

Managing Metadata with Oracle Data Integrator

*An Oracle Data Integrator Technical Brief
Updated December 2006*

Managing Metadata with Oracle Data Integrator: An Oracle Data Integrator Technical Brief

Metadata—the description of data itself, its purpose, how it is used, and the systems used to manage it—is the foundation of all information management initiatives aimed at increasing efficiency across an enterprise.

INTRODUCTION

Metadata—information that describes data—is the foundation of all information management initiatives aimed at increasing efficiency across an enterprise. It is the cornerstone for data integration and master data management (MDM). This technical brief defines metadata and describes how Oracle Data Integrator can help companies manage it more efficiently.

ABOUT METADATA

What Is Metadata?

Metadata is often generally described as “information about data.” More precisely, metadata is the description of the data itself, its purpose, how it is used, and the systems used to manage it. This usually includes

- Data model definitions
- Data transformation descriptions
- Process and flow information, such as schedules and logs
- Business intelligence systems description, such as cubes, dimensions, and metrics
- Any other type meaningful information, such as semantics

Metadata not only consists of technical information but also includes information that makes business users aware of the data’s purpose and use. In addition, it is not related only to data warehousing and business intelligence; it comprises all four categories of the entire enterprise architecture: business, application, data, and technology.

Where Is Metadata?

Metadata is scattered across the information system

- As data model descriptions in database dictionaries
- In Extensible Markup Language (XML), Document Type Definition (DTD), and XML Schema Definition (XSD) files

- In proprietary systems
- As pieces of code that perform data processing tasks
- In unstructured documents, such as Microsoft Word documents and spreadsheets

Metadata is also part of the knowledge held by a company's staff.

Why Is Metadata Critical?

You can think of metadata as the lines and points that draw the plan of the enterprise architecture. Accurate and comprehensive metadata enables organizations to evaluate their architecture's efficiency and eases the process of changing this architecture. Metadata is important for certification purposes (ISO-9002), and it becomes critical when regulatory compliance needs (Basel II, SOX) arise.

Metadata is now critical to information management initiatives such as MDM, customer data integration (CDI), product information management (PIM), and product data management (PDM). Such initiatives focus on creating and maintaining a clean, accurate view of corporate reference data that is shared across all operational and analytic systems. They consequently require the entire company to have clean, accurate metadata. For example, to create a unique view of customers, you must have a clear, accurate, and precise definition of a customer; its attributes; which processes and people can create, alter, or use customers; and so on.

MANAGING METADATA WITH ORACLE

Oracle Data Integrator provides all the features you need to retrieve, enrich, extend, and leverage existing metadata for agile corporate enterprise architecture.

The Oracle Data Integrator Repository

Oracle Data Integrator is built around an open, comprehensive metadata repository, which can be installed on relational databases such as Oracle, Microsoft SQL Server, IBM DB2 UDB, IBM DB2/400, Informix, Sybase AS Anywhere, Sybase AS Enterprise, and Sybase ASIQ. Because metadata is stored in database tables, the repository is completely open and can be used as a source by any reporting system.

The repository includes all the domains used in data integration:

- **Security information**—including users, profiles, and access privileges for the Oracle Data Integrator platform
- **Topology information**—including technologies, definitions of servers, schemas, contexts, and languages

Metadata is critical to information management initiatives such as MDM, CDI, and PIM/PDM, which focus on creating and maintaining a clean, accurate view of corporate reference data shared across operational and analytic systems.

- **Data models**—including descriptions of schemas, datastore structures, fields and columns, data integrity rules, cross-references, data lineage, and so on
- **Integration projects**—including declarative rules, transformation rules, mappings, packages, procedures, folders, knowledge modules, variables, and so on
- **Execution**—scenarios, scheduling information, and logs

This repository also includes information about data warehousing and business intelligence systems. For example, it includes information on slowly changing dimension behavior for columns or OLAP (online analytical processing) type (dimension, fact, time) for datastores. Users can further extend and customize the repository to include other information relevant to the enterprise.

Reverse Engineering

Oracle Data Integrator provides a reverse engineering process that populates the repository with metadata from the information system. Reverse engineering retrieves metadata from data storage (databases, XML files, and so on) and loads it into the repository. (For example, reverse engineering detects the description of tables, columns, datatypes, constraints, and comments from a database to load the repository.) Reverse engineering supports database management systems, files of many diverse formats, and COBOL CopyBooks through wizards.

Users can customize the reverse engineering process to retrieve nonstandard metadata from the databases or from proprietary repositories such as ERP/CRM (enterprise resource planning/customer relationship management) systems (SAP, Oracle's Siebel, Oracle Applications, salesforce.com, and so on). Customization is done through Oracle Data Integrator's reusable reverse engineering knowledge modules (RKMs). Oracle Data Integrator comes with pre-built RKMs for the major technologies.

Oracle's reverse engineering process populates the repository with metadata from the information system. This process retrieves metadata from data storage and loads it into the repository.

Enriching Metadata

Users can enrich and develop new metadata through the Oracle Data Integrator graphical modules. This metadata includes the integration processes designed with Oracle Data Integrator. Users can also enrich or create data models directly in Oracle Data Integrator, by using the Common Format Designer feature, and can implement the data models in the source data structures.

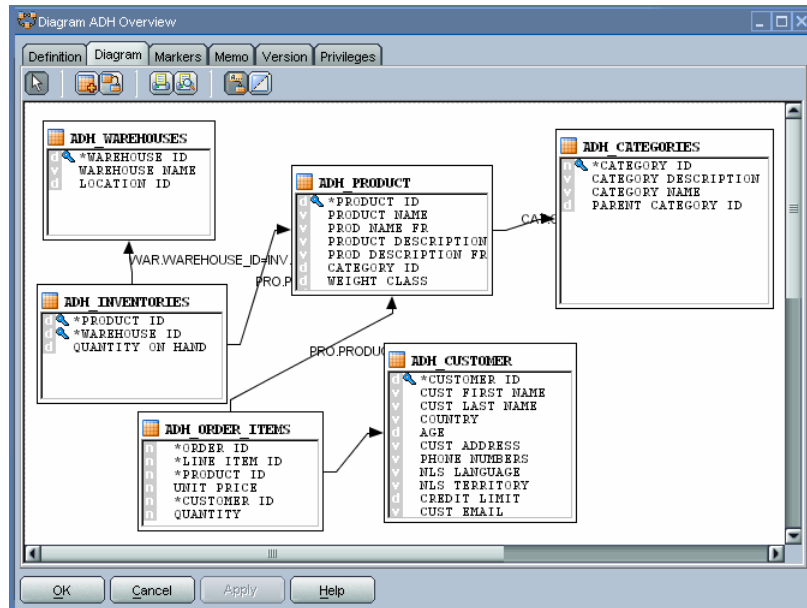


Figure 1: Data model design using Common Format Designer

Extend the Repository with FlexFields

With Oracle Data Integrator, users can extend any object with user-defined FlexFields. Like predefined fields, FlexFields are used seamlessly in Oracle knowledge modules.

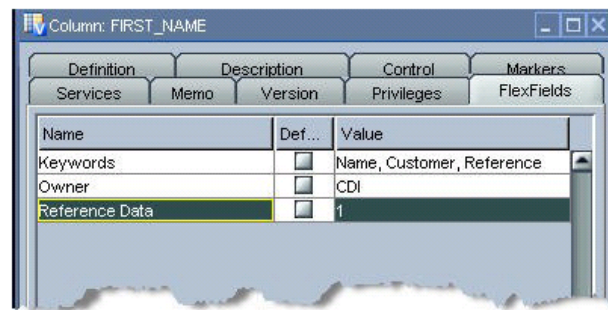


Figure 2: User-defined FlexFields in a column

LEVERAGING METADATA

Oracle Data Integrator is a metadata-based code generator. The Oracle knowledge modules use the metadata that's stored in the repository to create the appropriate code and run it on existing systems. Metadata, at the center of the integration and data integrity check processes, makes maintaining these processes faster.

Oracle Data Integrator can automatically generate rich Adobe PDF reports using the repository contents to document the enterprise architecture and integration processes.

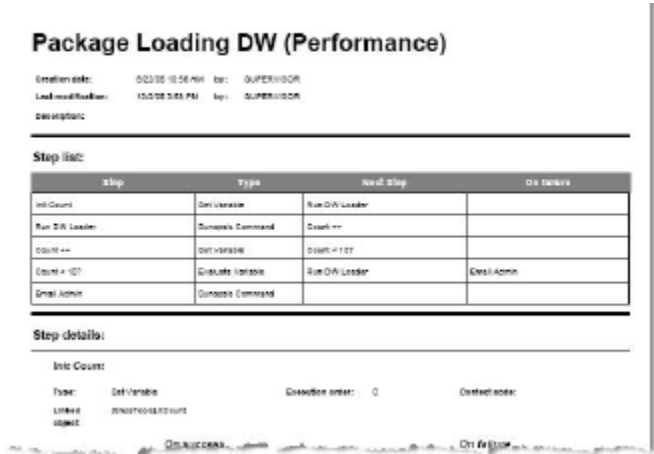


Figure 3: Repository metadata automatically generates documentation.

The Oracle Data Integrator Metadata Navigator web application provides access to repository metadata from any web browser. Users can browse the entire repository metadata and graphical views of the data flows. With the Data Lineage feature, they can easily track data to the source items from which they were derived and subsequently integrated by Oracle Data Integrator.

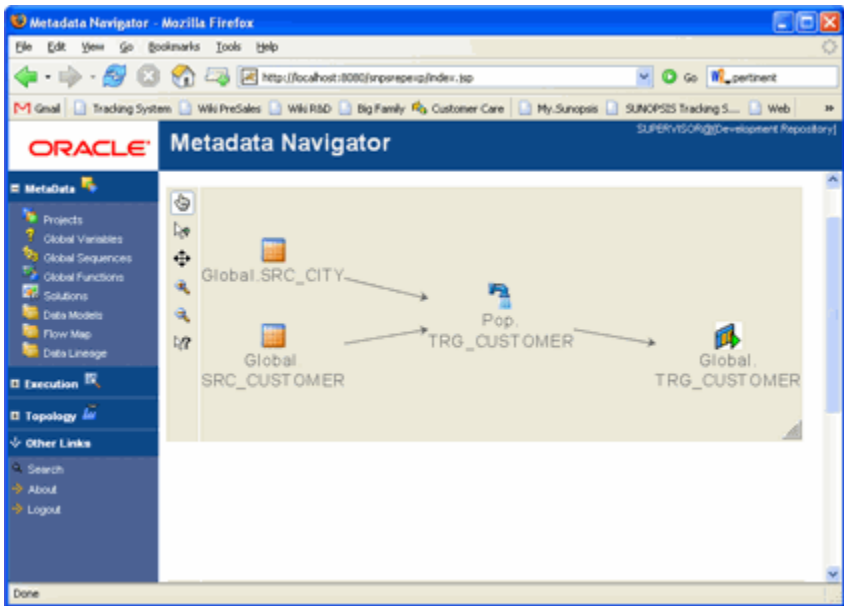


Figure 4: Data Lineage tracks and displays the source of data.

CONCLUSION

Oracle Data Integrator places metadata at the center of its architecture. It provides features for retrieving, enriching, extending, managing, and leveraging this metadata as a foundation for meeting business needs.



Managing Metadata with Oracle Data Integrator: An Oracle Data Integrator Technical Brief
Updated December 2006

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

Copyright © 2007, Oracle. All rights reserved.

This document is provided for information purposes only and the contents hereof are subject to change without notice.

This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission. Oracle, JD Edwards, PeopleSoft, and Siebel are registered trademarks of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.