

Oracle Data Integrator

¿Que es Integración de Datos?

La integración de datos la podemos definir como el proceso de combinar datos que residen en diferentes fuentes y permitirle al usuario final tener una vista unificada de todos sus datos. La habilidad de transformar datos inter-departamentales de fuentes heterogéneas en un plan de acción que se convertido en un reto y en una ventaja competitiva para compañías que requieran la integración de datos.

La integración de datos es un elemento fundamental y crítico en la variedad de tecnologías incluyendo Data Warehouse, aplicaciones de inteligencia de negocio, arquitecturas orientada a servicio, aplicaciones MDM y arquitecturas data-centric.

Oracle conociendo la necesidad de la integración de datos para muchas empresas y distintos tipos de industria, tiene una solución innovadora conocida como Oracle Data Integrator.

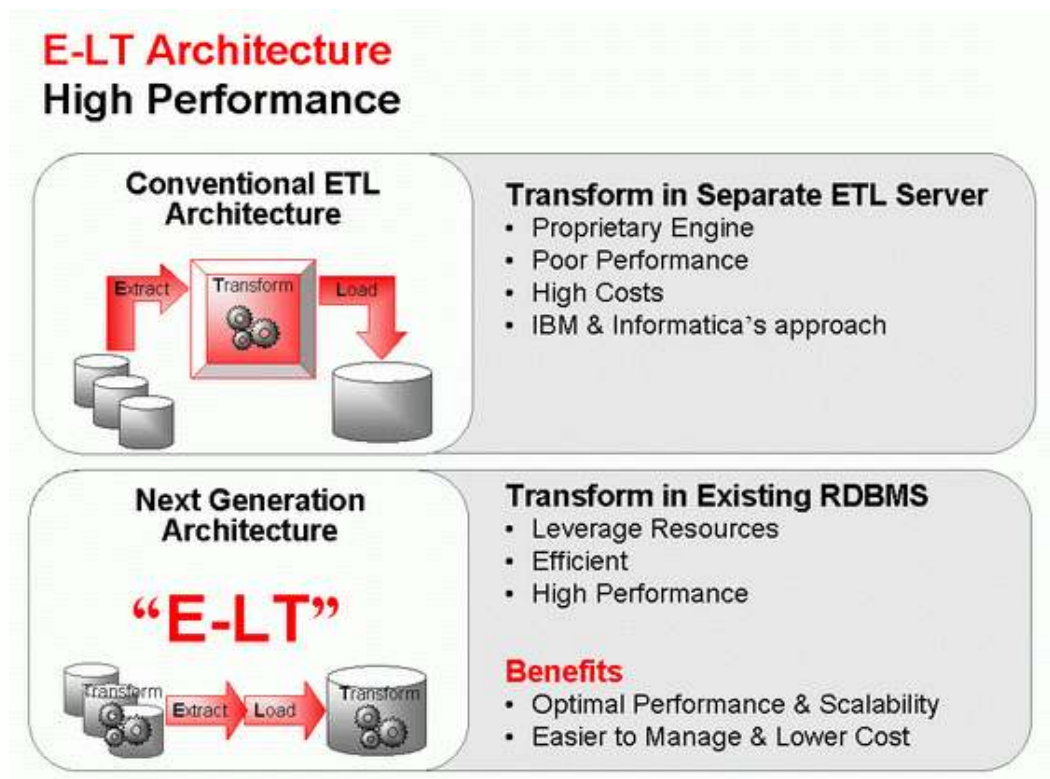
¿Que es el Oracle Data Integrator (ODI)?

Oracle Data Integrator es una plataforma de integración completa que cubre los requisitos de integración de datos. Maneja alto volumen, provee lotes de alto desempeño a procesos dirigidos a eventos, a servicios de integración basados en una arquitectura orientada a servicios y con la capacidad de procesar eventos en tiempo real.

Oracle Data Integrator maneja múltiples necesidades empresariales referentes a la integración de datos:

- **Data Warehousing e Inteligencia de Negocios -**
Tiene la capacidad de manejar grandes volúmenes de datos con un desempeño óptimo para cargar Data Warehouse y Data Mart. Maneja cargas incrementales, integridad de datos, reglas de negocio y consistencia
- **Arquitectura Orientada a Servicios –**
Provee la funcionalidad de invocar servicios externos para propósitos de integración e implementar servicios de integración y transformación integrados a una arquitectura orientada a servicios.
- **Master Data Management –** es una combinación de aplicaciones y tecnologías que consolidan, limpian, mejora los datos maestros de la empresa y los sincroniza con aplicaciones, procesos de negocio y herramientas analíticas como Oracle BIEE+.
- **Migración –** Provee cargas masivas eficientemente de datos históricos, incluyendo transformaciones complejas de sistemas legacy a sistemas nuevos.

El Oracle Data Integrator va mas allá de un ETL tradicional, donde se extraen los datos de diferentes fuentes, se transforman en el proceso y luego se cargan a su destino. La funcionalidad ELT cambia en el donde y como la transformación se ejecuta y se apalanca del motor de base de datos. ELT direcciona la transformación de datos a la base de datos destino, cambiando el orden de operación a extraer los datos de las distintas fuentes, cargarlos en el servidor de destino y luego hacer la transformación utilizando operadores SQL nativos del RDBMS.



¿Quién necesita ODI?

Toda empresa que necesite de sus datos para la toma de decisiones y la consolidación de estos datos de diferentes fuentes de información más que una oportunidad o un reto debería ser una acción a tomar.

Preguntas a considerar relacionadas a la integración de sus datos

- ¿Su empresa maneja procesos de carga definidos?
- ¿Sus iniciativas de integración de datos han fracasado o sean a atrasado debido a la poca confiabilidad de sus datos?
- ¿Siempre se envía la información correcta al lugar correcto, en el momento correcto?
- ¿Cuanto tiempo se demora en rehacer el trabajo para corregir data incorrecta, poco precisa o errónea?
- ¿Su empresa tiene dificultades para cumplir con políticas y/o regulaciones?

- ¿El nivel de satisfacción de sus clientes continúa disminuyendo?
- ¿Desconfía de la información extraída para que le apoye en la toma de decisiones crítica?
- ¿Necesita información tiempo real?
- ¿Es posible hoy tener una visión única de sus clientes?

Estas y muchas mas interrogantes pueden ser claves para determinar la importancia de una estrategia de integración junto con una herramienta los suficientemente ágil y capaz para cubrir todas aquellas necesidades relacionadas a la integración de datos

Drivers

- Aumentar la Agilidad –
Es importante para las empresas incorporar los procesos de negocio con la integración de datos para poder manejar eficientemente los procesos de cambio
- Consolidación –
Una de las grandes barreras para la integración de datos es la fragmentación. Aplicaciones independientes y soluciones que no trabajan en conjunto o no se integran a una arquitectura orientada a servicio o una plataforma de BI. Como resultado en el esfuerzo de integración se crean mapas entre los sistemas que no están integrados.

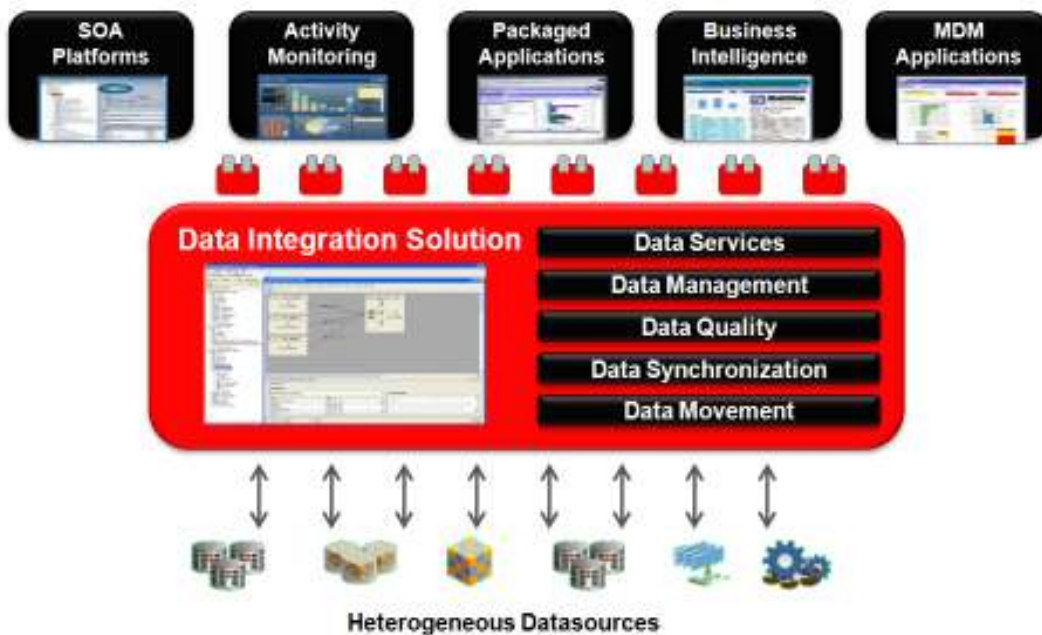


- Demanda Tiempo Real –
Para muchas empresas el tener su información en tiempo real sería una ventaja extraordinaria ante la competencia ya que le provee la facilidad de ver la dirección y le permite tomar decisiones basadas en resultados recientes.

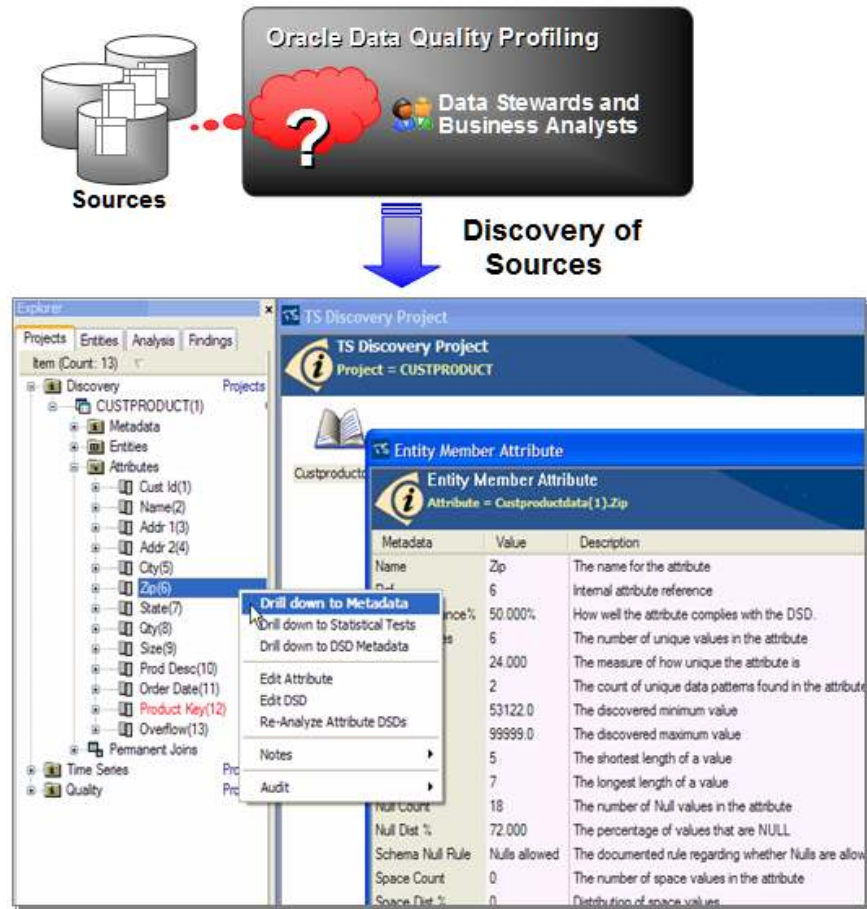
Capacidades Funcionales

Oracle Data integrator ofrece las siguientes capacidades funcionales:

- Data movement – Extraer datos y transferirlos de forma masiva
- Data synchronization - Capacidad de replicar y extraer los datos que sufrieron algún cambio
- Data quality – Limpieza de datos y a tono con las reglas de negocio
- Data management – Manejo de metadata y modelamiento de datos
- Data services - Plataforma para una implementación SOA para cubrir el gap entre los procesos de negocio y la infraestructura de aplicaciones



Data Profiling



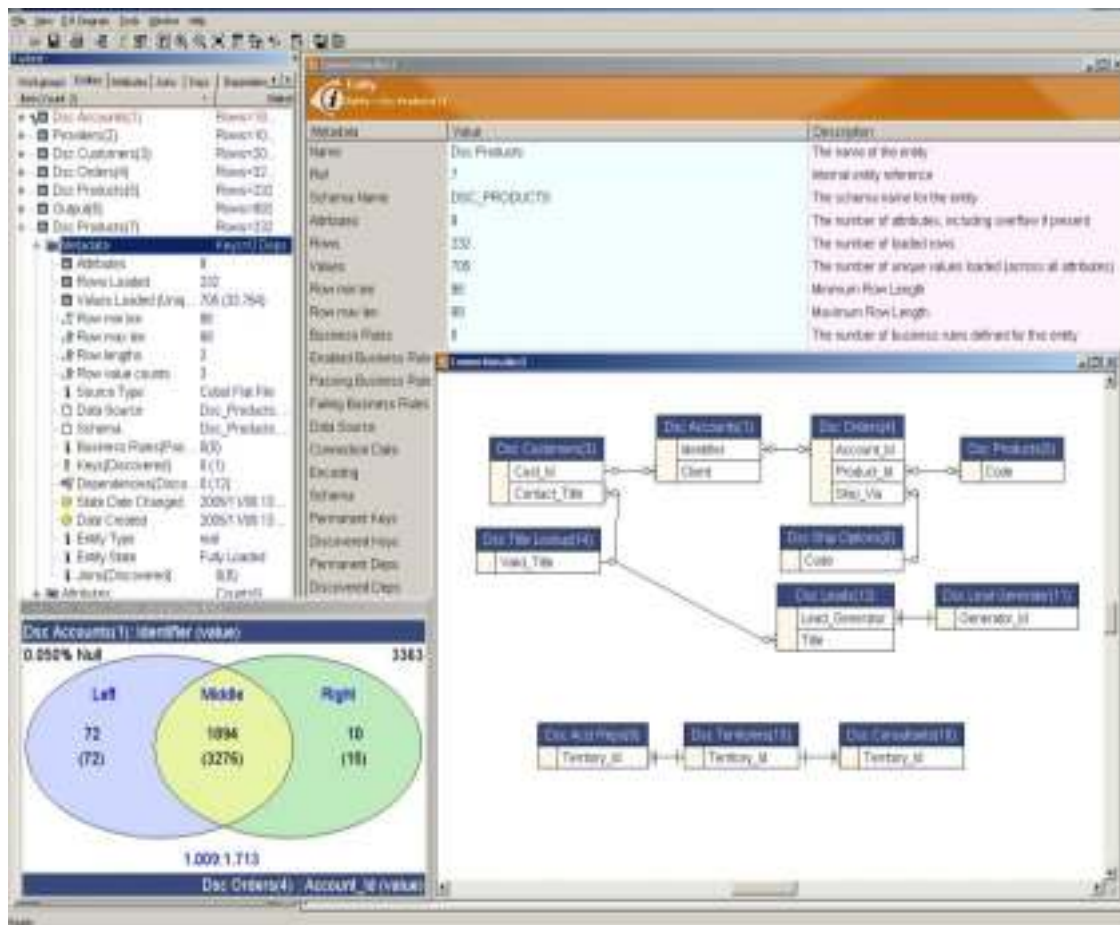
Oracle Data Profiling es una herramienta de análisis, investigación y monitoreo de calidad de los datos. Permite a los usuarios de negocio evaluar la calidad de sus datos a través de métricas, descubrir o inferir reglas basadas en los datos con la capacidad de monitorear la calidad de los datos a través del tiempo.

Su funcionalidad incluye:

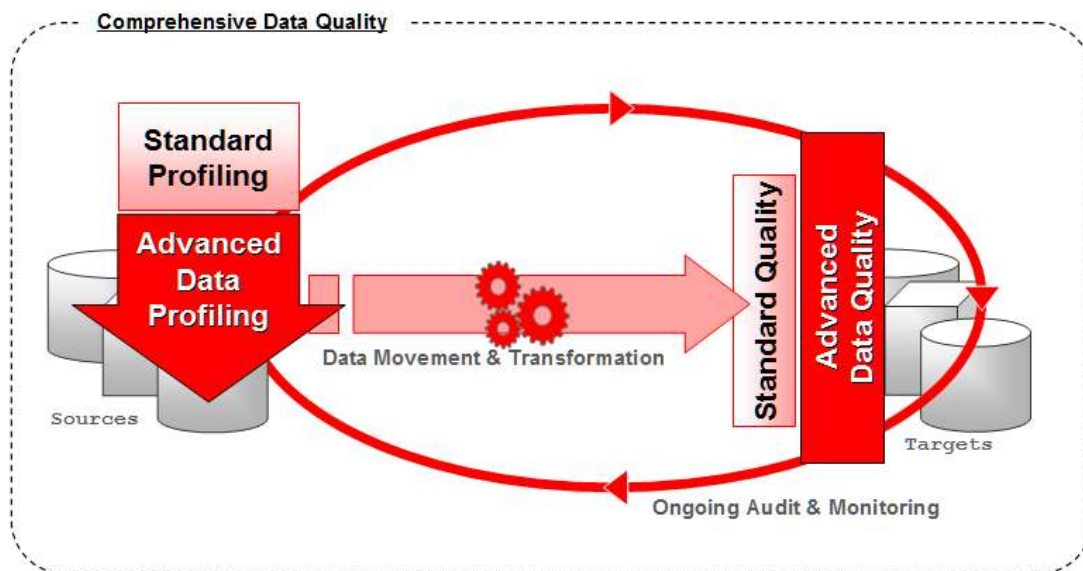
- Descubrimiento y análisis –
Oracle Data Profiling colecta los metadatos y datos de las fuentes y analiza la información, estadísticas, largo de atributos, valores máximos y mínimos, distribución, patrones, tipos de datos, etc. De esta forma aplica técnicas avanzadas para identificar problemas potenciales en la integridad de los datos
- Natural Drill Down –
Te permite profundizar a un nivel de detalle para evaluar los resultados de una forma gráfica evaluando detalles como estadísticas y metadatos

- Claves, dependencias funcionales y análisis –
Oracle Data profiling identifica duplicados y otras inconsistencias, detecta dependencias funcionales entre los atributos en una entidad

Oracle Data Profiling permite a los especialistas de datos documentar y anotar los hallazgos, definir reglas de negocio, crear reportes y exportar datos y metadatos para el uso de alguna aplicación. Oracle Data Profiling también utiliza notificaciones vía e-mail para alertar al usuario cuando los niveles de servicio no hayan sido cumplidos.

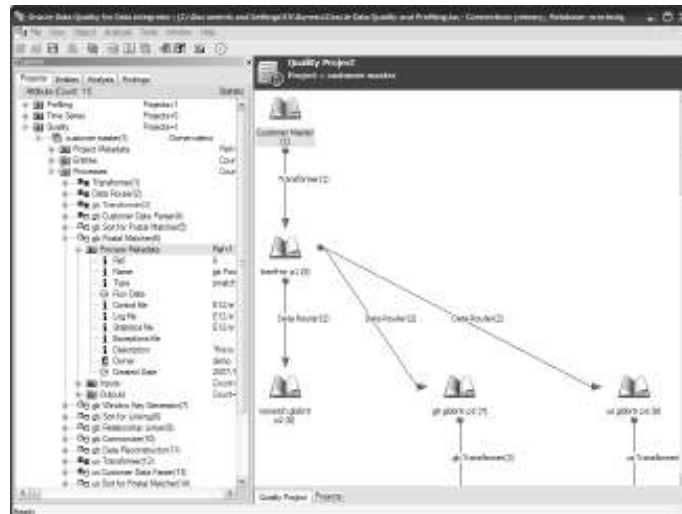


Data Quality

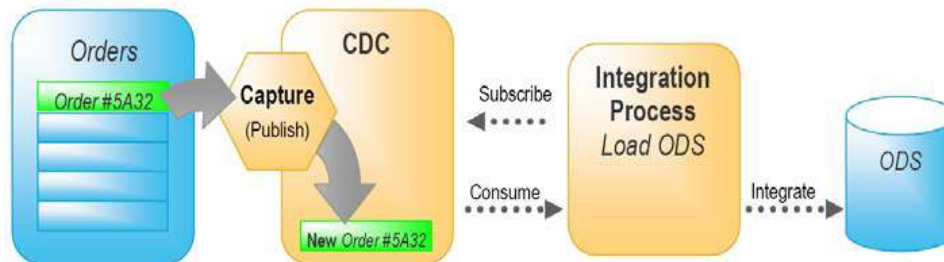


Muchas empresas sufren de poca calidad en sus datos. Datos inconsistentes, no precisos, incompletos y fuera de fecha son la raíz de muchos problemas operacionales, de incapacidad para optimización, insatisfacción de clientes entre otros.

Oracle Data Quality es una plataforma da calidad de datos que cubre hasta las necesidades mas complejas de calidad de datos. Es un motor basado en reglas, robusto y con una arquitectura escalable que apoya la estrategia de integración de datos. Oracle Data Quality dirige la calidad de datos de la empresa de acuerdo a la necesidad de los proyectos incluyendo Data Warehouse e Inteligencia de Negocios, Integración de Datos, migraciones e integración de procesos orientado a servicio.



Change Data Capture



Oracle Data Integrator's Changed Data Capture identifica y captura los datos como van a ser insertados, actualizados o borrados de la fuente de datos y hace que los datos que sufrieron cambios queden disponibles para los procesos de integración.

El modelo de publicar y subscribir trabaja en 3 pasos:

- Identifica el subscriptor, usualmente un proceso de integración, subscribe los cambios que pueden ocurrir en la fuente de datos.
- El Changed Data Capture framework captura los cambios en la fuente de datos y los publica para el subscriptor.
- El subscriptor puede procesar los cambios monitoreados en cualquier momento y consumir estos eventos

El Oracle Data Integrator puede procesar estos cambios de dos formas

- Regularmente en lotes – procesar las órdenes nuevas cada 5 minutos y cargarlas en la fuente de datos.
- Tiempo real - cuando los cambios ocurren

Módulos de Conocimiento

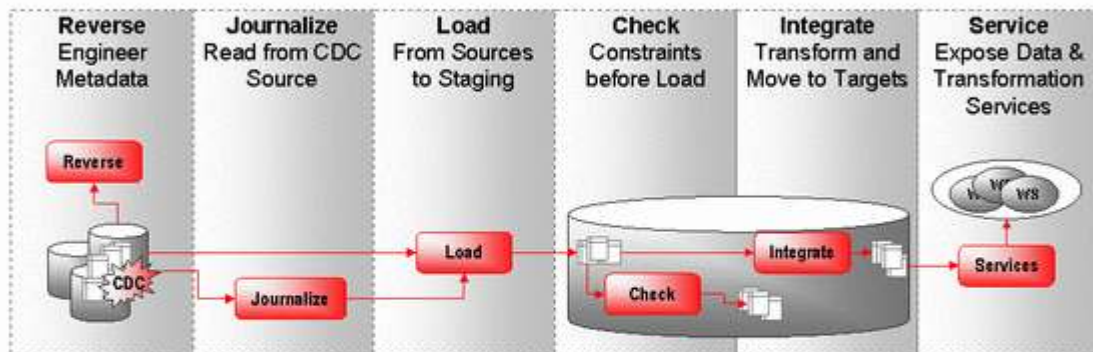
Los Módulos de Conocimiento son componentes de Oracle Data Integrator Open Connector. Estos módulos contienen el conocimiento requerido por el Data Integrator para desempeñar un set de tareas relacionadas a una tecnología en específico, como por ejemplo Oracle, SQL Server, DB2, Teradata, etc.

Estos Open Connectors contienen una combinación de:

- Estrategia de conexión(JDBC, utilidades de bd)
- Corrige la sintaxis o el protocolo de la tecnología asociada
- Controla la creación y la eliminación de tablas, vistas y triggers temporeros
- Estrategias de procesamiento y transformación
- Opciones de movimiento de datos (Crear tabla destino, insertar/borrar, actualizar, etc.)
- Administración de transacciones (commit/rollback), dependiendo de la tecnología

Knowledge Modules

Hot-Pluggable: Modular, Flexible, Extensible



Pluggable Knowledge Modules Architecture

Benefits

- Tailor to existing best practices
- Ease administration work
- Reduce cost of ownership

Diferentes tipos de Módulos de Conocimiento

- **RKM (Reverse Knowledge Modules)** are used to perform a customized reverse-engineering of data models for a specific technology.
- **LKM (Loading Knowledge Modules)** are used to extract data from the source database tables and other systems (files, middleware, mainframe, etc.).
- **JKM (Journalizing Knowledge Modules)** are used to create a journal of data modifications (insert, update and delete) of the source databases to keep track of the changes.
- **IKM (Integration Knowledge Modules)** are used to integrate (load) data to the target tables.
- **CKM (Check Knowledge Modules)** are used to check that constraints on the sources and targets are not violated.
- **SKM (Service Knowledge Modules)** are used to generate the code required for creating data services.

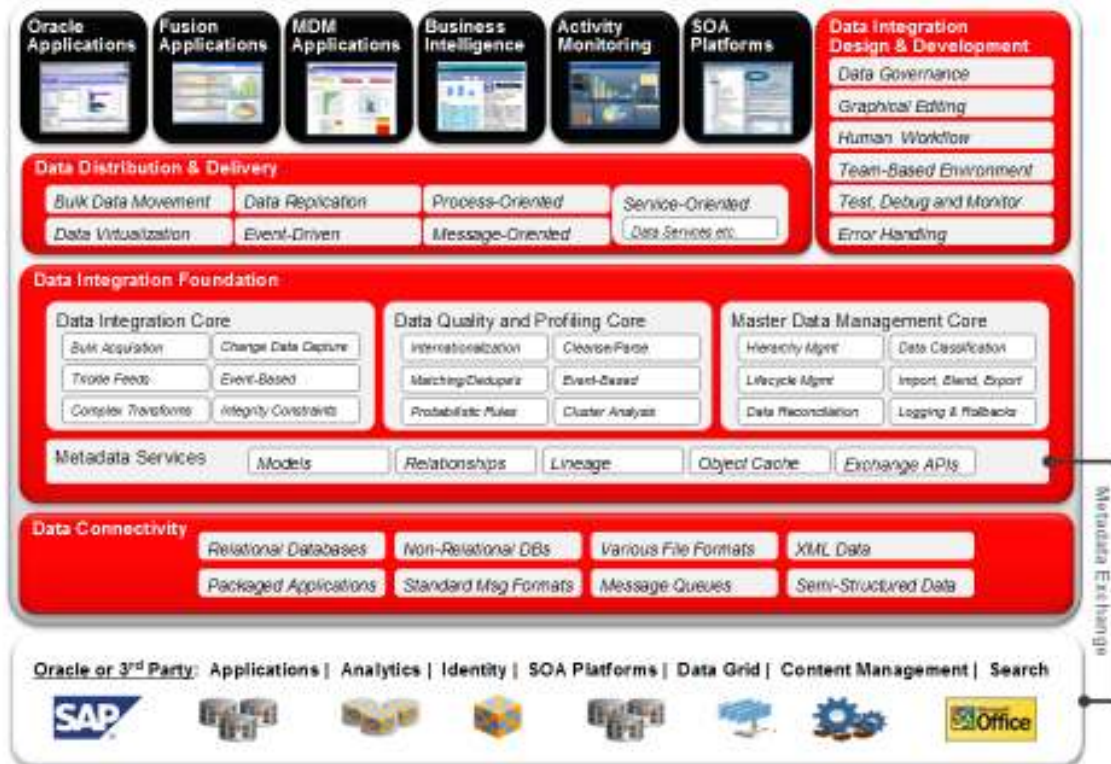
Ejemplo de los Módulos de Conocimiento

Sample of available Knowledge Modules:

- | | |
|--------------------------|------------------------------------|
| • General SQL DB | • Teradata V2R5.0 |
| • Oracle DB 9i | • Teradata V2R6.0 |
| • Oracle DB 10g | • Teradata 12 |
| • Oracle DB 10g XE | • Netezza Performance Server 2.2.1 |
| • Oracle DB 11g | • Hyperion Essbase* |
| • IBM DB2/400 | • PostgreSQL 8.1 |
| • IBM DB2/UDB | • MySQL 4.0 |
| • IBM Informix SE | • MySQL 5.0 |
| • IBM LDAP Server | • Oracle BI Suite 10g |
| • MS SQL Server 2000 | • Oracle BAM 10g |
| • MS SQL Server 2005 | • Oracle Internet Directory 9i |
| • MS SQL Server 2005 SE | • OpenLDAP 2.3 |
| • MS Office Access 2000 | • Oracle Siebel CRM 7.8* |
| • MS Office Excel 2000 | • Oracle JD Edwards* |
| • MS Active Directory | • Oracle PeopleSoft* |
| • Sybase ASA 8.x & 9.x | • Oracle E-Business Suite* |
| • Sybase ASE 11.9 & 12.5 | • Oracle AQ 10g |
| • Sybase IQ 12.x | • Oracle SOA Suite / ESB 10g |
| • Sonic MQ v7.0 | • Salesforce.com AppExchange |

Oracle Data Integration

Comprehensive Driven Data Integration Offering from Oracle



Competencia

Competitive Summary for Data Services

Feature	Oracle	IBM Info Server	Informatica	Business Objects EIM
Performance	✓	ODI Beats in transforms (ELT)	ODI Beats in transforms (ELT)	ODI Beats in transforms (ELT)
Data Services (SOA)	✓	✓	X	X
Data Movement & Delivery	✓	Process/Message Not Included	No Process/Message Integr.	X
Metadata Management Services	✓	✓	✓	✓
Master Data Management Services	✓	Hierarchy Mgmt Not Included	X	X
Data Integration Designer	✓	✓	✓	✓
Scheduling, Workflow, Error Mgmt	✓	✓	✓	✓
Data Quality & Profiling	✓	✓	✓	✓
Object Cache / Data Grid	✓	X	X	X
Packaged Adapters and Connectivity	✓	Strong in Mainframe	✓	✓

Detalles a tomar en consideración en una prueba de concepto de ODI

- Identificar la tabla destino
- Identificar las tablas fuentes
- Identificar las tablas de Referencia (Lookup)
- Seleccionar e importar los módulos de conocimiento para la extracción
- Identificar la estrategia de Integración
- Verificar los pareos de campos(mapping)
 - Pareos Automáticos
 - Columnas no nulas
 - Añadir columnas adicionales
- Probar regularmente la extracción
- En las transformaciones
 - Identificar, verificar y validar las condiciones
 - Verificar y validar campos y funciones para convertir formatos de fecha
 - Verificar tamaños de columnas para no truncar los datos extraídos o que de algún tipo de error
 - Verificar los tipos de datos(Datatype)
 - Verificar las secuencias

Conclusión

Oracle Data Integrator provee una plataforma de integración con capacidad de alto desempeño y productividad el cual provee un alto grado de flexibilidad y modularidad. El Oracle Data Integrator cumple con todas aquellas necesidades asociadas a la integración de datos incluyendo data Warehouse e inteligencia de negocios, integración de procesos, migraciones y todas aquellas iniciativas donde se requieran los datos correctos, en el lugar correcto en el momento correcto. El Oracle Data Integrator pudiera ser la plataforma de apoyo en la estrategia de consolidación de datos para su empresa.



Oracle Data Integrator
Febrero 2009
Author: Angel Rios

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 © 2008, Oracle Corporation and/or its affiliates. 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 is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.