An Oracle White Paper
October 12th, 2018

Oracle Metadata Management v12.2.1.3.0 New Features Overview

ORACLE

Disclaimer

This document is for informational purposes. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described in this document remains at the sole discretion of Oracle.

Table of Contents

Executive Overview

The Oracle Metadata Management (OMM) solutions include two products:

- the Oracle Metadata Management for Oracle Business Intelligence (OMM4OBI)
- and the Oracle Enterprise Metadata Management (OEMM)

Oracle Metadata Management for Oracle Business Intelligence is a software package for metadata management of Oracle environments. Oracle Metadata Management for Oracle Business Intelligence includes the following metadata management features:

- Metadata Harvesting from Oracle technologies
- Metadata Configuration and Stitching
- Metadata Browsing, Search and Reporting
- Metadata Collaboration (external URL, tagging, comments and review)
- Data Flow Lineage & Impact Analysis
- Metadata Explorer (simplified metadata user interface for business users)

Oracle Enterprise Metadata Management is a software package for metadata management of multi-vendor environments along with support for data governance. Oracle Enterprise Metadata Management includes all features of Oracle Metadata Management for Oracle Business Intelligence with the following extra metadata management features:

- Metadata Harvesting from multi-vendor technologies
- Metadata Version and Configuration Management (change management)
- Data Model Diagram Visualizer and Navigator
- Business Glossary for Data Governance
- Semantic Lineage & Impact Analysis
- Semantic Mapping Editor
- Data Flow Mapping Specifications Editor
- Data Documenter

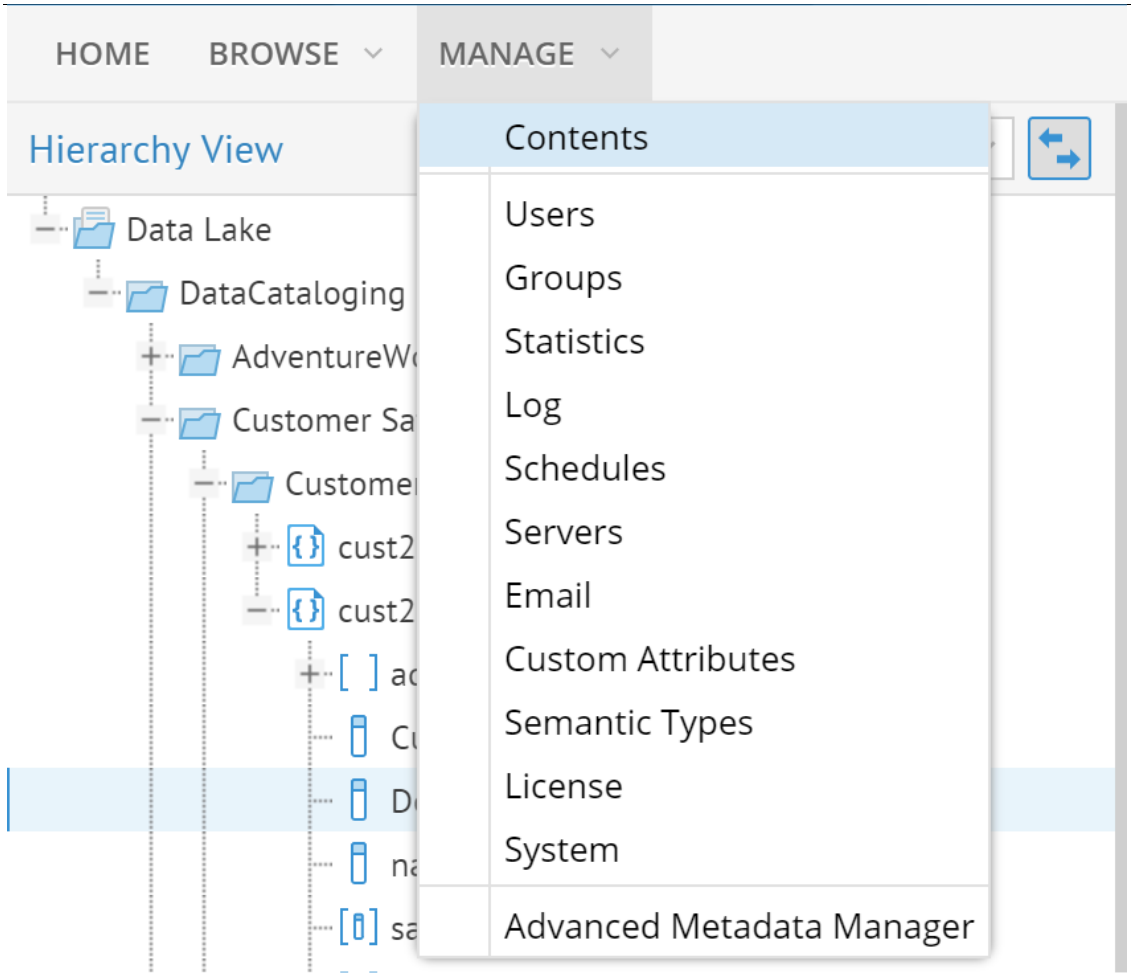
This whitepaper describes in detail some of the new features and capabilities offered in Oracle Metadata Management with the release of version 12.2.1.3.0.

Oracle Metadata Management 12.2.1.3.0

METADATA MANAGER VS METADATA EXPLORER UI

In previous OMM versions, the OMM Web Application Server has offered two different User Interfaces (UI) targeting different user communities. The original Metadata Manager UI was designed for the advanced technical users with a traditional development tool layout including multiple panels: tree structure on the left, multi-tab windows in the middle, attributes on the right, and log activities at the bottom. The Metadata Manager UI also presents the highest level of details and complexity of all harvested metadata. The Metadata Explorer UI was initially introduced as a read only UI with simpler metadata for business users offering an easy to use layout for multiple devices, including tablets. The Metadata Explorer became the new UI platform for all new editing capabilities such as the business glossary or data modeling.

With OMM v12.2.1.3, all other editing capabilities are now available in the Metadata Explorer UI, including data mapping, enterprise data architectures (Configuration editor and model stitching), and even the Administration features like Custom Attributes which are now under are now available in the Metadata Explorer UI > Manage > Custom Attributes. Consequently, the Metadata Manager UI is now only necessary (and therefore available) in the OMM Advanced Editions for repository management (with multi version and configuration management). The OMM Standard Edition v12.2.1.3 is now fully implemented in the Metadata Explorer UI where Manage > Content allows users to directly create models to the default single configuration, import metadata, stitch models (connections), and trace lineage right away.

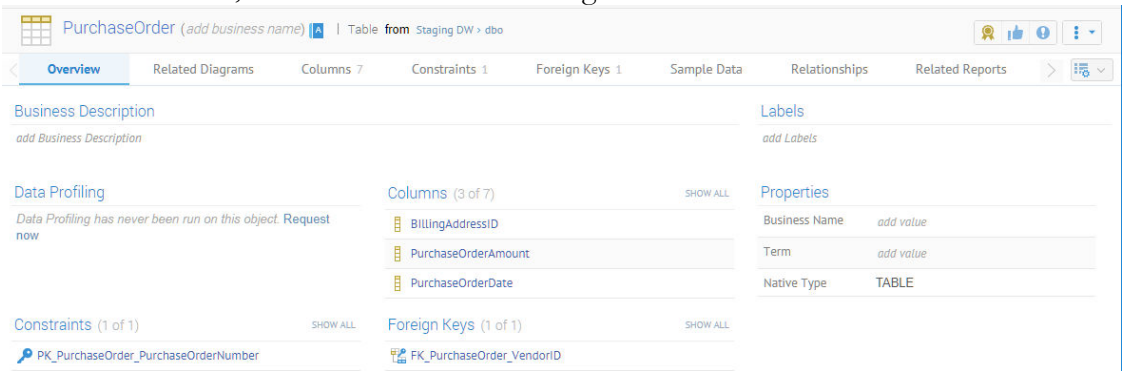


METADATA HOME PAGES

New metadata home pages with multiple top tabs offer quick access to all key information:

- The first tab is always the Overview tab which provides a dashboard to all critical information and properties.
- The next set of tabs are specific (metamodel / profile driven) to the type of object, for example:
  - Database Table objects have tabs for Columns and Constraints.
  - BI Reports (like Tableau Workbook) objects have tabs for Dashboards, Worksheets, and Data sources.
- The next set of tabs are for the common critical metadata analysis:
  - DATA FLOW for data lineage and impact analysis.
  - RELATIONSHIPS for detection, management, and curation of relationships (see new features below)
  - SEMANTIC FLOW for definition and usage perspectives (see new features below).

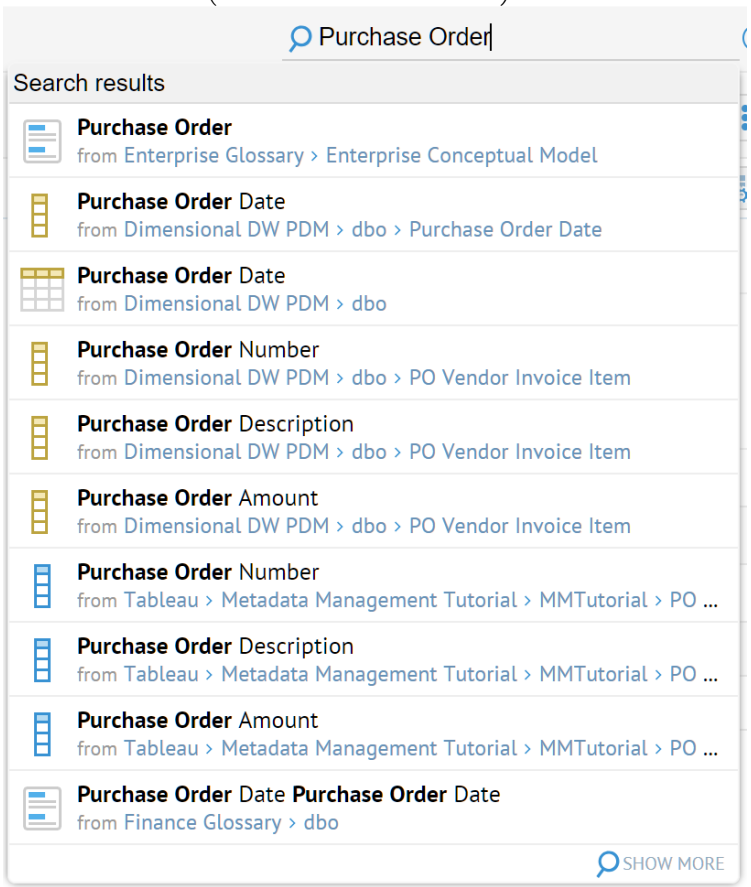
- The last set of tabs are for common documentation and administration like: Comments, Attachments and Audit Log.



METADATA QUICK ACCESS

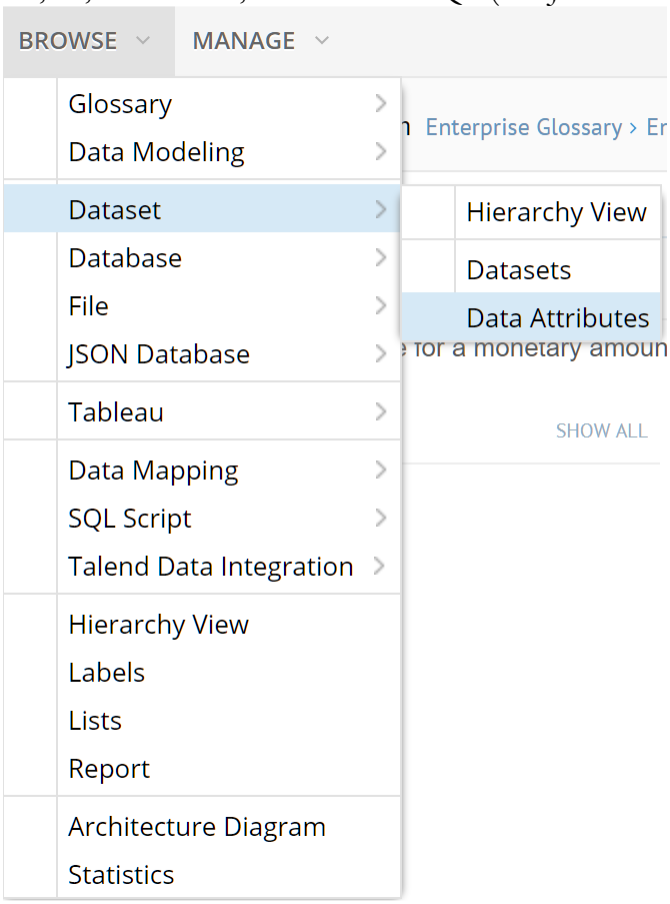
Much improved ways to quickly access the right metadata:

- SEARCH has been massively improved in both real time performance (now based on Lucene) and in functionality as a metadata driven search with natural language semantic search (see new features below)

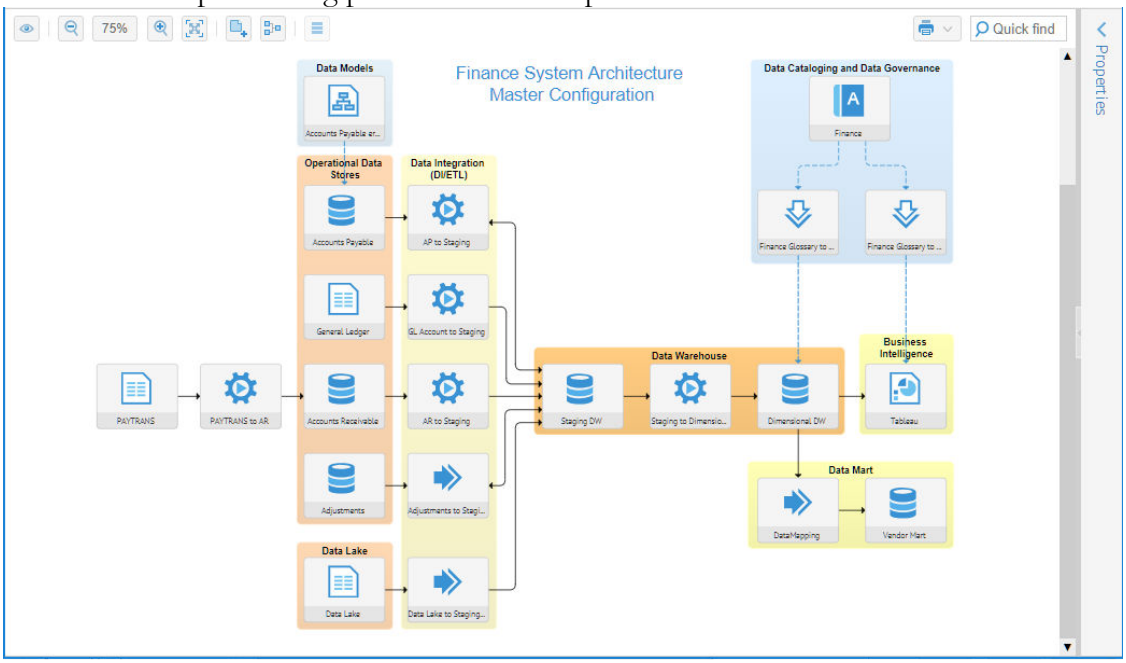


- BROWSE has also been massively improved in both performance (now also Lucene based) and in functionalities as a metadata asset type driven browser with

support for hierarchical display at all levels of any data sources including database, DI, BI, Data Lakes, and even No SQL (like JSON hierarchal structures)



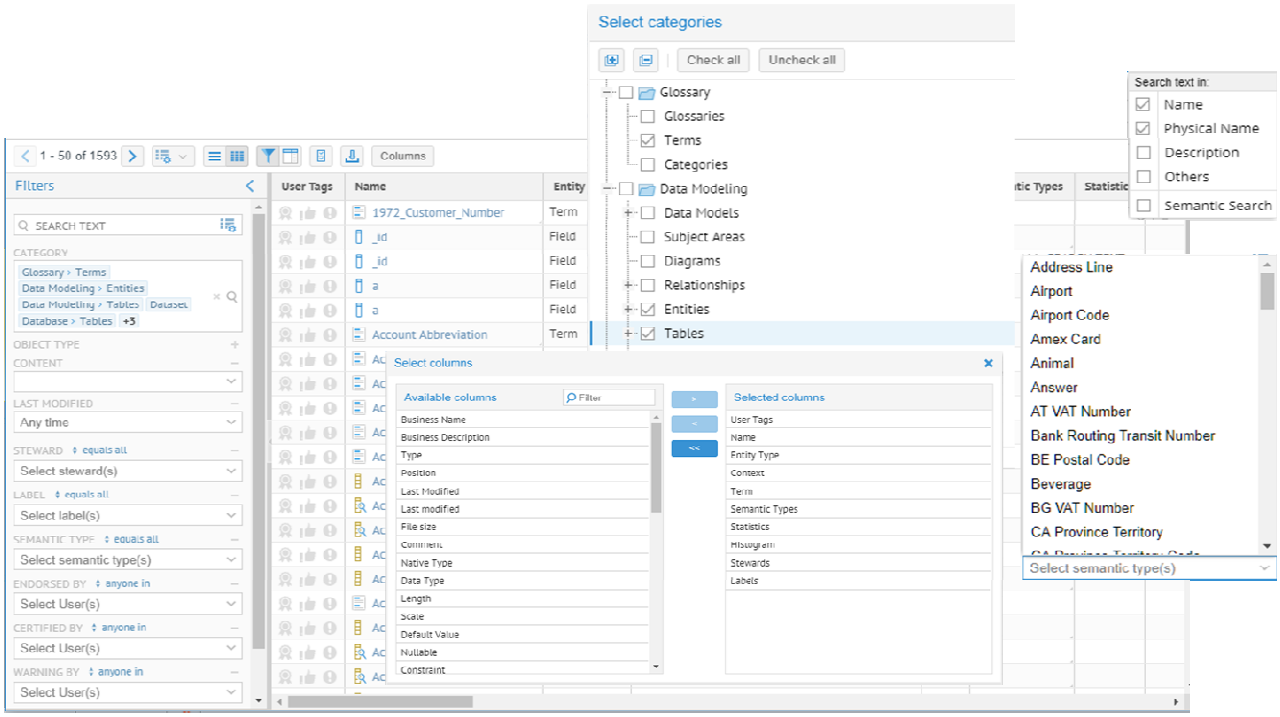
- Enterprise ARCHITECTURE driven graphical navigation allows users to drill down from a top down big picture of the enterprise architecture.



METADATA REPORTING

Brand new powerful unified metadata reporting capabilities where both search and browse end up to the same reporting page which is also directly available at Browse > Report. Starting from search simply predefines the text filtering (e.g. customer), while browsing predefines a category (e.g. database / tables), and direct access to reporting does not predefine anything.

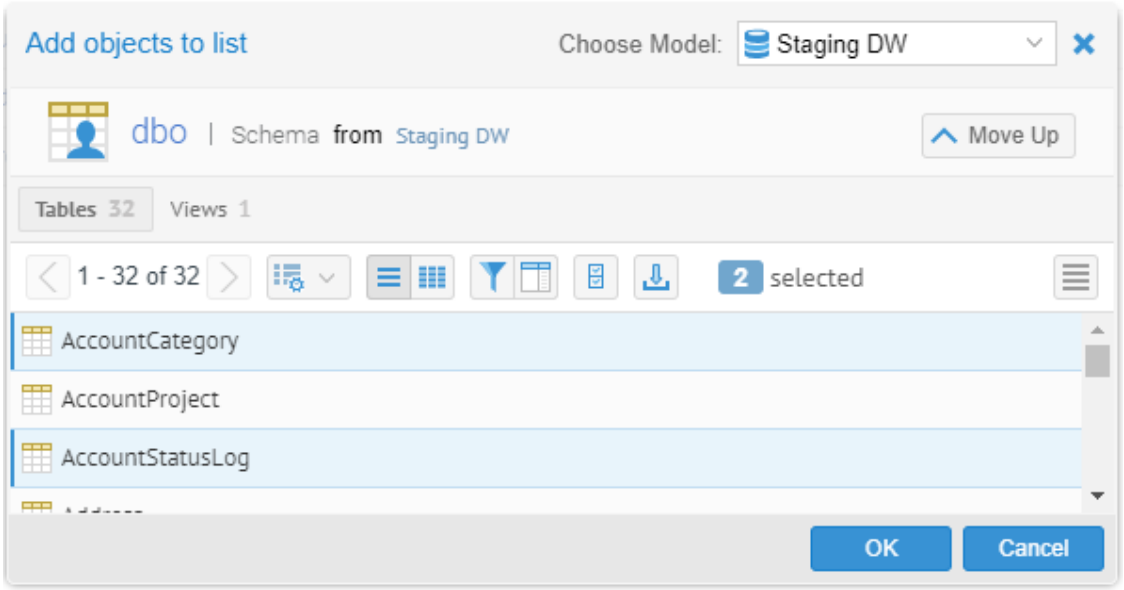
- The reporting capabilities offers to select multiple categories (e.g. database / tables + Flat files) and subset by content (My Data lake + Sales DW database) before drilling down with the following filters:
- Then filtering is available for Last Modified, Stewards, Labels, Semantic Types, Endorsed By, Certified By, Created By, Warning By, and Commented By.
- Finally, more custom filtering per attribute (including custom attributes) common to the metadata subset (e.g. SecurityLevel = Orange).
- Reports can be reused by saving the URL as favorites (further versions will support full report management within the application)



METADATA USER LISTS

Brand new user list management feature allows users to define and mange lists of metadata objects. Just like labels, lists are available anywhere in the UI to add/remove objects, bulk editing, and management. Lists can contain any type of metadata such my favorite list of terms, tables, or reports. Lists can also can contain multiple type of content such as my to do list with terms, tables, and reports in that list. Lists can be shared with

other users when marked as public, such as our quarterly review list. Note that lists are flat, therefore not hierarchical and with no sub-list or include concepts.



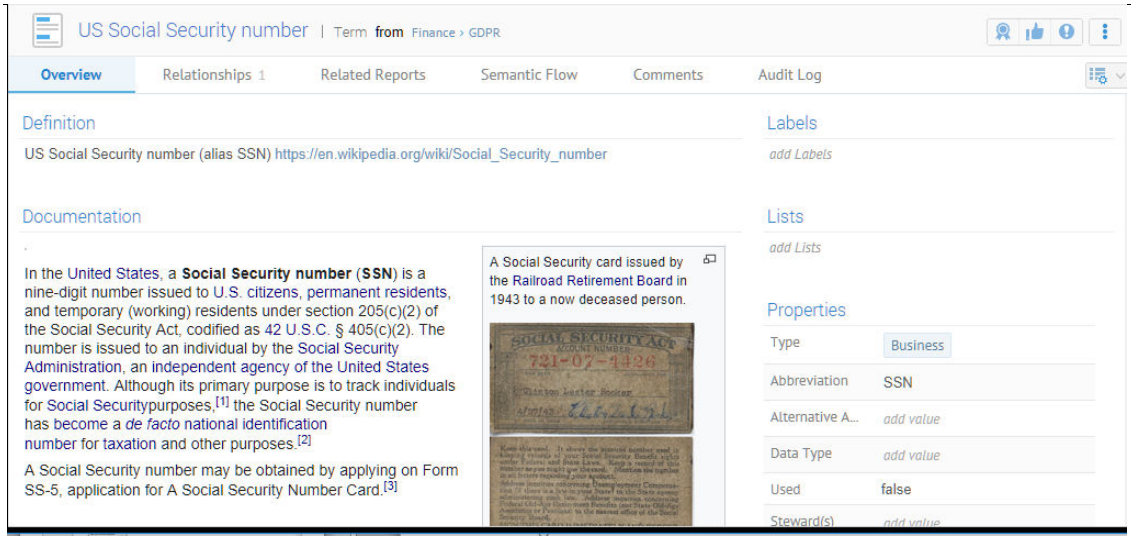
METADATA TAGGING WITH LABELS

The metadata tagging with labels has been much improved to be harmonized with the brand-new list management experience in order to facilitate adding/removing objects anywhere, grid editing, and more.

METADATA DOCUMENTATION

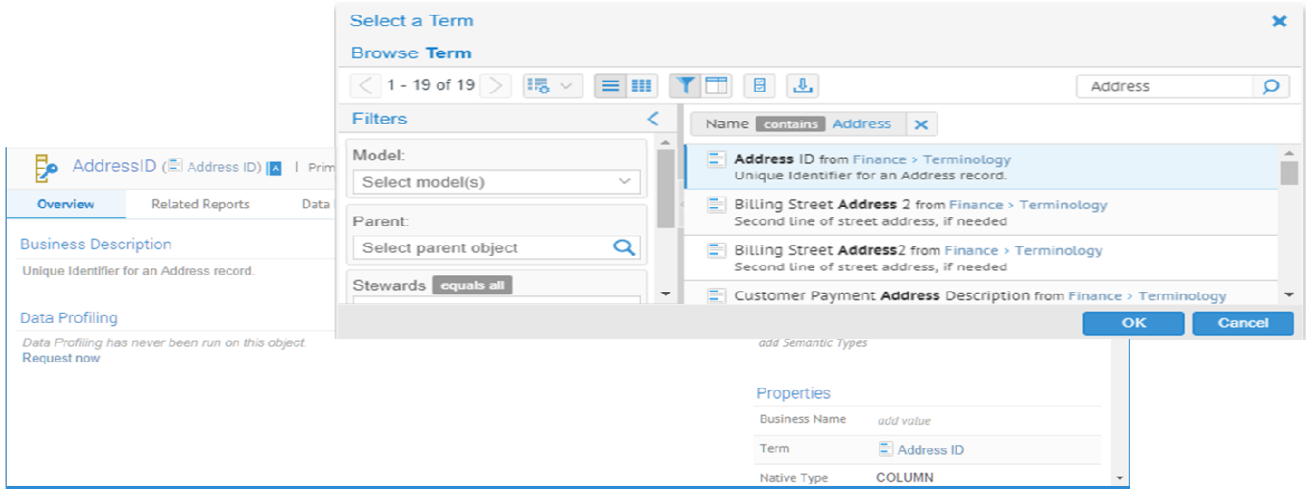
Much improved ways to document metadata:

- MULTI-LINE TEXT has been introduced (in addition to the previous single line Text for better formatting and layout. In addition, Multi-Line text has been enhanced with support for URL links and embedded image attachments using a JIRA like syntax. Multi-Line Text is not only the default format for all Descriptions and Comments, but is now also available as a new type of Custom Attribute that can be applied to any metadata for documentation.
- RICH TEXT Documentation with (WYSIWYG) Visual Edition is not only the default medium for Glossary Term documentation, but is now also available as a new type of Custom Attribute that can be applied to any metadata for documentation.
- SQL TEXT of SQL View, Stored Procedures and more are now better presented with colored syntax and optional reformatting. Note that this is not a new type of custom attribute but any predefined attribute with SQL is better formatted.
- ATTACHMENTS (such as pictures, documents, etc.) have been enhanced as part of its integration with the new Metadata Explorer, including Management (Drag and Drop), Preview, and Thumbnails that can be embedded in the Text (and Multi-Line text) descriptions, comments and custom attributes.



DATA CATALOGING

- Brand new Data Cataloging applications well integrated with the existing Data Governance (DG) capabilities, and based upon the solid Metadata Management (MM) foundations with full data linage and powerful metadata version and configuration management.
- Managing both modern cloud based data lakes and classic Data Warehouse (DW) Enterprise Architectures
- Harvesting metadata from both modern (XML, JSON, Avro, Parquet, ORC) files, Hive tables, and Kafaka messages), and classic (relational tables / CSV files) data technologies



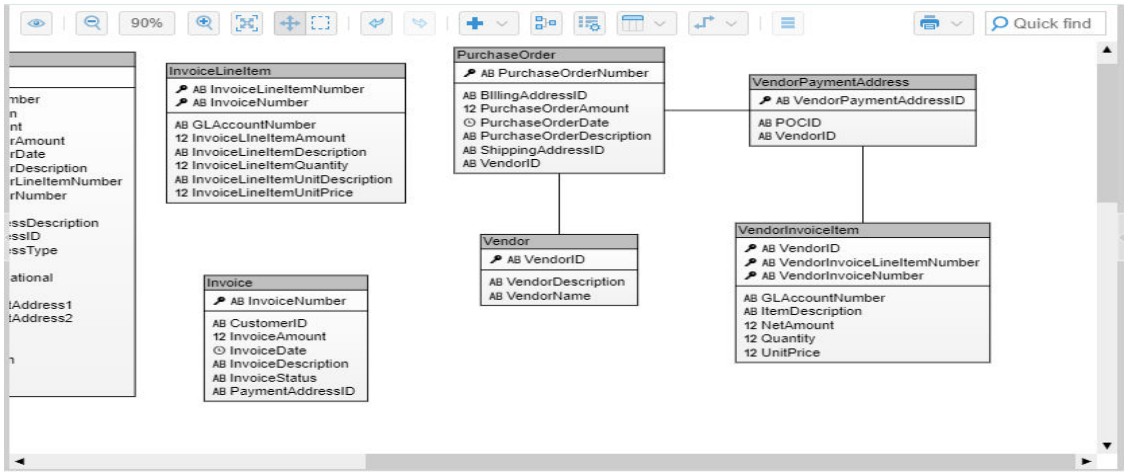
DATA MODELING AND DOCUMENTING ANY HARVESTED METADATA

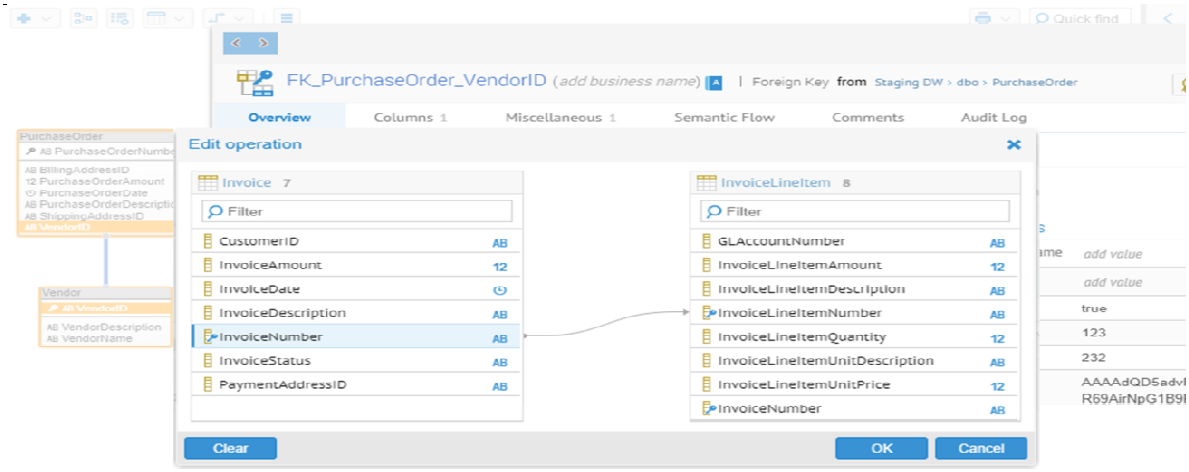
In OMM v12.2.1.2, existing data stores such as RDBMS could be harvested as a Physical Data Model (PDM) instead of a simple Model, in order to offer full documentation including business glossary term reuse based upon automatic semantic links, reverse

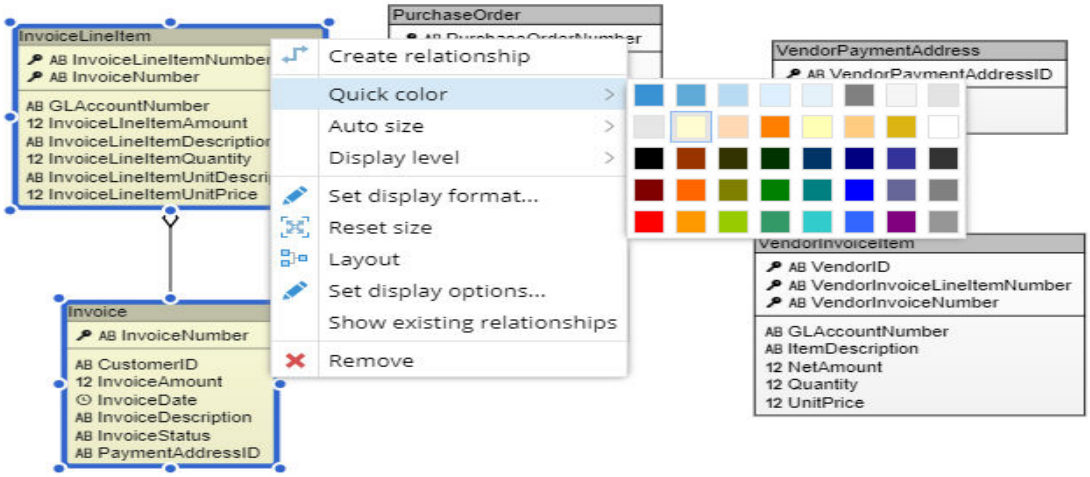
engineering based upon naming standards, data modeling with diagramming, and of course automatic change management (re-harvest/compare/merge).

- In OMM v12.2.1.3, all the above capabilities are now available on any harvestable model content without having to create a PDM. In other words, any data integration, business intelligence, reports, data stores (relational, hierarchical, NoSQL, files, etc.) can be documented at as needed, including support for relational data models. Consequently, all existing PDM in OMM v12.2.1.2 may be converted to Models in OMM v12.2.1.3 without loss of any existing documentation (including diagrams).
- The documentation (business names and definitions) process has been improved allowing any object (e.g. table, column, report field) to be quickly and easily:
    - "Classified" with a local semantic link to a glossary term, without having to use an intermediate Semantic Mapping content, or associating the Model to a Glossary as with the PDM.
    - "Documented" with a local business name and definition overwriting any Semantic link (Classified, Mapped or Inferred)

  Furthermore, this documentation process is also dramatically enhanced through the integration with a new "Semantic Flow" tab acting as an interactive dashboard on finding the right definitions (see below).
- When harvested databases that are already documented in data modeling tools (e.g. Erwin), such data models can be imported as a separate model and automatically stitched directly to its matching harvested database (without using any semantic mapping model). The semantic stitching is automatically maintained as both the database and its associated data model are independently re-imported/refreshed on regular basis (the stitching will report inconsistencies). From the user perspective, the documentation (business name, descriptions, relationships, diagrams) of any harvested database table / column is automatically inherited from its associated data model.

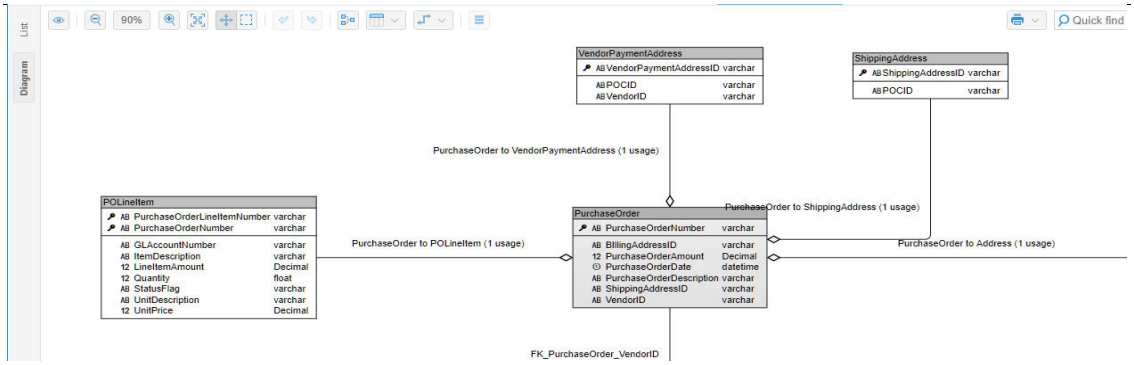






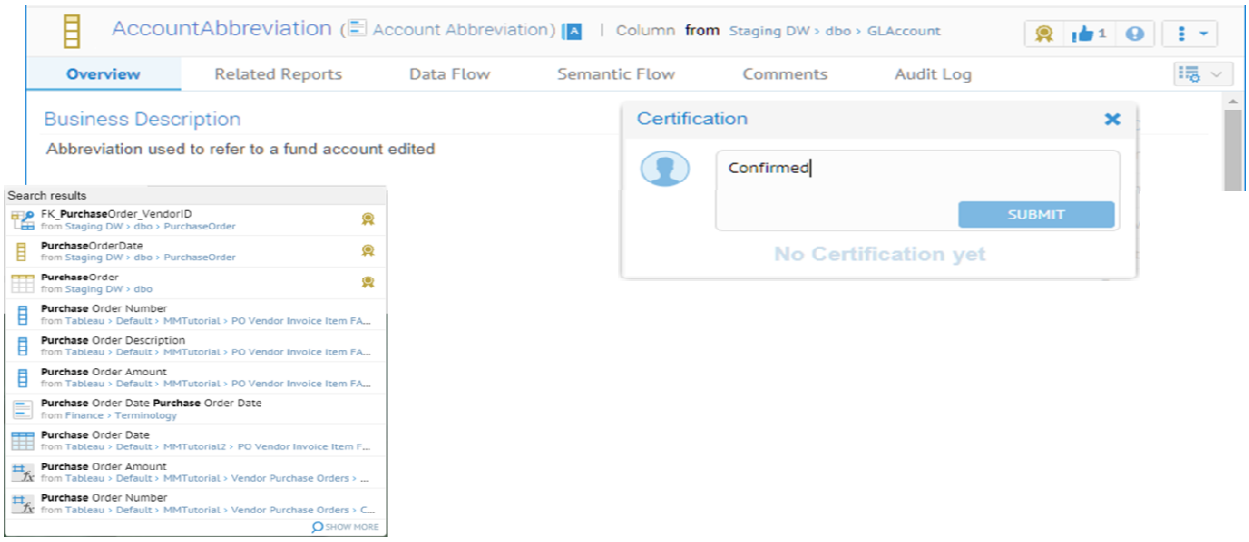
RELATIONSHIPS Discovery & Management

- Relationship Discovery using the following methods:
  - Automatically "Inferred" based on:
    - Metadata Usage Driven: using the surrounding data flow usage such as joins in DI (ETL Tools, SQL Scripts or Data Prep) and BI (traditional or self-service) activities.
  - On Demand "Detected" based on:
    - Metadata Name Matching: for example PurchaseOrder.SKU = Product.SKU or Customer.AccountId = Account.Id)
    - Semantic Definition Matching: classified by users to the same glossary term.
- Relationship Management with user defined relationships and social curation (e.g. endorsed or certified joins)
- Dynamic Data Model diagram generation from Relationships surrounding any object (e.g. table or file).



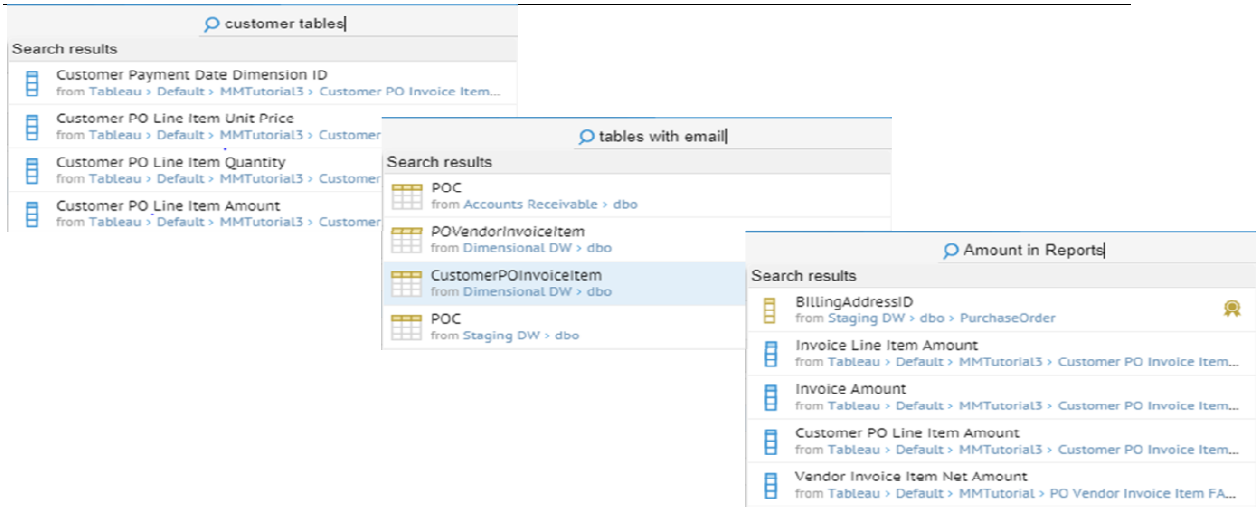
SOCIAL CURATION

Endorsement, warnings, certifications with impact on search ranking.



SEMANTIC SEARCH

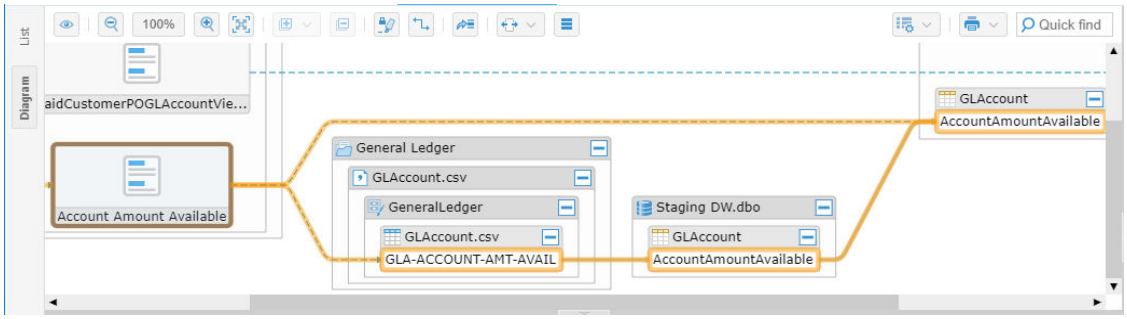
Metadata driven search language such as "Customer tables" for any tables with Customer in the name, "tables with SSN" for any table with a SSN column (e.g. for GDPR), or "ROI in Reports" for any reports containing ROI.



SEMANTIC MAPPING

Major improvements in semantic mapping including in place semantic mapping via two approaches:

- Top-down from Business Glossary Term or Data Model Entity/Attribute,
- Bottom-up from Data Store Tables/Columns or Report Fields.

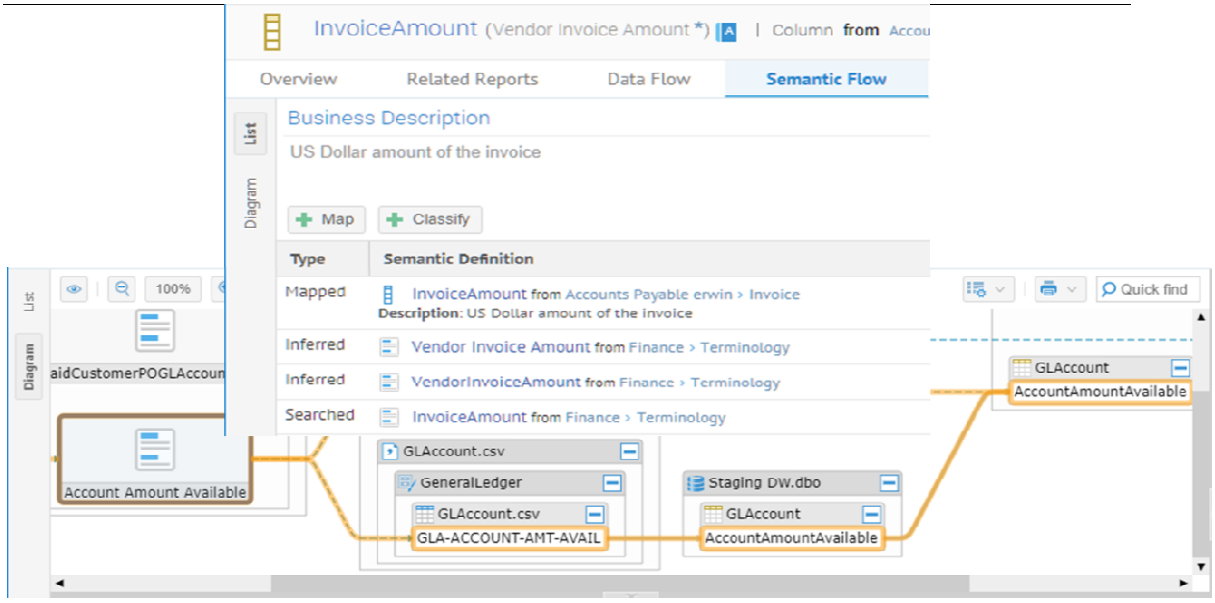


SEMANTIC FLOW

Major improvements on the semantic flow analysis now also supporting the documentation process acting as an interactive dashboard for finding definitions that are:

- "Local" (within the Model) that has been either "Imported" (metadata harvesting) or locally "Documented" (edited description overwrite),
- locally "Classified" (within the Model) to an external glossary term,
- directly "Mapped" via a Semantic Mapping Model or direct stitching (e.g. between a database and its data model),
- indirectly "Inferred" through complex data flow pass through and semantic flow (which can be graphically analyzed in the data flow diagram), or
- "Searched" for by name in all glossaries.

Any of the Searched, Inferred or Mapped definitions may quickly (in place) be reused/promoted as a Classified or Mapped definition.

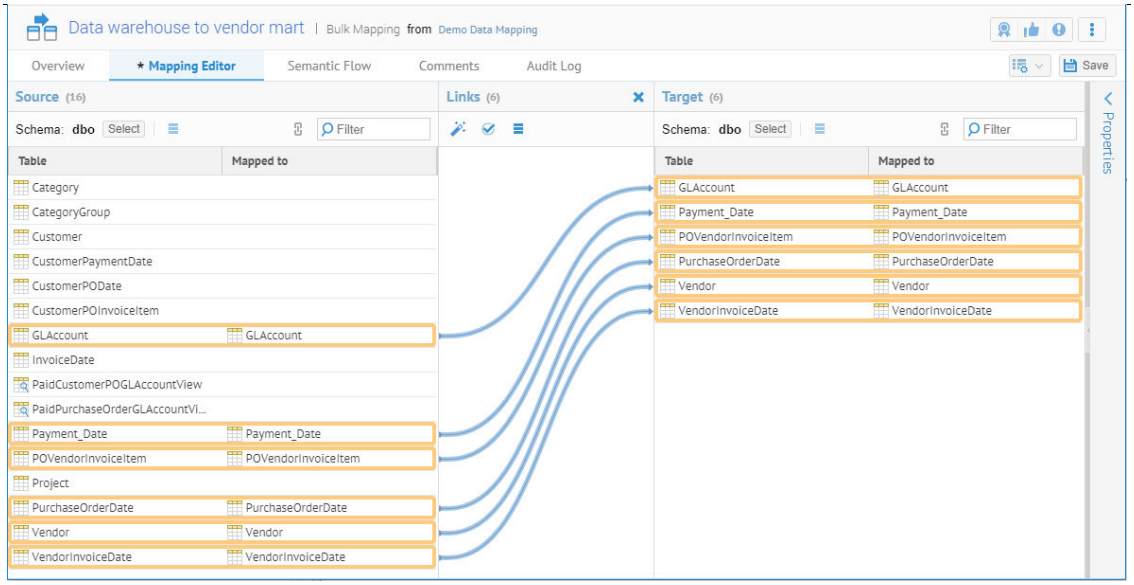


DATA CONNECTIONS / METADATA STITCHING

- Complete support for file format harvesting and stitching.
- Connection pool factorization (e.g. from DI and BI servers) to minimize the number and complexity of stitching connections.

DATA MAPPING Specifications & Design

- The Data Mapping Specifications and the Data Mapping Designs have been fully resigned and merged into brand new Data Mappings that can be used for multiple purposes, including capturing data flow mapping requirements.
- The Data Mapping tool allows for the mapping of multiple source data stores into a target data store in multiple steps with (schema or table level) Bulk Mappings and (column/field level) Query Mappings. The data mapping tool offers new graphical mapping visualization as you map, and new expression syntactical editors when designing joins, lookups, filters, etc.



Metadata Harvesting (Model Bridges) Improvements

- DATA STORES (Flat files, RDBMS, Cloud Databases, Big Data / Hadoop):
  - DATA CATALOGING OF THE DATA LAKE on local file servers (POSIX file systems), over the cloud (e.g. Amazon S3), on big data clusters (e.g. Hadoop distributions), on messages databases (e.g. Kafka), or on NoSQL databases (e.g. MongoDB):
    - with high level data lake directory inventory import bridges such as:
      - File System (for Linux and Windows) import more
      - Amazon S3 (Simple Storage Service) import more
      - Apache Hadoop Distributed File System (HDFS) (for Cloudera, Hortonworks, and MapR) import more
      - Microsoft Azure Blob Storage import more
      - Microsoft Azure Data Lake Storage import more
      - OpenStack Swift (Object Storage) import more
      - Apache Kafka Cluster import more
      - MongoDB Database (JSON) import more
      - Apache CouchDb Database (JSON) import more
      - MarkLogic NoSQL Database (for JSON only) import more
    - automatically crawling and discovering lower level metadata from flat file, NoSQL, and other file import bridges such as:
      - Flat File (CSV) import more
      - Open Office Excel (XSLX) File import more
      - JSON (JavaScript Object Notation) File import more
      - Apache Avro (JSON) File import more
      - Apache Parquet File import more
      - Apache ORC File import more
      - W3C XML File import more
  - DATA SERVICES (REST API, Web Services):
    - New Bridges:
      - Open API Specifications (OAS) import more
  - OTHER DATA STORES:

        - New Bridges:
            - Apache Hadoop HBase Database (Java API) import more
            - Apache Hadoop Hive Database (Hive Metastore via JDBC) import more
            - Microsoft Azure SQL Database import more
            - Microsoft Azure SQL Data Warehouse import more
            - SAP HANA import more
            - SAS Library import more
- DATA INTEGRATION (DI/ETL/ELT/Scripts):
    - New Bridges:
        - Apache Sqoop import more
        - Informatica Developer import new support for DI/ETL more
        - SAP Data Services (BODI/BODS) import more
        - SAS base SAS code import more
        - SAS Data Integration (via SAS Metadata Server) import more
        - Talend Data Preparation import more
    - Improved Bridges:
        - Microsoft SQL Server Integration Services (SSIS) import support for up to version 14.0 (2017) more
- BUSINESS INTELLIGENCE (BI/OLAP):
    - New Bridges:
        - SAS BI Visual Analytics (via SAS Metadata Server) import more
        - Microsoft Azure Power BI import more
        - TIBCO Spotfire Server import more
        - TIBCO Spotfire Document (File) import more
        - TIBCO Spotfire Binary Data (SBDF Format) (File)r import more
        - TIBCO Spotfire Text Data (STDF Format) (File) import more
    - Improved Bridges:
        - Microsoft SQL Server Analysis Services (SSAS) import support for up to version 14.0 (2017) and specifically its new JSON serialization of the Tabular Model more
        - Qlik Sense import more
- DATA MODELING (DM) & other Modeling Technologies:
    - New Bridges:
        - W3C Semantic Web Ontology (OWL/RDF) import more
- METADATA MANAGEMENT (MM) & DATA GOVERNANCE (DG):
    - Improved Bridges:
        - Apache Atlas (for HortonWorks) import more
        - Cloudera Navigator import more

ODI 12c New Features Overview
October 2018
Author: ODI Product Management

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200
oracle.com

Oracle is committed to developing practices and products that help protect the environment

SOFTWARE. HARDWARE. COMPLETE.