



U.T. 1: Introducción a la Gestión de Datos.

I.E.S. ANTONIO SEQUEROS

ADMINISTR. DE SISTEMAS INFORMÁTICOS EN RED

GESTIÓN DE BASES DE DATOS

CURSO 2022/2023

Índice

1.CONCEPTO DE SISTEMA DE INFORMACIÓN.....	1
2. COMPONENTES DE UN SISTEMA DE INFORMACIÓN.....	2
3. SISTEMAS DE INFORMACIÓN PARA LA GESTIÓN Y SISTEMAS DE INFORMACIÓN PARA LA AYUDA A LA DECISIÓN.....	3
4. FICHEROS. CONCEPTO Y CLASIFICACIONES.....	4
5. DE LOS SISTEMAS TRADICIONALES DE FICHEROS A LAS BASES DE DATOS	6
6. VENTAJAS E INCONVENIENTES DE LAS BASES DE DATOS FRENTE A LOS FICHEROS CLÁSICOS.....	8

Bibliografía

DE MIGUEL y PIATTINI (1999). *Fundamentos y Modelos de Bases de Datos*. Ed. Ra-Ma

DE MIGUEL y PIATTINI (1993). *Concepción y Diseño de Bases de Datos*. Ed. Ra-Ma

DATE, C.J. (1993) *Introducción a los sistemas de Bases de Datos, vol. I*. Ed. Addison Wesley Iberoamericana

1. CONCEPTO DE SISTEMA DE INFORMACIÓN

Uno de los pilares de cualquier organización es la información que necesita para su funcionamiento. Esta información se ha de transmitir entre sus elementos y, generalmente también, desde y hacia el exterior del sistema.

Asimismo, una de las ocupaciones básicas de la organización es el tratamiento de dicha información, que ya sea de forma manual o automática, tiene como objeto proporcionar la información correcta en el momento y lugar adecuados a las personas autorizadas con un coste mínimo.

Sistema de información (SI):

Es un conjunto de elementos ordenadamente relacionados entre sí de acuerdo a ciertas reglas, que aportan a la organización a la que sirven la información necesaria para el cumplimiento de sus fines.

Como funciones fundamentales de un SI se pueden destacar:

- la recogida
- procesamiento
- y almacenamiento de datos

facilitando además:

- la recuperación
- elaboración
- y presentación de los mismos

Un SI mecanizado es aquel que está soportado por un ordenador. A partir de ahora, SI siempre hará referencia a un SI mecanizado.

Entre el SI y el organismo donde está inserto existe una mutua y estrecha interrelación; en realidad, el SI no es otra cosa que un subsistema de los varios que integran la organización. Es imprescindible tener esto muy presente, ya que si no existe la debida interacción y se produce un desfase entre ambos, el SI no podrá cumplir los objetivos para los que fue diseñado. La falta de adaptación entre el SI y el organismo es causa del fallo de muchos sistemas que prometían ser eficaces, estando demostrado que las causas de estos fracasos se encuentran más frecuentemente en los aspectos sociales y humanos que en el diseño tecnológico.

2. COMPONENTES DE UN SISTEMA DE INFORMACIÓN

Los componentes básicos de un SI son los siguientes:

- ❖ contenido: los datos y su definición.
 - Referenciales
 - Factuales
 - Estructurados
 - No estructurados
- ❖ equipo físico: el ordenador soporte de información
- ❖ equipo lógico: sistema de comunicaciones, sistema de gestión de base de datos, etc.
- ❖ administrador: persona o equipo de personas responsables de asegurar la calidad y disponibilidad de los datos
- ❖ usuarios

Es importante distinguir entre dos tipos de datos: *referenciales* y *factuales*. Los sistemas de información referencial contienen referencias bibliográficas de los documentos dónde se puede encontrar la información, pero no la información en sí misma. En cambio los sistemas de tipo factual devuelven la información buscada, la cual puede ser directamente utilizada.

Otra clasificación aplicable a los datos *factuales*, se refiere a su formato, según la cual los datos pueden ser estructurados (los campos ocupan determinadas posiciones fijas, ej.:DNI primero y a continuación el nombre y dirección) o no estructurados (formato no fijo, ej.: textos, datos multimedia).

Se suele distinguir entre dos tipos de sistemas de gestión:

Sistemas de Gestión de Bases de Datos (SGBD): tratamiento de datos estructurados.

Sistemas de Recuperación de Información (SRI): tratamiento de datos no estructurados.

Actualmente se tiende a una convergencia entre ambos tipos de tratamiento, de modo que ambas funcionalidades sean ofrecidas por un único sistema.

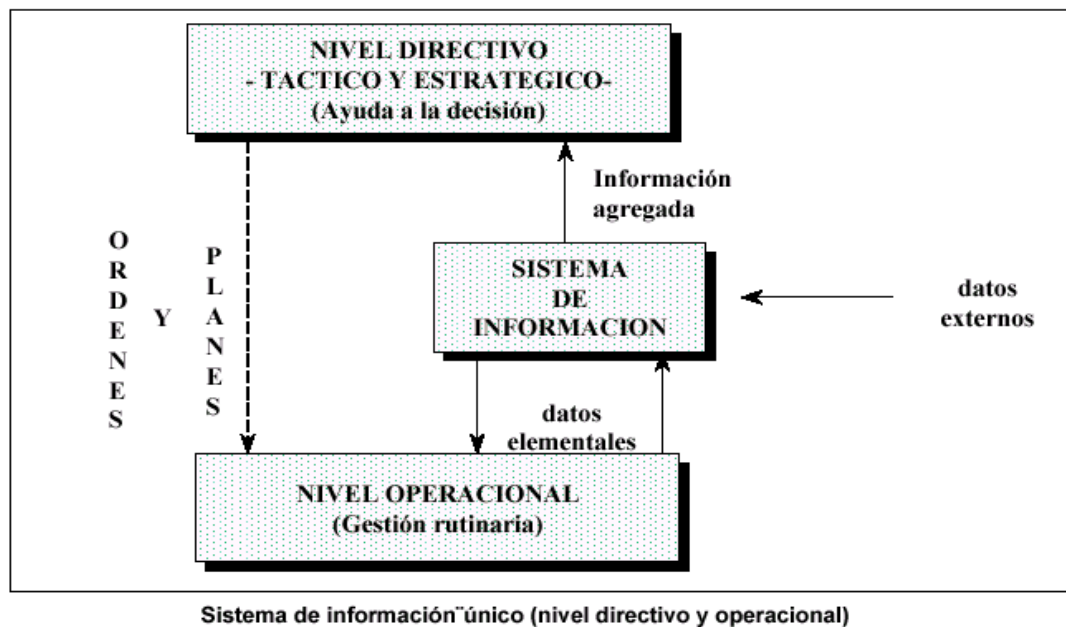
3. SISTEMAS DE INFORMACIÓN PARA LA GESTIÓN Y SISTEMAS DE INFORMACIÓN PARA LA AYUDA A LA DECISIÓN

La aplicación de los ordenadores en las empresas comenzó con el tratamiento administrativo de sus datos operacionales; para llevar a cabo las tareas rutinarias (nóminas, contabilidad,...). Estos **sistemas de información para la gestión** tienen las siguientes características:

- Datos voluminosos
- Propios, elementales y homogéneos
- Pocas interrelaciones y simples
- Muchas salidas normalizadas
- Procesos sencillos y periódicos
- Predomina el tratamiento secuencial y por lotes.

Posteriormente hemos asistido al nacimiento y crecimiento de los **sistemas de información para la toma de decisiones**, dirigidos a los directivos de las empresas. Características:

- Datos muy poco voluminosos
- Propios y ajenos, agregados y muy heterogéneos
- Muchas interrelaciones complejas
- Pocas salidas, con información significativa, oportuna y fácil de interpretar.
- Tratamiento no secuencial e interactivo.



4. FICHEROS. CONCEPTO Y CLASIFICACIONES.

Un **fichero** es la representación de una entidad, en cada fichero se guarda la información correspondiente a una de las entidades representadas.

Podemos clasificar los ficheros de 3 formas distintas:

Ficheros internos y externos:

Los ficheros pueden guardarse en memoria externa puesto que sirven para almacenar datos que deben permanecer en el sistema aunque éste se apague: son los **ficheros externos o permanentes**.

Sin embargo, durante la ejecución de un programa se trabaja con unas estructuras de datos análogas a los ficheros, pero que se encuentran en la memoria interna, y sirven para almacenar aquellos datos que van a ser inmediatamente utilizados para desaparecer a continuación (o se volverán a grabar en la memoria externa), son los denominados **ficheros internos o temporales**.

De todas formas, nosotros nos centraremos en los ficheros externos.

Clasificación de ficheros según su uso:

Según el uso que se les vaya a dar, los ficheros pueden ser:

- **De constantes:** Este tipo de ficheros contiene información que no varía en el tiempo, o lo hace muy de cuando en cuando. Por ello, se modifican en contadísimas ocasiones, o nunca. Ejemplos de ficheros constantes pueden ser los de las provincias de un país, o los de códigos postales y poblaciones.
- **Maestros:** Este tipo de ficheros contiene información que necesita actualizaciones frecuentes (altas, bajas, modificaciones). Por ejemplo, ficheros de clientes, de profesores o de alumnos, entre otros.
- **Históricos:** Se trata de ficheros que guardan información sobre hechos pasados e inamovibles, por lo que sólo se utilizan para consultas y estadísticas; así pues, no sufren cambios, excepto la incorporación de nuevos datos. Ejemplo, historial con todas las notas de años anteriores, al que cada año se añadirán las notas del curso que acaba.

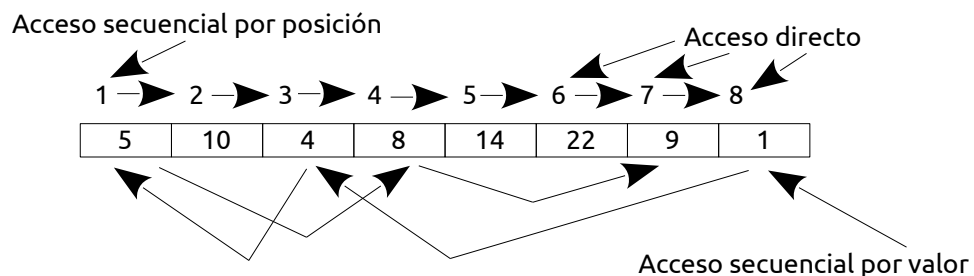
Clasificación de ficheros según su organización y acceso:

Otra clasificación importante es la que hace referencia a la forma de acceso a los ficheros.

Hay que tener presente que un fichero está formado, por registros. Estos registros están ordenados físicamente, pero ese orden puede ser aleatorio o puede estar en función de alguno de los campos, por ejemplo, personas ordenadas alfabéticamente por su apellidos.

Los métodos de acceso a un fichero, son los siguientes:

- **Acceso secuencial por posición:** El fichero debe leerse siempre empezando por el registro que ocupa físicamente el primer lugar y continuando de manera secuencial, de uno en uno, hasta llegar al registro que se busca, o bien al final físico del fichero.
- **Acceso directo por posición:** Si se conoce cuál es la posición exacta que ocupa un registro dentro del fichero, puede accederse directamente a dicho registro solicitando, por ejemplo, ir al alumno que ocupa la posición 20 en el fichero.
- **Acceso secuencial por valor:** Siguiendo el orden de alguno de los campos, por ejemplo apellidos o código, se puede empezar a leer el primer registro y, a continuación, el resto de registro uno por uno siguiendo el mismo orden hasta llegar al final del fichero o hasta encontrar el registro requerido.
- **Acceso directo por valor:** El fichero está organizado de manera que dando el valor de un atributo o conjunto de ellos, generalmente formando una clave de tipo ir al alumno con expediente nº 123123, se pueda acceder directamente al registro correspondiente.



Según el tipo de acceso que permita cada fichero, se puede hacer la siguiente clasificación, aunque está obsoleta:

- **Secuenciales:** Se trata de ficheros que sólo permiten el acceso secuencial por posición, de manera que cuando se abren únicamente es posible leer el primer registro y, a continuación, el registro que ocupa la posición siguiente, y así sucesivamente hasta el final del fichero. Los registros nuevos se insertan siempre al final del fichero; por tanto, sólo hay un orden físico, sin ningún orden lógico, o por valor, equivalente. Ejemplo cintas magnéticas.
- **Relativos:** Son los ficheros donde cada registro se identifica por su posición relativa respecto al inicio del fichero y, por este motivo, se permite el acceso directo por posición si ésta es conocida. En caso contrario siempre puede utilizarse el acceso secuencial por posición, que también está admitido.
- **Calculados:** Se trata de una variante de los ficheros relativos donde la posición del registro viene determinada por el valor de algún atributo, de varios atributos o de un cálculo realizado a partir de dichos valores. En estos casos, se puede realizar tanto el acceso secuencial por valor como el acceso directo por valor. En los ficheros calculados es más problemático el acceso directo por posición, aunque es posible.

- **Indexados:** Se trata de un tipo especial de ficheros que se presenta como una mezcla de los anteriores. Aparte del fichero que contiene los datos existe otros fichero, o a veces más de uno, denominado índice. El **fichero índice** guarda información relacionada con la posición que ocupan algunos registros, o a veces todos, y facilita la búsqueda de los mismos. Básicamente su funcionamiento es como el del índice de un libro, donde los registros se corresponderían, por ejemplo, con los capítulos, y el número de página, con su posición.

Ahora ya sabemos cómo organizar la información en un soporte informático para que éste nos resulte útil y podamos acceder fácilmente a él. Sin embargo, los programas informáticos deben conocer la ubicación exacta de cada fichero, al igual que de cada uno de sus registros, si se trata de un fichero secuencial o no, el tipo de accesos que dicho fichero permite, etcétera.

5. DE LOS SISTEMAS TRADICIONALES DE FICHEROS A LAS BASES DE DATOS.

Si analizamos la situación de algunos SI antiguos, encontramos sistemas en los que existe una gran cantidad de ficheros, específicos de una determinada aplicación. Por ello, se recogen los datos varias veces y se encuentran repetidos en distintos archivos. Este planteamiento produce una serie de problemas:

Ocupación inútil de memoria secundaria.

Aumento del tiempo de proceso, al repetirse las mismas operaciones en los distintos ficheros.

Inconsistencias en los datos, al intentar actualizar datos repetidos en varios ficheros.

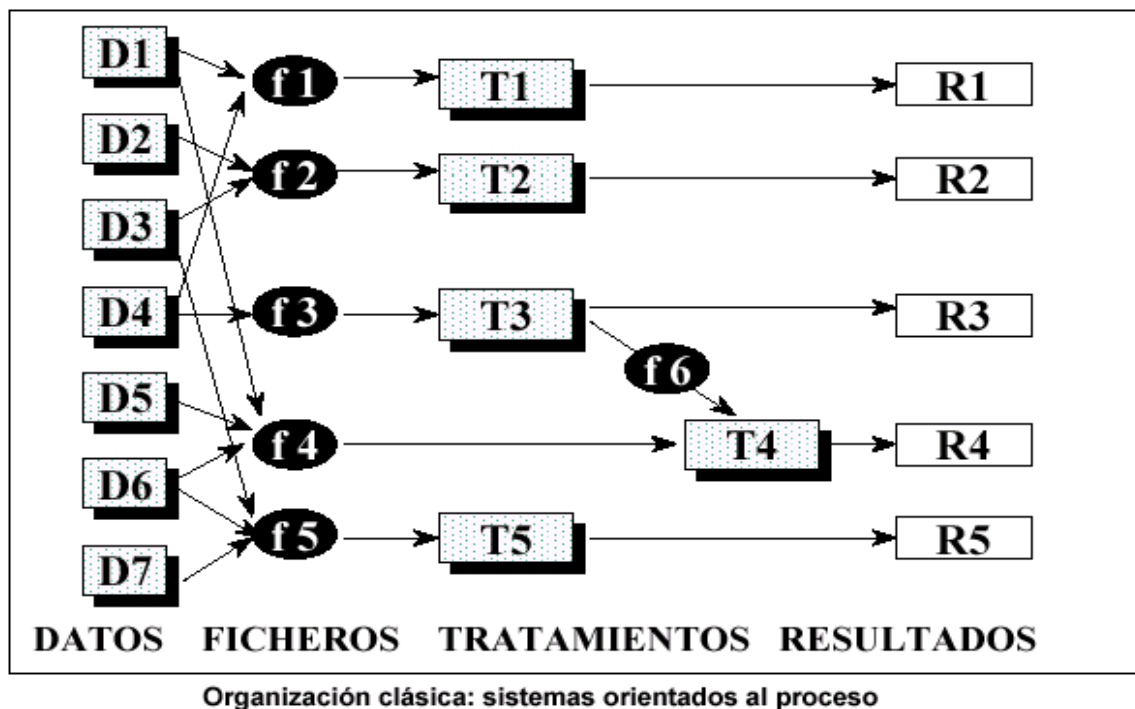
Falta de adaptabilidad a los cambios debido a la dependencia de los datos respecto al soporte físico y a los programas.

Imposibilidad de implementar un sistema de ayuda a la toma de decisión con esta configuración.

Anomalías con accesos concurrentes a los ficheros de datos.

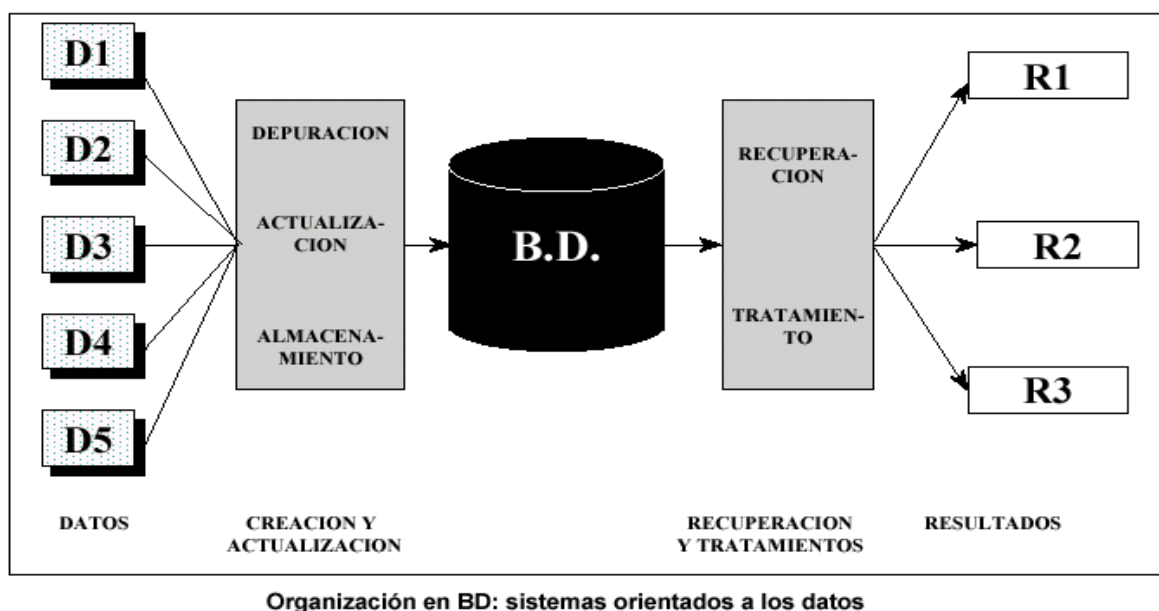
Costosa recuperación de ficheros por antigüedad de copias de seguridad.

Problemas de seguridad y confidencialidad: No todos los usuarios deben tener acceso a todos los datos. Como mucho ofrecen sistemas de protección a nivel de aplicación, pero nunca a nivel de usuario.



De este análisis se deduce claramente la necesidad de una gestión más racional del conjunto de datos, surgiendo así un nuevo enfoque que se apoya sobre una **base de datos**, en la cual los datos son recogidos y almacenados una sola vez, sin estar diseñados para una aplicación concreta, sino que tiende a satisfacer las necesidades de información de toda la organización.

Estos sistemas orientados hacia los datos van sustituyendo a los anteriores sistemas orientados hacia el proceso.



6. VENTAJAS E INCONVENIENTES DE LAS BASES DE DATOS FRENTE A LOS FICHEROS CLÁSICOS

Las bases de datos, surgidas como respuesta al nuevo planteamiento de los sistemas orientados hacia los datos para mejorar la calidad de las prestaciones de los sistemas informáticos y aumentar su rendimiento, presentan una multitud de ventajas frente a los sistemas clásicos de ficheros.

Debemos señalar que las bases de datos no son la panacea universal que soluciona todos los problemas que la información plantea a las empresas.

Las ventajas de los sistemas de bases de datos son, entre otras, las siguientes:

1) Reducción del espacio de almacenamiento

La desaparición (o disminución) de las redundancias, así como la aplicación de técnicas de compactación, conlleva una menor ocupación de almacenamiento secundario -disco magnético-. Se ha de tener presente, sin embargo, que los elementos del sistema (diccionario, referencias, punteros, ficheros invertidos, etc.) ocupan bastante espacio.

Como hemos señalado anteriormente, las bases de datos no sólo presentan ventajas, sino que también tienen posibles inconvenientes que es necesario valorar antes de tomar una decisión relativa a un cambio en la orientación del SI.

2) Mayor eficiencia en la recogida, validación e introducción de los datos en el sistema

Al no existir apenas redundancias, los datos se recogen y validan una sola vez, aumentando así el rendimiento de todo el proceso previo al almacenamiento.

3) Coherencia de los resultados

Debido a que la información de la base de datos se recoge y almacena una sola vez, en los tratamientos se utilizan los mismos datos, por lo que los resultados de todos ellos son coherentes y perfectamente comparables. Además, al no existir (o al menos disminuir en gran medida) la redundancia en los datos, desaparece el problema que se presentaba en el enfoque clásico de que el cambio de un dato obligaba a actualizar una serie de ficheros. De esta forma se elimina también el inconveniente de las divergencias en los resultados debidas a actualizaciones no simultáneas en todos los ficheros.

4) Independencia de los datos respecto a los tratamientos y viceversa

La mutua independencia de datos y tratamientos lleva a que un cambio de estos últimos no imponga un nuevo diseño de la base de datos. Por otra parte, la inclusión de nuevas informaciones, desaparición de otras, cambios en la estructura física o en los caminos de acceso, etc. no deben obligar a alterar los programas.

La flexibilidad que proporciona la independencia de los datos y programas es muy importante para conseguir sin excesivos costes la continua adaptación del sistema de información a la evolución de las organizaciones. Aunque en la práctica dicha independencia nunca es absoluta, los sistemas de gestión de bases de datos actuales van respondiendo cada vez mejor a las exigencias de los usuarios relativas a la independencia y flexibilidad de los datos.

5) Mejor disponibilidad de los datos para el conjunto de los usuarios

Cuando se aplica la metodología de bases de datos, cada usuario ya no es propietario de los datos, puesto que éstos se comparten entre el conjunto de aplicaciones, existiendo una mejor disponibilidad de los datos para todos los que tienen necesidad de ellos, siempre que estén autorizados para su acceso. Hay también una mayor transparencia respecto a la información existente, ya que todos los datos que se encuentran en la base se deben relacionar en un catálogo o diccionario que puede ser ampliamente difundido y accedido por medios informáticos.

6) Mayor valor informativo

Puesto que la base de datos ha de ser reflejo del mundo real, en ella se recogen las interrelaciones entre los datos. Por lo que el valor informativo del conjunto es superior a la suma del valor informativo de los elementos individuales que lo constituyen.

7) Mejor y más normalizada documentación de la información, la cual está integrada con los datos

En el enfoque clásico los datos se encuentran separados de su contenido semántico; los primeros se almacenan en ficheros y su descripción se hace mediante un lenguaje de programación que se encuentra en los programas. La documentación de los datos, realizada por el analista o programador, es en general insuficiente, y a veces incluso inexistente. Además, por lo común la estandarización brilla por su ausencia. Este problema se atenúa en gran medida en las bases de datos, ya que en la misma base se incluyen no sólo los datos, sino también la semántica de los mismos.

8) Mejoras en el control, seguridad y tolerancia a fallos

La base de datos centraliza toda la información del sistema y proporciona mecanismos para que los usuarios utilicen los datos. De esta forma es posible controlar todos los accesos y operaciones que se realizan, permitiendo elaborar estadísticas de utilización de los datos (auditoria), controlar los permisos de acceso de cada usuario o grupo de usuarios cuando realizan una petición de operación sobre los datos e instaurar mecanismos generales de protección de fallos y recuperación de errores para todas las operaciones. En el enfoque clásico, la implantación de cualquiera de estos controles obligaba a que cada usuario o aplicación realizara por sí mismo las tareas asociadas y resultaba prácticamente imposible en sistemas complejos o de gran tamaño debido al volumen de usuarios que se debían coordinar. Por ejemplo: para obtener estadísticas fiables de uso de los datos en un sistema de ficheros clásico es necesario modificar todas las aplicaciones que acceden a los datos para que registren sus operaciones (una sola aplicación que no anota sus accesos invalidaría la estadística) y volver a modificarlas cuando ya no sea necesario realizar la auditoria.

Entre los inconvenientes es preciso destacar:

1) Instalación costosa

La implantación de un sistema de bases de datos puede llevar consigo un coste elevado, tanto en equipo físico (nuevas instalaciones o ampliaciones), como en el lógico (sistemas operativos, programas, compiladores, etc. necesarios para su uso), además del mismo coste de adquisición y mantenimiento del SGBD.

2) Personal especializado

Los conocimientos, que resultan imprescindibles para una utilización correcta y eficaz y sobre todo para el diseño y administración de las bases de datos, implican una necesidad de personal especializado.

3) Implantación larga y difícil

Debido a las causas apuntadas anteriormente, la implantación de una base de datos puede convertirse en una tarea larga y laboriosa. Las dificultades que van apareciendo a lo largo de su desarrollo llevan en general a que se superen ampliamente los plazos inicialmente previstos.

4) Falta de rentabilidad a corto plazo

La implantación de un sistema de bases de datos, tanto por su coste en personal y en equipos como por el tiempo que tarda en estar operativo, no resulta rentable a corto plazo, sino a medio o, incluso, a largo plazo.

5) Escasa estandarización

Un problema muy importante que se pone de manifiesto en el momento de la creación de una base de datos, es la falta de estandarización que facilite a los usuarios el manejo de los sistemas de bases de datos. Empieza, sin embargo, a observarse ya una preocupación por este tema, y han aparecido estándares, sobre todo en el campo de las bases de datos relacionales (SQL92 y posteriores, la última SQL 2016).

6) Desfase entre teoría y práctica

Al existir un considerable avance de la teoría en relación con la práctica, en muchas ocasiones los usuarios, especialmente los directivos, se engañan respecto a las prestaciones reales que pueden proporcionarles los SGBD actuales, creyendo que constituyen ya una realidad ciertos aspectos que todavía son sólo teóricos.

Existe un evidente riesgo de frustración por parte de los directivos y usuarios de sistemas de bases de datos que puede hacer olvidar las grandes posibilidades reales que estos sistemas nos ofrecen. Las bases de datos no constituyen únicamente una nueva tecnología, más o menos avanzada, pero tecnología al fin y al cabo, sino que nacen de una concepción distinta del SI., por lo que han de tener una influencia decisiva en las estructuras y organización de su entorno. Si esto no se tiene bien presente, muchas de las posibles ventajas de las bases de datos no se harán realidad y, en cambio, se acentuarán sus inconvenientes y problemas.

EJERCICIOS DE REPASO

1. Define, diferenciándolos, los conceptos de información factual e información referencial.
2. Pon algunos ejemplos de datos no estructurados.
3. ¿Piensas que un sistema de bases de datos debe ser utilizado para una sola aplicación manteniendo aquellas que ya están programadas y en funcionamiento sobre una estructura clásica de ficheros? Supón que las aplicaciones tienen relación entre sí.
4. Los primitivos sistemas de información eran “sistemas orientados al proceso”. ¿Cómo han evolucionado estos sistemas y por qué?
5. ¿Qué relación existe entre el sistema de información y la base de datos?
6. ¿De qué partes está compuesto un sistema de información mecanizado?
7. Pon ejemplos de al menos 3 Sistemas de Gestión de Bases de Datos y 1 Sistema de Recuperación de la Información que utilicemos habitualmente o que estén implantados en nuestra sociedad.
8. Asocia las siguientes características a los Sistemas de Información para la Gestión o a los S.I. para la toma de decisiones:
 - a. Tratamiento no secuencial.
 - b. Muchos formatos de salida distintos y estandarizados.
 - c. Gran cantidad de datos.
 - d. Trabajamos con pocos datos pero con interrelaciones muy complejas.
 - e. Procesamiento secuencial.
 - f. Están dirigidos a los directivos de las empresas.
 - g. La aplicación de los ordenadores a las empresas surgió con este tipo de S.I.
 - h. Son procesos simples y repetitivos.
9. ¿Cuál es la principal diferencia entre un fichero maestro y uno de constantes?
10. ¿Cuáles son las principales propiedades de un fichero secuencial? ¿Para qué tipo de ficheros piensas que son más útiles? Razona tu respuesta.

11. ¿Qué es un índice? ¿Qué utilidad tiene?
12. Elige un libro cualquiera. Míralo atentamente y, a continuación, indica al menos una característica que el libro tenga en común con:
 - a. Los archivos secuenciales.
 - b. Los archivos relativos.
 - c. Los archivos indexados.
13. Es posible que tengas un teléfono móvil y, en ese caso, seguro que dispones de una agenda con los números de amigos y personas conocidas. Atendiendo a la manera de acceder a los datos, ¿a qué tipo de fichero crees que se parece más? Razona la respuesta.
14. Cita al menos 3 ventajas de las Bases de Datos pero relacionándolas con la forma de trabajar de los antiguos sistemas de ficheros clásicos.
15. Escribe una definición intuitiva de Base de Datos, nombra asimismo los componentes que piensas que la forman.