neos

Neos ODI to Collibra Integration Module

Standalone CLI program used to ingest and harvest ODI ETL/ELT modules into Collibra DGC

WWW.NEOS.HR • COLLIBRA@NEOS.HR

NEOS ODI TO COLLIBRA INTEGRATION MODULE: AT A GLANCE

NAME	Neos ODI to Collibra Integration Module (aka NeoHarvester)
CATEGORY	Integration
USE CASE	BI & Analytics
TARGET AUDIECE	• Architect • Collibra Admin • System Engineer • Integration Engineeer
WHO CAN SET IT UP	• Collibra Admin • System Engineer
TARGET BUSINESS FUNCTIONS AND INDUSTRIES	• Any business function • Any industry
LICENSE REQUIREMENTS	Collibra Catalog
DEPENDENCIES	Oracle Data Integrator Studio, OpenJDK 17

neos
ODI TO COLLIBRA INTEGRATION MODULE

SEAMLESS DATA INTEGRATION WITH ODI TO COLLIBRA INTEGRATION MODULE

The primary objective of this program is to ingest and harvest Oracle Data Integrator (ODI) ETL/ELT modules effectively. It is designed to create assets, attributes, and various types of relationships in such a manner that data lineage can be visualized through a Traceability Diagram.

The harvester constructs a hierarchical structure of the ODI Project Folder, incorporating Collibra assets and relationships, and then generates ODI Mapping Assets complete with their main attributes. These are placed in the correct location utilizing the Collibra relationships feature. For each ODI Mapping Asset, complex relationships are established so that the majority, if not all, of the target columns are connected to their upstream counterparts, where they are involved in transformations.

Before the harvest run, it is essential to ingest the physical assets used in ODI modules (databases, schemas, tables, columns) via Edge Site.

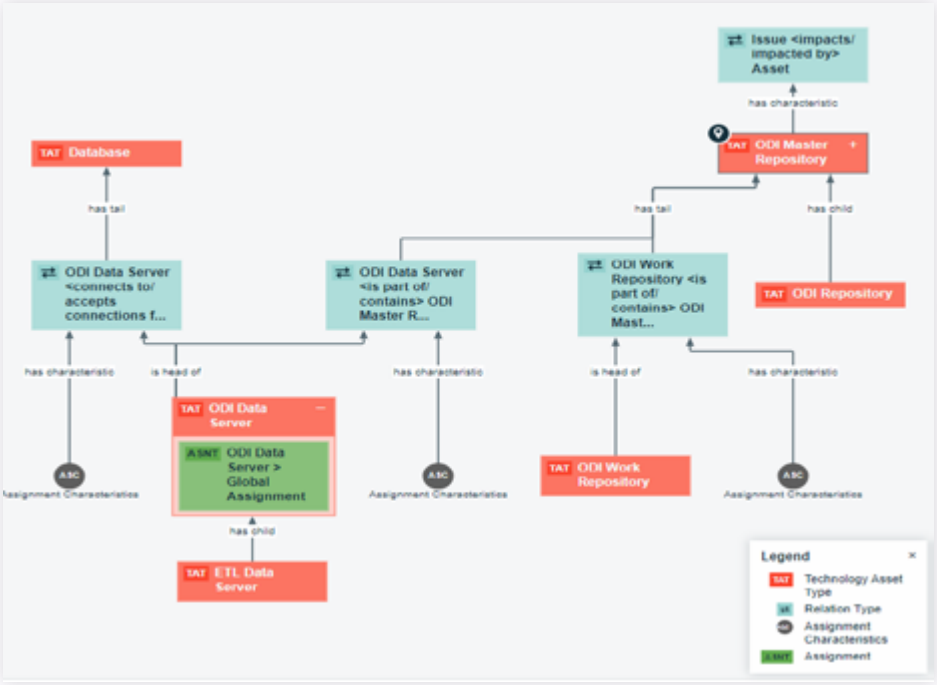
USE CASES

- For all user roles, the ODI is showcased in a comprehensive overview. It details the sources it utilizes, how data is transferred from a source to the initial stage, and subsequently to other stages, as well as identifying the ultimate targets (databases/schemas/tables/columns). It describes which sources are consumed and the impact on the target data structures.
- For technical users, the section provides information about the name of the ODI module responsible for transformations, including an overview of version details and its location within the ODI Repository.
- For ODI users, the document offers insights into impact analysis and considerations for further development and change analysis.

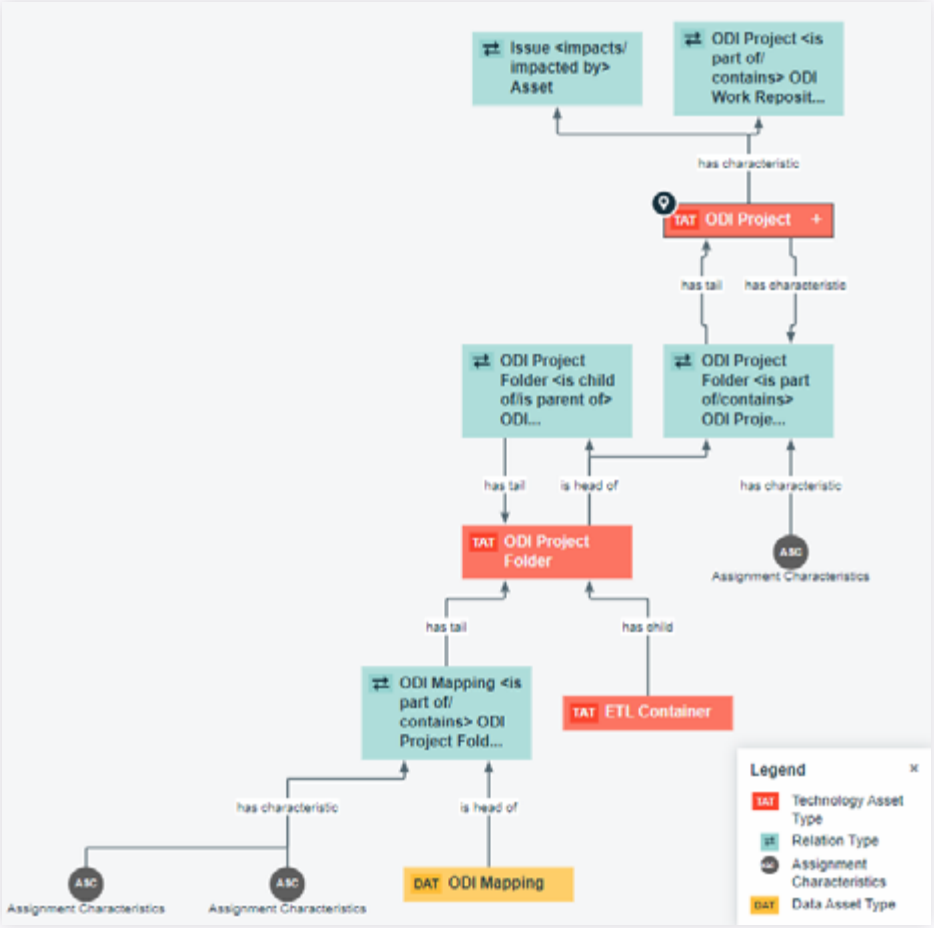
FUNCTIONAL DESIGN: INSIDE THE INTEGRATION MODULE

COLLIBRA METAMODEL

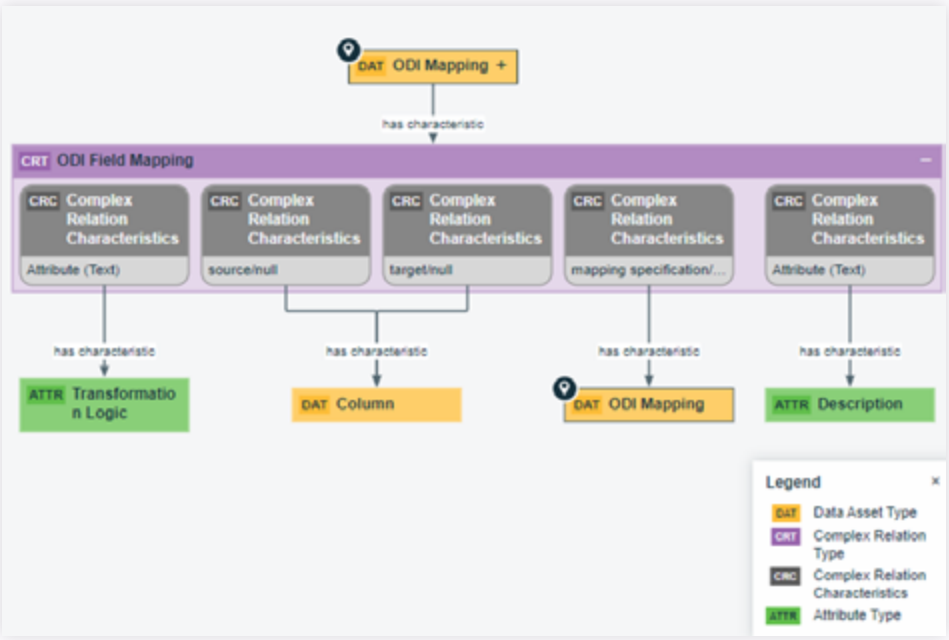
The following diagrams provide an overview of the asset types that are created on the Collibra Platform instance and how they are related.



ODI Physical Topology: metamodel



ODI Project Structure: Metamodel



ODI ETL module (ODI Mapping) complex relation: Metamodel

DOMAIN TYPES

One new domain is added to the OPMD.

Name	Type	Description
ODI Catalog	Technology Asset Domain	Domain used for Neos ODI to Collibra Integration Module assets.

ASSET TYPES

This section provides an overview of the out-of-the-box Asset types used in the application.

Name	Parent	Description
Database	Technology Asset	A collection of data that is systematically organized or structured in order to make it easy to create, update and query the information.
Schema	Data Structure	An organized structure described in a formal language supported by implementing technology that defines the objects in the technology assets (Table and columns in a relational database, fields in a file).
Table	Data Structure	An implementation of Data Entities in columns and rows, in a given database system. It is the basic structure of a relational database.
Column	Data Element	An atomic unit of data that can be stored in a database table. Examples: FST_NM, EMPID.

This section provides an overview of the new Asset types needed to be added in OPMD.

Name	Parent	Description
ETL Repository	Technology Asset	
ODI Repository	ETL Repository	Oracle Data Integrator objects container.
ODI Master Repository	ODI Repository	Data structure containing information on the topology of the company's IT resources, on security and on version management of projects and data models.
ODI Work Repository	ODI Repository	This is a data structure containing information about data models, projects, and their use.
ETL Data Server	System	The link between a ETL System and an external system.
ODI Data Server	ETL Data Server	The link between a ODI ETL System and an external system.
ETL Container	Technology Asset	
ODI Project	ETL Container	Oracle Data Integrator project is a collection of ODI objects created by users for a particular functional domain.
ODI Project Folder	ETL Container	Oracle Data Integrator project folder is a collection of ODI objects.
ETL Modules	Data Asset	
ODI Mapping	ETL Modules	Mapping in Oracle Data Integrator (ODI) is the logical and physical organization of your data sources, targets, and the transformations through which the data flows from source to target.

ATTRIBUTES

After new assets are created, we need to define any new attributes which are not present out-of-the-box.

We will keep it to the minimum, so new attributes will be created only for the ODI Mapping.

Attribute	Kind	Type	Description
ODI Created by	Text	Plain	ODI user who created this object.
ODI Created on	Text	Plain	When was the ODI object created in the ODI repository.
ODI Description	Text	Plain	ODI object description from ODI repository.
ODI Updated by	Text	Plain	ODI user who was last updated this object.
ODI Updated on	Text	Plain	When was the ODI object updated in the ODI repository.

RELATION TYPES

The following new relations are introduced in OPMD.

Head	Role	Corole	Tail
ODI Data Server	connects to	accepts con from	Database
ODI Data Server	is part	contains	ODI Master Repository
ODI Mapping	is part of	contains	ODI Project Folder
ODI Project	is part of	contains	ODI Work Repository
ODI Project Folder	is part of	contains	ODI Project
ODI Project Folder	is child of	is parent of	ODI Project Folder
ODI Work Repository	is part of	contains	ODI Master Repository
Data Structure	target of	target	ODI Mapping
Data Structure	source of	source	ODI Mapping

COMPLEX RELATION TYPES

The new complex relation type for ODI Mapping asset.

Name	Description			
ODI Field Mapping	Complex mapping between two or more data fields			
	Relations			
	Role	**Asset type**	**Min**	**Max**
	source	Column	1	
	target	Column	1	
	mapping specification	ODI Mapping	1	1
	Attributes			
	Description		0	1
	Transformation Logic		0	1

IMPLEMENTATION STRUCTURE

CLI APP consists of the following Files and folders.

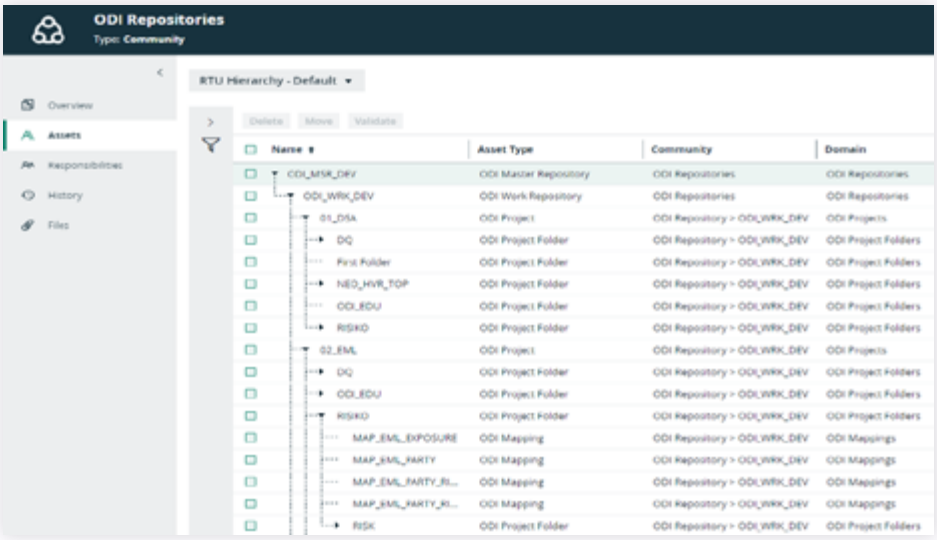
File	Description
neo-harvester<version>.jar	Bootable JAR file used for harvesting. Human readable version text can be written instead of <version> placeholder e.g. neo-harvester-v1.0.0-48491d767321c886407b1a7327ee27de82eee97e.jar.
odi-connector<version>.jar	JAR file used to connect to ODI.
config.xml	XML file with harvesting configuration.
pwd.xml	XML file with encrypted passwords for Collibra and ODI. This file is optional, if it is missing, NeoHarvester will request from user through terminal corresponding passwords for Collibra and ODI usernames.

Folder	Description
data	Folder contains H2 database files. This is applications internal repository. Do NOT edit or delete files or folder. Do not copy folder to other location or environment since internal repository represents your ODI repository and Collibra DGC instance.
out	Folder used for JSON files temporary location. If JSON file is successfully imported in Collibra via REST API it is deleted afterwards, otherwise it remains in this folder for analysing or debugging purposes until the next run of the application.
log	Contains log files in ISO extended timestamp format. See below heading Log Folder about logging behaviour.

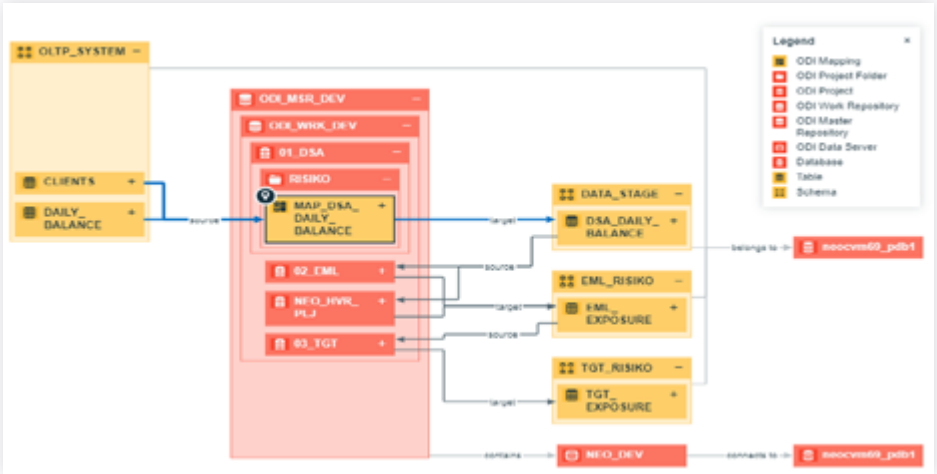
TECHNICAL DESIGN: ODI TO COLLIBRA METADATA INTEGRATION

This integration extracts metadata from an Oracle Data Integrator repository and imports it into Collibra Data Governance Center (DGC).

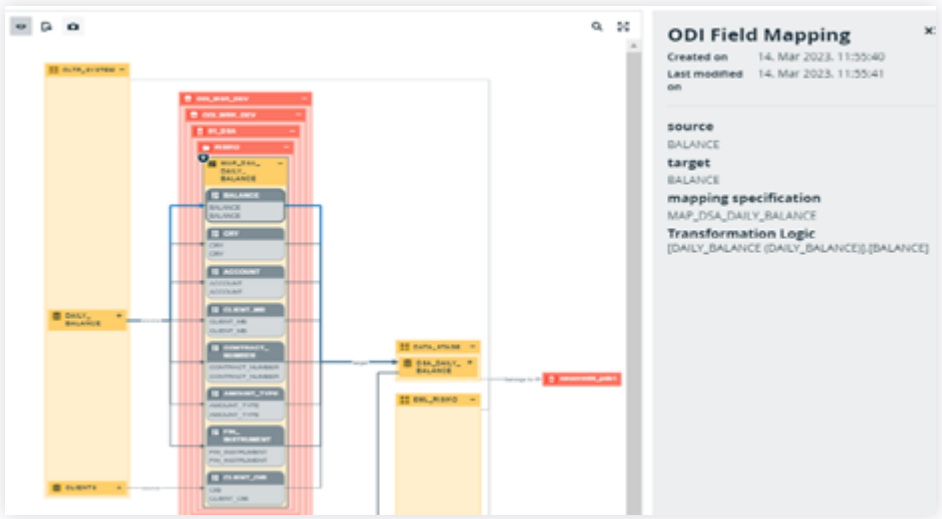
The images below display typical examples of views and diagrams as they are presented within the Collibra Platform.



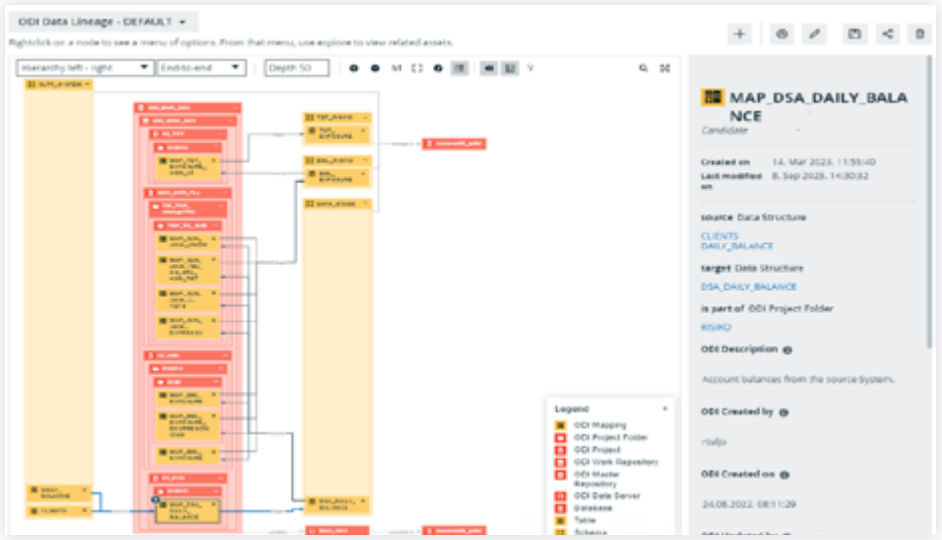
All Assets Hierarchy View



Data Lineage Diagram for MAP_DSA_DAILY_BALANCE: Collapsed



Data Lineage Diagram: Expanded Mapping



Data Lineage Diagram: Expanded All

INQUIRIES AND SUPPORT

If you require further information or have additional questions, feel free to reach out via email at collibra@neos.hr.