

# Introducción a las Bases de Datos

Mariano Beiró

Dpto. de Computación - Facultad de Ingeniería (UBA)

23 de agosto de 2022

# Temas

- 1 Bases de datos
- 2 Sistemas de Gestión de Bases de Datos
- 3 Arquitectura de 3 capas ANSI/SPARC
- 4 Funciones de los SGBDs
- 5 Bibliografía

# 1 Bases de datos

## 2 Sistemas de Gestión de Bases de Datos

## 3 Arquitectura de 3 capas ANSI/SPARC

## 4 Funciones de los SGBDs

## 5 Bibliografía

# Definición de base de datos

## Base de datos

Una base de datos es un conjunto de datos interrelacionados.

Pero, ¿qué es para nosotros un dato?

## Dato

Un dato es un hecho que puede ser representado y almacenado de alguna forma, y que tiene un sentido implícito.

Ejemplos:

- La mesa 5 consumió 2 milanesas napolitanas y 1 botella de vino.
- Alberto Malatesta tiene turno para oftalmología el 15/03.
- El Nissan GT-R alcanza una velocidad máxima de 315 km/h.
- 100 gramos de chocolate poseen 546 calorías.

# Bases de datos tradicionales vs. no tradicionales

## Bases de datos tradicionales. Predicados

Las bases de datos tradicionales almacenan datos de texto o numéricos, que pueden enunciarse a través de *proposiciones*.

Por ejemplo:

- Juan Martín Del Potro ganó el Abierto de Estados Unidos en 2009.
- Gabriela Sabatini ganó el Abierto de Estados Unidos en 1990.
- Roger Federer ganó el Abierto de Australia en 2018.
- ...
- *[persona]* ganó *[torneo]* en *[año]*

Un conjunto de proposiciones que tienen la misma estructura puede tipificarse a través de un *predicado*.

## Predicado

Es una función que toma uno o más argumentos y devuelve un valor de verdad.

En el caso anterior, podemos definir la función *GanadorDelGrandSlam(persona, torneo, año)*. Entonces:

- $\text{GanadorDelGrandSlam}(\text{Roger Federer}, \text{Abierto de Australia}, 2018) = V$
- $\text{GanadorDelGrandSlam}(\text{Marin Čilić}, \text{Abierto de Australia}, 2018) = F$

Atención: Las bases de datos sólo almacenan proposiciones verdaderas.

Lectura sugerida: "What a database really is: Predicates and propositions", H. Darwen, Warwick University, 1994.

<http://www.dcs.warwick.ac.uk/~hugh/M359/What-a-Database-Really-Is.pdf>

# Bases de datos tradicionales vs. no tradicionales

## Bases de datos no tradicionales

Actualmente, las bases de datos también almacenan tipos de datos más complejos como imágenes, audio, video, o datos geoespaciales.

En este curso centraremos nuestra atención en las bases de datos tradicionales.

1 Bases de datos

2 Sistemas de Gestión de Bases de Datos

3 Arquitectura de 3 capas ANSI/SPARC

4 Funciones de los SGBDs

5 Bibliografía



# Sistemas de Gestión de Bases de Datos

## Un poco de historia

[SILB19 1.9]

### ¿Cómo organizar el acceso a los datos?

Hasta los años '50, la forma más común de almacenamiento de datos eran los *ficheros manuales*.



Title: <b>ALICE IN WONDERLAND</b>		#C-5	TUNE-DEX (Trade Mark) Reg. U. S. Pat. Off.	
Performing Rights licensed through	<b>ASCAP</b>	Date of this issue		<b>May 1951</b>
Writers:	<b>W: Bob Hilliard</b>			
	<b>M: Sammy Fain</b>			
Copyright				<b>1951 by Walt Disney Productions - Copyright assigned to Walt Disney Music Company, R.K.G. Bldg., *</b>
Dance ork arr by		Price		
Vocal orks in (keys)		Price		
Piano copies in (keys)	<b>Eb</b>	Price		<b>35¢</b>
Other information	<b>*Radio City, New York, N.Y.</b>			
<b>From Walt Disney's "ALICE IN WONDERLAND"</b>				

Int'l copyr't secured. All rights reserved, inc. the right of arrangement & public performance for profit.  
 Warning! This copy is intended for PROFESSIONAL use ONLY, and anyone found reselling it will be prosecuted under the copyright law by the copyright owners.  
 Litho'd by Tune-Dex, Inc. 1619 B'way, N. Y. 19, N. Y. by permission of the copyright owners.  
 Copyright 1951 George Goodwin, 1619 Broadway, New York City Litho'd in U.S.A.



# Sistemas de Gestión de Bases de Datos

## Un poco de historia



- El procesamiento de las cintas (ordenamiento, *merge*, búsqueda) se realizaba típicamente en COBOL.
- Sin embargo, las cintas imponían fuertes restricciones en cuanto a la forma serial de acceso para lectura y escritura, ya que sólo permitían *acceso secuencial*.

# Sistemas de Gestión de Bases de Datos

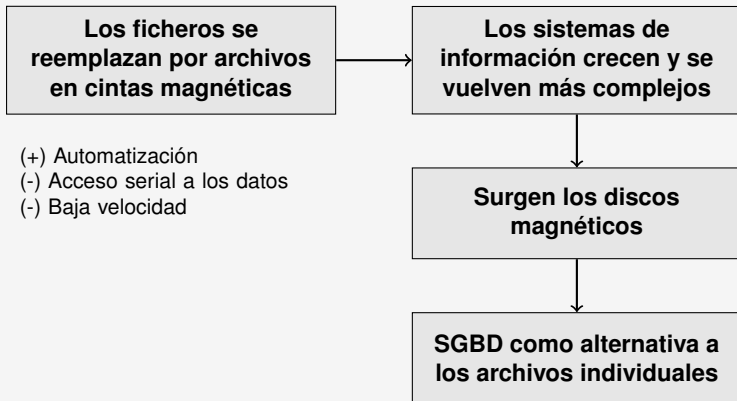
## Un poco de historia

- En los **años '60** surgen los discos magnéticos, y con la posibilidad de *acceso directo* a los datos aparecen los primeros SGBD's:
  - El **IMS** (Information Management System) de IBM, que utilizaba un *modelo de datos jerárquico*.
  - El **IDS** (Integrated Data Store) de General Electric, que usaba un *modelo de datos en red*.
- Si bien el IMS fue superado por los SGBDs relacionales en cuanto a sus limitaciones, aún se encuentra presente en numerosos mainframes, ATM's, y bancos.

# Sistemas de Gestión de Bases de Datos

## Resumen

(-) Se hace evidente que la relación directa entre programas y datos es una gran desventaja



# Sistemas de Gestión de Bases de Datos

## Definición

¿Qué es entonces un Sistema de Gestión de Bases de Datos (SGBD) o *Database Management System (DBMS)*?

### Sistema de Gestión de Bases de Datos (SGBD)

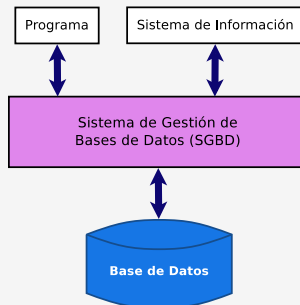
Es un conjunto de programas que gestiona y controla la creación, manipulación y acceso a la base de datos.

El SGBD provee un nivel de abstracción entre los programas o sistemas de información y los datos, resolviendo el problema conocido como *dependencia de datos*.

# Sistemas de Gestión de Bases de Datos

## Independencia de datos

[ELM16 2.2]



## Independencia de datos

Es la propiedad del SGBD consistente en que cambios en la estructura de la base de datos no repercutan en los programas o sistemas de información que la utilizan.

# Sistemas de Gestión de Bases de Datos

## Algunos hitos

[ELM16 1.7]

- 1964 • General Electric crea el IDS (Information Data Store), un SGBD con un *modelo de red*.
- 1966 • IBM crea el IMS (Information Management System), basado en un *modelo jerárquico*.
- 1970 • Codd –trabajando en IBM– propone el *modelo relacional*, que se impondrá en la década siguiente.
- 1975 • El ANSI/SPARC propone una *arquitectura de 3 capas* para los SGBDs.
- 1976 • Chen propone el modelo Entidad-Interrelación que dominará el diseño conceptual de bases de datos.
- 1980 • Surgen SGBDs comerciales basados en el modelo relacional, como DB2 (IBM), y Oracle.
- 1990s • Surgen los primeros SGBDs orientados a objetos.
- 1992 • Se estandariza el lenguaje SQL (estándar SQL-92).
- 1993 • Surgen los Data Warehouses y OLAP (Online Analytical Processing) como sistemas de soporte para la toma de decisiones.
- 1998 • Se estandariza XML como lenguaje de intercambio de datos.
- 2000 • Los ERPs (Enterprise Resource Planning) aparecen en escena como una capa de negocio por sobre el SGBD.  
*Ejemplos:* SAP ERP, Oracle ERP.
- 2005 • De la mano del *cloud computing* se crean SGBDs distribuidos de alta performance.  
*Ejemplos:* BigTable (Google, 2005) y Cassandra (Facebook, 2008) → Movimiento NoSQL.
- 2010 • Surge el concepto de Database-as-a-Service.  
*Ejemplos:* Microsoft Azure SQL (2010), Google Cloud Datastore (2013), Google Cloud SQL(2014).



1 Bases de datos

2 Sistemas de Gestión de Bases de Datos

3 Arquitectura de 3 capas ANSI/SPARC

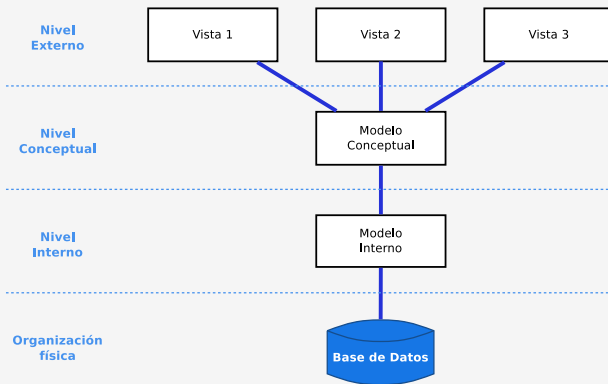
4 Funciones de los SGBDs

5 Bibliografía

# Arquitectura de 3 capas ANSI/SPARC

[ELM16 2.2]

- El ANSI-SPARC propuso en 1975 una arquitectura en 3 niveles de abstracción para la descripción/representación de los datos de una base de datos.



# Arquitectura de 3 capas ANSI/SPARC

- **Modelo interno:** Representa la forma en que los datos se almacenan utilizando estructuras de datos y organizaciones de archivos. Representa cómo perciben los datos el sistema operativo y el SGBD.
- **Modelo conceptual:** Describe la semántica de los datos, abstrayéndose de su implementación física. Describe entidades, tipos de datos, operaciones y restricciones de seguridad y de integridad.
- **Modelo externo:** Representa la forma en que los usuarios perciben los datos.

Esta arquitectura asegura la *independencia de datos*, tanto física como lógica.

- 1 Bases de datos
- 2 Sistemas de Gestión de Bases de Datos
- 3 Arquitectura de 3 capas ANSI/SPARC
- 4 Funciones de los SGBDs**
- 5 Bibliografía

# Funciones de los SGBDs

[ELM16 1.6, SILB19 1.2]

- Almacenamiento y Consulta
  - Ofrecer estructuras eficientes.
  - Ofrecer un lenguaje de consulta (aumenta la productividad).
- Integridad
  - Asegurar la integridad de datos a través de restricciones.
- Seguridad
  - Evitar accesos no autorizados.
- Concurrencia
  - Permitir el acceso en simultáneo de muchos usuarios.
- Recuperación
  - Ofrecer herramientas para la recuperación ante fallas.
- Soporte transaccional

- 1 Bases de datos
- 2 Sistemas de Gestión de Bases de Datos
- 3 Arquitectura de 3 capas ANSI/SPARC
- 4 Funciones de los SGBDs
- 5 Bibliografía**

# Bibliografía

[ELM16] Fundamentals of Database Systems, 7th Edition.

R. Elmasri, S. Navathe, 2016.

Capítulo 1, Capítulo 2.2

Será nuestra referencia principal.

[SILB19] Database System Concepts, 7th Edition.

A. Silberschatz, H. Korth, S. Sudarshan, 2019.

Capítulo 1

[GM09] Database Systems, The Complete Book, 2nd Edition.

H. García-Molina, J. Ullman, J. Widom, 2009.

Capítulo 1.1

[CONN15] Database Systems, a Practical Approach to Design, Implementation and Management, 6th Edition.

T. Connolly, C. Begg, 2015.

Capítulo 1, Capítulo 2.1, 2.3