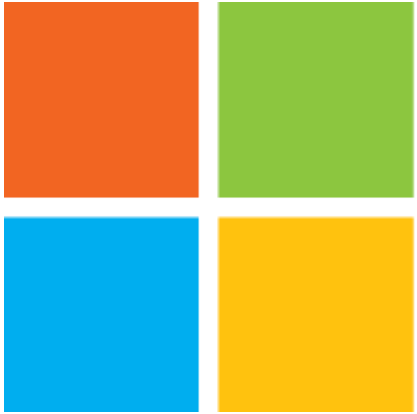
You Rock!

Gold

Silver

Bronze

Thank you






Microsoft

For the Venue



Benni De Jagere?

- Senior Data Insights Consultant
- Realdolmen, a Gfi company 
- **dataMinds**.be, Co-leader
-  @BenniDeJagere
-  /bennidejagere
- #TeamOxfordComma



Dataflows?

~~(Google) Cloud Dataflow~~

~~TensorFlow Dataflow~~

~~SSIS Data Flows~~

~~Microsoft Flow~~

~~ADFv2~~

~~Mapping Data Flows~~

~~Wrangling Data Flows~~

~~PowerApps Dataflows~~

Power BI Dataflows!





Who?

Who should be using Dataflows

(Financial) Power Users

Business Analysts

Tech savvy Users

Everyone else 😊





What?

Only Entity, Many Datasets

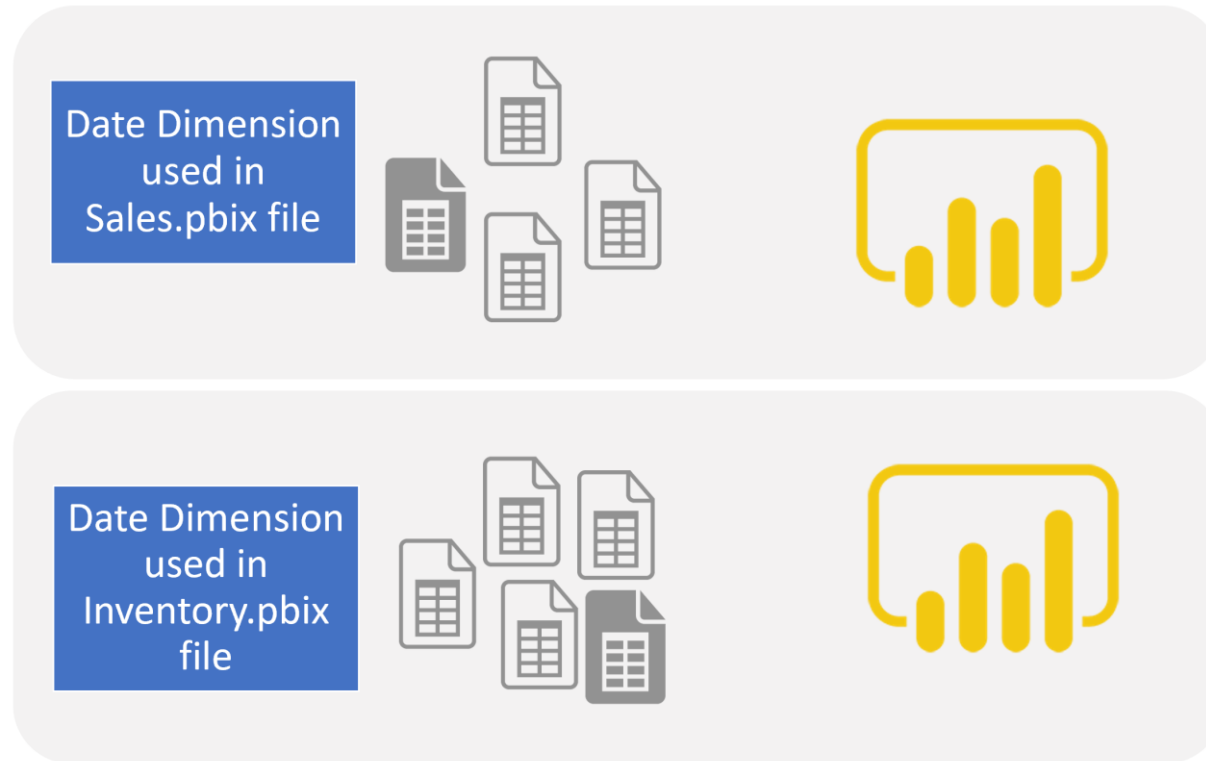
Lack of semantic model will create many versions

Re-use of queries across datasets is limited

Shared Datasets is all or nothing approach



Only Entity, Many Datasets



Date Dimension transformation executed multiple times,
when only once is needed

Only Entity, Many Datasets



Date Dimension transformation executed only once, and then used multiple times



How?

Enter, Power BI Dataflows!

Preview Late 2018, GA April 2019

Cloud based PowerQuery experience

Self-Service Data Preparation

Non-technical user personas



Enter, Power BI Dataflows

Build reusable Data Entities

Not a replacement for a DWH Solution

Manageable and securable like all Power BI Objects

Managed CDM Folders





Demo: Dataflows

New York Citibikes

www.citibikenyc.com/system-data

Public Open Data

Starts June 2013

Information about every trip

Longer than 60 seconds

Starts at public station

Masterdata



<https://i0.wp.com/thenypost.files.wordpress.com/2013/12/citibike1.jpg?quality=90&strip=all&ssl=1>





How?

Common
Data Model

Common Data Model

Unified metadata layer for Cloud Apps and Services
Shared

- Data Language

- Storage Format and Location

- Set of Default Entities

Extend Entities for all consumers

Simplified integration and disambiguation of data

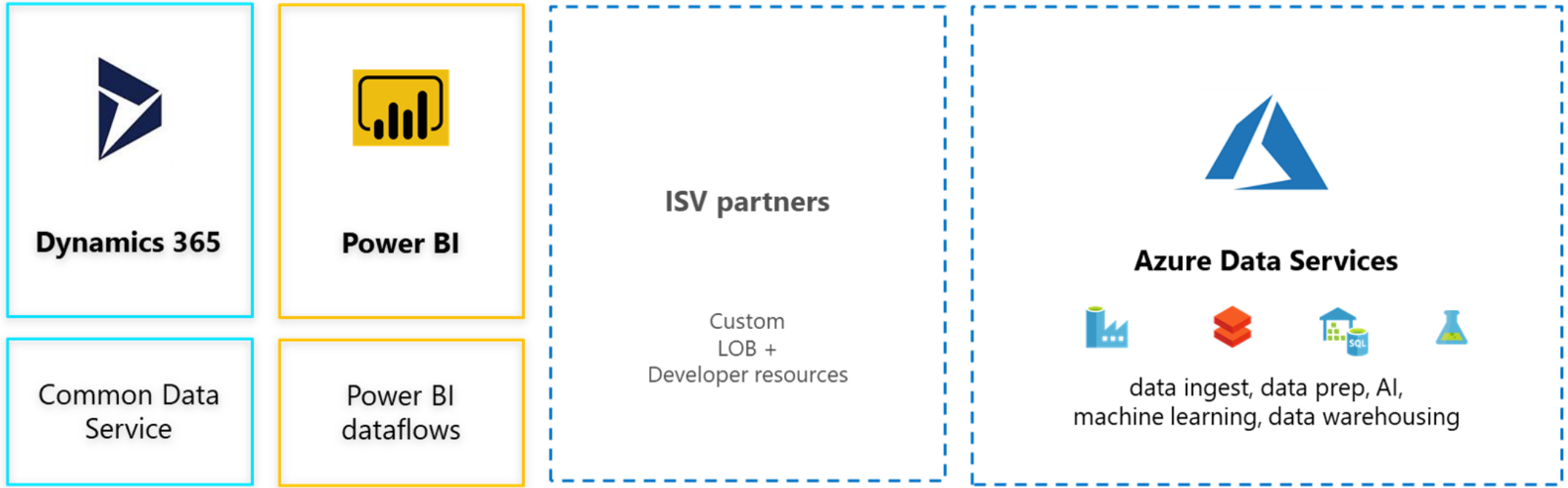


Decision Makers, Business analysts

No code, low code

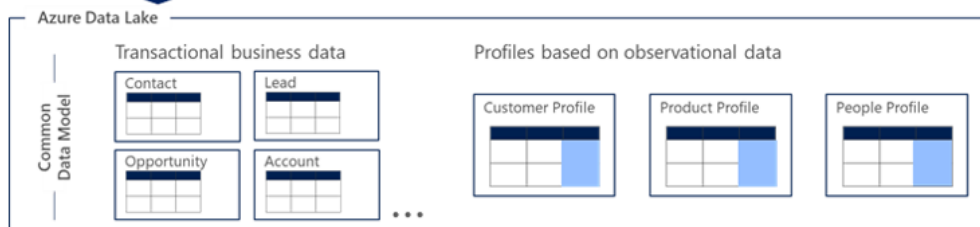
Data scientists, Data engineers

Low to high code

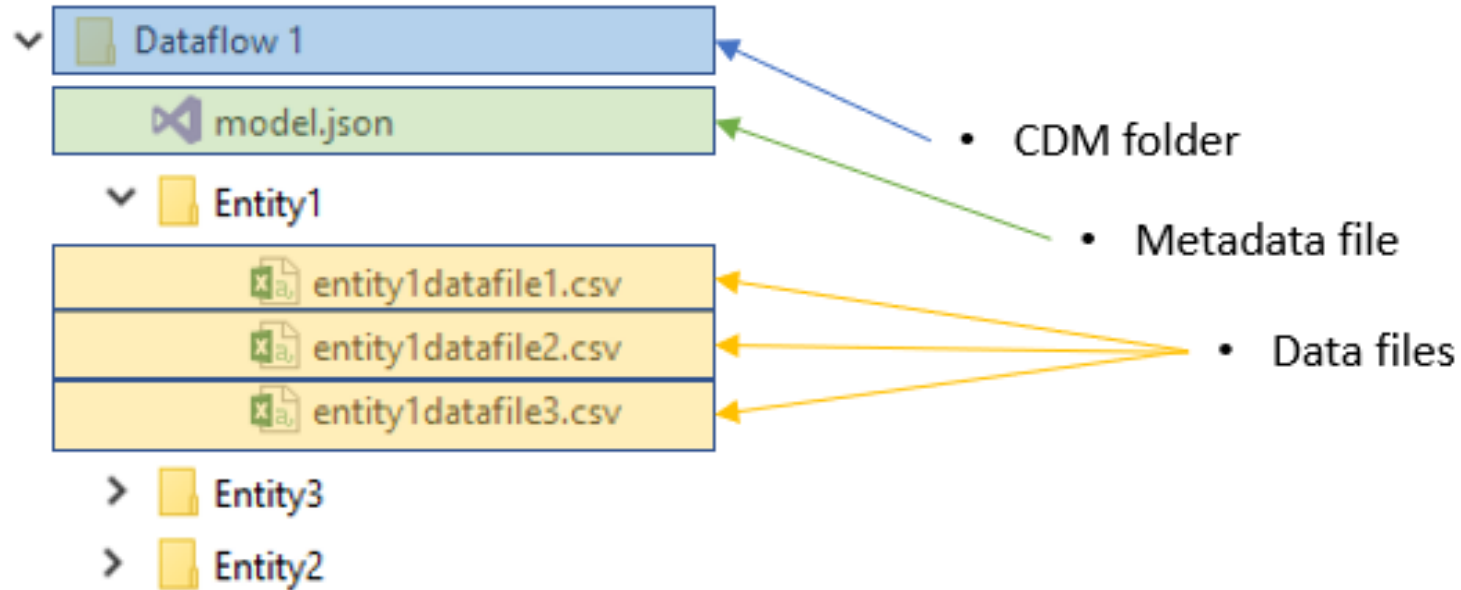


Azure Data Lake Storage Gen2

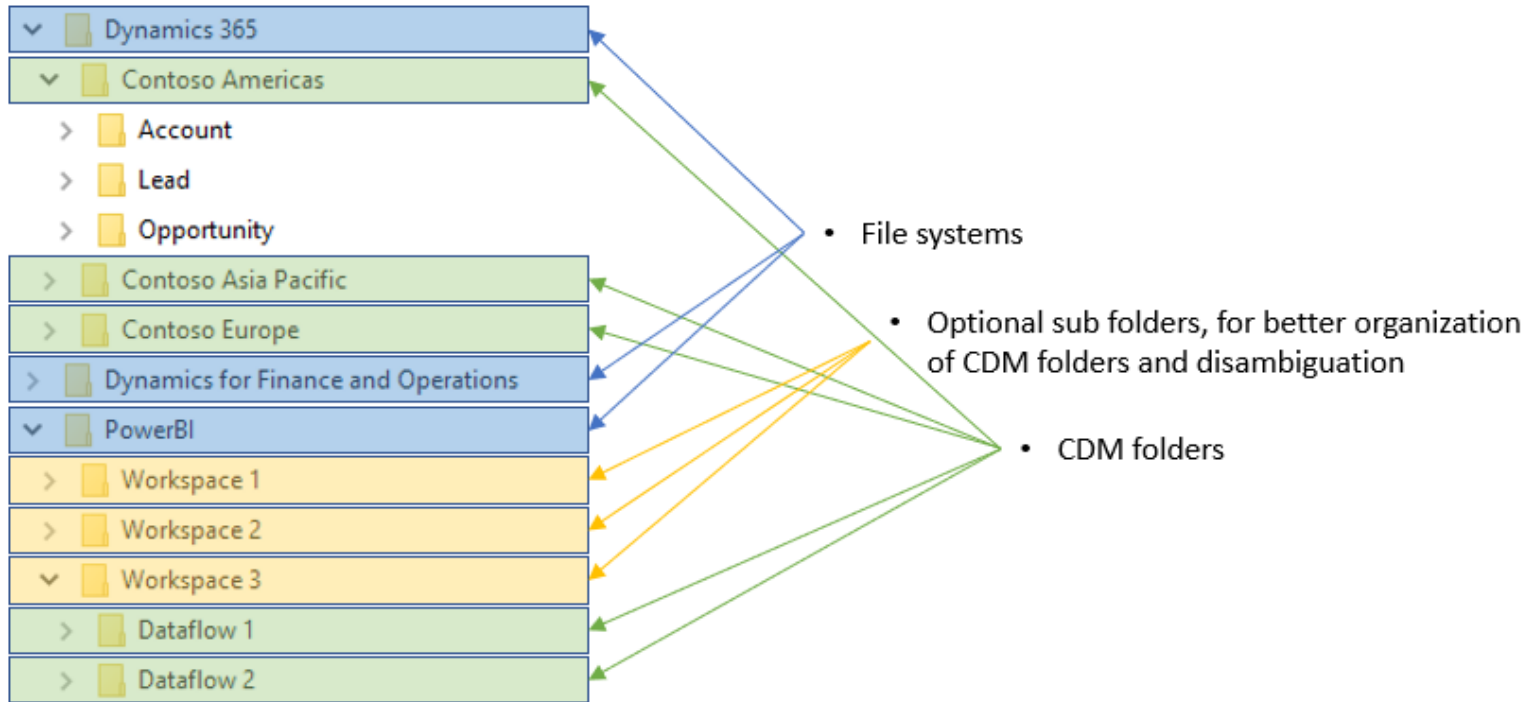
“CDM folders”



Common Data Model



Common Data Model





How?

Data Refresh

Data Refresh

Trigger Manual refresh in the UI

Scheduled refresh comparable to Power BI Datasets

Refresh History offers an interesting view

Automatically refreshes any dataflows it is dependent on

Incremental Refresh is Premium Only (for now?)



A large, stylized teal graphic on the left side of the slide, consisting of several overlapping curved lines that form a partial circular shape.

Demo : Scheduled Refresh



Query Reusage

Query Reusability

No developer experience in Power BI Desktop

- Copy/Paste queries between environments

Linked/Computed Entities (Premium Only)

- In-lake Compute, which requires dedicated capacity

- Can Reference any Entity in any Workspace

- Data is only stored in the original Entity

Reference Query, Append or Merge

- Create a query, and mark 'Disable for load'

- Reference this query, in the same workspace

- Logic is reused, and data is duplicated every time the query is referenced



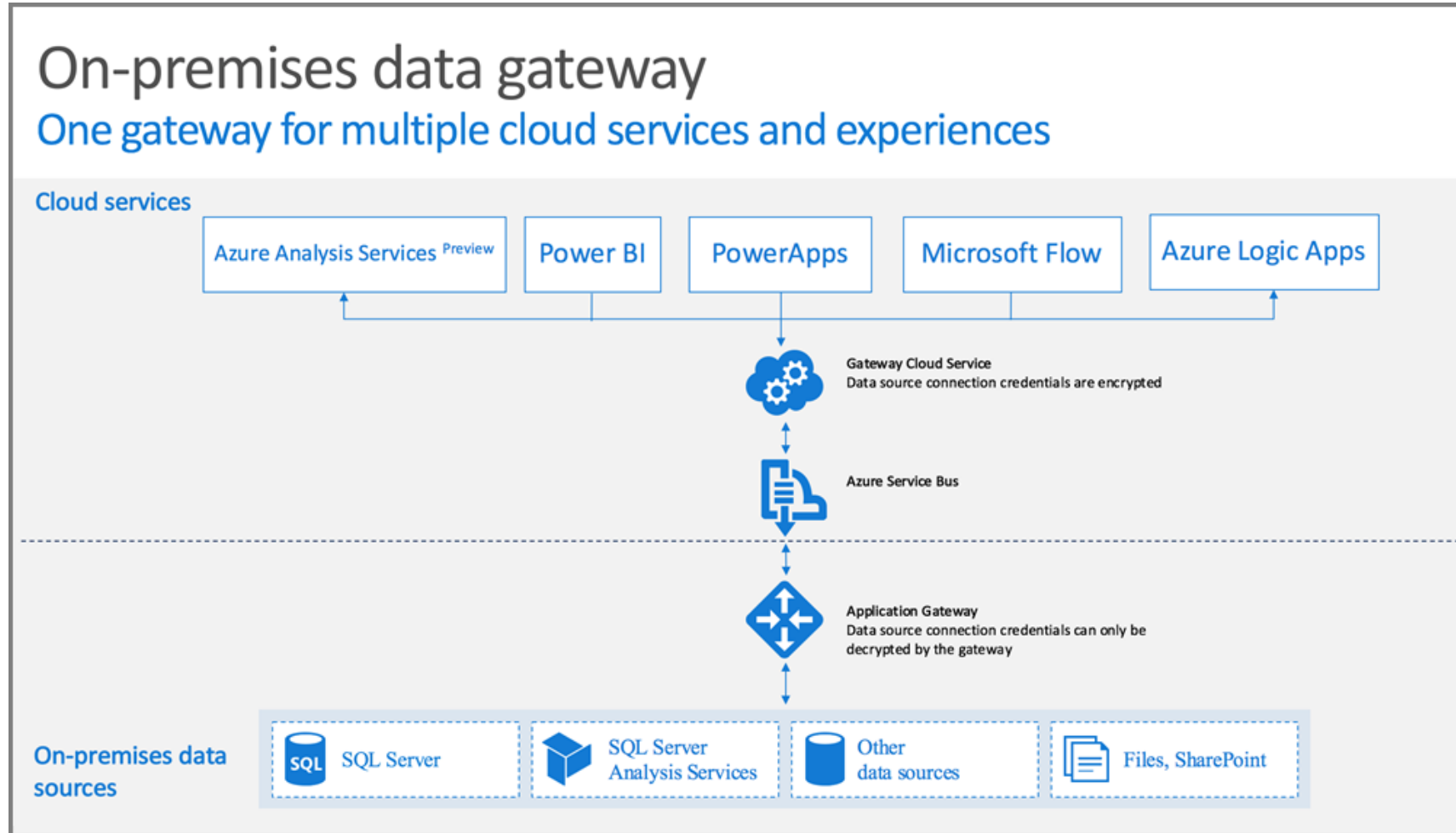
A large, stylized teal graphic on the left side of the slide, resembling a thick, curved arrow or a stylized letter 'C' that points towards the right.

Demo : Query Reusage

A large, stylized teal graphic on the left side of the slide, resembling a thick, curved arrow pointing right, composed of several parallel lines.

Gateway & Security

On-Premises Datasources



Securing Dataflows

Dataflows reside in Workspaces

Secure/Hide Linked Entities by disabling load

Datasets using Dataflows will always be Import Mode

Security is the responsibility of the dataset designer!



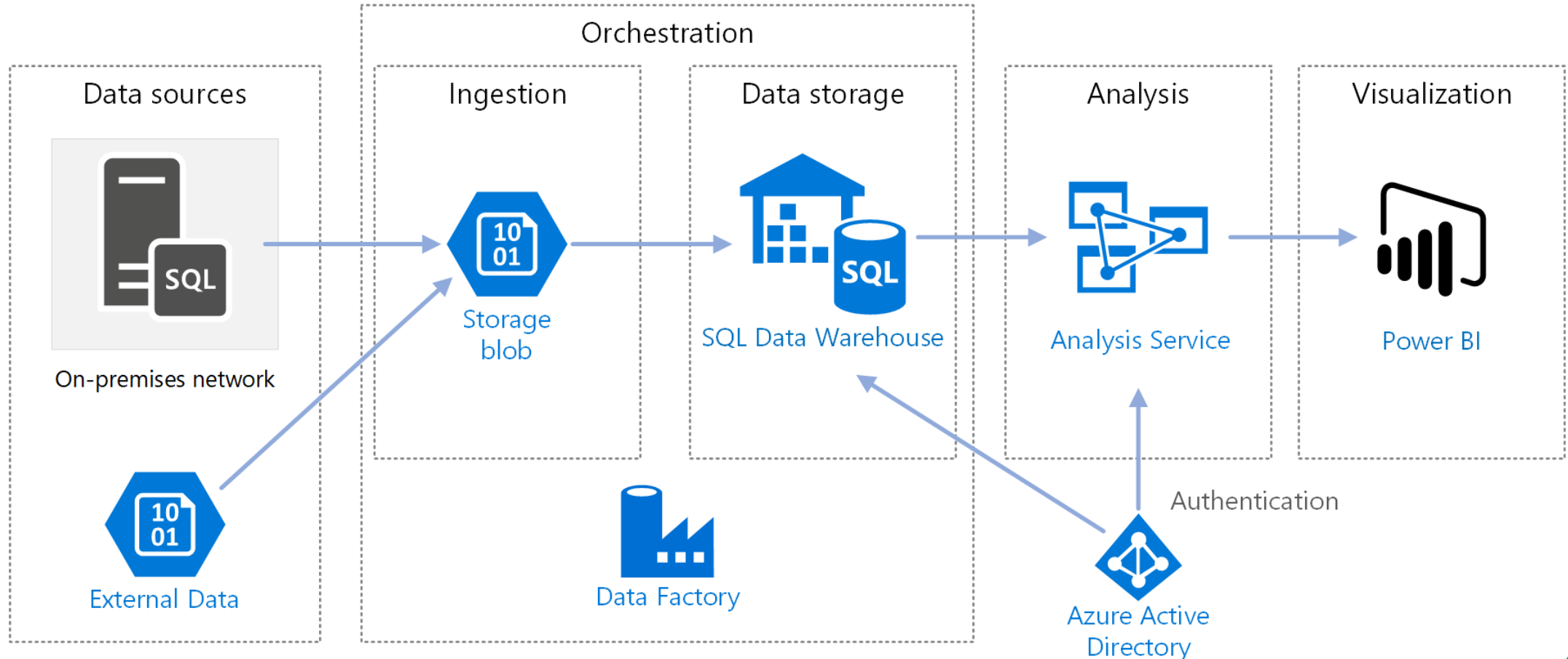
A large, stylized teal graphic on the left side of the slide, consisting of several overlapping curved lines that form a partial circular shape.

Demo : Gateway & Security

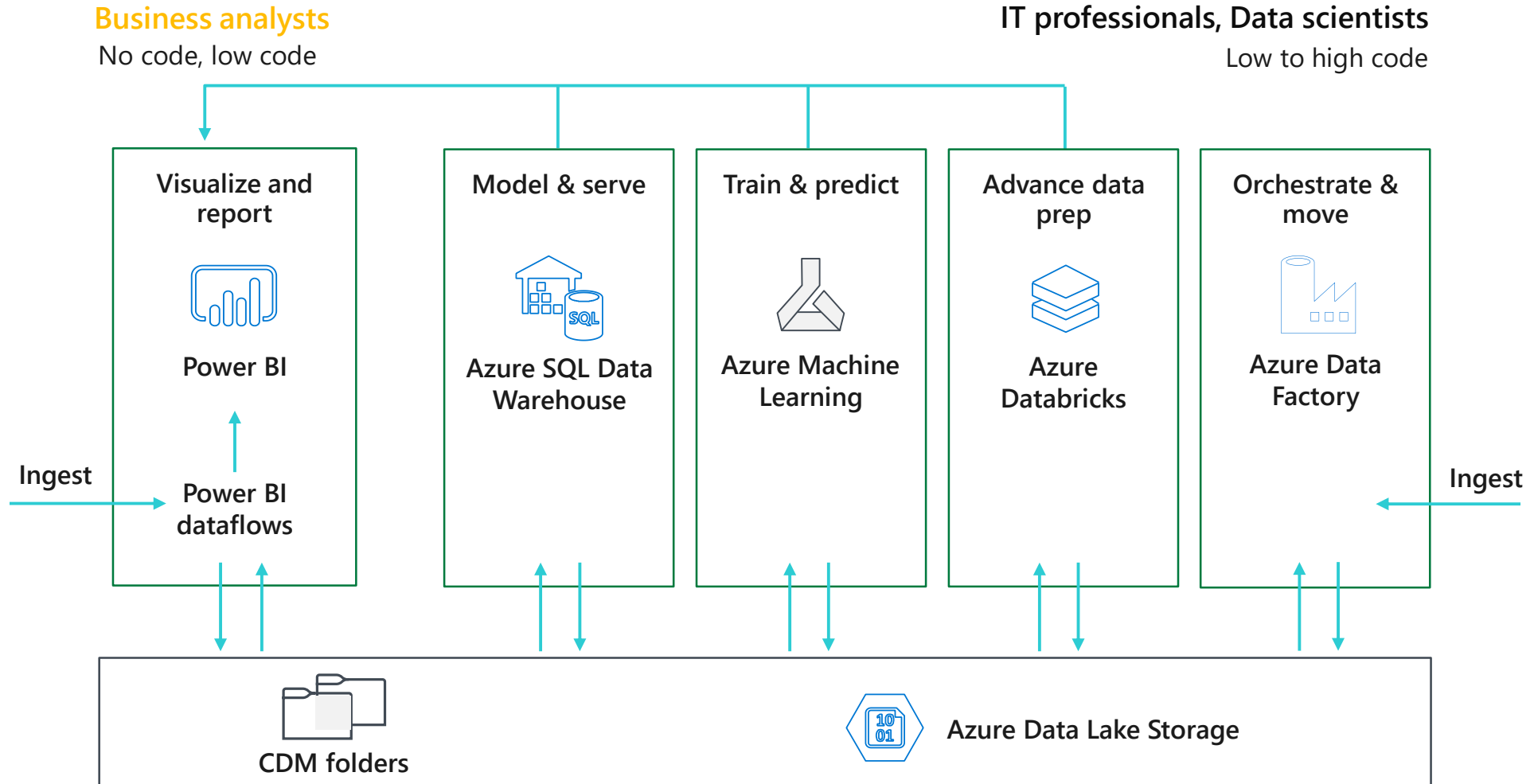


The Big Picture

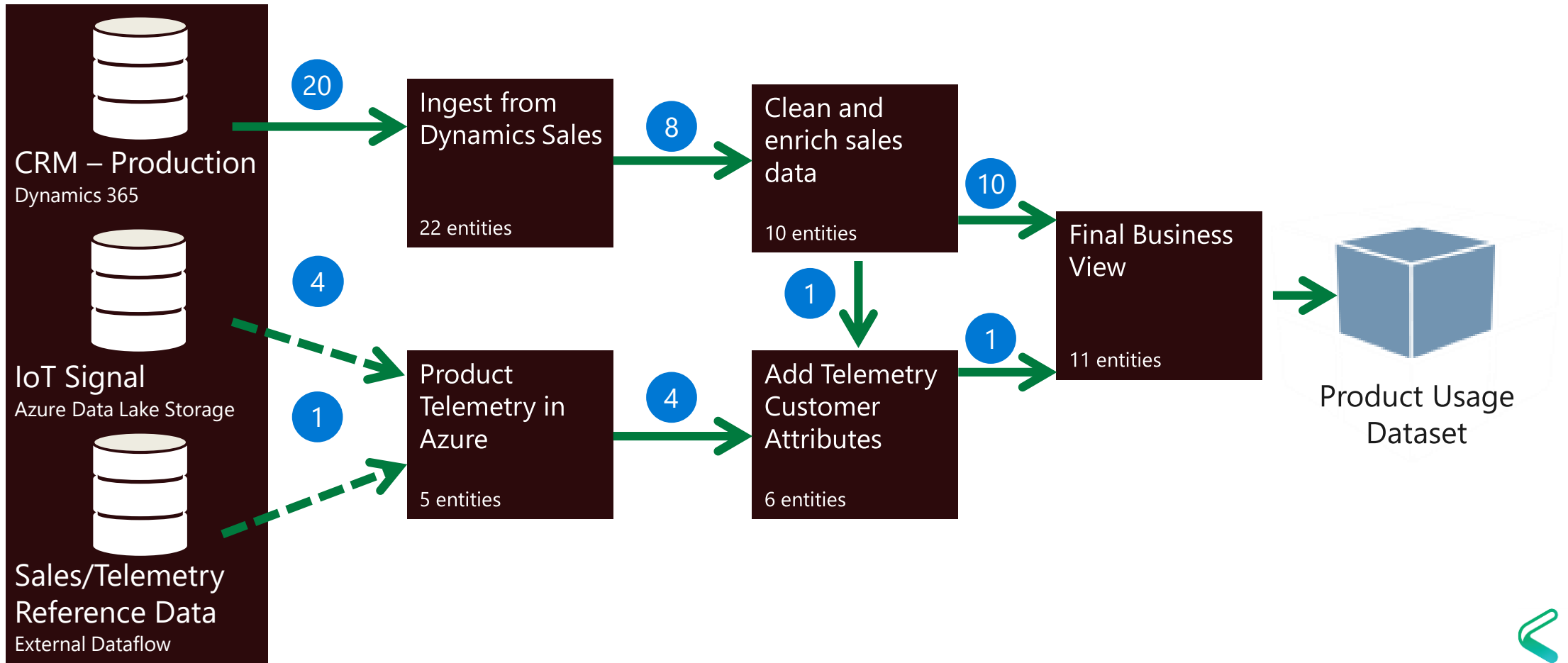
The Bigger Picture



The Bigger Picture



The Bigger Picture





Why?

Where do Dataflows fit in?

Without dataflows, users might

- Be blocked on IT involvement

- Use Excel and manual processes

- Require 3rd party data preparation tools

Dataflows are for Power Query users

- Easy to build reusable data entities

- Easier to compose and orchestrate

- Multi-stage, multi-user data prep workflows

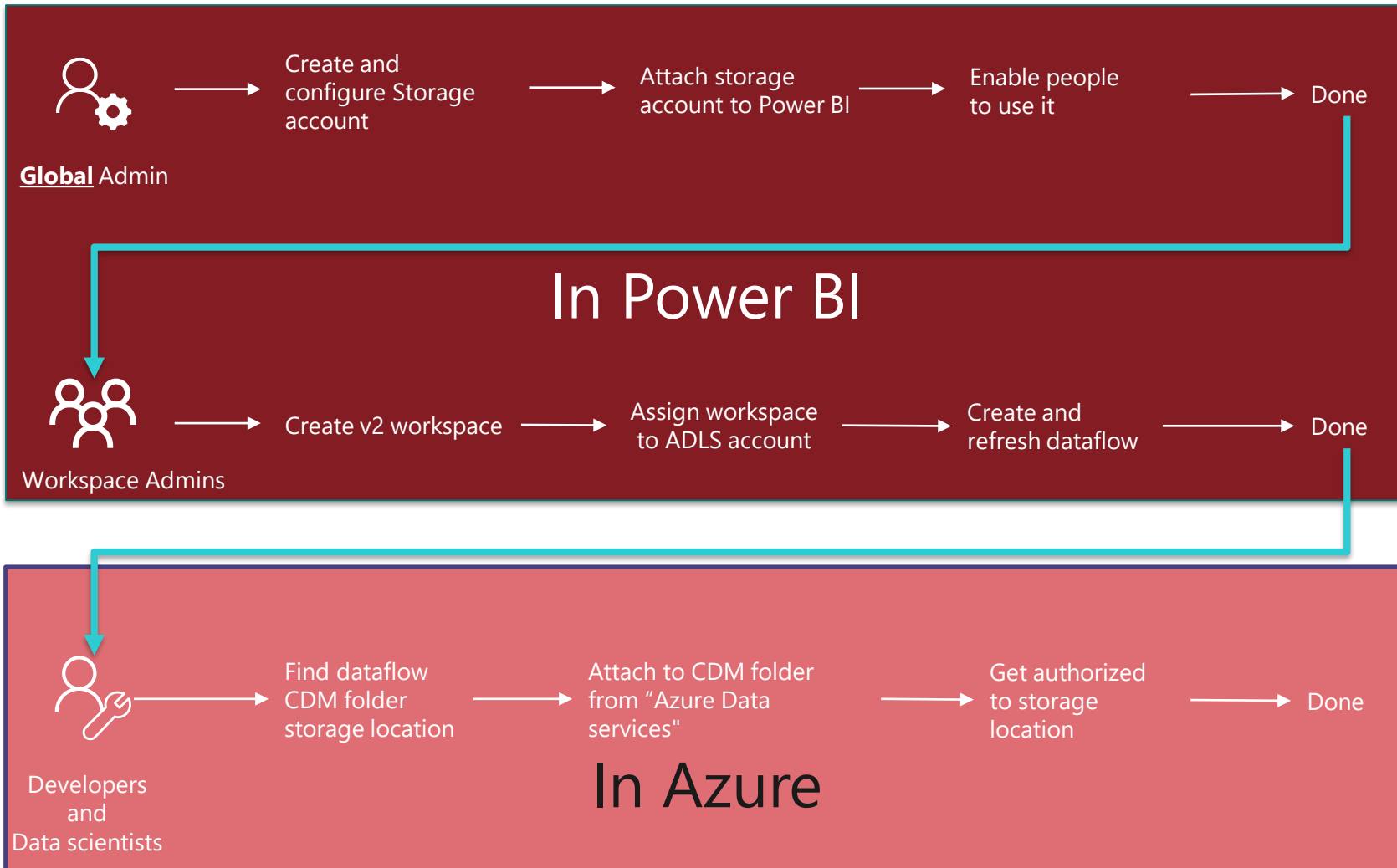
- Dataflows fill a self-service gap in the end-to-end story



A large, stylized teal graphic on the left side of the slide, resembling a thick, curved arrow or a stylized letter 'C' that points towards the right.

Extending Dataflows

Bring your own Storage Account



What "done" means:

Workspace Admins can now configure workspace to store its dataflows in ADLS account Global Admin configured

Dataflow CDM folder is stored in ADLS and is accessible by its creator ONLY.

Other people in workspace can get data if they are authorized to the CDM folder outside of Power BI

Now data can be read from CDM Folder via Azure Data Services or LOB solutions that are CDM folder aware



Bring your own Storage Account

Use your organizational ADLSg2 resource and attach it to Power BI

A Global Admin and an Azure subscription is required

Once a dataflow storage location is configured, it cannot be changed

By default, only owners of a dataflow stored in Azure Data Lake Storage Gen2 can access its data

To authorize additional users to the dataflows stored in Azure, you must add them to the dataflow's CDM folder

Creating dataflows with linked entities is only possible when they are stored in the same storage account



External CDM Folders

CDM Folders is an open standard

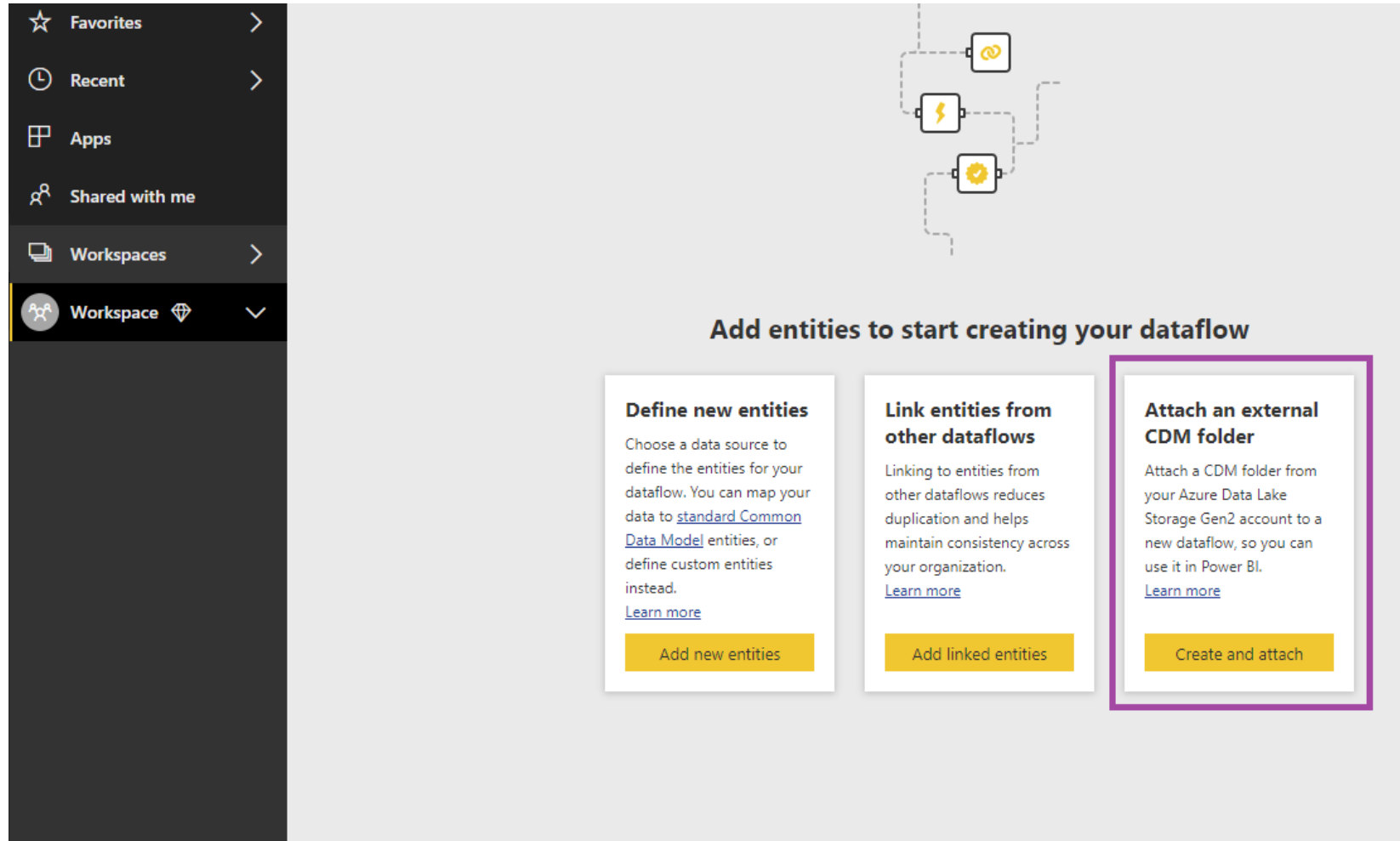
Any tool that can access ADLSg2, and output JSON

Link External CDM Folder in Power BI

Done!



External CDM Folders



The screenshot displays the Azure Data Factory 'Add entities to start creating your dataflow' screen. On the left is a dark sidebar with navigation options: Favorites, Recent, Apps, Shared with me, Workspaces, and Workspace (selected). The main area features a diagram of a dataflow with three tasks (Source, Transformation, Sink) connected by dashed lines. Below the diagram, the heading 'Add entities to start creating your dataflow' is followed by three cards:

- Define new entities**
Choose a data source to define the entities for your dataflow. You can map your data to [standard Common Data Model](#) entities, or define custom entities instead.
[Learn more](#)
Add new entities
- Link entities from other dataflows**
Linking to entities from other dataflows reduces duplication and helps maintain consistency across your organization.
[Learn more](#)
Add linked entities
- Attach an external CDM folder** (highlighted with a purple border)
Attach a CDM folder from your Azure Data Lake Storage Gen2 account to a new dataflow, so you can use it in Power BI.
[Learn more](#)
Create and attach



External CDM Folders

Attach an external CDM folder to a new dataflow

Name *

Description

CDM folder path *

Create and attach Cancel

Define

Choose to define the dataflow data to source from [Data Model](#) or from external data source. You can define a dataflow instead.

[Learn more](#)

Add new entities Add linked entities Create and attach





BONUS!

Power BI dataflows sort-of-FAQ

- ✓ • A capability for self-service data preparation in Power BI
- ✓ • Delivered in a familiar Power Query experience
- ✓ • Built on the foundation of Azure Data Lake Storage gen2
- ✓ • Utilize the CDM folder format for data storage
- ✓ • A tool for business users to drive data reuse without requiring IT involvement
- ✓ • Enable Excel-like data lineage and orchestration

- ✗ • NOT a replacement for datasets
- ✗ • NOT a replacement for a data warehouse
- ✗ • NOT a replacement for Azure Data Factory or SSIS
- ✗ • NOT a Premium-only feature
- ✗ • NOT an additional cost or fee
- ✗ • NOT spelled with a space or any capital letters





Recap

Why you should use it!

Business Driven Initiatives

Fill existing gaps

Quick prototyping

Strive for IT handover to operationalize later on

Configure your own ADLSg2 Storage



The Self-Service BI Hoax

“Products and technology don’t enable Self-Service BI, people and methodologies do”

[Nimrod Avissar on LinkedIn Pulse](#)



Resources

<https://docs.microsoft.com/en-us/power-bi/service-dataflows-best-practices>

<https://powerbi.microsoft.com/en-us/blog/power-bi-dataflows-and-azure-data-lake-storage-gen2-integration-preview/>

<https://docs.microsoft.com/en-us/power-bi/service-dataflows-overview>

<https://docs.microsoft.com/en-us/azure/architecture/reference-architectures/data/enterprise-bi-adf>

<https://docs.microsoft.com/en-us/common-data-model/>



Resources

<https://ssbipolar.com/2018/10/23/dataflows-in-power-bi/>

<https://www.sqlchick.com/entries/2019/3/28/three-ways-to-use-power-bi-dataflows>

<https://prologika.com/dataflows-the-good-the-bad-the-ugly/>

<https://excleratorbi.com.au/power-bi-dataflows-and-why-you-should-care/>

<https://www.jamesserra.com/archive/2018/08/power-bi-dataflows/>

<https://ssbipolar.com/2019/10/14/power-bi-dataflows-vs-datasets/>

<https://radacad.com/what-are-the-use-cases-of-dataflow-for-you-in-power-bi>



Resources

[Dataflows Whitepaper – Amir Netz](#)

[Release Notes & Roadmap](#)

[CDM Tutorial](#)

[CDM Documentation](#)

[Dataflows and CDM Sessions at MBAS](#)



Evaluations

Please rate this session!

Hardware provided by:

DYMATRIX
we know your customers.

Power BI Dataflows? Why you need to implement it!
Room: New York, Timeslot: 09:15 - 10:15

Excellent	Good	Fair	Weak
-----------	------	------	------

Current time: 10:39 Next presentation starts: 10:45



Sponsors



business.
people.
technology.



Many thanks to our sponsors, without whom
such an event would not be possible.






You Rock!

Gold

Silver

Bronze

Benni De Jagere?

- Senior Data Insights Consultant
- Realdolmen, a Gfi company 
- **dataMinds**.be
-  @BenniDeJagere
-  /bennidejagere
- #TeamOxfordComma



PASS Deutschland e.V.

For further information about future events, visit our
PASS Deutschland e.V. booth in the exhibitor area.

