Collibra – Power BI Integration

Overview

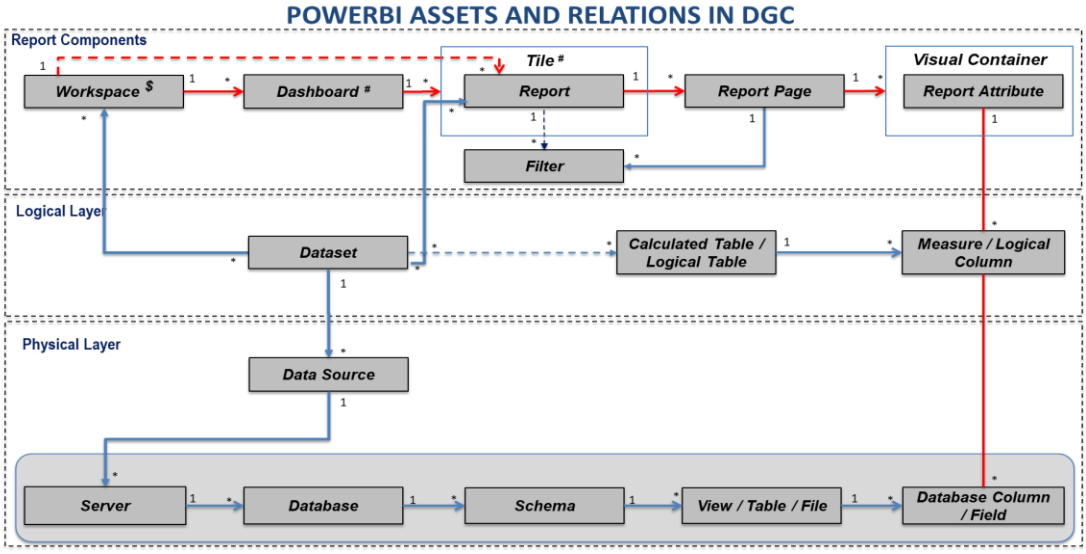
Lucid's Collibra Power BI Integration template loads the Power BI metadata into Collibra DGC. The Power BI REST APIs as well as .pbix archive are used to extract the Power BI assets like Workspace, Dashboard, Report, Report Pages, Report attributes etc. and the relations supporting them. It supports both Power BI service (Cloud) and Power BI Report Server (On-prem).
Some key features:

- Detailed metadata from Dashboards and Reports are extracted by parsing the Power BI (.pbix) archive
- Asset extraction can be filtered based on Workspace, Dashboard and Report in Power BI Service
- Asset extraction can be filtered based on Folder and Report in Power BI Report Server
- Target community in DGC for loading the assets can be specified
- Source metadata is mapped to DGC asset types as indicated in the diagram below
- Modification and Delete of the source (Power BI) objects are automatically captured and can be used to trigger actions in DGC such as asset status change, asset management workflows, etc.
- Support for full load and incremental load. Incremental loads can load changes since the last successful execution or from a specific date passed as a parameter
- Metadata loads can be scheduled or triggered on-demand
- Connectivity Check template available for environment verification
- Support for Mule clusters

Power BI versions supported:

- Power BI Service (April 2020 release) * Power BI Desktop Dec 2020 release is supported with the relevant feature exclusions listed in 'Unsupported Features' section
- Power BI Report Server (May 2020 release)

Assets and relations that will be loaded into DGC:



Note: The Physical Layer assets (indicated in the shaded area - namely Schema, View /Table/Column, File/Field) will not be loaded by the PowerBI template. It must be loaded separately using Collibra Catalog or via Lucid''s Database templates. Lines in red will be shown as part of the Traceability view in DGC.
- Loaded only for PowerBI Service and not for the PowerBI Report Server. **$** - For Report Server, Folders are loaded as Workspaces.

Key Use case(s)

- Traceability/Lineage – provide ability to trace the provenance of a report attribute to its source
- Ability to do impact analysis of changes to the Power BI objects

Version

2.6

Release Notes

2.6

- Support for handling renames of Workspaces, Dashboards, Tiles, and Folders assets without their deprecation in Collibra
- Separate end point to delete Workspaces, Apps, and Folders
- Attributes “Is Hidden”, “IS SQL Attached?” captured for Power BI reports
- Support for alias names for report attributes on Hierarchies
- SQLs as part of Direct Query will be extracted as is from the metadata source if available and attached to the Dataset (table level asset) that is created. The interim steps will not be captured as transformation logic

2.5

- Support for Collibra DGC Connector version 1.6.2 that provides improved load performance
- REST APIs secured with Basic authentication
- All configurable properties for the integration managed in DGC using custom configuration assets
- A flag - "Is Partially Loaded", added to the assets to denote if the ingested assets’ mandatory characteristics are available or not
- Email notification for Execution Summary including status and asset counts

1.3

- Support for incremental load of Power BI Service
- Support for metadata capture via custom attributes/annotations added in Power BI

1.2

- Support for Power BI Apps
- Support for metadata capture via custom attributes/annotations added in Power BI

1.1

- Support for Kerberos authentication
- Asset extraction filters on Folder and Reports for Report server
- Support for Mule clusters
- Location URL for Power BI Report, Workspace, Dashboard & Folder
- Additional Metadata Captured: Visual Container Titles

1.0

- Support for extracting and loading the following assets in Power BI:
    - Workspace (supported only for Power BI service)
    - Dashboards (supported only for Power BI service) - selection of report pages for visualization
    - Reports - collection of report pages
    - Report Pages - collection of visualization, each fitting into a Tile
    - Tiles
    - Report attributes – fields used in a Report Page that maps to one-or-more data elements from one or more data sets
    - Report-level Filters
    - Datasets - collection of data for analysis coming from one or more data sources
    - Data Sources (published to server) – metadata as captured from Power BI
- Link to underlying physical data assets as captured from Power BI metadata - actual metadata of the underlying database/ physical data source should be loaded using Collibra Catalog or Lucid's database integrations or other integration solutions
- Data Sources supported: SSAS Tabular View, relational databases, Flat files / Excel, other ODBC sources
- Calculations used in a report attribute – simple calculations will be parsed but not complex Data Analysis eXpressions. Data Analysis eXpressions will also be captured as-is as a definition. Dependencies from Data Analysis eXpressions to underlying columns and measures will be extracted to the extent available in the underlying metadata in the .pbix archive
- M Code (Power Query Language Transformations) will be extracted as is from the metadata source if available and attached to the Dataset (table level asset) that is created. The interim steps will not be captured as transformation logic.
- Alignment to Collibra Catalog asset nomenclature (including Full Name and Display Name)
- Asset extraction can be filtered based on Workspace, Dashboard and Report
- Target community in DGC for the assets can be specified during load
- Modification and Deletes of the source (Power BI) objects are captured in DGC
- Support for full loads as well as incremental loads. Incremental loads can load changes since the last successful execution or from a specific date passed as a parameter
- Metadata loads can be scheduled or triggered on-demand

Unsupported Features:

- Support for Power BI reports created using the 'Enhanced Dataset Metadata' option in the Power BI Desktop
- Paginated reports (.RDL), mobile reports, and KPIs in Power BI Report Server
- Content Packs
- Every intermediate step of the Data Transformations/Manipulations built using PowerQuery / Query Editor in Power BI will not be captured / parsed for transformation logic
- Calculated Tables
- Parsing of Data Analysis eXpressions Functions released after Nov 2018
- SQL Parsing can be supported using 3rd party tools as a custom extension

- Support for Power Query M functions
- Data Sources
    - Odata Feed
    - Online Services – Salesforce (SFDC), ServiceNow, etc.
    - Azure Cloud platform
    - SSAS OLAP (Multi-dimensional Model)
    - DirectQuery with HANA

Editions and feature comparison

Integration is available on an annual subscription basis. A free trial[#] is also available.

Feature	Trial	Basic
Licensing	Free	Annual Subscription
Validity Period	15 days	1-year
Support for Functional Features	All listed in Release Notes	All listed in Release Notes
Number of deployments allowed	1 Non-Prod	1 Prod + 1 Non- Prod
Limit on number of assets loaded	Yes	No
Documentation (1) Installation & Configuration Guide (2) Support Guide	Yes	Yes
Training included: 2 hours of virtual sessions for install/config (or) troubleshooting	No	Yes
Helpdesk/Support	No	Email-based support (Request logged through Support Portal - https://lucidtechsol.freshdesk.com/support/login)
Support response SLA	NA	1 business day
Custom Extensions	NA	Not supported*
Pricing	NA	Contact Lucid

** No custom extensions requests are supported as part of the subscriptions, however, can be separately considered for a one-time fee, with on-going support part of the subscription.*

[#] Trial subscriptions are available on Docker containers for sites with no Mule Server

Prerequisites / Dependencies

Prerequisites for this integration with the product versions currently supported:

- Collibra Data Governance Centre v5.6.x, v5.7.x
- Collibra Connect with DGC Connector v1.4.5 - Mule ESB v3.9.1 or v3.9.3 (recommended) (Only the Mule standalone, packaged as part of Collibra Connect, is supported)
- Access to Postgres Database (with server version 9.6) Need an empty schema with privileges to drop and create tables.
- DGC migration feature - for DGC v5.6.x, v5.7.x

Note: *The integration has been tested and certified by Lucid to work against specific versions of Collibra and the Metadata source system. As newer versions of Collibra and the metadata source system are released, Lucid will test and certify the integration as part of support. Please note if Collibra or the metadata source system fundamentally change their approach on how metadata is exposed, extracted or ingested, enhancing the integration to support these newer approaches might not always be possible, in which case the current subscription may not automatically support the new version. However, Lucid will make all efforts to provide an upgrade option to support the new versions at the earliest.*

To Subscribe

- Visit http://www.lucidtechsol.com/downloads/

Demo video

- Visit https://youtu.be/ZS1x-a6oqAk

To know more / request a demo

- Email us at : collibrainfo@lucidtechsol.com