

TEMA 1. INTRODUCCIÓN A LAS BASES DE DATOS.....	1
1. CUALIDADES DE LA INFORMACIÓN.....	1
2. SISTEMAS DE INFORMACIÓN.....	2
2.1. <i>Componentes de un sistema de información.....</i>	<i>2</i>
3. NIVELES DE GESTIÓN DE UNA ORGANIZACIÓN.	3
4. EVOLUCIÓN: DE LOS SISTEMAS TRADICIONALES DE FICHEROS A LAS BASES DE DATOS.....	4
4.1. <i>Inconvenientes de los sistemas tradicionales de ficheros.....</i>	<i>4</i>
5. CONCEPTO DE BASE DE DATOS.....	6
5.1. <i>Niveles de abstracción.....</i>	<i>6</i>
5.2. <i>Independencia de los datos:.....</i>	<i>6</i>
5.3. <i>Usuarios de un sistema de Base de Datos. La función de administración.....</i>	<i>7</i>
5.4. <i>El Sistema de Gestión de la Base de Datos.....</i>	<i>8</i>
5.5. <i>Ventajas del uso de bases de datos.....</i>	<i>8</i>
5.5.1. <i>Reducción de la redundancia.....</i>	<i>8</i>
5.5.2. <i>Puede evitarse la inconsistencia.....</i>	<i>9</i>
5.5.3. <i>Los datos pueden compartirse.....</i>	<i>9</i>
5.5.4. <i>Pueden hacerse cumplir las normas establecidas.....</i>	<i>9</i>
5.5.5. <i>Restricciones de seguridad.....</i>	<i>9</i>
5.6. <i>Inconvenientes de las Bases de Datos.....</i>	<i>9</i>
BIBLIOGRAFÍA.....	10

TEMA 1. INTRODUCCIÓN A LAS BASES DE DATOS

Las bases de datos constituyen una parte fundamental de los sistemas de información en los que están integrados.

El estado actual de la tecnología de bases de datos es el resultado de la evolución que a lo largo de décadas ha tenido lugar en el procesamiento de los datos y en la gestión de la información. La tecnología de gestión de datos se ha desarrollado desde los métodos primitivos de los años cincuenta hasta los potentes sistemas de hoy en día, empujada por un lado por la demanda y las necesidades de la gestión de la información y restringida por otro por las limitaciones de la tecnología.

Las necesidades de administración crecen paralelamente a la evolución de la tecnología. Los primeros sistemas de procesamiento de datos resolvían tareas administrativas para reducir el papeleo. Estos sistemas han ido evolucionando hacia la gestión y la producción de información, que se ha convertido en un recurso vital para las compañías. En la actualidad, la función más importante de los sistemas de bases de datos es servir de fundamento a los sistemas de información para la gestión corporativa.

1. Cualidades de la información

[Adaptado de DEMIG97] La creciente necesidad de información y la mayor disponibilidad de ésta puede conducir a un deterioro de la calidad de la misma si no se toman las medidas necesarias para evitarlo.

Si pierde sus cualidades, la información, lejos de cumplir sus objetivos, es decir, aportar un conocimiento, puede perder su valor informativo siendo incluso perjudicial para los destinatarios de la misma.

Las cualidades que hacen de la información un recurso fundamental para las organizaciones y los usuarios individuales son, básicamente:

- **Precisión:** Porcentaje de información correcta sobre el total de información almacenada en el sistema.
- **Oportunidad:** Tiempo transcurrido desde que se produjo el hecho que originó el dato hasta el momento en el que la información se pone a disposición del usuario. También se mide según el tiempo transcurrido desde que el dato tendría que estar disponible o por el desfase que produce su procesamiento por ordenador. Dependiente de la finalidad de los datos. Por ejemplo, para la elaboración de un censo de población de, por ejemplo, varios millones de datos de carácter bastante estable, un tiempo de proceso de meses no le resta oportunidad a la información. En cambio, esta demora en otro tipo de valores como, por ejemplo, las cotizaciones en bolsa, sería inadmisibles, despojando a los datos de su valor informativo. Aunque en general el valor informativo de los datos pierde valor con el transcurso del tiempo, en los sistemas de soporte a la toma de decisiones es al contrario.
- **Compleción:** Ha de ser completa para cumplir los fines para los que se gestiona. Por ejemplo, si se mantiene un sistema de información como sistema de apoyo a la toma de decisiones, la información proporcionada por éste ha de contener todos los elementos informativos necesarios para dar apoyo a la toma de decisiones.
- **Significado:** El contenido semántico de los datos ha de ser máximo. Al diseñar un sistema de información debemos adecuar el volumen de información suministrada por el sistema a

los objetivos propuestos. Más información no significa *mejor* información, ocurriendo que un volumen excesivo de información dificulta su asimilación por parte del usuario.

- **Coherencia:** La información contenida en el sistema debe ser coherente en sí misma, además de consistente con las restricciones semánticas del mundo al que ha de representar lo más fielmente posible. La literatura sobre bases de datos suele referirse a esta cualidad como *integridad*.
- **Seguridad:** Debe prestarse especial atención a la protección de la información tanto frente a su deterioro físico o lógico como a los accesos no autorizados. El problema de la seguridad abarca la confidencialidad, la disponibilidad y la integridad.

2. Sistemas de Información

Toda organización necesita para su funcionamiento un conjunto de informaciones que han de transmitirse entre sus diferentes elementos, así como desde y hacia el exterior de la propia organización.

Un sistema de información se diseña con el fin de satisfacer las necesidades de información de una organización, en la que está inmerso. El sistema de información toma datos del entorno (tanto de la organización como de fuentes externas) y los resultados de las operaciones sobre esos datos serán la información que dicha organización necesita para su gestión y toma de decisiones.

2.1. Componentes de un sistema de información

En todo sistema de información encontramos los siguientes componentes:

Contenido (Datos: hechos conocidos con significado implícito, que pueden ser almacenados) es el centro del sistema de información. Los datos contenidos en un sistema de información pueden ser de tipo *referencial* o de tipo *factual*.

Datos de tipo referencial son aquellos que contienen información acerca de dónde se encuentra la información buscada.

Los datos de tipo factual son aquellos que contienen la información en sí. A su vez pueden ser estructurados o no estructurados. Los primeros son aquellos con una estructura definida. Los no estructurados son aquellos que, por sus características, no tienen una estructura definida (sonido, vídeo, etc.).

Equipo físico. Ordenadores y periféricos.

Soporte lógico. Incluye todo el *software* necesario para la implantación del sistema de información (Sistema Operativo, sistema de base de datos, *software* de comunicaciones y otros programas para tratamientos específicos).

Administrador. Los datos y las informaciones manejadas por nuestro sistema de información han de ser gestionadas por las personas adecuadas. Tendremos, por un lado los responsables de tomarlas decisiones estratégicas y de política con respecto a la información de la empresa y por otro los responsables de dar apoyo técnico para poner en práctica esas decisiones.

Usuarios. Los datos han de ser puestos a disposición de los usuarios del sistema. Podemos clasificarlos en usuarios informáticos, aquellos que realizan las aplicaciones que manejarán los datos almacenados en el sistema y usuarios no informáticos, los cuales se ocuparán de realizar las tareas administrativas con los datos y/o serán los destinatarios de las informaciones extraídas del sistema.

3. Niveles de gestión de una organización. Sistemas de información para la gestión y sistemas de información para la ayuda a la decisión.

En toda organización suelen distinguirse tres niveles de gestión (operacional, táctico y estratégico) por lo que el sistema de información ha de ser diseñado para satisfacer las necesidades y proporcionar las informaciones adecuadas a cada uno de los niveles.

En el plano operacional, los usuarios manejan datos puntuales o elementales que describen los sucesos que caracterizan las actividades de la organización. Esta información, compuesta por datos totalmente desagregados (microdatos) es necesaria para los procesos comúnmente denominados administrativos (tareas diarias y de rutina) y el volumen de datos manejado será muy grande.

Los niveles táctico y estratégico, cuyos usuarios tienen necesidades de información muy distintas, obtendrán del anterior, mediante procesos de elaboración adecuados (generalmente de agregación) junto con datos provenientes del exterior las informaciones necesarias para ayuda a la decisión, que exige prestaciones muy diferentes en la que los datos están agregados (macrodatos) y cuya elaboración es mucho más compleja.

Como ya se ha apuntado, la aplicación de los ordenadores en las organizaciones comenzó con el tratamiento administrativo de sus datos operacionales (nóminas, contabilidad, etc.) La creciente potencia de los ordenadores los hizo aptos para intervenir en otros niveles de la empresa, ayudando a la sistematización de las funciones de dirección y constituyendo un elemento activo en el proceso de toma de decisiones. Los sistemas de soporte a la toma de decisiones (DSS, *Decision Support Systems* o EIS, *Executive Information Systems*) tienen como componente principal una base de datos. Del mismo modo, los datos provenientes del plano operacional almacenados en bases de datos así como en otros soportes, organizados en almacenes de datos (*Data Warehouses*) sirven de base para la extracción y el descubrimiento de conocimiento en bases de datos (KDD, *Knowledge Discovering in Databases*) y para la minería de datos (*Data Mining*).

Nivel Estratégico

- Objetivos generales
- Elaboración de planes

Nivel Táctico

Ordenes y planes

- Objetivos específicos
- Control de gestión

Información

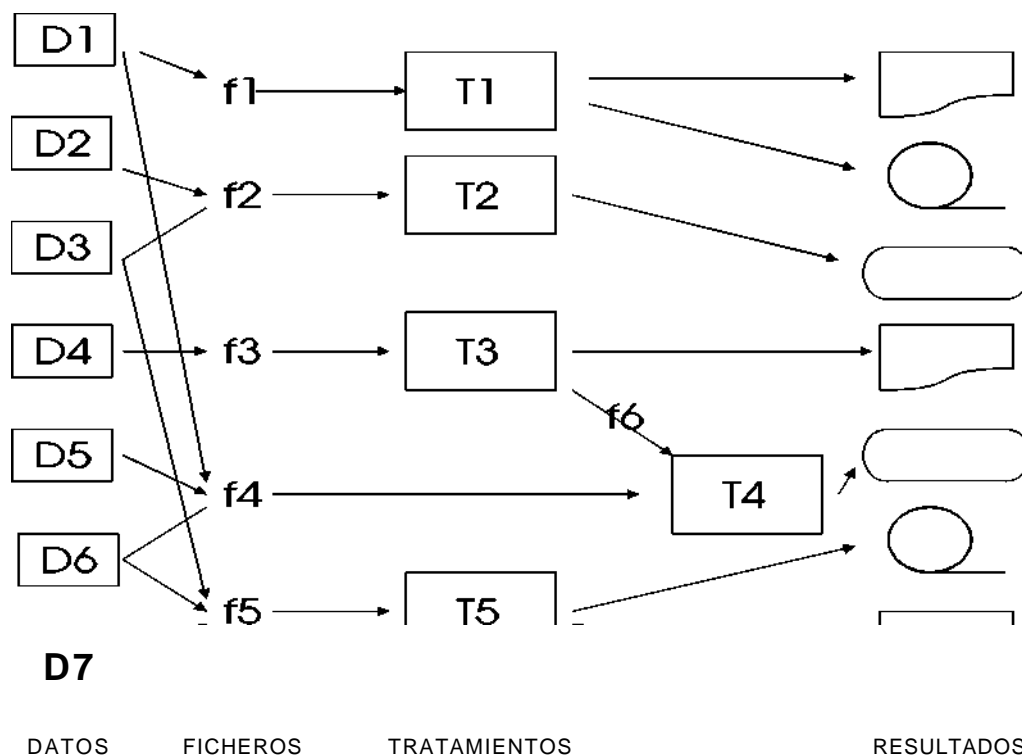
Nivel Operacional

- Tareas administrativas

4. Evolución: De los sistemas tradicionales de ficheros a las bases de datos

4.1. Inconvenientes de los sistemas tradicionales de ficheros

Sistemas tradicionales de ficheros: Sistemas orientados hacia el proceso (se pone énfasis en el tratamiento que reciben los datos, los cuales se almacenan en archivos diseñados para una aplicación.) Las aplicaciones se diseñan e implantan independientemente unas de otras y los datos se duplican si las diferentes aplicaciones los necesitan, en lugar de transferirse entre ellas.



Redundancia: Duplicidad innecesaria de información.

Mal aprovechamiento del equipo de almacenamiento, como consecuencia inmediata de la redundancia.

Aumento de los tiempos de proceso. Se repiten los mismos controles y operaciones en los distintos ficheros, con lo que se consume más de tiempo de CPU del necesario. En el caso de modificar un campo hay que hacerlo en todos los registros de todos los ficheros que lo contengan.

Inconsistencia de la información por la alta redundancia. Si se deja de actualizar un dato en uno de los archivos donde aparece, la información proporcionada por ese dato se vuelve inconsistente.

Aislamiento de los datos. Cada archivo pertenece a un programa y no es posible que éstos sean usados por nuevos programas. Un nuevo programa necesitará su(s) propio(s) archivo(s) de datos que habrán de crearse aunque parte de los datos ya existan en otros archivos de otros programas, contribuyendo a aumentar la redundancia y las consecuencias de ésta.

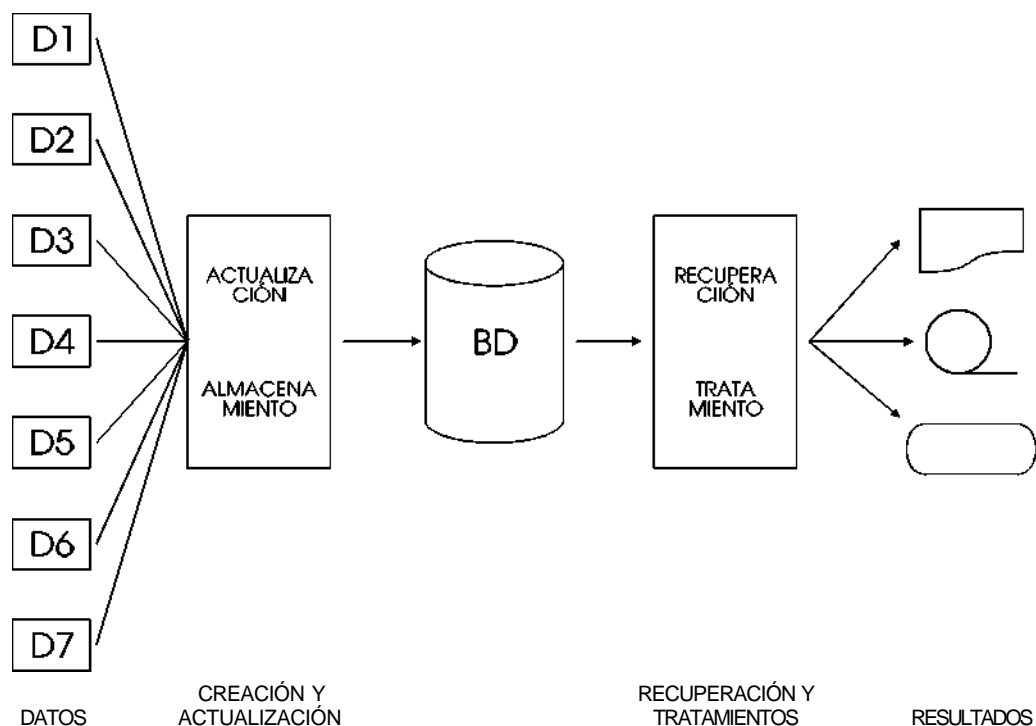
Imposibilidad de responder a demandas inesperadas de información. Los sistemas tradicionales de archivos son inoperantes para conseguir un sistema de información orientado a la toma de decisiones.

Dependencia total entre los programas y la estructura física de los datos. No es posible modificar las características físicas (estructura y métodos de acceso) de los archivos sin afectar a los programas que los usan. Conseguir la independencia entre datos y aplicaciones va a ser uno de los principales objetivos de los sistemas de bases de datos.

Surge la necesidad de una gestión más racional del conjunto de los datos que no presente los inconvenientes anteriormente descritos. Surge así un nuevo enfoque que se apoya sobre una *base de datos* en la que éstos son recogidos y almacenados una sola vez con independencia de los tratamientos que se van a aplicar sobre ellos.

De esta forma, la información contenida en una base de datos está *integrada* y *compartida*. Integrada porque puede considerarse como una unificación de varios archivos de datos de los que hemos eliminado la redundancia y compartida porque los programas que antes accedían a los archivos individuales acceden ahora al depósito común de datos, por lo que cada usuario o aplicación tendrá acceso a un subconjunto de los datos y como consecuencia diferentes usuarios verán de formas muy diferentes la misma base de datos. Es importante tener en cuenta que los subconjuntos de datos a los que acceden las diferentes aplicaciones o usuarios no tienen por qué ser disjuntos por lo que usuarios o aplicaciones distintas pueden acceder a la misma parte de la base de datos para utilizarla con propósitos diferentes.

Estructura con un sistema de base de datos:



5. Concepto de Base de Datos

Existen numerosas definiciones de base de datos. Algunas de ellas:

- Colección de datos interrelacionados.
Elmars, R, Navathe, S.B. 1989
- Colección no redundante de datos que son compartidos por diferentes programas de aplicación.
Howe, 1983
- Conjunto de datos de la empresa memorizado en un ordenador, que es utilizado por numerosas personas y cuya organización está regida por un modelo de datos.
Flory, 1982
- “Una Base de Datos es un conjunto de información almacenada en memoria auxiliar que permite acceso directo, y un conjunto de programas que manipulan esos datos.”
- “Una Base de Datos es un conjunto exhaustivo, no redundante de datos estructurados, organizados independientemente de su utilización y su implementación en máquina, accesibles en tiempo real y compartibles por usuarios concurrentes que tienen necesidad de información diferente y no predecible en el tiempo.”
- “Colección o depósito de datos integrados, con redundancia controlada y con una estructura que refleje las interrelaciones y restricciones existentes en el mundo real; los datos, que han de ser compartidos por diferentes usuarios y aplicaciones, deben mantenerse independientes a éstas, y su definición y descripción, únicas para cada tipo de dato, han de estar almacenados junto con los mismos. Los procedimientos de actualización y recuperación, comunes y bien determinados, habrán de ser capaces de conservar la integridad, seguridad y confidencialidad del conjunto de los datos.”

5.1. Niveles de abstracción

La mayoría de las aplicaciones son dependientes de los datos; la organización del almacenamiento y los modos de acceso dependen de los requerimientos de la aplicación y el conocimiento de la organización física de los datos y las técnicas de acceso forman parte de la lógica de la aplicación. La aplicación es dependiente de los datos, porque no se puede modificar la estructura de almacenamiento o los modos de acceso sin afectar a la aplicación.

Como hemos visto, los sistemas de bases de datos se plantean los siguientes objetivos:

Independencia de la base de datos de los programas para su utilización.

Proporcionar a los usuarios una visión abstracta de los datos. El sistema esconde los detalles de almacenamiento físico (cómo se almacenan y se mantienen los datos) pero éstos deben extraerse eficientemente.

5.2. Independencia de los datos:

ANSI: “La independencia de los datos es la capacidad de un sistema para permitir que las referencias a los datos almacenados, especialmente en los programas y en sus descriptores de los datos, estén aislados de los cambios y de los diferentes usos en el entorno de los datos, como pueden ser la forma de almacenar dichos datos, el modo de compartirlos con otros programas y como se reorganizan para mejorar el rendimiento del sistema de Bases de Datos.”

Para conseguir esta independencia entre los datos y las aplicaciones es necesario separar la representación física y lógica de los datos, distinción que fue reconocida oficialmente en 1978, cuando el comité ANSI/X3/SPARC propuso un *esqueleto* generalizado para sistemas de bases de datos. Este esqueleto propone una arquitectura de tres niveles, los tres niveles de

abstracción bajo los que podría verse una base de datos: el nivel interno, el nivel conceptual y el nivel externo.

NE1 NE2 NE3 NIVEL EXTERNO

NIVEL CONCEPTUAL

NIVEL INTERNO

Los tres niveles de la arquitectura ANSI/X3/SPARC

- **NIVEL INTERNO:** En el se define la estructura física de la base de datos: dispositivos de almacenamiento físico, direcciones físicas, estrategias de acceso, relaciones, índices, apuntadores, etc. Es responsabilidad de los diseñadores de la base de datos física. Ningún usuario, en calidad de tal, tiene conocimiento de este nivel.
- **NIVEL CONCEPTUAL:** Contiene el diseño conceptual de la base de datos, que implica el análisis de las necesidades de información de los usuarios y las clases de datos necesarias para satisfacer dichas necesidades. El resultado del diseño conceptual contiene la descripción de todos los datos y las interrelaciones entre ellos, así como las restricciones de integridad y de confidencialidad.
- **NIVEL EXTERNO:** Visión que de la base de datos tiene un usuario o aplicación en particular. Habrá tantas vistas de la base de datos como exijan las diferentes aplicaciones. Las vistas se derivan directamente del esquema conceptual, o de otras vistas, y contienen una descripción de los elementos de datos y sus interrelaciones orientadas al usuario o aplicación y de las que se compone la vista. Una misma vista puede ser utilizada por varias aplicaciones.

Esta arquitectura de tres niveles nos proporciona la deseada independencia, que definiremos como capacidad para cambiar el esquema en un nivel sin tener que cambiarlo en ningún otro nivel. Distinguimos entre independencia física y lógica:

- **Independencia lógica de los datos:** Cambio del esquema conceptual sin cambiar las vistas externas o las aplicaciones.
- **Independencia física de los datos:** Cambio del esquema interno sin necesidad de cambiar el esquema conceptual o los esquemas externos.

5.3. Usuarios de un sistema de Base de Datos.

La función de administración

Podemos clasificar en tres clases amplias los usuarios de un sistema de base de datos:

- Por un lado los programadores de aplicaciones, que escriben los programas que utilizan los datos almacenados en la base de datos.

- Por otro los usuarios finales, que interactúan con el sistema a través de las utilidades que éste proporciona para ello.
- En tercer lugar, el administrador del sistema. Es el responsable del control general del sistema desde el punto de vista técnico. Entre sus funciones cabe destacar:
 - Definir el esquema conceptual.
 - Definir el esquema interno.
 - Establecer las restricciones de seguridad, integridad y confidencialidad.
 - Definir los procedimientos de copia de seguridad y recuperación.
 - Supervisar el rendimiento del sistema y responder a los cambios en los requerimientos.

5.4. **El Sistema de Gestión de la Base de Datos**

En un sistema de bases de datos, debe existir una capa intermedia entre los datos almacenados en la base de datos, las aplicaciones y los usuarios del mismo. Se trata del Sistema de Gestión de la Base de Datos (SGBD). Actúa de intermediario entre los usuarios y aplicaciones y los datos proporcionando medios para describir, almacenar y manipular los datos, y proporciona herramientas al administrador para gestionar el sistema, entre ellas las herramientas de desarrollo de aplicaciones, generadores de informes, lenguajes específicos de acceso a los datos, como SQL (*Structured Query Language*) o QBE (*Query By Example*) (en bases de datos relacionales).

Sistema de Gestión de Base de Datos: Se puede definir como un conjunto coordinado de programas, procedimientos, lenguajes, etc. que suministra, tanto a los usuarios no informáticos como a los analistas, programadores o el administrador, los medios necesarios para describir, recuperar y manipular los datos almacenados en la base, manteniendo su integridad, confidencialidad y seguridad.

Entre sus funciones están:

- **Definición y control centralizado de los datos.** Definición de todos los elementos de datos en la base de datos en los tres niveles de abstracción definidos anteriormente (interno, conceptual y externo). Descripción de los datos (campos, grupos, registros, tablas), interrelaciones entre las diferentes estructuras de datos. Es una parte de la base de datos. La base de datos es autodescriptiva: contiene información que describe su estructura (metadatos), los cuales están disponibles para su consulta. A estos metadatos se les llama *Catálogo* o *Diccionario de datos*.
- **Manipulación de los datos:** el SGBD debe ser capaz de atender las solicitudes del usuario para extraer, modificar o añadir datos a la base de datos.
- **Seguridad e integridad:** Proporciona los medios para definir y gestionar las autorizaciones de acceso, ya sea mediante claves de acceso al sistema, o mediante la definición de vistas externas de usuario.
Proporciona así mismo los medios para garantizar la integridad y la consistencia de los datos definiendo (en el diccionario de datos) restricciones sobre los valores que pueden tomar y proporcionando capacidades de recuperación ante fallos y de copia de seguridad.
- **Garantiza la disponibilidad de la información** asegurando el acceso concurrente a varios usuarios simultáneos.

5.5. **Ventajas del uso de bases de datos**

5.5.1. **Reducción de la redundancia**

En los sistemas tradicionales de ficheros cada aplicación tiene sus datos privados, lo que provoca una alta redundancia y desaprovechamiento del espacio en disco.

La redundancia debe minimizarse y controlarse. Aunque se mantenga cierto grado de redundancia por motivos de rendimiento u otros, el sistema proporciona mecanismos para garantizar la consistencia.

5.5.2. Puede evitarse la inconsistencia

Se controla la redundancia garantizando que los datos redundantes se actualicen de forma automática.

5.5.3. Los datos pueden compartirse

Las necesidades de datos de nuevas aplicaciones pueden atenderse con los ya existentes sin tener que almacenar nuevos datos.

5.5.4. Pueden hacerse cumplir las normas establecidas

Al tener un control centralizado de la base de datos, el administrador (a instancias del responsable de la información de la organización) puede garantizar que se observen las normas de la empresa aplicables a la representación de los datos.

5.5.5. Restricciones de seguridad

El administrador puede *asegurar* que el único modo de *acceso sea* a través de los canales establecidos, y en consecuencia definir controles de autorización que pueden afectar a cada modo de *acceso* (modificación, inserción, borrado o lectura).

Sin estos controles de seguridad, es mucho más sensible una BD que los archivos.

5. 6. Inconvenientes de las Bases de datos

- Instalación costosa
- Necesidad de personal especializado
- Implantación larga y difícil
- Falta de rentabilidad a corto plazo
- Escasa estandarización
- Desfase entre teoría y práctica

BIBLIOGRAFÍA

- [DATE86] C. J. DATE. *Introducción a los sistemas de bases de datos*. Tercera edición. Addison-Wesley Iberoamericana, 1986
- [DATE93] C. J. DATE. *Introducción a los sistemas de bases de datos*. Vol. 1 Quinta edición. Addison-Wesley Iberoamericana, 1993
- [DEMIG93] ADORACIÓN DE MIGUEL Y MARIO PIATTINI. *Concepción y Diseño de bases de datos*. Ed. Ra-Ma. 1993.
- [DEMIG97] ADORACIÓN DE MIGUEL Y MARIO PIATTINI. *Fundamentos y modelos de bases de datos*. Ed. Ra-Ma. 1997.
- [HANSEN97] HANSEN GW, HANSEN JV. *Diseño y Administración de Bases de Datos*. 2^a Edición. Prentice Hall, Madrid, 1997