

Profesora: Olga Cuervo Miguélez

TEMA 1

SISTEMAS DE ALMACENAMIENTO DE LA INFORMACIÓN

Módulo: **Bases de Datos**

Ciclo: **DAM**



En este tema aprenderás lo que es una base de datos y que funciones y componentes tiene, valorando su utilidad. Se trata de la primera unidad de trabajo del módulo y con ella se pretende que conozcas los sistemas de almacenamiento de la información que se han venido utilizando hasta ahora y sus inconvenientes, para que puedas comprender mejor la importancia de los sistemas de bases de datos actuales.

1. Introducción

Supongamos un empresario que trata de operar en su empresa sin saber quiénes son sus clientes, qué productos vende, quién trabaja para él, quién le debe y a quien le debe dinero, toda esta información y mucha más debe estar disponible cuando se necesite.

El propósito de los **sistemas de información** es ayudar a usar la información como un recurso organizacional lo implica capturar, almacenar, agregar, modificar y administrar los datos.

Según el tipo de sistema de información y las características del negocio, estos datos pueden variar desde unos pocos de megabytes hasta miles de terabytes. Podemos considerar el caso clásico de Google, se estima que la compañía responde a más de 63.000 consultas de búsquedas al segundo lo que equivale a unos 5.600 millones de búsquedas por día unos 2 billones de búsquedas al año en un conjunto de datos de varios terabytes. Es impresionante que hablando de estas cifras se pueda disponer de los resultados casi de forma instantánea.

¿Cómo puede esta compañía procesar esa enorme cantidad de datos? ¿Cómo puede guardar y después recuperar de forma casi instantánea?. La respuesta es que usan bases de datos. Las bases de datos son estructuras especializadas que permiten a los sistemas guardar, manejar y recuperar datos con gran rapidez.

2. Sistemas lógicos de almacenamiento de la información

Los términos **datos** e **información** a menudo se usan indistintamente. Un dato puede ser un número, un símbolo, un carácter, una palabra que, si no se pone en un contexto dado, no significa nada para los humanos.

La información son datos puestos en contexto. La información son datos organizados y procesados que se presentan en un contexto dado y son útiles para los humanos. La información depende de los datos.



Los valores históricos de temperaturas en un lugar determinado en los últimos 100 años son datos. Si los datos se organizan y analizan para descubrir que la temperatura global va aumentar, tenemos información.

Formulario de contacto

Nombre y Apellido (*)

E-mail (*)

Género
☒ Hombre ☐ Mujer

Intereses
☐ Deporte ☐ Política ☐ Ciencia

Comentarios

Campos requeridos (*)

Enviar datos Borrar campos

Alumnos - Netscape

Archivo Edición Ver Ir Comunicador Ayuda

file:///D:/Mis documentos/Alumnos_1.html

Trabajo Coches Windows 95 Hobbies Internet. HTML Personales

Alumnos

DNI	NOMBRE	TELEFONO	DIRECCION	CIUDAD	FECH_NA
52302456	Berenguer Iglesias, Juana Mª	953134	Dpto. Biología Animal	Alicante	3/5/190
52302457	Berenguer Navarro, Vicente	213359	Avda. Madrid, 57	Alicante	3/5/190
52302458	Berenguer Ros, María Dolores	396755	Dpto. CC. Medioambientales	Alicante	4/5/190
52302459	Bermejo Garcia, Rosario	309210	Valencia, 314, 4º. 1a	Alicante	5/5/190
52302460	Bermudez Aznar, Agustín	342277	Dpto. CC. Agroforestales. Av. Madrid, 57	Alicante	6/5/190
52302461	Bermudez Bellido, María Dolores	552322	Cdad. Univ. Meléndez. A.A. 25360	Alicante	7/5/190
52302462	Berna Galiano, Jose Angel	337339	San Roque, 16	Alicante	8/5/190
52302463	Bernabe Espinosa, María Josefa	944483	Km. 2.5, otra. Antigua a Coatepec	Alicante	9/5/190
52302464	Bernabe Gil, David	664325	Luis Fontes Pagán, s/n	Alicante	10/5/190
52302465	Bernabe Gil, David	743117	Tembleque, 16, 3º. D	Alicante	11/5/190
52302466	Bernabe Pons, Luis Fernando	224536	Paseo del Bosque, s/n	Alicante	12/5/190

Formulario

Datos sin refinar

ESTADO	TOTAL
SIN EMPEZAR	22
SIGUIENTE	6
INICIADO	3
CON AVANCE	1
TERMINADO	182
CANCELADO	5
TOTAL TAREAS	223

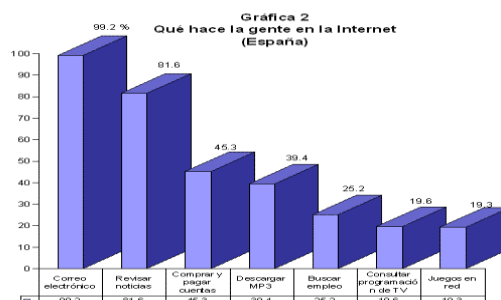
VENCIMIENTO	TOTAL
RETRASO	0
HOY	1
MAÑANA	1
Faltan 2 días	1
Faltan 3 días	1
Faltan 4 días	0
Faltan 5 días	4
Faltan 6 días	0
Faltan 7 días	0
Sin fecha	21
TOTAL	32

Prioridad (Importancia y Urgencia)	MUY URGENTE	URGENTE	POCO URGENTE	TOTAL
MUY IMPORTANTE	0	3	0	3
IMPORTANTE	0	9	8	17
POCO IMPORTANTE	0	0	12	12
TOTAL	0	12	20	32

Vencimiento y Estado	SIN EMPEZAR	SIGUIENTE	INICIADO	CON AVANCE	TOTAL
RETRASO					1
HOY	1				1
MAÑANA			1		1
Faltan 2 días		1			1
Faltan 3 días				1	1
Faltan 4 días					
Faltan 5 días		3	1		4
Faltan 6 días					
Faltan 7 días					
Sin fecha	21	2	1		24
TOTAL	22	6	3	1	32

Contexto y Estado	SIN EMPEZAR	SIGUIENTE	INICIADO	CON AVANCE	TOTAL
Oficina	16	3	1		20
Casa	1	1	1		3
Personal	5	2	1	1	9
Clientes					
Recados					
Teléfono					
TOTAL	22	6	3	1	32

Información en formato resumen



Información en formato de gráfica

2.1. Sistemas de información

2.1.1. Introducción

En la actualidad, los Sistemas de Información juegan un papel cada vez más importante en organizaciones empresariales, hasta el punto de condicionar su éxito o fracaso en un entorno económico y social tan dinámico y convulso como el que caracteriza al mundo actual.

Nuevos fenómenos como la globalización o la transición hacia una economía más basada en el conocimiento han inducido cambios importantes en las organizaciones empresariales. En este nuevo contexto, los Sistemas y Tecnologías de la Información y la Comunicación (TIC) se han convertido en un elemento esencial como motor de cambio y fuente de ventajas competitivas.

La información y las nuevas tecnologías que la soportan aparecen como un nuevo factor productivo que se suma a los factores tradicionales, trabajo y capital, diferenciándose de estos últimos por su carácter intangible, dando lugar a un nuevo escenario en el que la información y el conocimiento son básicos para las empresas. Los Sistemas y Tecnologías de la Información se convierten, por tanto, en una pieza fundamental para impulsar este cambio y este nuevo escenario al que nos referimos.

Dentro de una organización, el Sistema de Información actúa como el "sistema nervioso" ya que este es el encargado de hacer llegar a tiempo la información que necesitan los diferentes elementos de la organización empresarial (departamentos, áreas funcionales, equipos de trabajo, delegaciones, etc.), permitiendo así una actuación conjunta y coordinada, ágil y orientada a la consecución de los objetivos de la organización.

Los Sistemas de Información han adquirido una dimensión estratégica en las empresas del nuevo milenio que dejaron de ser considerados una simple herramienta para automatizar procesos operativos, para convertirse en una pieza clave a tener en cuenta a la hora de formular la estrategia de negocio para llevar a cabo su implementación y llevar a cabo el control de administración.

2.1.2. Elementos de un sistema de información

En general, podemos definir un **sistema** como un *"conjunto de elementos que interactúan entre sí, orientados hacia el logro de un fin común"*.

Cualquier empresa u organización es en sí misma un sistema. Su entorno es el sistema productivo en el que se inserta, del que recibe una serie de entradas, en forma de demandas, y al que entrega una serie de productos, u ofertas.

A su vez, dentro de la empresa u organización existen varios subsistemas conformados por diferentes departamentos o áreas en que se divide la empresa. Ejemplo: Sistema productivo de la provincia de Pontevedra - > Faro de Vigo -> Editorial -> Editorial deportiva.

Podemos definir un sistema de información como un *"conjunto de elementos que interactúan entre sí para recopilar, gestionar, controlar y difundir información de la empresa u organización"*.

Los elementos esenciales de un sistema de información que deben interactuar entre sí son:

- ✓ **Hardware.** Se refiere a la tecnología de almacenamiento, comunicaciones, entradas y salidas de datos.
- ✓ **Software.** Son programas destinados a la recolección, almacenamiento, procesamiento y análisis de datos, generando información para el consumidor final.
- ✓ **Datos.** Elemento fundamental del sistema.
- ✓ **Procedimientos.** Son las políticas de negocio que aplican a los procesos de la empresa.
- ✓ **Usuarios.** Son los encargados de interactuar con la información obtenida de los datos, siendo el componente esencial para que cualquier proyecto empresarial tenga éxito o fracase.



2.1.3. Estructura de un sistema de información

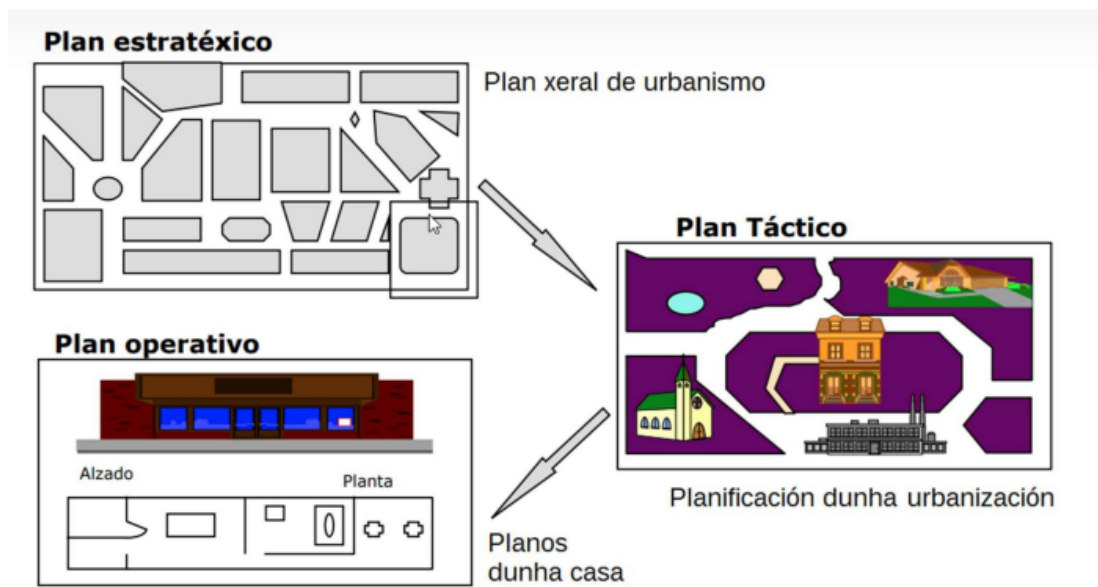
En la mayoría de las organizaciones hay tres niveles de gestión que ayudan a planificar líneas principales que moverán la actividad empresarial.

Cada nivel cumple una serie de tareas que se complementan para asegurar el éxito de toda la organización. A la estructura jerárquica según las responsabilidades en la toma de decisión se llama la pirámide organizacional.



Los niveles de la pirámide organizacional son los siguientes:

- ✓ Nivel **operativo**. Este nivel incluye el procesamiento de actividades diarias o transacciones, los eventos rutinarios que afectan a la organización: facturación, pagos, etc. Se analizan los resultados de la organización, fundamentalmente con respecto a los recursos consumidos en las transacciones para la toma de decisiones de corto plazo y de consecuencias limitadas.
- ✓ Nivel **táctico**. Este nivel se ocupa de la asignación efectiva de recursos a medio plazo para mejorar el desempeño de la empresa.
- ✓ Nivel **estratégico**. Se trata de la visión que impulsa las acciones de la empresa. Establece los objetivos a cumplir y las pautas para alcanzarlos. La dirección juega un papel principal a la hora de definir la estrategia, por lo que debe actuar con extrema precisión para que toda la organización entienda su visión. En este nivel, la información se utiliza para decidir las directrices a largo plazo de la organización.



2.1.4. Clasificación de los sistemas de información

El sistema de información existe en todas las empresas. Hasta la aparición de las computadoras el sistema de información trabajado de forma manual. Cuando las computadoras participan en el funcionamiento del sistema de información, hablamos de un sistema de información automatizado.

En general, las clasificaciones más extendidas de los Sistemas de Información se agrupan estos dependiendo de su propósito.

De forma muy global se puede considerar que existen dos funciones básicas para los sistemas:

- ✓ **Soporte a las actividades operativas**, que da origen a los sistemas de información para agilizar las actividades de la empresa, tanto las más estructuradas (aplicaciones de contabilidad, nóminas, pedidos y, en general, lo que se denomina "gestión negocio"), así como sistemas que permitan la gestión de la información menos estructurada: aplicaciones ofimáticas, programas técnicos para funciones de ingeniería, etc.
- ✓ **Soporte a las decisiones y control de gestión**, que puede ser proporcionado desde las propias aplicaciones de gestión empresarial (a través de salidas de información existentes) o a través de aplicaciones específicas.

El soporte de las decisiones requiere consultar los datos de una manera menos predecible que las aplicaciones de gestión. Por esta razón, la tecnología de bases de datos es de gran importancia para hacer posible este tipo de aplicaciones. De hecho, a menudo se considera que la tecnología de Bases de Datos permitió la aparición del llamado Sistema de Información para la Gestión (Sistema de Información Gerencial - MIS, por sus siglas en inglés), que permite abordar problemas no estructurados.

Los MIS utilizan los datos almacenados en los sistemas informáticos de la empresa para generar informes que permitan a los gerentes mejorar el control de gestión de las diferentes áreas funcionales de la empresa. De esta manera, es posible acelerar la proceso de toma de decisiones, proporcionando la información necesaria de forma rápida, precisa y fiable.

La dirección necesita herramientas para diagnosticar problemas (análisis) y elegir la mejor alternativa (simulación, planificación...). Este tipo de herramientas requieren el manejo de información histórica y no estuvieron disponibles hasta principios de los noventa, dando lugar a un conjunto de aplicaciones cubiertas por términos como "aplicaciones de apoyo a la toma de decisiones" (DSS), "software de apoyo a la gestión" (EIS, ESS), "Datawarehousing and Datamining Systems" o, más genéricamente, Sistemas de "Inteligencia" de Negocios" (Business Intelligence).

2.1.5. Sistemas de datawarehousing

Los avances en las Tecnologías de la Información han hecho posible el desarrollo de Sistemas de Datawarehousing/Datamining, que constituyen el núcleo de las aplicaciones de Business Intelligence (Inteligencia de Negocio').

Estos sistemas constan de tres elementos principales:

- ✓ **Recopilación y gestión de grandes volúmenes de datos:** Tecnología Datawarehousing.
- ✓ **Análisis de datos:** Tecnología OLAP y herramientas de 'Minería de Datos' (Datamining)'.- ✓ **Software de consulta amigable e intuitivo**, accesible al usuario final.

Un Datawarehouse constituye el elemento fundamental de un sistema de soporte para la toma de decisiones en la empresa.

Podemos considerar un Datawarehouse como un gran almacén de datos, que integra datos de varias fuentes:

- ✓ **Datos de las diferentes aplicaciones de gestión de la empresa** (distribuidos por diferentes departamentos: administración, marketing, producción, etc.).
- ✓ **Datos de fuentes externas.**

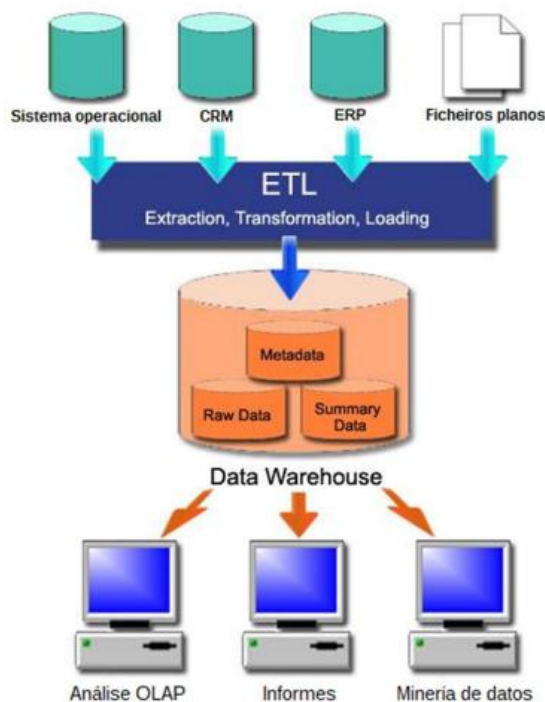
Las características de los datawarehouse incluyen:

- ✓ Son de **sólo lectura**, los datos almacenados en ellas no se borran ni modifican.
- ✓ **Su diseño está orientado a facilitar la lectura y análisis de los datos** (OLAP, análisis multidimensional). Esto hace que su estructura no está normalizada y admita redundancias.
- ✓ La estructura de los almacenes de datos distingue entre "hechos" y "dimensiones". Los hechos vendrían a ser los datos en si que queremos analizar: ventas, pedidos de clientes, etc. Las dimensiones serían los datos que organizan esos hechos según las características relevantes en el análisis: las provincias en las que se hacen las ventas, la fecha de las ventas, el vendedor que las hizo...
- ✓ Los datos almacenados en el datawarehouse reflejan la evolución temporal de los "*hechos*".
- ✓ Algunas de las técnicas utilizadas guardan datos agregados, estos son datos que corresponden a grupos de "*hechos*". Los grupos se forman de acuerdo a los valores correspondientes a las dimensiones, por lo que cuando se combinan varios, el número de los grupos se vuelven muy grandes.
- ✓ Estas redundancias pueden hacer que el tamaño del datawarehouse crezca exponencialmente, si se almacenan datos agregados correspondientes a muchos grupos.
- ✓ Un aspecto importante en la práctica son los **sistemas y técnicas para capturar y guardar la información en el almacén de datos.**

Además, un Datawarehouse *debe tener una gran capacidad de almacenamiento*, ya que los datos pertenecen a largos periodos de tiempo (varios años económicos de empresa).

Por otro lado, en estos sistemas, el *diccionario de datos* contiene metadatos referentes a la descripción y procedencia de los datos, las transformaciones que han sufrido, la periodicidad de su actualización, etc.

En estos sistemas también se utiliza el concepto de **Data Mart** para referirse a un almacén de datos más pequeños, generalmente restringidos a un área temática o un departamento de la empresa.



Las herramientas de **datamining** (minería de datos) son técnicas avanzadas que le permiten detectar y modelar relaciones entre datos y obtener información no obvia: patrones de consumo, predicción del comportamiento del cliente, asociaciones de productos, etc.

Con los Sistemas de Datawarehousing/Datamining, los gerentes pueden tener la información necesaria en muy poco tiempo y con el mínimo esfuerzo. De esta forma, los directivos pueden dedicarse más al análisis de la información obtenida y no tanto a su búsqueda.

Con estas herramientas aplicadas al marketing se profundiza en el conocimiento del comportamiento del cliente, lo que permite personalizar la oferta y conseguir una mayor efectividad de acciones comerciales.

2.1.6. Ejemplos de aplicación de los sistemas de Datawarehousing

Los sistemas de almacenamiento de datos se pueden usar en cualquier lugar donde haya una gran cantidad de datos y deseamos ver resultados estadísticos que ayuden en la elaboración de decisiones.

Ejemplos de aplicaciones de Datawarehouses son los siguientes:

- ❖ **Sitios web de redes sociales:** sitios web de redes sociales como Facebook, Twitter, LinkedIn, etc. se basan en el análisis de grandes conjuntos de datos. Estos sitios recopilan datos relacionados con miembros, grupos, localizaciones, etc. y almacenarlos en un único repositorio central.
- ❖ **Banca:** La mayoría de los bancos utilizan los Datawarehouse para ver los patrones de gastos de los titulares de la cuenta. Utilizan esta información para ofrecerle ofertas especiales, promociones, etc.

- ❖ **Educación:** Las universidades utilizan Datawarehouse para extraer información para comprender la demografía de sus estudiantes, las variaciones en la inscripción y para la gestión de los recursos humanos.
- ❖ **Seguros:** Los almacenes de datos se utilizan principalmente para analizar patrones de datos y tendencias de los clientes.

Vídeos:

https://www.youtube.com/watch?v=ElqUy_7eYLQ

<https://www.youtube.com/watch?v=S5tKZjVA0cU>

2.1.7. Big Data

El término **Big Data** se refiere al almacenamiento de grandes cantidades de datos y a los procedimientos utilizados para encontrar patrones dentro de esos datos.

El Big Data está compuesto por cinco dimensiones que lo caracterizan, conocidas como las 5 Vs del Big Data. Estos 5Vs son:

- ❖ **Volumen.** Tradicionalmente, los datos se generaban manualmente. Ahora vienen de máquinas o dispositivos y se generan automáticamente, por lo que el volumen a analizar es enorme. Estos grandes volúmenes de datos que se producen en cada momento suponen importantes retos técnicos y analíticos para las empresas que los gestionan.
- ❖ **Velocidad.** En el entorno de Big Data, los datos se generan y almacenan a una velocidad sin precedentes. Los datos pueden volverse obsoletos rápidamente y pierden su valor cuando aparecen otros nuevos. Por lo tanto, las empresas deben reaccionar muy rápidamente para poder recogerlos, almacenarlos y procesarlos. El reto del área de tecnología es almacenar y gestionar grandes cantidades de datos que se generan continuamente. El resto de las áreas también deben trabajar a una velocidad vertiginosa para convertir esos datos en información útil antes de que pierdan valor.
- ❖ **Variedad.** El origen de los datos es muy heterogéneo. Proviene de múltiples soportes, herramientas y plataformas: cámaras, smartphones, coches, sistemas GPS, redes sociales, registros de viajes, movimientos bancarios, etc. Los datos recopilados pueden ser estructurados (más fáciles de administrar) o no estructurados (en forma de documentos, videos, mensajes y correo electrónico, etc.). En función de esta diferenciación, cada tipo de información será tratada de manera diferente a través de herramientas específicas. La esencia del Big Data reside en, posteriormente, combinar y configurar unos datos con otros. Por esta razón, aumenta el grado de complejidad en los procesos de almacenamiento y análisis de los datos.
- ❖ **Veracidad.** Esta característica de Big Data es probablemente la más desafiante. El gran volumen de datos que se genera puede hacernos dudar del grado de veracidad de todos ellos, ya que la gran variedad de datos hace que muchos de ellos estén incompletos o son incorrectos. Los datos deben ser limpiados y analizados. La incertidumbre sobre la veracidad de los datos puede generar ciertas dudas sobre su calidad y su disponibilidad en el futuro.
- ❖ **Valor.** Esta característica representa el aspecto más relevante de Big Data. El valor que generan los datos, una vez convertidos en información, puede considerarse el aspecto más importante. Con este valor, las empresas tienen la oportunidad de sacar el máximo partido a los datos para introducir mejoras en su gestión,

definir estrategias más óptimas, obtener una clara ventaja competitiva, realizar ofertas personalizadas a los clientes, incrementar la relación con el público, y mucho más.

Vídeos:

<https://www.youtube.com/watch?v=w4vsFKMO7XA>

https://www.ted.com/talks/kenneth_cukier_big_data_is_better_data?language=es

https://www.ted.com/talks/tricia_wang_the_human_insights_missing_from_big_data?language=es

ÍNDICE

1. INTRODUCCIÓN	1
2. SISTEMAS LÓGICOS DE ALMACENAMIENTO DE LA INFORMACIÓN	1
2.1. SISTEMAS DE INFORMACIÓN	2
2.1.1. INTRODUCCIÓN.....	2
2.1.2. ELEMENTOS DE UN SISTEMA DE INFORMACIÓN	3
2.1.3. ESTRUCTURA DE UN SISTEMA DE INFORMACIÓN.....	4
2.1.4. CLASIFICACIÓN DE LOS SISTEMAS DE INFORMACIÓN	5
2.1.5. SISTEMAS DE DATAWAREHOUSING	6
2.1.6. EJEMPLOS DE APLICACIÓN DE LOS SISTEMAS DE DATAWAREHOUSING.....	7
2.1.7. BIG DATA	8