Benni De Jagere

POWER BI DATAFLOWS?
WHY YOU NEED TO IMPLEMENT IT!

WHO AM I

Power BI CAT

dataMinds.be Member

@BenniDeJagere

/bennidejagere
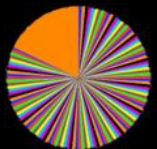
/bennidejagere

/bennidejagere

#SayNoToPieCharts

DATAFLOWS?

(Google) Cloud Dataflow

TensorFlow Dataflow

SSIS Data Flows

Microsoft Flow

ADFv2

    Mapping Data Flows
    Wrangling Data Flows

PowerApps dataflows

Power Platform dataflows

DATAFLOWS?

(Google) Cloud Dataflow

TensorFlow Dataflow

SSIS Data Flows

Microsoft Flow

ADFv2

Mapping Data Flows

Wrangling Data Flows

PowerApps dataflows

Power Platform dataflows

Power BI dataflows!

WHO?

WHO SHOULD USE IT?

POWER USERS

BUSINESS ANALYSTS

TECH SAVVY USERS

EVERYONE ELSE ☺

WHAT?

Date Dimension used in Sales.pbix file

Date Dimension used in Inventory.pbix file

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

ONE ENTITY, MANY DATASETS

Date Dimension processed in a **Dataflow**

Sales.pbix file

Inventory.pbix file

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

ONE ENTITY, MANY DATASETS

ONE ENTITY, MANY DATASETS

Lack of semantic model will create many versions

Re-use of queries across datasets is limited

Shared Datasets is an all or nothing approach

Will "Direct Query over Power BI datasets and Azure Analysis Services" change the way we work?

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 (my 2 cents)

Manageable and securable like most Power BI artefacts

Managed CDM Folders

LET'S TAKE IT FOR
A SPIN!
DEMO

DEMO: USE CASE

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

HOW:
COMMON DATA MODEL

ENTER, POWER BI DATAFLOWS!

Unified metadata layer for Cloud Apps and Services
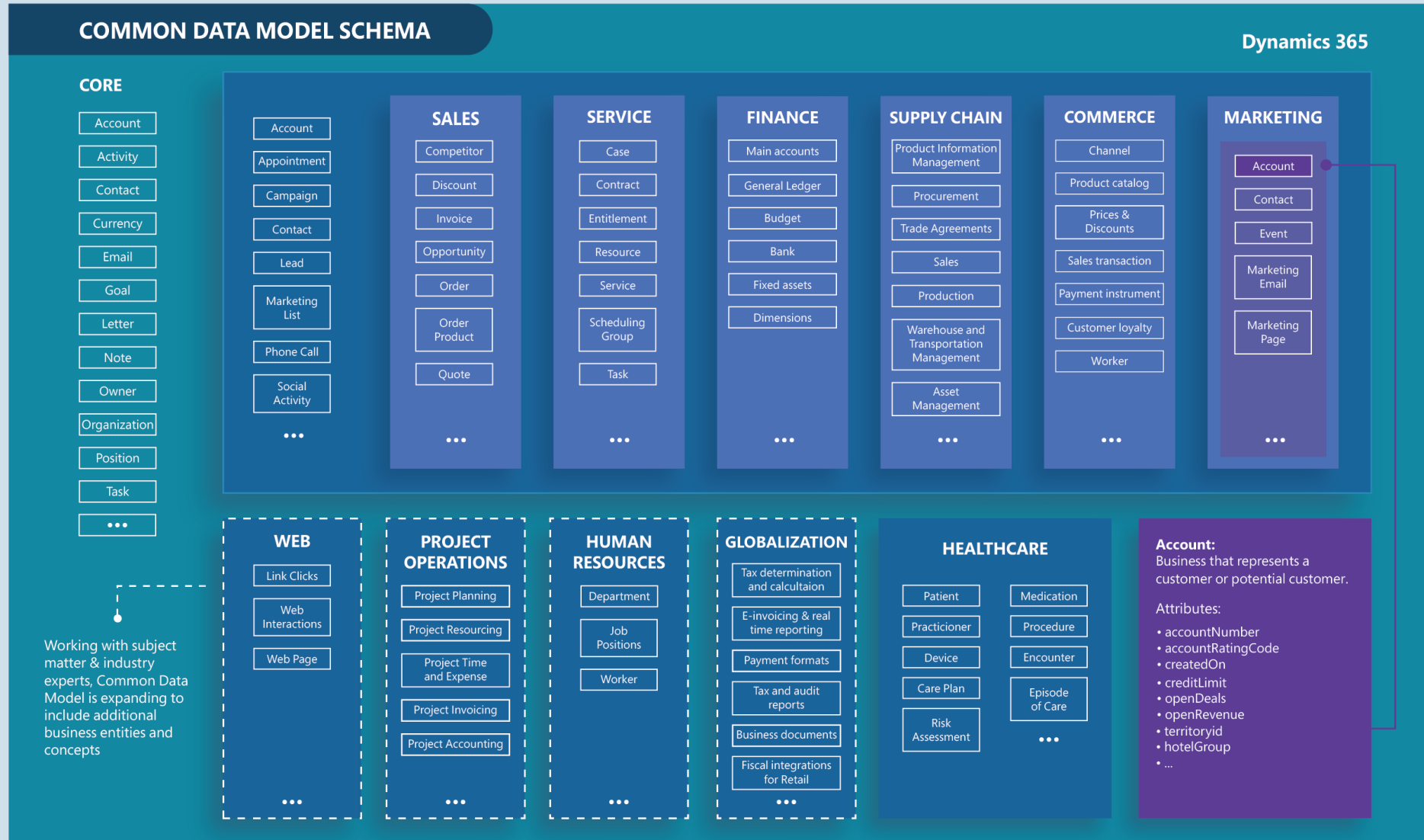
Shared

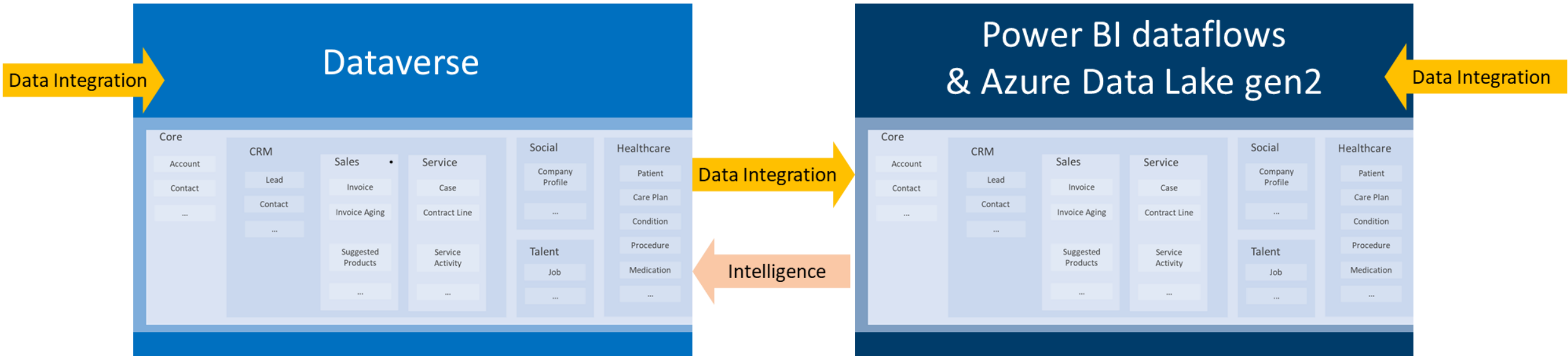   Data Language

    Storage Format and Location

    Set of Default Entities

Extend Entities for all consumers
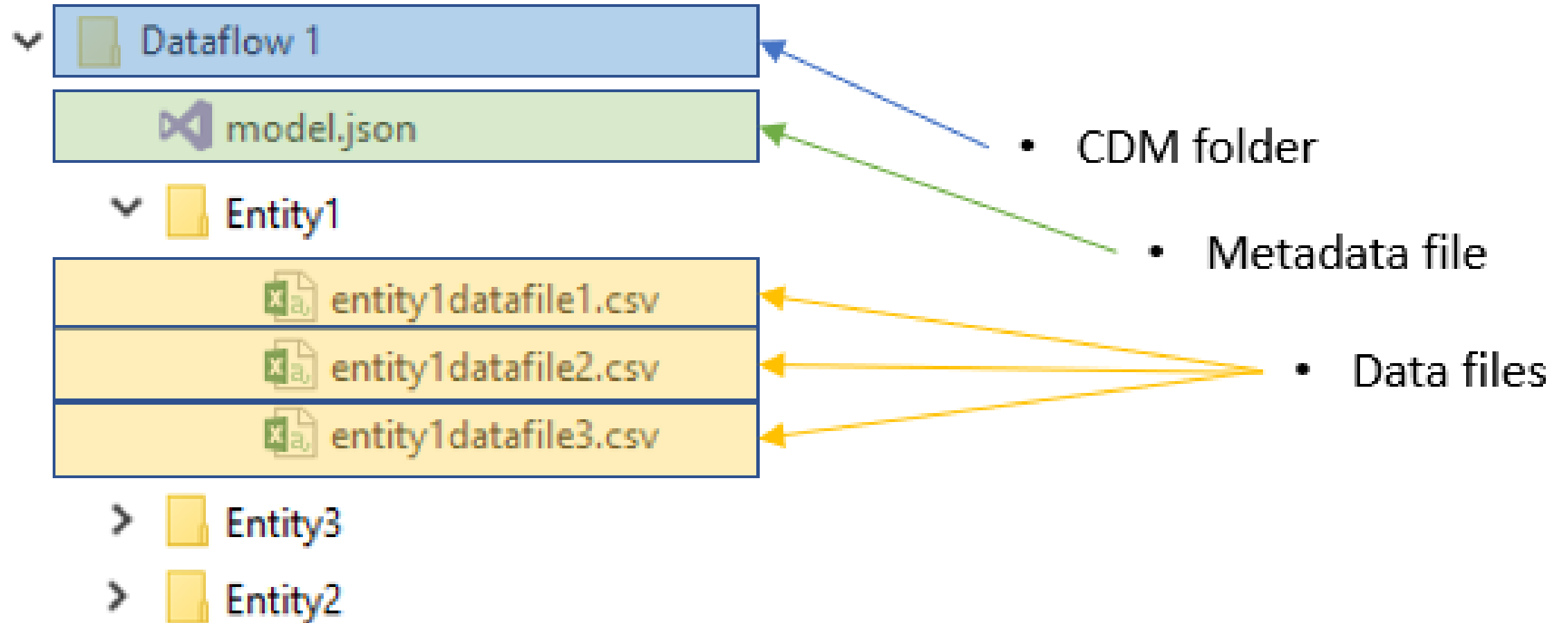
Simplified integration and disambiguation of data

INTEGRATE & DISAMBIGUATE DATA WITH
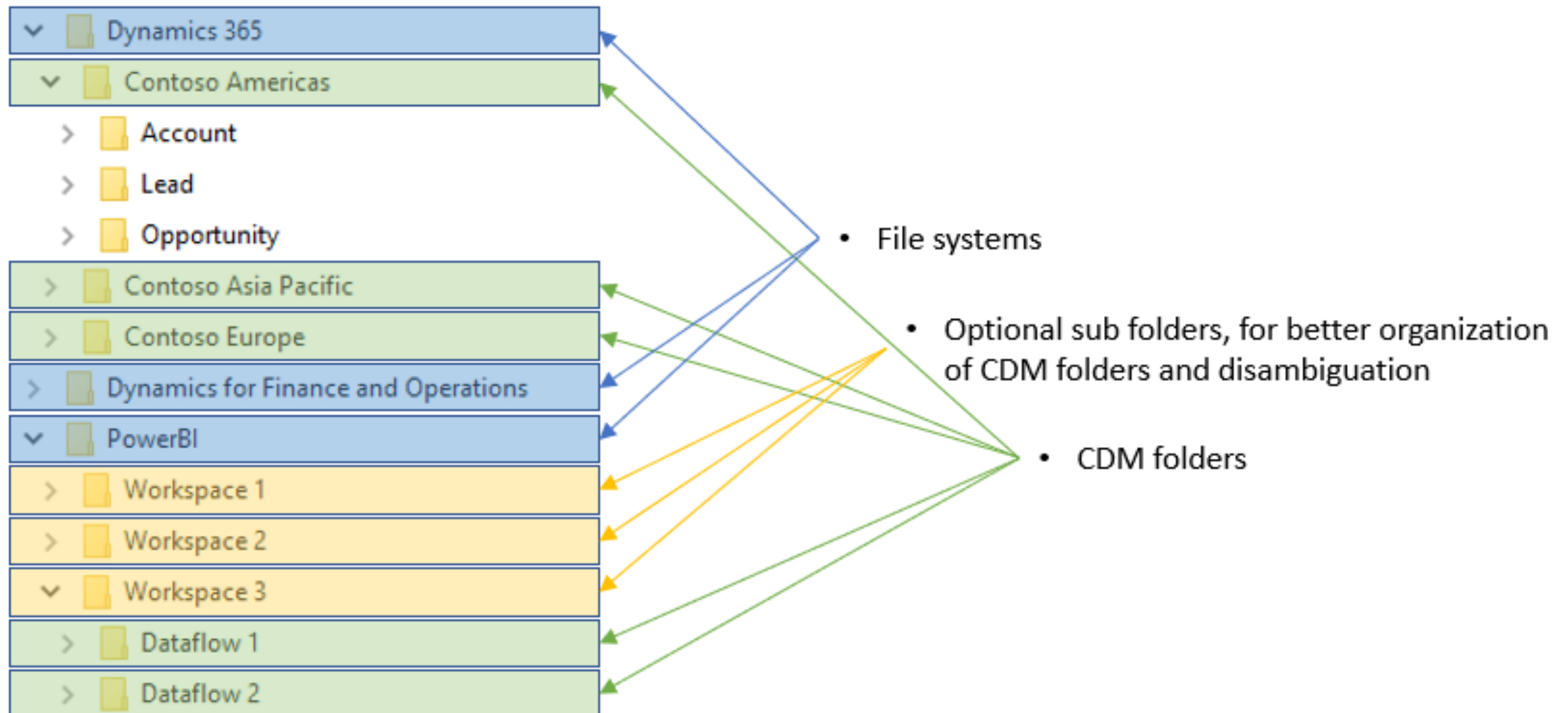
THE COMMON DATA MODEL

COMMON DATA MODEL SCHEMA

Dynamics 365

CORE

- Account
- Activity
- Contact
- Currency
- Email
- Goal
- Letter
- Note
- Owner
- Organization
- Position
- Task
- ...

Account	
Appointment	
Campaign	
Contact	
Lead	
Marketing List	
Phone Call	
Social Activity	
...	

SALES
- Competitor
- Discount
- Invoice
- Opportunity
- Order
- Order Product
- Quote
- ...

SERVICE
- Case
- Contract
- Entitlement
- Resource
- Service
- Scheduling Group
- Task
- ...

FINANCE
- Main accounts
- General Ledger
- Budget
- Bank
- Fixed assets
- Dimensions
- ...

SUPPLY CHAIN
- Product Information Management
- Procurement
- Trade Agreements
- Sales
- Production
- Warehouse and Transportation Management
- Asset Management
- ...

COMMERCE
- Channel
- Product catalog
- Prices & Discounts
- Sales transaction
- Payment instrument
- Customer loyalty
- Worker
- ...

MARKETING
- Account
- Contact
- Event
- Marketing Email
- Marketing Page
- ...

WEB
- Link Clicks
- Web Interactions
- Web Page
- ...

PROJECT OPERATIONS
- Project Planning
- Project Resourcing
- Project Time and Expense
- Project Invoicing
- Project Accounting
- ...

HUMAN RESOURCES
- Department
- Job Positions
- Worker
- ...

GLOBALIZATION
- Tax determination and calcultaion
- E-invoicing & real time reporting
- Payment formats
- Tax and audit reports
- Business documents
- Fiscal integrations for Retail
- ...

HEALTHCARE
- Patient
- Practicioner
- Device
- Care Plan
- Risk Assessment
- Medication
- Procedure
- Encounter
- Episode of Care
- ...

Working with subject matter & industry experts, Common Data Model is expanding to include additional business entities and concepts

Account:
Business that represents a customer or potential customer.

Attributes:
- accountNumber
- accountRatingCode
- createdOn
- creditLimit
- openDeals
- openRevenue
- territoryid
- hotelGroup
- ...

COMMON DATA MODEL

COMMON DATA MODEL

HOW:
COMMON DATA MODEL
DEMO

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 related dataflows

Incremental Refresh is Premium Only

HOW:
DATA REFRESH DEMO

HOW:
QUERY REUSAGE

QUERY REUSABILITY (PRO)

No developer experience in Power BI Desktop

Copy/Paste queries between environments

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

QUERY REUSABILITY (PREMIUM / PPU)

Linked/Computed Entities (Premium/PPU Only)

In-lake Compute, which requires dedicated capacity

Can Reference any Entity in any Workspace

Data is only stored in the original Entity

HOW:
QUERY REUSAGE DEMO

HOW:
GATEWAYS

GATEWAYS

On-premises data gateway
One gateway for multiple cloud services and experiences

Cloud services

| Azure Analysis Services Preview | Power BI | PowerApps | Microsoft Flow | Azure Logic Apps |
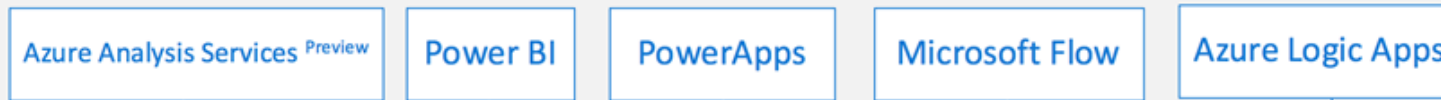
Gateway Cloud Service
Data source connection credentials are encrypted

Azure Service Bus

Application Gateway
Data source connection credentials can only be decrypted by the gateway

On-premises data sources

| SQL | SQL Server | SQL Server Analysis Services | Other data sources | Files, SharePoint |

HOW:
SECURITY

SECURING DATAFLOWS

Dataflows reside in Workspaces

"Secure"/Hide Linked Entities by disabling load

Datasets using Dataflows will likely be Import Mode

Security is the responsibility of the dataset designer!

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

WHERE DO DATAFLOWS FIT IN?

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

EXTENDING
DATAFLOWS

EXTENDING DATAFLOWS

Dataflows

Datasources   Azure Data Lake Storage Gen2   Datasets   Reports   Dashboards

Link

Link   Compute

Data is stored as entities in Microsoft
Common Data Model compliant folders.

BRING YOUR OWN STORAGE ACCOUNT

https://powerbi.microsoft.com/en-us/blog/power-bi-dataflows-multiple-data-lakes-support/

BRING YOUR OWN STORAGE ACCOUNT

EXTERNAL CDM FOLDERS

CDM Folders is an open standard

Any tool that can:

Access ADLSg2

Output JSON

Create the correct manifest format

Link External CDM Folder in Power BI

Done!

EXTERNAL CDM FOLDERS
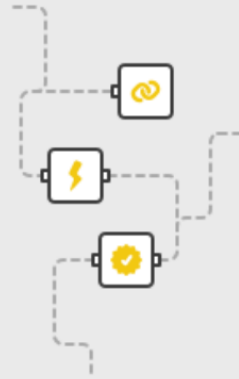
Start creating your dataflow

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

Import Model

Choose a dataflow model to import into your workspace.

Learn more

Import model

Attach a Common Data Model folder (preview)

Attach a Common Data Model 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

WHAT'S THE CATCH?

EXTERNAL CDM FOLDERS

The cool stuff is Premium / PPU only

Validation of queries can take a loooooonng time

Modifying a schema will drop all data

Dataflows can only be modified by their owners

Dataflows using gateway data sources do not support multiple credentials

When working with APIs (for CI/CD) the story isn't smooth yet

Exporting/Importing creates a new ID

Importing will not fix existing references to linked entities

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

@SQLAllfather – Matthew Roche

TAKEAWAYS

TAKEWAYS

Business Driven Initiatives

Fill existing gaps

Quick prototyping

Strive for IT handover to operationalize later on

Configure your own ADLSg2 Storage

POWER BI DATAFLOWS – APRIL 2021 UPDATE

Coming Soon

Careful reading might have noticed a few teasers. Indeed, we have some exciting features that will be available in the coming months ahead. Currently, we are laser focused on **performance and reliability improvements**, and **delivering on YOUR top asks, including**:

- **Background Validation**
  - With this update, users will be able to Save a dataflow, leave the screen, and continue working. Validation now takes place in the background so users can save the dataflow without friction.

- **Multiple Contributors to Dataflows**
  - We're working to allow non-owners to edit Dataflows, see the applied steps, and manage them – making Dataflows more collaborative so that teams can be productive.

- **Notifications to Non-Owners**
  - Sending refresh failure emails to users or groups, just like Datasets

- **Deployment Pipelines**
  - We'll bring support for Deployment Pipelines to Dataflows, allowing you to deploy and update Dataflows across the lifecycle stages and make them more enterprise grade.

- **Updated Guidance**
  - Advice for common scenarios and prescriptive how-to's

RESOURCES

RESOURCES

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

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://docs.microsoft.com/en-us/power-bi/transform-model/dataflows/dataflows-azure-data-lake-storage-integration

https://docs.microsoft.com/en-us/power-bi/transform-model/dataflows/dataflows-premium-features

https://docs.microsoft.com/en-us/power-bi/transform-model/dataflows/dataflows-machine-learning-integration

https://docs.microsoft.com/en-us/power-bi/transform-model/dataflows/dataflows-features-limitations

RESOURCES

https://docs.microsoft.com/en-us/power-bi/transform-model/dataflows/dataflows-directquery

https://docs.microsoft.com/en-us/power-bi/transform-model/dataflows/dataflows-understand-optimize-refresh

https://docs.microsoft.com/en-us/power-bi/transform-model/dataflows/dataflows-streaming

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://exceleratorbi.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-bite-dataflows-vs-datasets/

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

https://ssbipolar.com/2020/09/23/session-resources-patterns-for-adopting-dataflows-in-power-bi/

RESOURCES

[Dataflows Whitepaper – Amir Netz](#)

WHO AM I
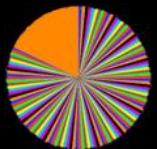
Power BI CAT

.be Member

@BenniDeJagere

/bennidejagere

/bennidejagere

/bennidejagere

#SayNoToPieCharts

THANK YOU!

HTTPS://SQLB.IT/?7007