

ORACLE DATABASE 11G ORACLE WAREHOUSE BUILDER DATA QUALITY OPTION

CARACTERÍSTICAS Y BENEFICIOS

Oracle Warehouse Builder 11g Release 1 es una completa herramienta para todos los aspectos de administración de datos y metadatos. Esta brinda características para asegurar la calidad de datos, que integran el modelado relacional y multidimensional y que permiten la administración de todo el ciclo de vida de datos y metadatos.

Data Quality Option

La opción *Data Quality* para Warehouse Builder es una de las opciones que pueden ser adquiridas con Oracle Warehouse Builder como parte de la edición Enterprise del motor de base de datos de Oracle.

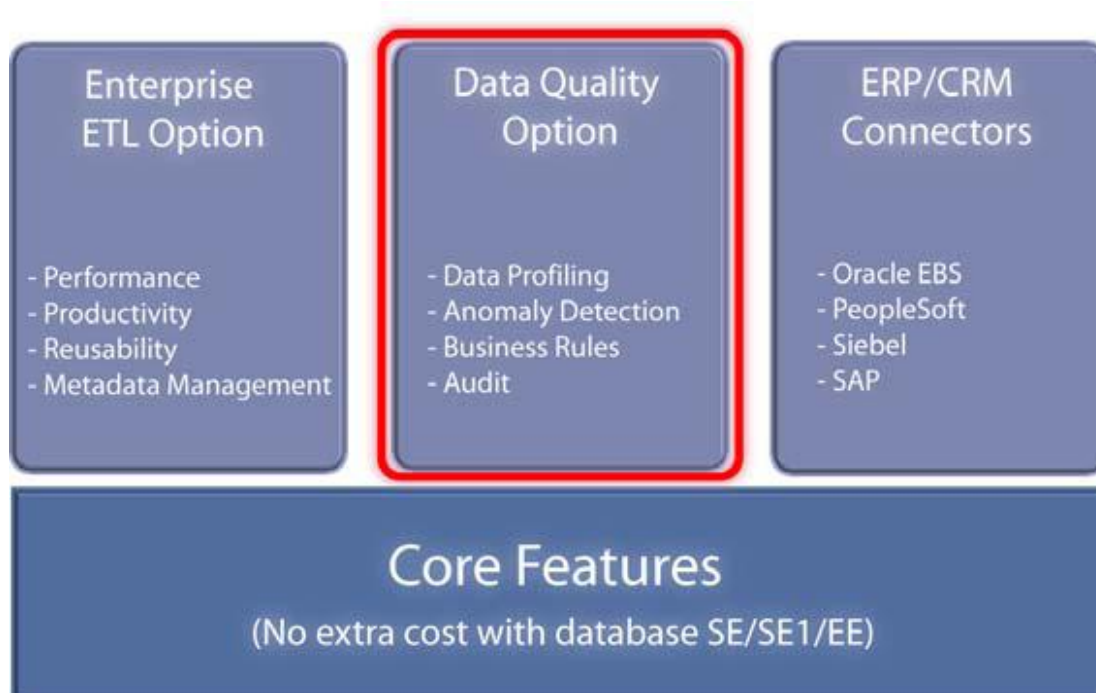


Figura 1 Opciones de Warehouse Builder

Este resumen ejecutivo cubre únicamente la opción Data Quality que se muestra en la figura 1.

Perforación de Datos y Detección de Anomalías

Una de las características más importantes en cuanto a la manipulación y administración que posee Warehouse Builder es la capacidad de evaluar la calidad de los datos de un sistema u organización. Esta característica se conoce comúnmente como Perfilamiento de Datos (Data Profiling). El componente de alto nivel para perfilamiento de datos de Warehouse Builder se muestra en la figura 2.

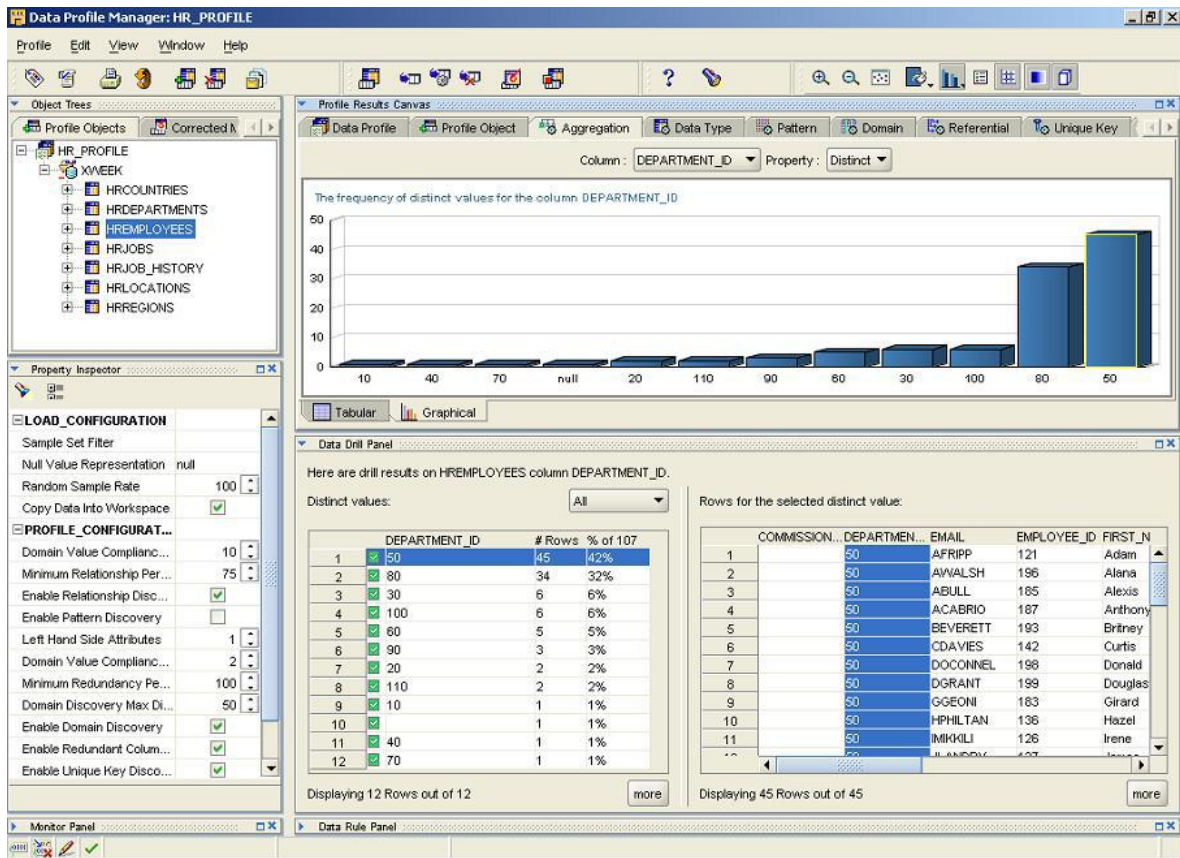


Figura 2. Perfilamiento de Datos en Warehouse Builder

Utilizando los resultados de la perfilamiento, Warehouse Builder le permite generar reglas de limpieza de datos, que usted puede aplicar a un conjunto de datos mediante rutinas de extracción, transformación o carga (ETL), llamadas *Correction Mappings*. Esta integración en la creación de los procesos ETL constituye el principal diferenciador de Warehouse Builder frente a las herramientas clásicas de perfilamiento de datos. Adicionalmente, esta integración permite que la obtención y publicación de información de alta calidad se efectúe de manera inmediata de acuerdo a los requerimientos del negocio.

Reglase de Negocio

Con el correcto conjunto de herramientas y la correcta aproximación a la integración de datos, los desarrolladores pueden de manera ágil producir datos integrados de alta calidad. Para garantizar que el sistema sea flexible y evitar grandes y complejos bloques de código, los desarrolladores pueden hacer uso de los metadatos para encapsular reglas de negocio de manera más efectiva. Warehouse Builder posee además una gran cantidad de reglas de negocio y lógica pre configurada que se puede implementar dentro de los procesos ETL, brindando mayor agilidad y flexibilidad en la etapa de desarrollo.

Otra vía para aprovechar las características de perfilamiento de datos es definir reglas asociadas a objetos de datos (por ejemplo a tablas). Esto permitirá manipular los errores y validaciones basados en una sola configuración. De esta manera el mantenimiento y actualización de estas reglas se puede hacer de una manera mucho más fácil.

Auditoria

Para asegurar que sus objetivos de calidad de datos no caen abajo de los objetivos corporativos de su empresa, Usted puede crear un Auditor de Datos, que continuamente monitorea la calidad de los datos durante las cargas.

Si el índice de calidad es menor a los umbrales especificados, el auditor de datos le enviará una alerta y usted podrá tomar las acciones apropiadas. Warehouse Builder a la vez almacena esta información en tablas de registro de ejecuciones, así usted puede revisar el progreso de carga y validación, y extraer tendencias relacionadas a las múltiples cargas a través del tiempo.

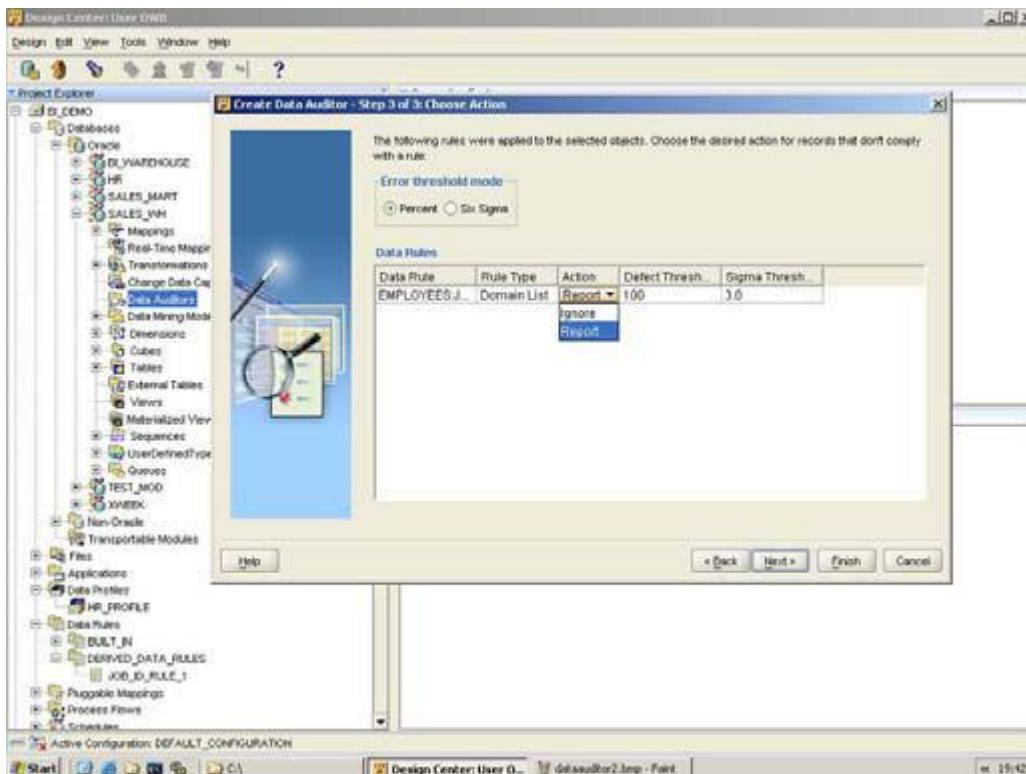


Figure 3 Creando un Auditor de Datos

Integrando Calidad de Información con ETL

¿Imagine que usted puede genera un esquema complete, basado en cómo sus datos se ven, después generar el mapeo de datos y el nuevo esquema, y finalmente corregir los errores en los datos, mientras estos fluyen hacia el esquema destino? Esto es realidad ahora con los módulos de corrección de Warehouse Builder.

En pocas palabras, así es como funciona: Basándose en los resultados del perfilamiento de datos, se pueden derivar y diseñar reglas que definan cómo se desean ver los datos. Después se seleccionan las correcciones (Correction Mappings) y los esquemas destino a los que se desean aplicar. Cuando se generan las correcciones, es posible escoger que métodos de limpieza se quieren aplicar para asegurar así que las reglas se apliquen correctamente durante el movimiento de datos dentro del esquema destino.

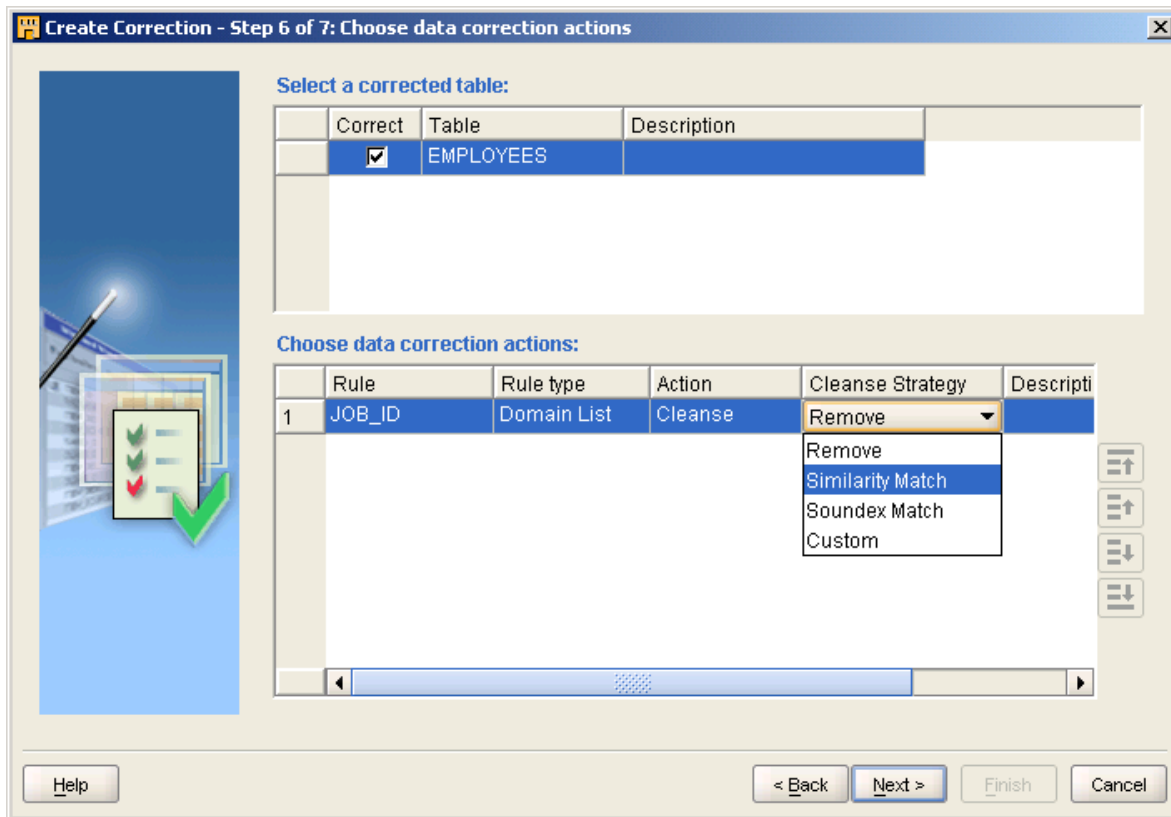


Figura 4 Limpiando Datos

Fuente Original:
<http://www.oracle.com/technology/products/warehouse/11gr1/datasheets/warehouse-builder-11g-data-quality-datasheet.pdf>