

Resumen de Data Warehouse

Data Warehouse: en palabras sencillas podemos decir que es un almacén muy grande de datos, el cual utilizan las empresas para tomar sus decisiones financieras.

Algunas de las características de un DW: integrados, no volátiles, actuales e históricos y organizados

Integrados:

Los datos almacenados en el Data Warehouse deben integrarse en una estructura consistente en la cual todos posean una relación lógica. La información suele estructurarse en diferentes niveles para adecuarse a las distintas necesidades de los usuarios.

No volátiles:

El almacén de información de un Data Warehouse existe para ser leído, no modificado.

La información es por tanto permanente, esto quiere decir que un Data Warehouse está abierto a recibir nueva información, pero no eliminar u modificar información previamente existente.

Actuales e Históricos:

El tiempo es parte implícita de la información contenida en un Data Warehouse. En los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente.

Esto es un recurso crucial en los Data Warehouse, ya que funcionan para realizar análisis de tendencias, permitiendo que se puedan plantear múltiples comparativas bajo esta misma variable.

Organizados:

A la hora de plantear un Data Warehouse es necesario tener en claro los datos que se utilizarán, los datos son necesarios para el proceso de generación del conocimiento del negocio se integran desde el entorno operacional. Estos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales.

Otra característica del Data Warehouse es que contiene datos relativos a los datos, concepto que se ha venido asociando al término de metadatos.

Los metadatos permiten mantener información de la procedencia de la información, la periodicidad de refresco, su fiabilidad, forma de cálculo, etc., relativa a los datos de nuestro almacén.

Estos metadatos serán los que permitan simplificar y automatizar la obtención de la información desde los sistemas operacionales a los sistemas informacionales.

Soportar al usuario final, ayudándole a acceder al Data Warehouse con su propio lenguaje de negocio, indicando qué información hay y qué significado tiene. Ayudar a construir consultas, informes y análisis, mediante herramientas de navegación

Procesos de un Data Warehouse

Extracción: es la obtención de información desde las distintas fuentes tanto internas como externas.

Elaboración: abarca tareas como filtrado, limpieza, depuración, homogeneización y agrupación de la información.

Carga: organización y actualización de los datos y los metadatos en la base de datos de forma ordenada y clara según se planteara previamente.

Explotación: extracción y análisis de la información en los distintos niveles de agrupación

Desde el punto de vista del usuario, el único proceso visible es la explotación del almacén de datos, aunque el éxito del Data Warehouse radica en los tres procesos iniciales que alimentan la información de este y suponen el mayor porcentaje de esfuerzo a la hora de desarrollar el almacén

Sistema Tradicional vs Data Warehouse

Sistema Tradicional

Importante del tiempo de respuesta de la transacción instantánea

Estructura relacional

Usuarios de perfiles medios o bajos

Explotación de la información relacionada con la operativa de cada aplicación

Data Warehouse

Importancia de la respuesta masiva

Visión multi dimensional

Usuarios de perfiles altos

Explotación de toda la información interna y externa relacionada con el negocio

Beneficios de utilizar Data Warehouse

Proporciona una herramienta para la toma de decisiones en cualquier área funcional, basándose en información integrada y global del negocio.

Facilita la aplicación de técnicas estadísticas de análisis y modelización para encontrar relaciones ocultas entre los datos del almacén; obteniendo un valor añadido para el negocio de dicha información.

Proporciona la capacidad de aprender de los datos del pasado y de predecir situaciones futuras en diversos escenarios.

Simplifica dentro de la empresa la implantación de sistemas de gestión integral de la relación con el cliente.

Supone una optimización tecnológica y económica en entornos de Centro de Información, estadística o de generación de informes con retornos de la inversión espectaculares

Data Mart

Otro tema a tener en cuenta son los Data Mart, los cuales, generalizando son o cumplen las funciones de un Data Warehouse a menor escala. Como los DW demandan mucho tiempo y recursos el desarrollarlos y mantenerlos en funcionamiento, los Data Mart sirven como una opción rápida y considerablemente menos costosa, para ya después dar el paso a un DW. Cabe aclarar que un Data Mart posee las mismas características de integración, no volatilidad, orientación temática e historicidad que tienen los Data Warehouse.

Aunque los Data Mart también tienen otras funciones prácticas, estos son particularmente apropiados cuando el Data Warehouse central crece muy rápidamente y los distintos departamentos requieren sólo una pequeña porción de los datos contenidos en él. En estos casos se puede aplicar una “Estructura descentralizada”.

Conclusión:

Los Data Warehouse son una herramienta de negocio, que no solamente se utiliza para almacenar datos sino que también para analizar los mismos, y ayudar a tomar decisiones mas certeras en las empresas.