Glosario de Términos

ADMINISTRADOR DE BASE DE DATOS O DBA

Experto técnico competente para uno (o varios) gestores de datos, capaz de escribir o de optimizar programas de extracción, de carga o de acceso en el lenguaje del motor de datos.

ADMINISTRADOR DE DATOS

Experto en el negocio que conoce la semántica de los datos de la empresa y se encarga del referencias de datos. Por ello, es capaz de arbitrar los conflictos inherentes a la constitución de definiciones únicas de los objetos de negocio de la empresa.

ANÁLISIS

Operación consistente en estudiar y descomponer los datos a fin de extraer de ellos los elementos esenciales y obtener un esquema de conjunto.

BASE DE DATOS (DATA BASE)

Conjunto de datos no redundantes, almacenados en un soporte informático, organizados de forma independiente de su utilización y accesibles simultáneamente por distintos usuarios y aplicaciones. La diferencia de una BD respecto a otro sistema de almacenamiento de datos es que éstos se almacenan en la BD de forma que cumplen tres requisitos básicos: no redundancia, independencia y concurrencia.

CATÁLOGOS

En ciertas herramientas clientes del Data Warehouse, estructura que permite al usuario trabajar sobre una vista lógica y orientada al negocio de los datos que desea ver.

CLIENTE/SERVIDOR

Arquitectura de sistemas de información en la que los procesos de una aplicación se dividen en componentes que se pueden ejecutar en máquinas diferentes. Modo de funcionamiento de una aplicación en la que se diferencian dos tipos de procesos y su soporte se asigna a plataformas diferentes.

CODIFICACION

a) Transformación de un mensaje en forma codificada, es decir, especificación para la asignación unívoca de los caracteres de un repertorio (alfabeto, juego de caracteres) a los de otro repertorio.

b) Conversión de un valor analógico en una señal digital según un código prefijado.

DATA WAREHOUSE

«Almacén de datos». Base de datos específica del mundo de la decisión destinada principalmente a analizar las palancas de negocio potenciales. Un Data Warehouse es (fuente.- Bill Inmon):

- integrado;
- orientado al tema;
- y contiene datos no volátiles.

DETECCION DE DESVIACION

Normalmente, para la detección de desviación en bases de datos grandes se usa la información explícita externa a los datos, así como las limitaciones de integridad o modelos predefinidos. En un método lineal por contraste, se enfoca el problema desde el interior de los datos, usando la redundancia implícita de los datos. Aquí se simula un mecanismo familiar a los seres humanos: después de ver una serie de datos similares, un elemento que perturba la serie se considera una excepción.

DICCIONARIO DE DATOS

Descripción lógica de los datos para el usuario. Reúne la información sobre los datos almacenados en la BD (descripciones, significado, estructuras, consideraciones de seguridad, edición y uso de las aplicaciones, etc.).

DIMENSIÓN

Eje de análisis asociado a los indicadores; corresponde normalmente a los temas de interés del Data Warehouse. ejemplo: dimensión temporal, dimensión cliente...

DIRECTORIO DE DATOS

Es un subsistema del sistema de gestión de base de datos que describe dónde y cómo se almacenan los datos en la BD (modo de acceso y características físicas de los mismos).

INCONSISTENCIA

El contenido de una base de datos es inconsistente si dos datos que deberían ser iguales no lo son. Por ejemplo, un empleado aparece en una tabla como activo y en otra como jubilado.

INTEGRIDAD

Condición de seguridad que garantiza que la información es modificada, incluyendo su creación y borrado, sólo por el personal autorizado.

INTERNET

Término usado para referirse a la red más grande del mundo, que conecta miles de redes con alcance mundial. Está creando una cultura que basándose en la simplicidad, investigación y estandarización fundamentado en usos de la vida real, está cambiando la forma de ver y hacer muchas de las tareas actuales. Mucha de la tecnología de punta en redes está proviniendo de la comunidad Internet.

INTRANET

Constituye un servicio de comunicación de los sistemas de información corporativos orientados a su personal, sobre el formato de los sistemas Web, operando sobre la red Internet. Ejemplo: El sistema contable de una empresa de ventas de productos de ferretería, tipo Home Center.

LIBRO ELECTRONICO

Guía electrónica. Documento realizado en un sistema informático, normalmente con características hipertexto y multimedia.

METADATO

Información que describe un dato. En un contexto de Data Warehouse, cualifica un dato precisando por ejemplo su semántica, las reglas de gestión asociadas, su fuente, etc.

MOLAP (MULTIDIMENSIONAL ON LINE ANALYTICAL PROCESSING)

Cf. OLAP

MPP (MASSIVE PARALLEL PROCESSING)

Caracteriza una Base de Datos dedicada a la decisión almacenando los datos en forma de tablamultidimensional. Estos SGBD son una alternativa a los SGBD relacionales.

MULTIDIMENSIONAL (SGBD)

Arquitectura de Hardware que hace colaborar a varios procesadores (hasta centenares) contando cada uno con su propia memoria.

OLAP (ON LINE ANALYTICAL PROCESSING)

Caracteriza la arquitectura necesaria para implementar un Sistema de Información de Decisión. Se opone a OLTP. El término OLAP designa a menudo las herramientas de análisis basadas en BD multidimensionales. Se habla también de herramientas MOLAP por oposición a las herramientas ROLAP.

OLTP (ON LINE TRANSACTIONNEL PROCESSING)

Tipo de entorno de tratamiento de la información en el que debe de darse una respuesta en un tiempo aceptable y consistente.

REDUNDANCIA

Repetición de los mismos datos en varios lugares.

REPOSITORIO

Base de datos central en herramientas de ayuda al desarrollo. El repositorio amplía el concepto de diccionario de datos para incluir toda la información que se va generando a lo largo del ciclo de vida del sistema, como por ejemplo: componentes de análisis y diseño (diagramas de flujo de datos, diagramas entidad-relación, esquemas de bases de datos, diseños de pantallas, etc.), estructuras de programas, algoritmos, etc. En algunas referencias se le denomina Diccionario de recursos de información.

ROLAP (RELATIONAL ON LINE ANALYTICAL PROCESSING)

Caracteriza la arquitectura necesaria para implementar un Sistema multidimensional basado en tecnologías relacionales.

SISTEMA DE GESTION DE BASE DE DATOS

Software que controla la organización, almacenamiento, recuperación, seguridad e integridad de los datos en una base de datos. Acepta pedidos de datos desde un programa de aplicación y le ordena al sistema operativo transferir los datos apropiados.

Cuando se usa un sistema de gestión de base de datos, SGDB, (en inglés DBMS), los sistemas de información pueden ser cambiados más fácilmente a medida que cambien los requerimientos de la organización. Nuevas categorías de datos pueden agregarse a la base de datos sin dañar el sistema existente.

SISTEMA DE INFORMACION (SI)

Conjunto de elementos físicos, lógicos, de comunicación, datos y personal que, interrelacionados, permiten el almacenamiento, transmisión y proceso de la información.

SMP (SYMETRIC MULTIPROCESSING)

Arquitectura de Hardware que hace colaborar varios procesadores (unas docenas) en una memoria compartida.

SQL (Structured Query Language)

Lenguaje de interrogación normalizado para bases de datos relacionales. El SQL es un lenguaje de alto nivel, no procedural, normalizado, que permite la consulta y actualización de los datos de BD relacionales. Se ha convertido en el estándar para acceder a BD relacionales. La primera versión se aprobó como norma ISO en 1987 y la segunda, conocida como SQL2 y vigente actualmente, en 1992.

Actualmente se trabaja en la norma SQL3 que soportará bases de datos orientadas a objeto y bases de datos activas. El SQL facilita un lenguaje de definición de datos y un lenguaje de manipulación de datos. Además, incluye una interfase que permite el acceso y manipulación de la BD a usuarios finales.

SPONSOR O PATROCINADOR

Individuo o grupo de individuos cuya función es obtener la adhesión y la implicación de todos los actores afectados por la implementación del Data Warehouse. Debe promover el proyecto y garantizar la sinergia entre los usuarios y los equipos informáticos. También debe gestionar los eventuales problemas políticos que la implementación de estos sistemas puede generar.

UNIX

Sistema operativo multiproceso, multiprograma y multiusuario. Software diseñado por AT&T para ingeniería de telecomunicación. Ha sido el primer sistema operativo concebido con independencia de los fabricantes. Posee una

gran facilidad para adaptarse a ordenadores con diferentes arquitecturas, siendo ampliamente autónomo respecto del hardware. Está escrito en lenguaje de alto nivel C.

Bibliografía

Libros

IDC - IDC Study on the Financial Impact of a Data Warehouse. Estudio respaldado por Andersen Consulting, Digital, Dun & Bradstreet Software & Pilot Software, EMC, Hewlett Packard, IBM, InfoiTnix, KPMG Peat Marwick, Mierosoft, NCR, Oracie, Prism, Pyramid, SAP, SAS, SGI, Software AG, Sun, Sybase, Unisys. Difundido por sus sponsors y disponible en Internet en la dirección:
http://direct.boulder.ibm.com/dss/solution/idc.html

EDS - El Data Warehouse El Data Mining, Septiembre 1997

EDS - Data Warehouse Primer, Enero 1995.

ETI - Anticipating the Cost of Maintenance, Julio 1995.

Red Brick - A Primer Information Technology and Business Management, Junio 1995.

Referencias

[DWHInsta] The Data Warehousing Institute Select Survey Results 1995-1996. Estudio basado en 6.214 respuestas.

[DWHInstb] Alan Paller.- Ten Mistakes to Avoidfor Data Warehousing Managers, Data Warehouse Institute.

[LOV96] Bruce Love.- S-trategie DSS Data Warehouse: A Case Study in Failure, Journal of Data Warehousing, julio de 1996.

[SKE96] Michael Haisten.- A History of Access and Analysis Tools, Journal ofdata Warehousing, julio de 1996.

[Inmon94] W.H. Inmon & Richard D. Hackathorn.- Using the Data Warehouse, Wiley-QED Publication, 1994.

[Mar91] James Martin - Rapid Application Development - Prentice Hall, 1991.

Internet

http://www.dw-institute.com/ - El Data Warehouse Institute es una asociación temática especializada en el Data Warehouse. Recoge más de 2.000 adheridos en 40 países.

http://www.inei.gob.pe/cpi-mapa/bancopub/libfree/lib619/INDEX.HTM
Documento HTML con información de Data Warehouse

http://www.intelligententerprise.com/ - Intelligent enterprise es la un sitio especializado en temas informáticos, el Data Warehouse, ERP, etc.

http://www.prometheus.eds.fr/ - El servicio en línea del Instituto Promethéus de EDS, Este servicio propone la base de direcciones INFOmédiaire, que contiene cerca de mil referencias a documentos o sedes WEB. Contiene también extractos de estudios, artículos, etc.

http://www.dwinfocenter.org/ - El Data Warehousing Information Center es un centro de información dedicado al Data Warehouse