



Bases de Datos

Tema 1: Conceptos básicos

Contexto

- Asignatura “Bases de datos”, Grado en Ingeniería Informática - UCA
- Competencias
 - CB2 Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio
 - C12 Conocimiento y aplicación de las características, funcionalidades y estructuras de las bases de datos, que permitan su adecuado uso, y el diseño y el análisis e implementación de aplicaciones basadas en ellos
 - CT1 Trabajo en equipo: capacidad de asumir las labores asignadas dentro de un equipo, así como de integrarse en él y trabajar de forma eficiente con el resto de sus integrantes
- Resultados del aprendizaje
 - R1 Conocer la importancia de la información y la necesidad de gestionarla de forma eficiente
 - R3 Capacidad para descubrir la necesidad de utilizar los sistemas de bases de datos y presentar las características diferenciadoras de los mismos
- Tiempo estimado: 2 horas

Contenidos

- Objetivos
- La información
- Sistemas de ficheros
- Referencias

Objetivos

- Contenidos:
 - La información
 - 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 clave desde tiempos inmemoriales
 - Teorías proponen 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 ;)
 - <https://pubmed.ncbi.nlm.nih.gov/25794684/>
 - La escritura la impulsó a un nuevo *status*



La información

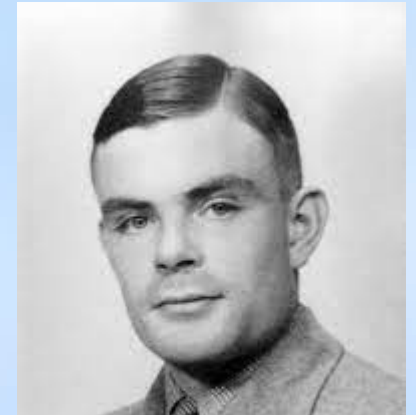


- Siglo II a.C. Sun Tzu escribe “El Arte de la Guerra”:
 - 4. *Por lo tanto, lo que permite al soberano sabio y el interés general de huelga y conquistar, y lograr cosas más allá del alcance de los los hombres ordinarios, es conocimiento previo*
 - 5. *Ahora conocimiento previo no puede ser suscitó de 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>
 - Posteriormente ... *El arte de la paz*

La información

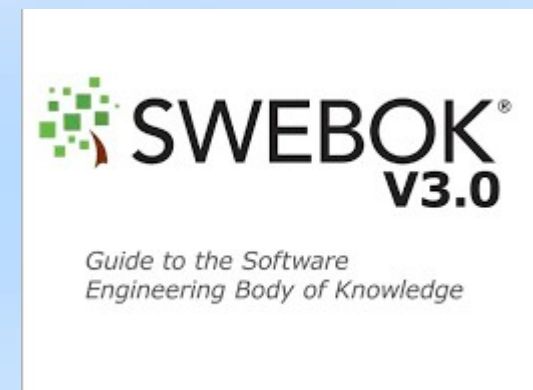
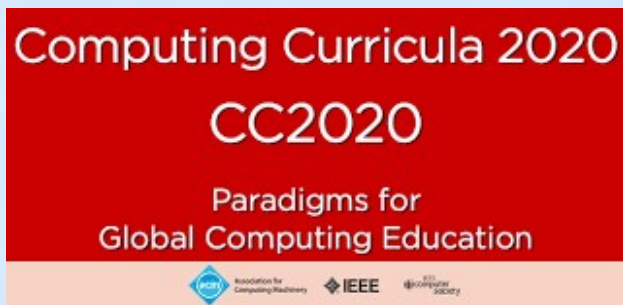


- Francis Bacon, Barón de Verulam (año 1597)
 - La información es poder / knowledge is power (Scientia potentia est)
- Decisiva en la WWII:
 - Máquina enigma
 - Turing (en Sevilla)
 - Mujeres en Bletchley Park
 - Criptonomicón



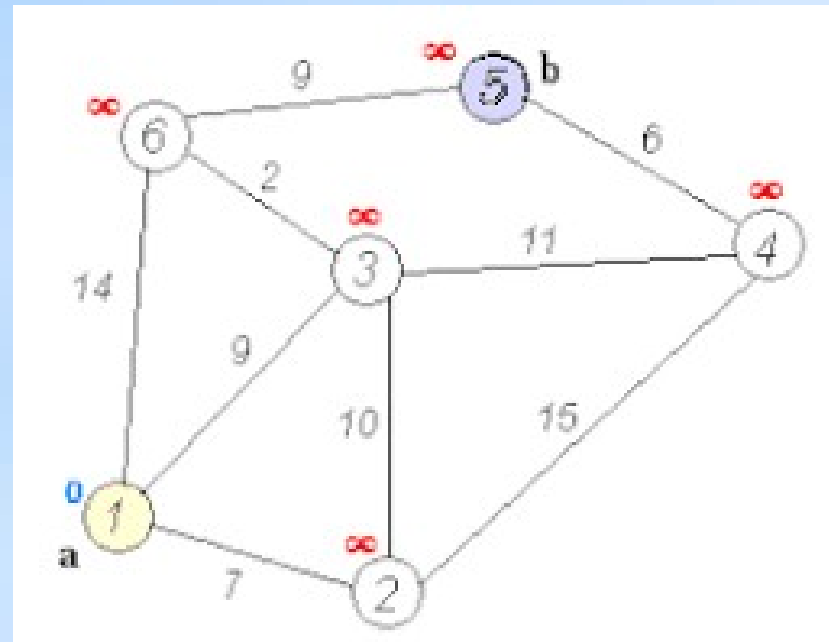
La información

- 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



Frases de Dijkstra

- «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

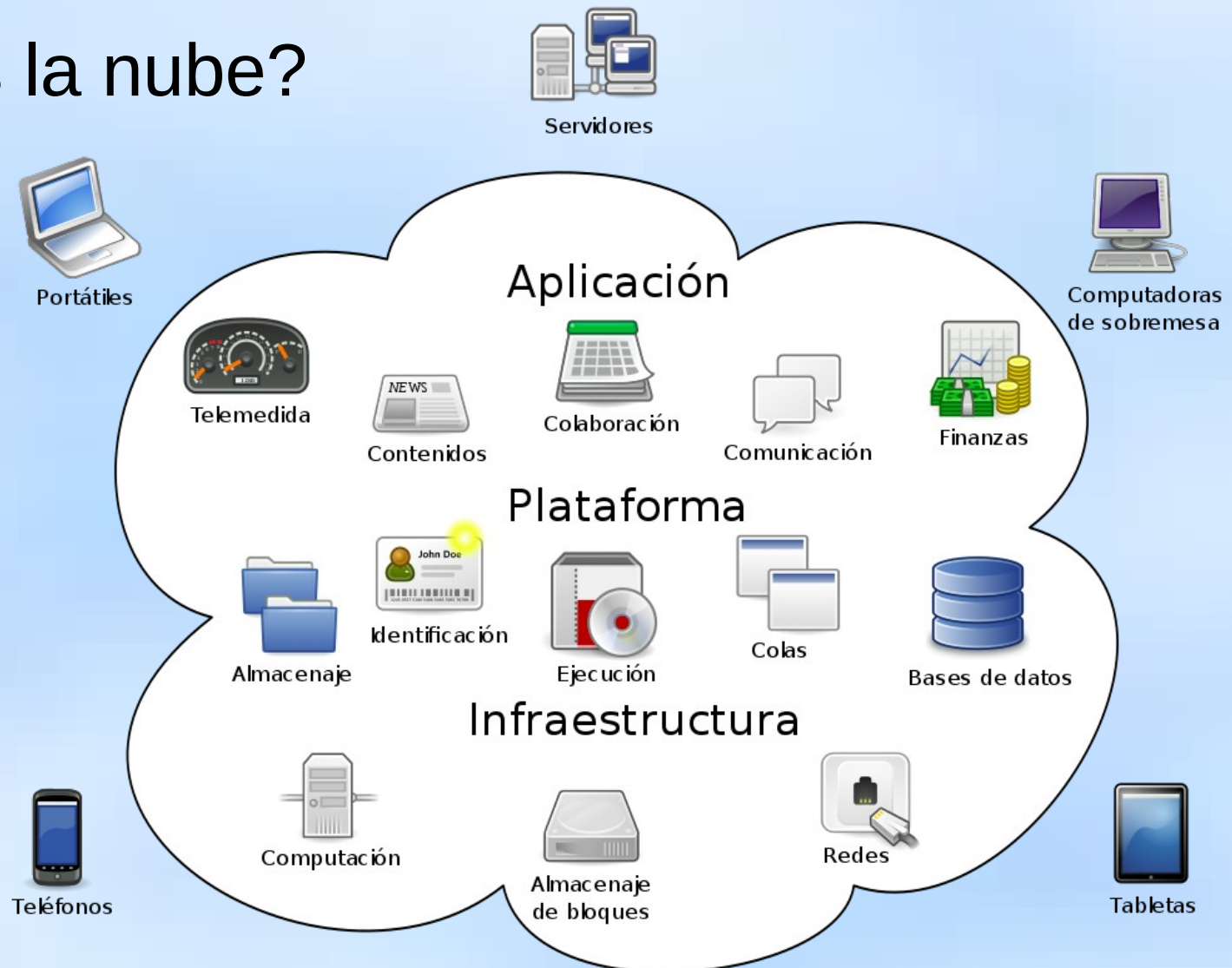
La información

- Otros ejemplos:
 - Sistemas “clásicos de gestión”
 - Información bancaria y financiera
 - Los seguros: automóvil, vivienda, etc.
 - Uso de carreteras, matriculaciones de vehículos, paseos de móviles por supermercados, etc.
 - 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

- ¿Qué es la nube?



Computación en la nube

There is NO CLOUD, just



other people's computers

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
 - “Si me traes datos hablaremos de los datos ... si me traes tu opinión te hablaré de la mía”
 - Richard Stallman: el aumento de capacidad de almacenamiento ...



La información

- La información se concreta en datos
 - El dato es el antecedente necesario para conocer alguna cosa
- 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/GDPR, ...)



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

S
I
S
T
E
M
A

D
E

I
N
F
O
R
M
A
C
I
Ó
N

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 almacenamiento

- ¿Qué sistemas software de almacenamiento de información habéis visto en el Grado hasta la fecha?
 - A.k.a. persistencia

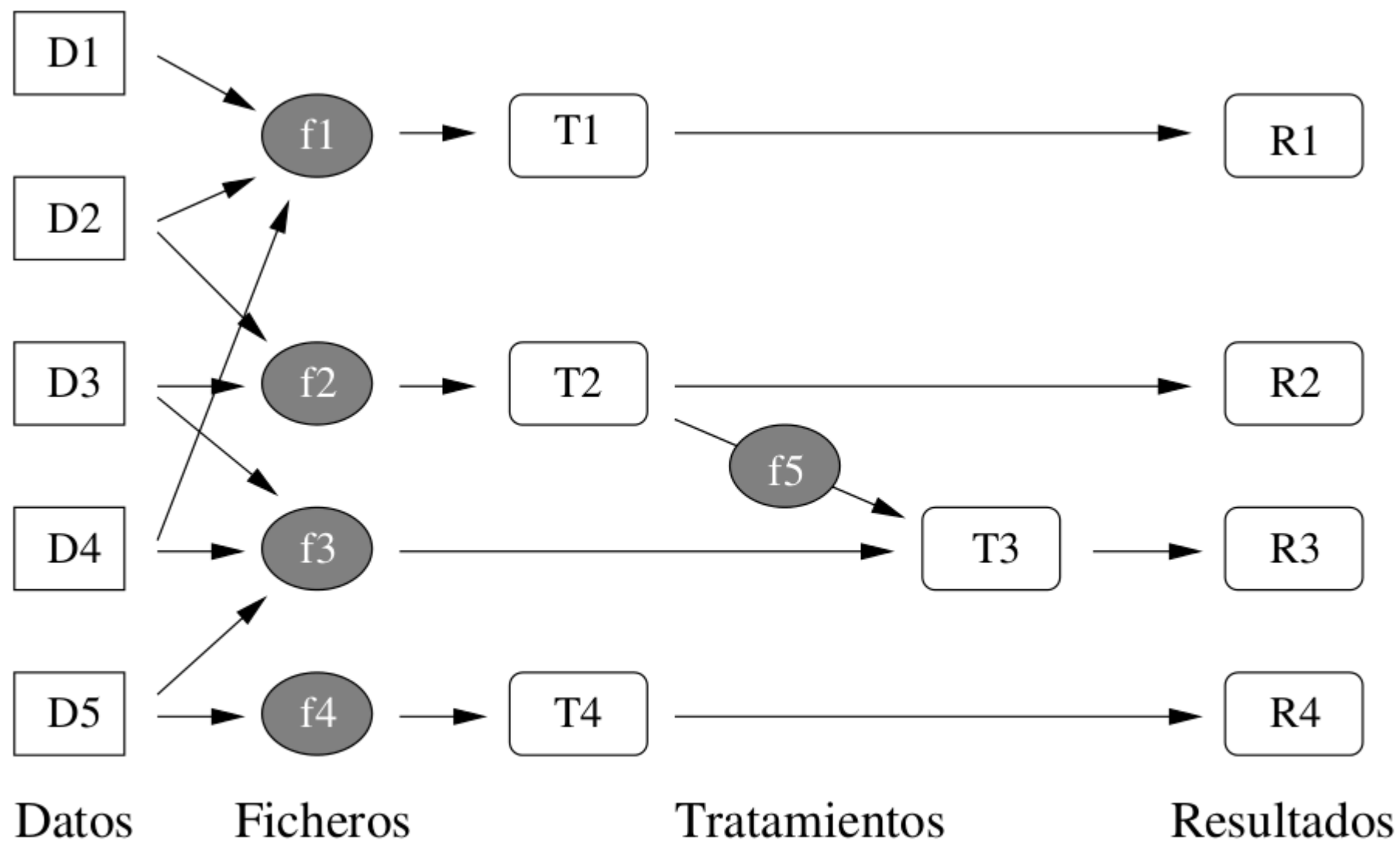


Sistemas de ficheros

- El sistema de almacenamiento trabajado hasta este momento en GII es el fichero:
 - Conjunto de bloques físicos relacionados en los que se puede almacenar información
 - Pueden ser textos, hojas de cálculo, ...
- 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 explotando ~igual el sistema pero se consigue mayor rendimiento*
- Es ~ similar a procesos industriales de fabricación transformando recursos



Organización clásica: sistema orientado al proceso

Sistemas de ficheros

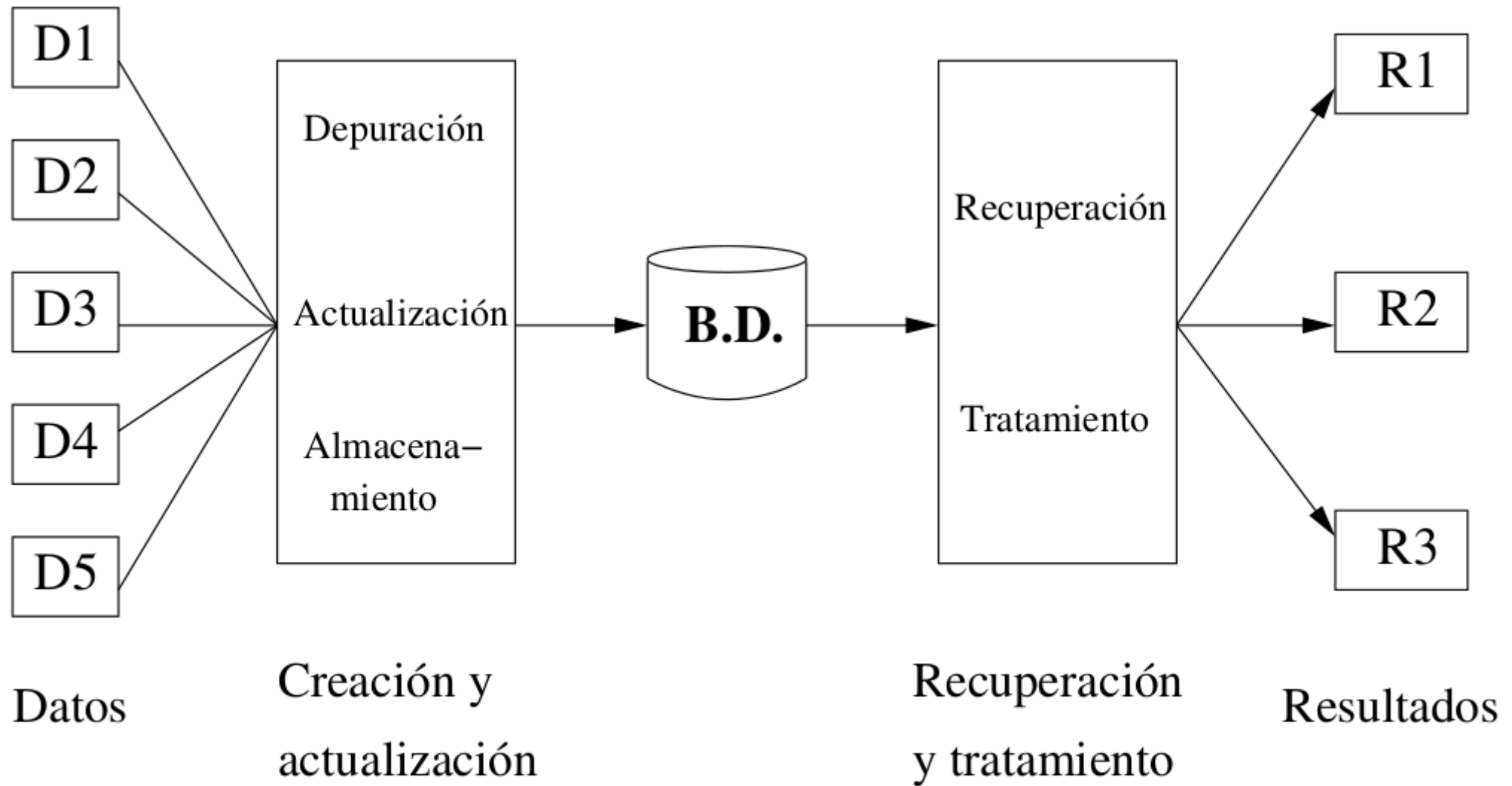
- Limitaciones de los sistemas de ficheros:
 - Dependencia del soporte físico
 - Dependencia de las aplicaciones específicas
 - Escasa flexibilidad y adaptabilidad frente a cambios
 - Bajo rendimiento del personal encargado de su desarrollo
 - Muy limitada interoperabilidad
 - Duplicación de datos: consumo 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
 - 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

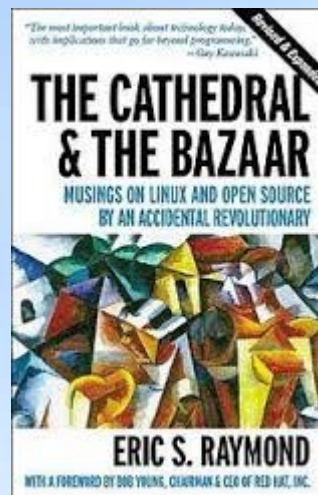
- Un Sistema de BD (SBD) es un sistema de mantenimiento de registros por ordenador, cuyo propósito general es registrar y mantener la información. Se compone de:
 - Los datos: información recopilada y almacenada que forma la BD
 - Software: conjunto de programas que mantendrán y actualizarán a los datos. 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

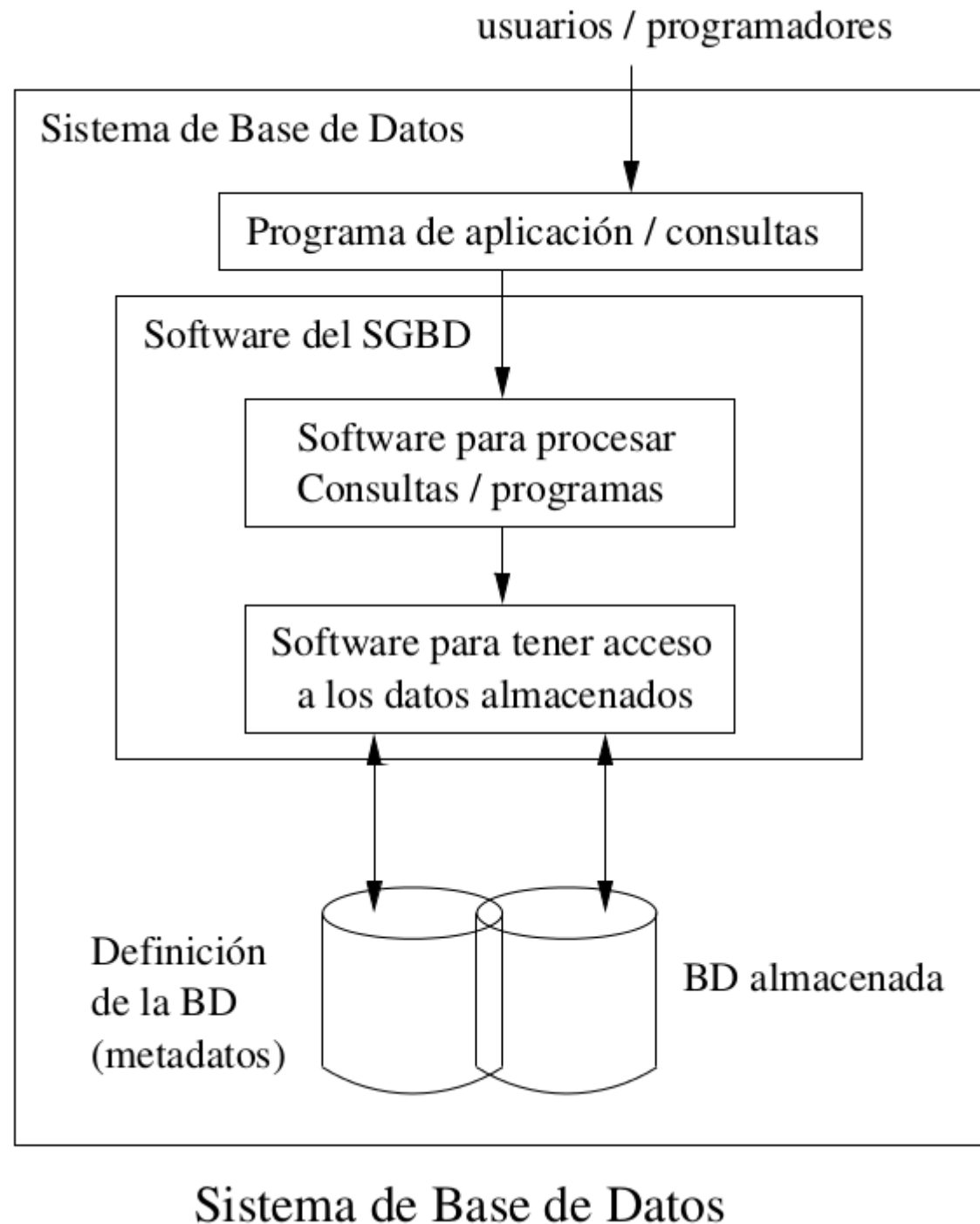
La importancia de los datos

- 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
 - De hecho, no se estudia en la asignatura
 - Necesidad de un personal cualificado para diseño, implantación, administración, ...
 - Dudosa rentabilidad a corto plazo o entornos peq.

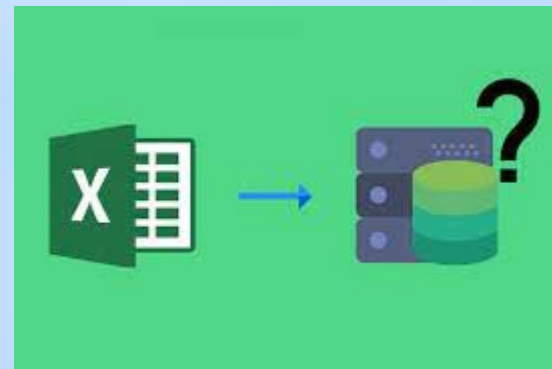
Sistemas de BD

- 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?



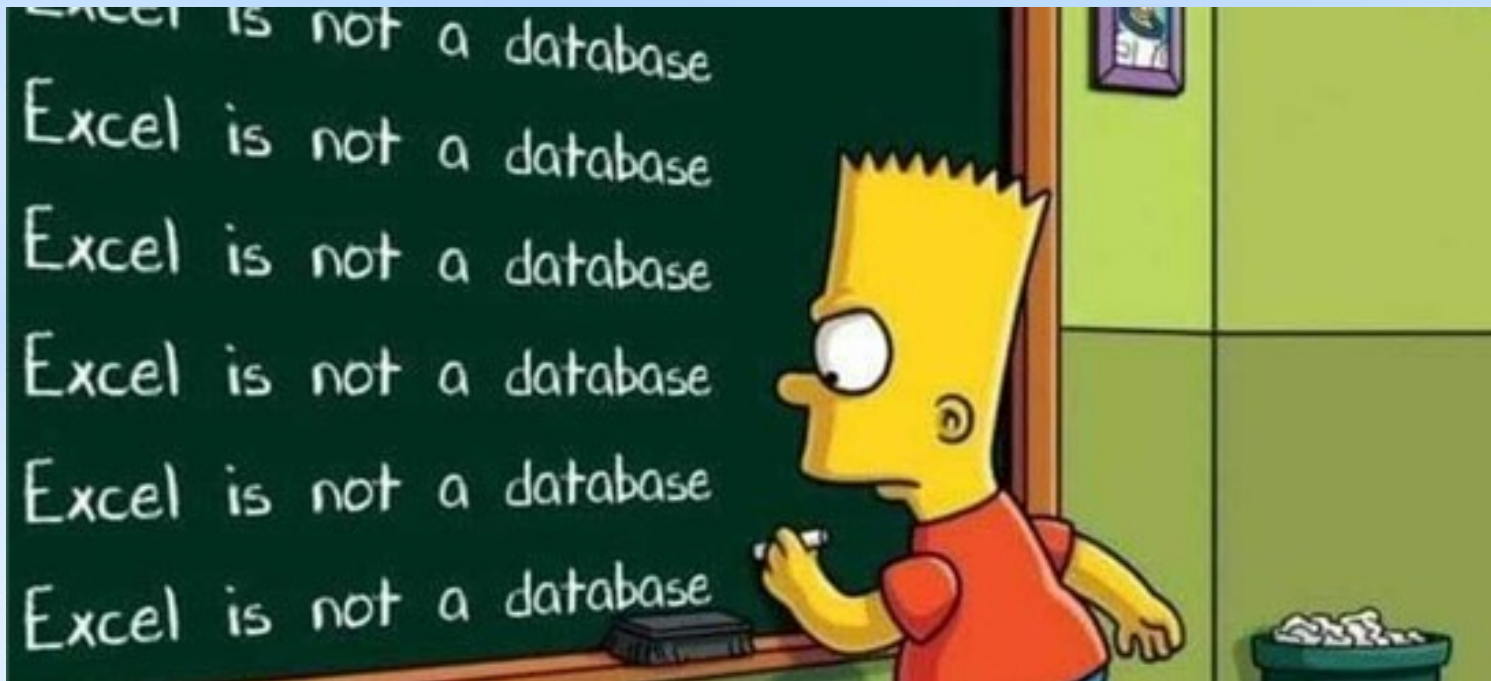
Sistemas de BD

- 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 que piensa crecer nos pregunta si tira por ficheros/Excel o una BBDD, ¿qué le recomendamos?

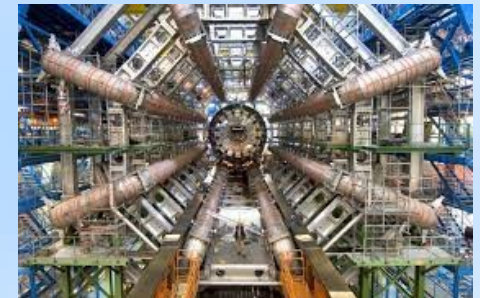
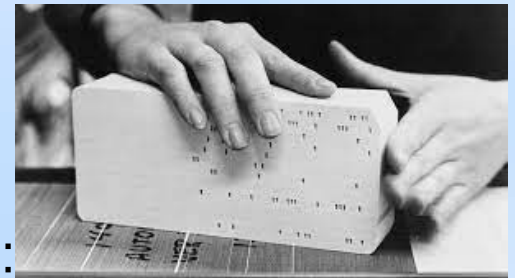


Sistemas de BD

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

Historia de los SGBD

- 1880 Hollerith: almacenamiento de 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 arquitectura tres niveles:
 - Externo, conceptual e interno
- 1970: Codd proponer las bases matemáticas para el modelo de datos 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
- 202X: Siguen haciendo falta ficheros (Arduino vs 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 el Internet masivo había una “killer app”
 - Pero después otras se han sumado al carro
 - ¿Conocéis nombres?



Sistemas de BD

- Vamos a hacer una prueba del Taller de Moodle
 - Cada alumno tiene 10 minutos para buscar 3 sistemas gestores de Bases de Datos
 - Por favor, ponedle un poquito de ganas e intentad no coger los 3 primeros de Google/Wikipedia
 - Después se subirán al taller Moodle correspondiente
 - Se evaluarán las subidas de compañeros
 - Se podrán consultar los resultados

Sistemas de BD

The Oracle logo, consisting of the word "ORACLE" in white, uppercase, sans-serif font, set against a solid red rectangular background.

- 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
- Access se abrió un cierto mercado en PCs
- 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
 - Aparece MariaDB, un fork de MySQL
 - Hay otros peces grandes (Google, Facebook, ...) alimentando a los chicos
 - Por diversas razones
- La vía 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
 - 152 páginas + 25 subcategorías (en 2021)



Tecnología concreta

- Si existen decenas de SGBD es porque existen decenas de perfiles distintos de necesidades
 - Algunos se usan porque son “los mejores” para el campo concreto
 - Otros porque son “gratis”
 - Otros porque son “los únicos que van bien con 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

- Vale, pero ¿se usan todos igual?
 - DB-engines.com
 - <http://db-engines.com/en/ranking>
 - DZone: Most Popular Relational Databases (2016 Edition)
 - <https://dzone.com/articles/most-popular-relational-databases-2016-edition>
 - The Most Popular Database (TOPDB Index) 2004 – 2019
 - <https://www.youtube.com/watch?v=6DrTKf8Vwps>
 - Stack Overflow 2022 Survey
 - <https://survey.stackoverflow.co/2022/#section-most-popular-technologies-databases>
 - Techradar
 - <https://www.thoughtworks.com/es-es/radar>

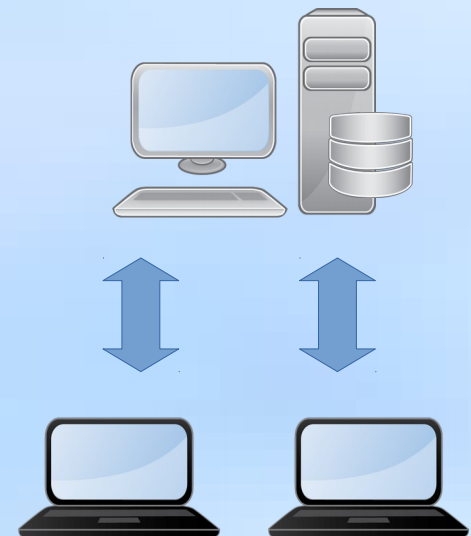
Fijaos desde hace cuánto tiempo se pide SQL y Oracle en las ofertas de empleo :-O

Tecnología concreta

- No siempre la tecnología más potente es la más adecuada
 - ¿Matamos moscas a cañonazos?
 - ¿Tenemos hardware que lo soporte bien?
 - ¿Podemos pagar los sueldos de un DBA experto?
 - ¿Está formado el equipo de desarrollo?
 - Ej:
 - <https://www.theguardian.com/info/2018/nov/30/bye-bye-mongo-hello-postgres>
 - <https://twitter.com/juanmacias/status/1357688686830321666?s=20>

Arquitectura C/S

- Se suele usar en los SGBD serios. Ej MySQL:
 - 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. Ej MySQL:
 - Servidor: mysqld
 - Cliente: mysql, phpmyadmin, u otros
- En entornos de producción el cliente y el servidor están en equipos distintos
- Pero en desarrollo se pueden tener juntos:



La información

- Otros aspectos en el tratamiento de información:
 - Seguridad en acceso
 - Recuperación ante pérdidas
 - Aspectos legales ¿cookies?
 - Límites éticos
 - Ej: ¿dónde van las grabaciones del micrófono de un robot de juguete?
 - https://elpais.com/tecnologia/2016/12/12/actualidad/1481533650_459698.html
 - Grabaciones de altavoces
 - <https://unaaldia.hispasec.com/2018/12/amazon-envia-por-error-1-700-grabaciones-de-alexa-de-otro-cliente.html>
 - Estás embarazada ... y no lo sabes ...
 - 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?