



Bases de Datos

Tema 1: Conceptos básicos

Contenidos

- Contenidos
- La información
- Sistemas de ficheros
- Referencias

Objetivos

- Contenidos:
 - Conceptos básicos
 - Introducción a los conceptos básicos de una base de datos
 - Sistemas de ficheros tradicionales
 - Limitaciones y diferencias entre los sistemas de ficheros y los sistemas de bases de datos
 - Breve descripción de la evolución histórica de las bases de datos

La información

- El almacenamiento y gestión de la información ha sido una tarea humana desde tiempos inmemoriales
 - Algunas teorías apuntan a que las pinturas rupestres eran “apuntes de caza” claves para la supervivencia del ser humano mediante el consumo de carne (que reducía el tiempo dedicado a buscar alimento frente a la recolección)
 - Aclaración: con esto no animo a la gente a comer carne ni a dejar de hacerlo ;)
 - La escritura la impulsó a un nuevo *status*

La información

- Siglo II a.C. Sun Tzu escribe “El Arte de la Guerra”:
 - 4. *Por lo tanto, lo que permite al soberano sabio y el interés general de huelga y conquistar, y lograr cosas más allá del alcance de los los hombres ordinarios, es conocimiento previo*
 - 5. *Ahora conocimiento previo no puede ser suscité de los espíritus, no puede ser obtenidos inductivamente de la experiencia, ni por cualquier cálculo deductivo*
 - 6. *El conocimiento de las disposiciones del enemigo sólo se puede obtener de otros hombres*
 - <http://es.allreadable.com/0b534xgu>

La información

- La información es poder / knowledge is power. (Scientia potentia est)
 - Francis Bacon, Barón de Verulam (año 1597)
- La informática es un término que proviene de INFORMación automÁTICA
 - Tiene por objetivo el estudio del tratamiento de la información
 - El tratamiento de la información (automático o automatizado) abre nuevas vías de utilización
 - Matiz versus *computer science* anglosajón

Frase de Djisktra

- «La ciencia de la computación no trata sobre las computadoras más de lo que la astronomía trata sobre los telescopios»
 - <https://www.computer.org/web/awards/goode-edsger-dijkstra>

La información

- Hoy en día la información está en casi todos los aspectos de nuestra vida:
 - Antes de nacer ya están recolectando y almacenando información nuestra: medidas del feto, aumento del peso de la madre, etc.
 - Y la están comparando con información previamente almacenada: valores medios, umbrales que indiquen malformaciones, etc.
 - Y en cuanto nacemos, después de la palmada en el culete nos miden y pesan, y regularmente nos sacan muestras de sangre, etc

La información

- Otros ejemplos:
 - Sistemas “clásicos de gestión”
 - Información bancaria y financiera
 - Los seguros: automóvil, vivienda, etc.
 - Uso de carreteras, matriculaciones de vehículos, ...
 - Información multimedia: imágenes, sonidos, música, vídeos, etc.
 - Información que generan las cosas
 - IoT: Internet of Things
 - Cuadros de mando/monitorización y sistemas Inteligentes
 - Predicción de comportamientos, etc

La información

- Usos de la información:
 - Preservar conocimiento
 - Transferir conocimiento entre personas
 - Apoyo a la toma de decisiones
 - Individuales
 - Globales (avances sociales). Ej: se tardaron décadas en demostrar cómo el humo del tabaco generaba células cancerígenas, pero ya antes se sabía que lo hacía
- La información se concreta en datos
 - El dato es el antecedente necesario para conocer alguna cosa

La información

- Fases del tratamiento de la información:
 - Recopilación de datos
 - Almacenamiento
 - Tareas de mantenimiento
 - Redundancia, inconsistencias, curación, etc.
 - Procesamiento
 - Mantenerlos seguros (Bloques Wikileaks, blockchain, ¿backups en La Luna?)
 - Actualizarlos
 - Eliminarlos (derecho al olvido, políticas internas, LOPD, ...)

La información

- Características de la información, niveles de:
 - Precisión
 - ¿La información imprecisa no interesa? → Metrología
 - Oportunidad temporal, de localización, etc
 - Completitud: ¿faltan datos?
 - Significativa: (subjetivo) entendible y útil
 - Coherencia
 - Integridad
 - Grado de confidencialidad
 - Disponibilidad. Ej: Fundación Joaquín Díaz
 - https://commons.wikimedia.org/wiki/Commons:Fundación_Joaquín_Díaz
 - Seguridad (fuera del ámbito de la asignatura)

Sistema de información

- **Sistema de información:** conjunto de elementos, ordenadamente relacionados entre sí de acuerdo con unas ciertas reglas, que aporta al sistema objeto la información necesaria para el cumplimiento de sus fines, para lo cual tendrá que recoger, procesar y almacenar datos, procedentes tanto de la misma organización como de fuentes externas, facilitando la recuperación, elaboración y presentación de los mismos.

Sistema de información

- Características:
 - Tecnológicas
 - Funcionales y semánticas
 - Económicas
 - Sociales

SISTEMA DE INFORMACIÓN

Contenido
dato

Referencial

Factual

Estructurados

No estructurados

Equipo físico
hardware

Unidad central de proceso

Equipo periférico

Soporte lógico
software

Sistema operativo

Gestión de datos (SGBD)

Control de comunicaciones

Tratamientos específicos

Administrador

Área de datos

Área informática

Usuarios

Informáticos

No informáticos

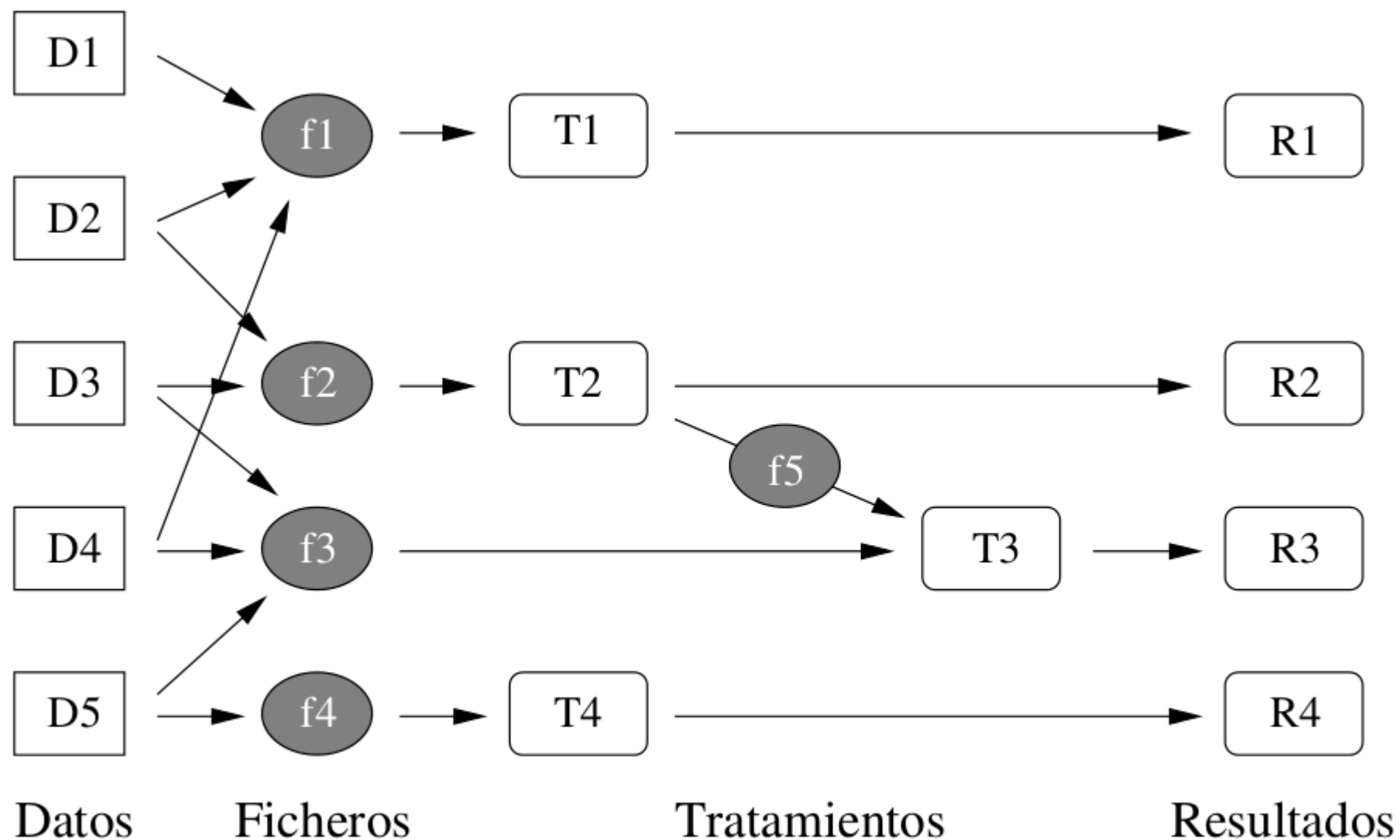
Componentes de un sistema de información

Sistemas de ficheros

- El sistema de información trabajado hasta este momento en el Grado en Ing. Informática es el fichero:
 - Conjunto de bloques físicos relacionados en los que se puede almacenar información
- Los sistema de ficheros (SF) facilitan el trabajo con ficheros individuales
- Sistema de gestión de ficheros: sistema que gestiona a un conjunto de ficheros específicos para una o varias aplicaciones

Sistemas de ficheros

- Los SF son la base de las aplicaciones de usuario final o administración de sistemas
 - Cada una con un fin (y desarrollo específico)
- Sistemas orientados hacia el proceso: los datos se almacenan en ficheros diseñados específicamente para cada aplicación
 - *En la Administración de Bases de Datos se llega a trabajar a nivel de fichero, pero se mantiene la capa de abstracción. Es decir, se sigue usando igual el sistema pero se consigue mayor rendimiento*



Organización clásica: sistema orientado al proceso

Sistemas de ficheros

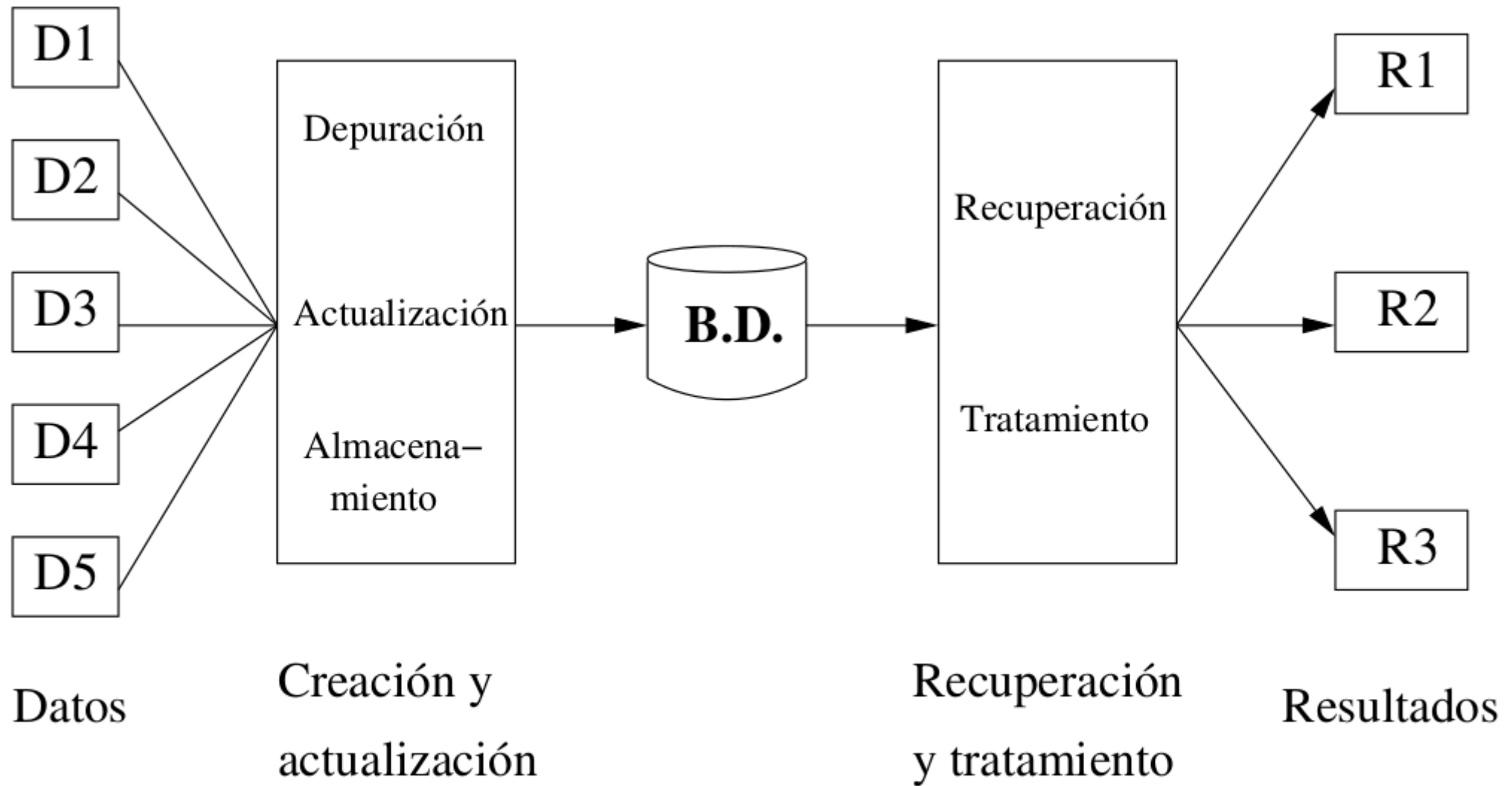
- Limitaciones de los sistemas de ficheros:
 - Dependencia del soporte físico
 - Dependencia de las aplicaciones específicas
 - Escasa flexibilidad y adaptabilidad frente a cambios
 - Bajo rendimiento del personal encargado de su desarrollo
 - Muy limitada interoperabilidad
 - Duplicación de datos: consumo de memoria, bajo rendimiento y posibles inconsistencias
- La “informática de gestión” se encontraba recurrentemente con estos problemas ...

Sistemas de ficheros

- Los sistemas de ficheros no son óptimos frente a:
 - Demandas inesperadas de información
 - Sistema orientados a la toma de decisión
- Necesidad de una gestión más racional de los datos, por lo que se requiere un nuevo enfoque: Sistemas de Bases de Datos.
 - Datos recogidos y almacenados una sola vez, e
 - Independientes de los tratamientos.
- Sistemas orientados a los datos: los datos se organizan y mantienen en un conjunto estructurado sin estar diseñados para una aplicación en concreto.
 - Satisfacen todas las necesidades de información de la empresa.

Sistemas de BD

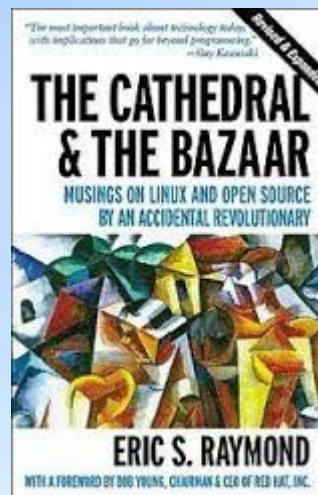
- Un Sistema de BD (SBD) es un sistema de mantenimiento de registros por ordenador, cuyo propósito general es registrar y mantener la información. Se compone de:
 - Los datos: información recopilada y almacenada que forma la BD
 - Software: conjunto de programas que mantendrán y actualizarán a los datos. Se denomina Sistema de Gestión de Bases de Datos (SGBD)
 - Hardware: componentes electrónicos que soportarán al software y el almacenamiento de los datos
 - Los usuarios: grupo de personas que interaccionan con el sistema de BD



Organización en BD: sistemas orientados a los datos

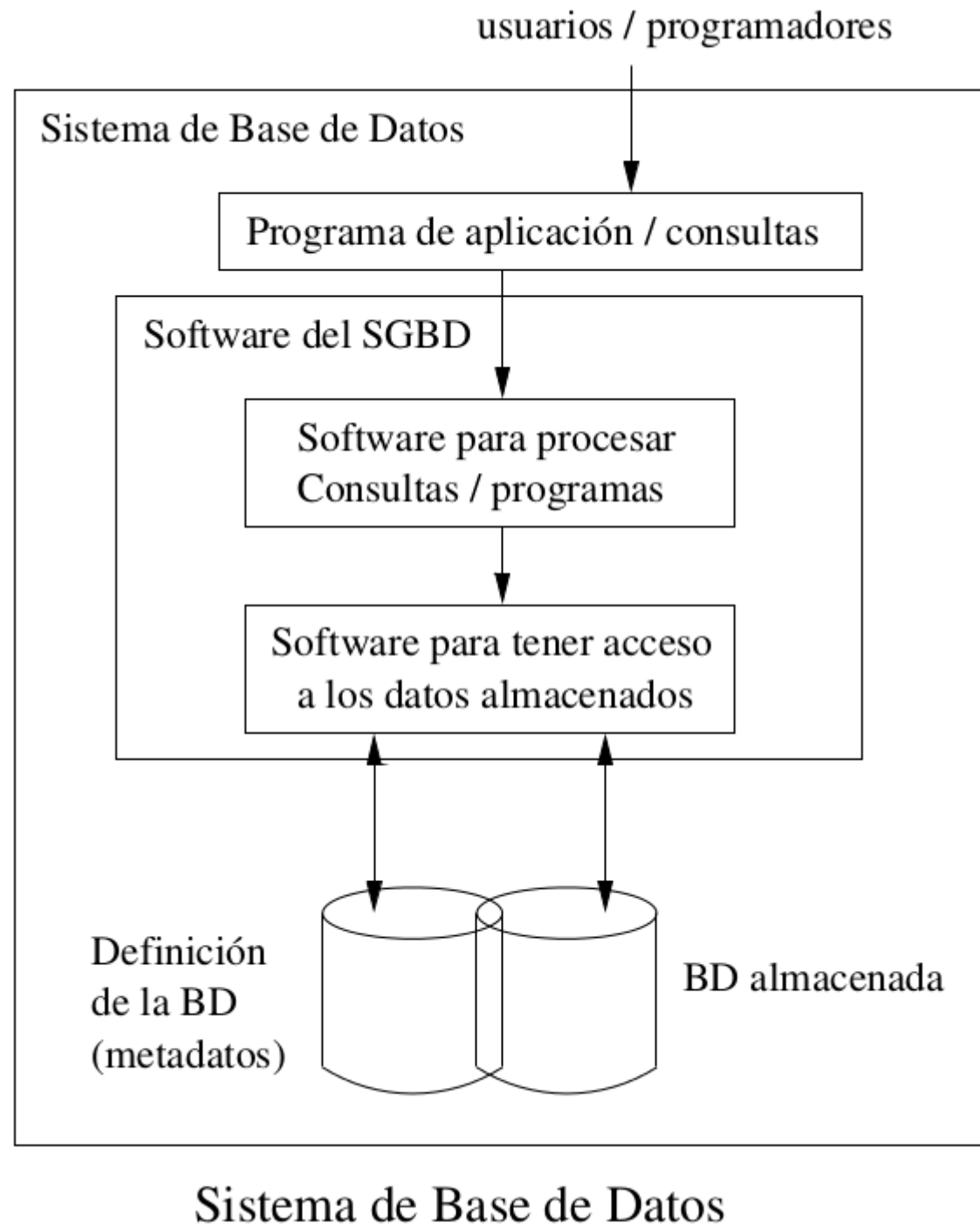
La importancia de los datos

- Eric S. Raymond
- Ensayo “La catedral y el bazar”
 - *Smart data structures and dumb code works better than the other way around*



Sistemas de BD

- Solución: las Bases de Datos, sistemas de información de soporte lógico (software)
 - Colección de datos interrelacionados almacenados en conjunto sin redundancias perjudiciales o innecesarias; su finalidad es servir a una aplicación o más, de la mejor manera posible; los datos se almacenan de modo que resulten independientes de los programas que los usan; se emplean métodos bien determinados para incluir nuevos datos y para modificar o extraer los datos almacenados. [Martin, 1975]



Sistemas de BD

- Ventajas de un SBD frente a un SF
 - En cuanto a los datos:
 - Independencia respecto al tratamiento que se les dé (y viceversa)
 - Mayor disponibilidad
 - Mayor eficiencia en la recogida, codificación e introducción en el sistema
 - En cuanto a resultados
 - Mayor coherencia
 - Información más normalizada
 - Información mejor documentada
 - Seguridad (*)

Sistemas de BD

- Ventajas de un SBD frente a un SF
 - En cuanto a los usuarios:
 - Acceso más rápido y sencillo a los datos
 - Facilidad para compartir datos
 - Adaptabilidad a cambios en requisitos
- Desventajas de un SBD frente a un SF:
 - Instalación costosa y (potencialmente) compleja
 - Necesidad de un personal cualificado para diseño, implantación, administración, ...
 - Dudosa rentabilidad a corto plazo o entornos peq.

Sistemas de BD

- SBD frente a un SF:
 - Pregunta: ¿cuántas PYMES de nuestro entorno creéis que almacenan datos en hojas de cálculo o en un programa X de una empresa concreta?
 - ¿Es eficiente?
 - ¿Es sostenible si la empresa crece?
 - ¿Qué pasa si el empleado “responsable” se da de baja?

Sistemas de BD

- SBD frente a un SF:
 - Pregunta: ¿cuántas PYMES de nuestro entorno creéis que almacenan datos en hojas de cálculo o en un programa X de una empresa concreta?
 - ¿Es eficiente?
 - ¿Es sostenible si la empresa crece?
 - ¿Qué pasa si el empleado “responsable” se da de baja?
 - Aún así, ¿les merecería la pena en todo caso poner un SBD?

Sistemas de BD

- Y si una empresa es pequeña pero piensa crecer nos pregunta si tira por ficheros/Excel o una BBDD, ¿qué le recomendamos?

Sistemas de BD

- Y si una empresa es pequeña pero piensa crecer nos pregunta si tira por ficheros/Excel o una BBDD, ¿qué le recomendamos?
 - Le diremos que esa decisión no es nuestra
 - Informaremos de ventajas e inconvenientes de una y otra alternativa
 - La gerencia debe decidir
 - Igual que con cualquier otro recurso. Por ejemplo: ¿dedica dinero a arreglar una oficina pequeña y económica (que no recupera si se muda a otra más grande) o se mete en la grande y cara directamente?

Historia de los SGBD

- 1880 Hollerith inventa un método para almacenar información en tarjetas perforadas
- 194X: la Segunda Guerra Mundial potencia el desarrollo de computadores
 - Procesos semi-automatizados hombre-máquina
- Década 1950: aparecen las cintas magnéticas
- 1960: concepto de Bases de Datos
- 1975: ANSI/X3/SPARC propone una arquitectura de tres niveles:
 - Externo, conceptual y interno

Historia de los SGBD

- 1970: Codd proponer las bases matemáticas para un nuevo modelo de datos, el relacional
- Década de 1970: aparece SEQUEL, que pasa a llamarse SQL
- 1980: Aparece Oracle, el primer SGBD
- 1990: Bases de Datos Orientadas a Objetos
- 2000: Bases de Datos Nosql, etc
- 201X: Siguen haciendo falta ficheros (CERN)

Sistemas de BD

- Para la implantación exitosa de SBD es clave la disponibilidad de Sistema de Gestión de Bases de Datos (SGBD) de calidad
 - Y el factor humano, claro está ;)
- Desde hace muchos años hay un desarrollo enorme de SGBD
 - Hasta la llegada de Internet había una “killer app”
 - Pero después otras se han sumado al carro
 - ¿Conocéis los nombres?

Sistemas de BD

- Oracle: SGBD por excelencia durante muchos años (época de la informática de gestión “dura”)
 - La mayoría de competidores que le salían sólo la mejoraban en precio y menores requisitos
- Con la llegada de Internet apareció un SGBD libre que lo mejoraba mucho en rendimiento (CGI) ante el uso masivo con pocos recursos: MySQL
 - ¿A cambio de qué mejoraba el rendimiento?
- El mercado se abrió a nuevas necesidades: inmediatez, interoperabilidad, sistemas distribuidos, etc: MongoDB, Neo4j, ...

Sistemas de BD

- Oracle vs MySQL vs MariaDB
 - El pez grande se come al chico
 - ¿Para qué?
 - ¿Igual intención que al adquirir OpenOffice vía Sun?
 - Julian Assange era colaborador de MySQL
 - Hay otros peces grandes (Google, Facebook, ...) alimentando a los chicos
 - Por varias razones
- El papel de postgresSQL

Tecnología concreta

- ¿Cuántos Sistemas Gestores de Bases de Datos existen en el mercado?

Tecnología concreta

- ¿Cuántos Sistemas Gestores de Bases de Datos existen en el mercado?
 - Database management system:
 - https://en.wikipedia.org/wiki/Category:Database_management_systems
 - 138 páginas + 24 subcategorías

Tecnología concreta

- Si existen decenas de SGBD es porque existen decenas de perfiles distintos de necesidades empresariales
 - Algunos se usan porque son “los mejores” para el campo concreto
 - Otros porque son “gratis”
 - Otros porque son “los únicos que van bien en mi sistema de gestión de X”
 - Otros porque son los más ligeros para un móvil
 - Otros ...
 - https://en.wikipedia.org/wiki/Category:Database_software_comparisons

Tecnología concreta

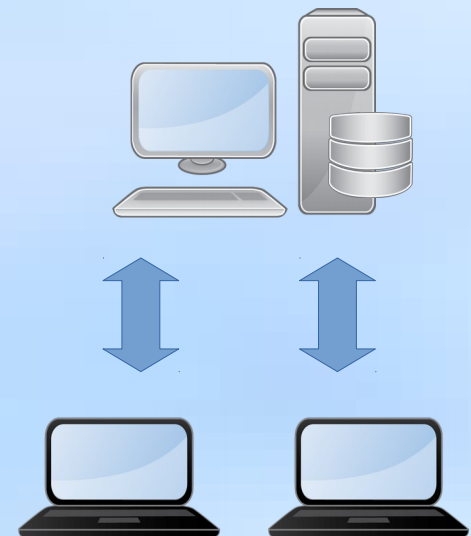
- DB-engines.com
 - <http://db-engines.com/en/ranking>

Tecnología concreta

- DZone: Most Popular Relational Databases (2016 Edition)
 - <https://dzone.com/articles/most-popular-relational-databases-2016-edition>

Arquitectura C/S

- Se suele usar en los SGBD serios
 - Servidor: mysqld
 - Cliente: mysql, phpmyadmin, u otros
- En entornos de producción el cliente y el servidor están en equipos distintos:



Arquitectura C/S

- Se suele usar en los SGBD serios
 - Servidor: mysqld
 - Cliente: mysql, phpmyadmin, u otros
- En entornos de producción el cliente y el servidor están en equipos distintos:
- Pero en clase los tendremos juntos:



La información

- Otros aspectos en el tratamiento de información:
 - Seguridad en acceso
 - Recuperación ante pérdidas
 - Aspectos legales ¿cookies?
 - Límites éticos
 - Ej: ¿robot de juguete que manda lo que registra su micrófono a un servidor?
 - Etc.

Cuestiones

- ¿Qué características de los SGBD encuentra más útiles? ¿por qué?
- ¿Y menos menos útiles?
- Indique 5 casos de la vida real en que crea que ha usado algún SBD
- Indique 5 aplicaciones informáticas en que crea que ha usado SBD
- ¿Realmente piensa que es necesario SQL? ¿O basta usar bibliotecas C/C++?
- Explique con ejemplos las diferencias entre la independencia de datos física y lógica.

Referencias

- Apuntes de bases de datos de la profesora Esther Gadeschi 2003/4
 - Apuntes de bases de datos del profesor Carlos Rioja 2006/7

Gracias por la atención
¿Preguntas?